universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Decomposing the Life Expectancy at Birth Racial Gap between Whites and Mixed & Blacks in São Paulo, Brazil

verfasst von / submitted by

Matheus Lobo Alves Ferreira

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 945
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Global Demography
Betreut von / Supervisor:	Hon.-Prof. Dipl.-Geogr. Dr. Marc Luy, Privatdoz.

Acknowledgments

Words cannot express my gratitude to all those who provided me with the possibility to complete this thesis. I would like to express my deepest appreciation to my academic advisor, Hon.-Prof. Dipl.-Geogr. Dr. Marc Luy, your feedback, patience, and support were fundamental throughout this journey.

I am also grateful to Professor Dr. Bernardo Lanza Queiroz. Thank you for encouraging me to pursue this degree and for all the advice you so generously gave me.

A special thanks goes to Christoph and Érika, who gave me my first home when I arrived in Vienna almost two years ago. Thank you for everything, my time here would not have been the same without your friendship.

I also want to extend my heartfelt thanks to my friends, the ones who are far away and those here in Vienna. Your unwavering support, patience, understanding, and encouragement – especially during the last weeks of writing have been invaluable.

Lastly, I would be remiss in not mentioning my family, who stood by me, unconditionally, from the day I told them I was moving to Vienna. Writing this thesis wouldn't have been possible without your support – *eu amo vocês*.

Table of Contents

List of Abbreviations

Abbreviation	Definition
CRVS	Civil Registration and Vital Statistics
CVD	Cardiovascular Diseases
DATASUS	Department of Informatics of the Unified Health System
DDM	Death Distribution Methods
GGB	Generalized Growth Balance
IBGE	Brazil's Statistical Office
ICD	International Classification of Diseases
OHCHR	United Nations Human Rights Office of the High Commissioner
PNAD-C	Comprehensive sample survey conducted quarterly by IBGE
SEG	Synthetic Extinct Generations
SIM	System of Mortality Information
SUS	Unified Health System

List of Figures

List of Tables

1 Introduction

According to a recent survey by institute PEREGUM and SETA about racism in Brazil, 81% of Brazilians believe that their country is racist. Furthermore, 44% of the respondents consider race to be the primary factor creating inequalities in Brazil. Other factors such as socioeconomic status (20%), place of residence (7%), and gender (6%) were also mentioned. These figures are significant and highlight the importance of addressing race and racism in the ongoing debate about inequalities in Brazil. Similarly, race and racism are also central issues within American society (Travassos & Williams, 2004); in the United States, there is a significant body of literature that investigates a wide range of dimensions from a racialized perspective. From a demographic point of view, research on health and mortality differentials by race stands out (Abedi et al., 2021; Dwyer-Lindgren et al., 2022; Johnson et al., 2022; Nelson, 2002).

Unfortunately, the same is invalid for Brazil, where demographic literature on the topic is far more incipient. One obstacle to advancing research on this topic is that IBGE, Brazil's statistical office, does not publish life tables by race. Moreover, racial classification is a complex process and, in Brazil, has been described as ambiguous (Telles, 2002, 2014). Nevertheless, studies that attempted to provide evidence of racial disparities in health and mortality show a systematic disadvantage experienced by Non-Whites compared to Whites (Cunha et al., 2016; Goes et al., 2020; Lotufo et al., 2007; Martins et al., 1991; Matijasevich et al., 2008; Paixão et al., 2010; Reichenheim et al., 2011; Wong et al., 2014).

It is vital to emphasize yet another challenge faced when conducting demographic analysis in Brazil – the country's regional diversity, which presents varying levels and patterns of mortality. The 2019 estimates of life expectancy at birth for females and males combined illustrate this diversity very clearly; while the national average was 76.6 years, life expectancy at birth was 71.40 years in *Maranhão*, a state in the Northeast and 79.90 in *Santa Catarina*, a state in the South. Data quality regarding population estimates and death counts also varies across regions; research shows that despite significant improvements, the South and Southeast regions, which have higher socioeconomic development, have better-quality data compared to the North, Northeast, and Midwest (Agostinho & Queiroz, 2016; Lima & Queiroz, 2014; Paes, 2005; Paes & Albuquerque, 1999; Queiroz et al., 2017, 2020). More recently, the COVID-19 pandemic added a new layer of complexity: mortality levels increased significantly since 2020 (Baptista et al., 2022; Castro et al., 2021; Lima et al., 2021).

Using population estimates and death counts for the 2010 to 2019 period in the state of São Paulo, Brazil, the main goal of this thesis is to measure the life expectancy at birth race gap for females and males separately; moreover, it also aims to evaluate at which ages do mortality conditions vary the most. Secondary research questions are necessary to achieve this goal. The first deals with analyzing the completeness of death counts by race, which will be evaluated using death distribution methods. The second part of the analysis involves using different life table methods to decide which approach yields the most reliable life expectancy at birth by race estimates.

In other terms, the main research question to be answered in this thesis is:

- Do mortality conditions, measured by life expectancy at birth, differ across racial groups in São Paulo between 2010 and 2019? If so, what are the age groups contributing the most to it?

Furthermore, the additional questions associated with the thesis' main objective are:

- Are death counts subjected to different completeness levels across racial groups?
- Do different life table methods yield different life expectancy at birth by race estimates?

I hypothesize that life expectancy at birth differs among races, with Non-White people facing a disadvantage regardless of gender. Additionally, I anticipate that adult mortality will have a more significant impact on male deaths than on female deaths, primarily due to external causes and violence-related deaths. Since this analysis is limited to São Paulo, where death counts are complete (Lima & Queiroz, 2014; Queiroz et al., 2017, 2020), I do not expect completeness to vary when evaluated for race-based population subgroups. Lastly, I expect different methods to estimate life expectancy at birth to yield consistent results.

This thesis is comprised of five chapters following this introduction. Chapter 2 reviews the literature on the mortality transition, racial differences in health and mortality, and the Brazilian racial classification system. Chapter 3 presents the data on population and mortality, highlighting their main features and discussing limitations that could impact life expectancy at birth estimates. Chapter 4 focuses on the methodological aspects of this research, with a detailed presentation of the methods chosen. Chapter 5 presents the main results obtained. Finally, in Chapter 6, the results are discussed, and research opportunities emerging from this thesis are presented.

2 Literature Review

The literature review presents an in-depth discussion of mortality from a racial perspective in Brazil. It starts by tracing the historical changes in mortality and morbidity patterns in the country and introduces the concept of epidemiologic polarization in relation to the demographic and epidemiologic transition theories and their updates. The review also explores the definition of race beyond genetic factors and its potential links to health and mortality. Additionally, it offers a more nuanced examination of racial issues in Brazil's and discusses racial classification methods.

2.1 The Mortality Transition

Between the end of the 19th and the beginning of the 20th century, the average length of life in Brazil was approximately 33.7 years. Over the next five decades, little increases in life expectancy were observed, an increase of roughly three years, reaching 36.5 years by 1930 (Souza 1978, as cited in da Silva Simões, 2016). However, estimates from the 1940 census started to show the effects of reduced mortality levels, with life expectancy reaching 41.5 years; since then, steady and accelerated increases in survivorship have been observed. In the following five decades, life expectancy increased an additional 25.4 years, 8.5 times the gains observed between 1890 and 1930. Official estimates for 2019 are of 76.6 years of life.

The mortality transition is the primary process behind the increase in life expectancy observed since the middle of the 20th century in Brazil. To better understand the drivers of said transition, one must look at the decline of mortality *levels* and the changes in its *pattern* (Horiuchi, 1997); the changes in the level of mortality are explained within the demographic transition theory framework (Davis, 1945; Notestein, 1945; Thompson, 1929), while the field of epidemiologic transition explains changes on the pattern of mortality (Omran, 1971).

Originally based on observations for England between the 18th and the early 20th centuries, the demographic transition describes the historical shift from high and unstable fertility and mortality rates to a new pattern characterized by low and stable rates experienced by populations worldwide. The classic formulation of the demographic transition theory is structured under four distinct stages, each describing changes in fertility and mortality regimes and their effect on the population size and structure.

According to its classical proposition, populations in the first stage of the demographic transition experience high and fluctuating fertility and mortality rates, characterizing a state of equilibrium without significant changes in the population size or structure. However, mortality would surpass fertility due to fluctuations, leading to a temporary decrease in population size. The transitional phase, known as the second stage of the demographic transition, starts with a decline in mortality levels accompanied by a sustained high fertility regime; this combination drives an increase in the total population size. The following stage starts when population growth reaches its peak; at this point, fertility levels start to decline rapidly,

and while the population continues to grow, it does so at a slower rate. The final stage is a new equilibrium between fertility and mortality, which have reached historically low and stable levels, changing the population's age structure, characterized by fewer births combined with a growing share of the elderly population.

The demographic transition theory offers a clear framework to understand historical changes in fertility and mortality and its implications on populations. Nevertheless, it paints only some of the picture required to understand the drivers that led mortality to fall; to gain a more comprehensive understanding of the mortality transition, it is necessary also to consider the role changes in the cause-of-death profile. In the epidemiologic transition theory, introduced by Omran (1971), the author formulates that the observed increase in life expectancy results from the replacement of communicable diseases by degenerative causes of death.

According to Omran's 1971 article, this process took place in three different stages: (1) the age of pestilence and famine, when death rates fluctuated at high levels, and the major causes of deaths affected the young; (2) the age of receding pandemics, when improvements in public health and standards of living contributed significantly to reduce mortality amongst the young, which led to increases in life expectancy at birth, and (3) the age of degenerative and man-made diseases, characterized by a plateau phase, with individuals dying at older ages, primarily due to chronic degenerative diseases (Omran, 1971).

The author has also described three distinct trajectories of the epidemiologic transition; a trajectory describes a linear and unidirectional path through the different stages of the epidemiologic transition. The first is called the classical (or Western) model – the transition pattern followed by Western European nations fits this model; the move from the age of pestilence and famine to the current era of degenerative diseases took place in a steady manner, and by the beginning of the 20th century it was complete. The second trajectory – the accelerated model- is reasonably similar to the classical model, with the notable difference that it happened much faster. Finally, the contemporary (or delayed) model describes an incipient incomplete process (Omran, 1971).

The original framework of the epidemiologic transition was, in some ways, a product of its time; it was criticized and modified through the years by different authors (Foz, 2021). The revisions to the theory include the addition of a fourth stage that captures the reductions in mortality brought by the cardiovascular revolution (Olshansky & Ault, 1986) and a complete review of the epidemiologic transition trajectories by Omran himself in 1998, now split into four groups: (1) the western model, (2) the non-western model, (3) the intermediate model, and (4) the slow model (Omran, 1998). The timing and speed at which countries go through each stage of the transition distinguish one trajectory from another; however, more is needed to account for the experience of developing countries. Frenk et al. (1991), when studying the epidemiologic transition in Latin America, argue that the main problem with Omran's framework is the assumption that countries would go through this process in a linear and unidirectional manner. The authors introduce a new

epidemiologic transition trajectory – the polarized and prolonged model; the key features of this model are (1) the overlap of stages, (2) a countertransition, (3) a prolonged transition, and (4) epidemiologic polarization.

The authors suggest that Brazil is experiencing a countertransition, where communicable, chronic, and degenerative diseases are coexisting simultaneously. This is evidenced by the reintroduction of diseases such as *dengue* and cholera and the resurgence of malaria, leprosy, and leishmaniasis. The countertransition nature of this phenomenon means that it is not unidirectional. This situation results in high morbidity and mortality for both patterns, with no clear resolution. The epidemiologic situations in different regions or amongst social groups within the same country become contrasting (epidemiologic polarization) (Schramm et al., 2004). Epidemiologic polarization is also present in research conducted by Vallin & Meslé (2004); the authors describe a divergence-converge system related to health improvements. Generally, changes in health improvements first benefit the more privileged segments of a given population to the detriment of the poorer segments (divergence); later, when advancements are disseminated and become accessible to the entire population, there is a convergence process.

2.2 Racial Differences in Health and Mortality

Science, while a powerful tool for understanding and improving the world, has unfortunately played a significant role in the perpetuation of racism throughout history. In recent editorials, The Lancet (2022) and Nature (2023) acknowledged their roles in using science to justify racist practices and beliefs. Moreover, The Lancet describes racism as a global public health crisis and emphasizes the need for continued research on the topic as the beginning of the path towards reducing inequalities in racial health outcomes; complementarily, Nature calls for a cautionary approach when attributing outcomes solely to race or ethnicity, highlighting the potential for inaccuracy and harm. The journal then offers authors dealing with the topic a three-step guideline; authors should: (1) explain the reasons behind using race in the study; (2) discuss the treatment given to other dimensions of social stratification that might act as confounding variables the analysis, such as SES; and (3) describe the racial classification methods used.

From a biological perspective, race is not a self-evident explanation for differences in health and mortality (Santos, 2011). Genetic heterogeneity does not align with socially defined racial categories; genetic diversity is more commonly observed within individuals than across said groups (Williams et al., 2016). However, as a social stratification dimension, race can explain health and mortality differences due to unequal access to resources and exposure to risk factors. Race must be understood as a risk marker rather than a risk factor (Travassos & Williams, 2004); race should be viewed as an indicator or signal rather than a direct cause.

The discussion about differences in health outcomes and mortality among racial groups is of particular relevance in Brazil, as one cannot neglect the relevance of the country's complex historical context of colonization and slavery. Brazil was the leading destination of slave trade ships between the 16th and 19th centuries: approximately 45% of the total number of Africans trafficked in the period were shipped to the country and forced into slavery during that time, which accounts for about 4.8 million people; such contingent is comparable to the total amount of Africans disembarked in the British Caribbean (2.3 million), the Spanish Americas (1.3 million), and the French Caribbean (1.1 million) (*SlaveVoyages*, 2008).

Centuries of colonial rule and the legacy of the transatlantic slave trade have left a deep-rooted mark on Brazil's society. Santos (2011) argues that race is a crucial dimension of the country's social stratification; according to the author, historically, white groups have predominantly wielded the power to make racial designations, structure social life in racial terms, and associate an "inferior" value with other categories. Moreover, he argues that these divisions have been used to interpret experiences, form social relationships, and organize individual and collective action. Thus, racial divisions constitute social groups typically defined as distinct based on physical characteristics considered common among their members.

Socioeconomic status works as an umbrella term that encompasses various economic and social factors, such as wealth, influence, and status, and is usually measured by income, education, occupation, or a combination of these variables (Travassos & Williams, 2004). SES impacts health outcomes differently throughout the life span and across different scales (Braveman et al., 2005). Research shows that individuals with higher income, educational attainment, or occupational standing tend to enjoy longer lives and experience reduced disease prevalence compared to those with lower social status (Williams et al., 1994); moreover, there is a strong racial pattern to SES variables (Williams et al., 2016). A 2022 report discussing racial inequalities in Brazil shows that the average per capita household income income fow Whites is twice that of Non-Whites (IBGE, 2022).

However, this racial pattern in SES serves only as a partial explanation for differences in health outcomes. Racial disparities in health typically persist, although at reduced levels, across all levels of SES (Santos, 2011; Travassos & Williams, 2004; Williams et al., 1994, 2016). The same report on racial inequalities in Brazil shows that income levels are higher for whites regardless of the educational attainment level and are the highest for the group with college-level education (IBGE, 2022). Thus, despite the relationship between race and SES, these are not interchangeable variables and represent their own stratification processes; each is likely to serve as a proxy for particular exposures influencing health outcomes. (Santos, 2011; Travassos & Williams, 2004; Williams et al., 2016).

Finally, Williams et al. (2016) offer four reasons why race matters regardless of its relationship to SES. Firstly, early life adversity affects long-term physical and mental health, which has been shown to vary both by race and SES. Secondly, SES indicators are not equivalent across racial groups. Thirdly, racism is a crucial social exposure for racial minorities, impacting various health outcomes. Lastly, racial minorities face a higher risk of exposure to crime, violence, and financial strain.

Previous Studies in Brazil

Despite the relevance of approaching health and mortality from a racialized perspective, a limited amount of studies investigated such differences (Chor & Lima, 2005); Brazilian literature on the topic focuses mainly on mortality differences between Whites and Non-Whites, comprised of Whites on one side and Mixed and Blacks on the other, for specific groups such as infants, children, and young adults. Only a few studies attempted to estimate life expectancy by racial groups, and a search of the relevant literature yielded no articles presenting an age decomposition of differences in life expectancy at birth by gender and race.

Batista et al. (2004) conducted a study to examine if specific causes of death are more commonly associated with certain racial groups. They used death records data from the state of São Paulo and divided racial categories into Whites and Non-Whites. The study revealed two distinct patterns. Non-White deaths are due to complications during labor and delivery, external causes, mental disorders, or ill-defined reasons. However, On the other hand, White deaths are attributed to neoplasms and diseases of the circulatory or respiratory systems.

The work by Batista et al. emphasizes key aspects regarding the dynamics between gender, race, and health and mortality; moreover, his results are consistent with the literature on the topic. Research provides evidence supporting systematic disadvantages for Non-Whites when compared to their White counterparts in infant and child mortality, maternal mortality, and stroke mortality; Non-Whites are also more likely to suffer from the negative effects of violence in particular males.

Matijasevich et al. (2008) used 1982, 1993, and 2004 birth cohorts from the South of Brazil to study racial disparities in infant mortality; the authors evaluated mortality trends among infants born to White and Non-White mothers. Results showed that infant mortality declined regardless of the mother's race: for infants born to White mothers, the rate dropped from 30.4/1,000 live births in 1982 to 19.9/1,000 in 2004 – approximately 54%; a smaller, but still significant decrease took place for infants born to Non-White mothers, 44% (53.8/1,000 in 1982 to 30.4/1,000 in 2004). It is essential to notice that despite the decrease observed for both groups, it took more than twenty years for mortality rates of infants born to Non-White mothers to reach the same level experienced by those born to White mothers. Moreover, the gap between infant mortality rates across racial groups increased between 1982 and 2004; while said rates were 1.8 times higher for children born to Non-White mothers in 1982, it rose to 2.2 by 2004.

Using census and vital statistics data, Wong et al. (2014) assessed trends in infant mortality by gender and race in Brazil. The results presented by the authors support the results described by Matijasevich et al.. Infant mortality decreased for females and males born to White and Non-White mothers. However, mortality for Non-White male infants is about 30% higher when compared to their White counterparts; a similar pattern is observed for females, with Non-White girls experiencing rates 25% higher. The comparison between genders paints a more positive outlook: infant mortality decreased more significantly for males, thus reducing the infant mortality gender gap.

Associated with infant mortality are the health conditions of mothers; different studies approached this topic in a variety of ways, but the evidence unequivocally points to worse health status and higher mortality for Non-White mothers. A study on induced abortion among low-income women showed that Non-White women are three times more likely to die due to complications during an abortion than White women (Martins et al., 1991). This increased mortality due to complications after abortion can be explained by racial discrimination in health services and the stigma associated with abortion, as shown by Goes et al. (2020).

Leal et al. (2005) also studied the effects of racial discrimination in health services in Brazil. In theory, universal access to free delivery care is guaranteed to all Brazilians; however, to test the universality of access to such services across racial groups, the authors studied trends in access and utilization of health care services by White and Non-White women. Their results show that Non-White women reported the lowest educational attainment levels, combined with the highest rates of adolescent childbirth and unemployment. Moreover, domestic violence, smoking, abortion rates, and increased challenges to being admitted to hospitals were also more frequent among non-white females. When women of similar educational attainment levels were compared, the same pattern persisted but at lower levels.

These studies serve as supporting evidence for racial inequalities in the access and use of health services by adult women, which poses challenges for their own lives but also for infants, in particular for those born to Non-White mothers, as previously shown. While these are the main challenges associated with adult female mortality, the critical dimension to be considered when discussing adult male mortality is deaths due to external causes – or violence-related deaths. Reichenheim et al. (2011) claim that violence ranks third among the causes of death for males; moreover, it is the sixth leading cause of hospital admissions. Furthermore, Non-White men with low SES are both the primary victims and perpetrators of violence, whereas Non-White women with low SES are the primary victims of domestic violence.

Recent data on deaths by homicide confirms the pattern of differences by gender and race: in 2020, the homicide rate per 1,000 people in Brazil was 11.5 for Whites, 21.9 for Mixed, and 34.1 for Blacks – a significant increase when compared to what was experienced by White people. When the statistics are further disaggregated and presented also by gender, the differences become even more striking: homicide rates for Mixed and Black males are two times higher than that for White males. When comparing by

gender, rates for White males are eight times higher than that of their female counterparts, while the same comparison for mixed and black males yields differences as high as fifteen times (IBGE, 2022).

It is worth noting that there are documented racial disparities in cardiovascular diseases (CVD) in the United States, as highlighted by Graham (2015). However, further investigations are needed to explore this in Brazil. Lotufo et al.'s study in São Paulo (Lotufo et al., 2007) analyzed the patterns of CVD among white and non-white males and females to assess potential inequalities in deaths related to race and gender. The results indicated that non-white men had 1.4 times higher age-adjusted mortality rates than white men, while women exhibited similar patterns, although with lower rates.

The evidence presented so far indicates a systematic disadvantage of Mixed & Blacks over Whites, with increased mortality rates at all age groups analyzed. Previous studies that attempted to estimate life expectancy in Brazil by racial groups show that this increased mortality reflects on life expectancy estimates. Cunha et al. (2016) applied a wide range of different methods in their estimation and regardless of differences in the size of the effect race had in life expectancy at birth estimates, White females, White males, Mixed & Black females, and Mixed & Black males were, for all but one method, ranked in this order.

2.3 Race and Skin Color: Classifying Individuals

"The statistical representation of race relations depends on classification, and this needs to be understood in Brazilian terms" (Telles, 2014).

There is no international standard for racial or ethnic classification: in different countries, race and ethnicity are defined by how the historical context of each society shapes the distinctions among social groups (Osório, 2003). Telles (2014) describes racial classification in Brazil as an ambiguous matter. The author argues that after slavery was abolished, towards the end of the 19th century, a project of whitening the population through miscegenation was put in place, which left classification to a matter of individual perception and established a dynamic where individuals would be classified as white or whiter than others, rather than having membership to a particular group assigned to them, in contrast to countries like the United States or South Africa, where race-based laws required highly specified classification systems.

Another fundamental aspect of racial and ethnical classification in Brazil is that race and color (skin color) carry the same meaning; that is, the idea of pertaining to a particular racial group is associated with one's phenotype: skin color, hair type, nose, and lip shape are characteristics related to membership to a racial group. This association is evident in the census, where individuals are asked: "what is your color or race?" (*a sua cor ou raça é?*), which demonstrates the perception of some racial groups as a part of a color gradient, from lighter-skinned individuals to darker-skinned ones, with some groups shading into one another (Telles, 2014).

Understanding the ambiguity surrounding racial classification in Brazil, from the post-abolition period to the present day, is fundamental to discussing racial classification as a system and how its two main components operate: the categories themselves and the criteria used to establish membership to one category over another.

Race Categories: The Census and The Black Movement Effect

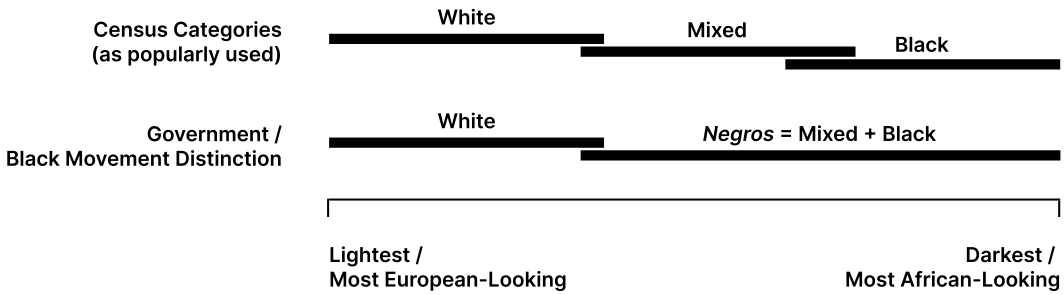
White (Branca), Mixed (Parda), Indigenous (Indígena), Asian (Amarela), and Black (Preta) are the five categories currently employed by IBGE in the census and in surveys that collect data on race. The story behind the terms used to describe racial categories dates back to the 19th century; at the time, White, Mixed, and Black were used to distinct free women and men from enslaved people (Osório, 2003). The first census in Brazil, held in 1872, adopted these three terms, together with *Caboclo* – a designation for the Indigenous population, as its racial groups.

The treatment given to racial categories in official statistics in Brazil changed over time. Despite including information on the population's racial composition in the first and second censuses (1872 and 1890, respectively), this information was ignored in the 1900 and 1920 censuses until it was reintroduced in the 1940s round. That year, the same categories used in the 1872 collection were adopted, and the Asian category was included as a result of international immigration flows, particularly those from Japan between 1908 and 1930 (Miranda, 2015; Osório, 2003). Since then, the only change happened in 1991, when the term Indígena replaced Caboclo to describe the indigenous population.

It is noteworthy how stable racial categories used by IBGE have been since its first round; apart from the inclusion of Amarela (Asian) and the replacement of Caboclo by Indígena, the remaining categories were, to some extent, always present. However, the views on race evolved significantly since the first census round. In particular, the emergence of the Black movement in Brazil reclaimed the term Negro, seeking to form a new Black identity where no distinction would be made between Mixed and Blacks (Telles, 2014).

Government reports, the media, and even academics adopt the concept of a racial group composed of Mixed and Blacks together; it is not uncommon to see articles in the media highlighting that more than half of the Brazilian population is *Negra*. However, IBGE has yet to reduce its racial categories from five to four. On the one hand, those in favor of the change argue that unifying the Mixed and Black categories will reduce much of today's ambiguity. On the other hand, those against it state that this change would represent the government, rendering invisible all Mixed individuals who would not consider themselves either White or Black (Telles, 2014).

Figure 2.3.1 – Color Gradient of Brazilian Racial Categories



Source: Adapted from Telles (2014).

Figure 2.3.1, adapted from Telles (2014), shows where the White, Mixed, Black, and Negro racial groups stand on a color gradient, moving from lighter-skinned (more European-looking) individuals on the left to darker-skinned ones (more African-looking) on the right. The categories adopted on the census and sample surveys conducted by IBGE are shown on top, followed by the dichotomous system with only two possible groups on the bottom. As previously discussed, the ambiguity of the race categories is represented by the overlap between Mixed with Whites and Blacks and between Negros and Whites.

The second mechanism of the system of racial classification that needs to be discussed is the methods used for racial classification. A method of racial classification is the process adopted for determining one's membership to a racial group; there are different approaches to determining racial membership; here, two methods will be presented, and their implications will be discussed: identification and categorization.

Methods of Racial Classification: Identification or Categorization

Racial membership is usually defined either by identification or categorization. When individuals provide information about their racial or ethnic identity, it is called identification. With categorization, on the other hand, a third party assigns a racial or ethnic category to the individual based on their observation or judgment. Therefore, racial classification can be heavily affected by the choice of method. There is no single "gold standard" for racial classification in data collection that is universally accepted. International institutes such as the OHCHR (United Nations Human Rights Office of the High Commissioner) have published guidelines arguing that identification should be preferred over categorization, as it is aligned with principles of autonomy and individual agency in expressing one's identity (RIVA & AG, 2018).

In Brazil, identification and categorization are widely used, depending on the data source. In the census IBGE's surveys, identification is the chosen method, except when a household member is either absent or unable to participate in the interview (Osório, 2003). Civil Registration and Vital Statistics (CRVS) systems

adopt categorization for birth and death records for obvious reasons. At the same time, identification is preferred if the individual can provide this information themself. Since this work combines census, survey, and CRVS data, it is fundamental to discuss how identification and categorization may affect the data.

Much of this discussion about the choice of classification method is related to the idea that one yields more accurate results than the other; usually, the argument is that categorization is an objective measure of race, while identification is subjective and, thus, less reliable. However, there are no guarantees that multiple interviewers, even if trained to identify physical traits specific to each racial category, will all agree on an individual's race. It would be more appropriate to look at both methods as subjective since identification and categorization lack definitive measures of objectivity; the only difference is that identification reflects the respondent's subjectivity and categorization of the interviewer (Osório, 2003). Moreover, identification is a process of understanding one's socialization, incorporating descent and cultural traits, in contrast with categorization, which tends to reflect other people's perceptions (Telles, 2014).

Adding to the debate about racial classification methods, studies in Brazil attempted to quantify the implications of using identification or categorization. Osório (2003) analyzes the main results from these studies and emphasizes key common features among all of them: (1) identification and categorization show a high overall level of agreement, (2) among whites, identification and categorization match for roughly 90% of individuals, (3) results from interviewers and interviewees diverge at significant levels for mixed and blacks, and (4) interviewers whitened respondents in every study.

The author emphasizes how a convergence of results between identification and categorization is not surprising; to some extent, as both interviewers and respondents are members of a single society, they share similar perspectives on race. It was also expected that Whites are not as often misclassified in contrast to Mixed and Blacks. However, the whitening effect produced on the population by categorization requires further discussions; the author suggests, based on previous research by Nogueira (1985, as cited in Osório, 2003), that this effect could be interpreted as a concession from the interviewers to the interviewees – given the low value associated with blackness, to categorize Blacks as Mixed, and Mixed as Whites becomes almost a "generous" act.

In 2013, IBGE published an extensive report on the racial profile of the Brazilian population; as a part of this report, they presented the results from a 2008 survey on the same topic. The survey data made allows, among other investigations – to assess if the features described by Osório are also evident when more recent data is used. Table 2.3.1 summarizes the comparison between race, when assigned by identification and categorization; as expected, when all races are accounted for, both methods yield similar results, around 77.5% of individuals were consistently classified. Moreover, in accordance with the existing research on the topic, categorization leads to whitening. The marginal distribution from Table 2.3.1 highlights this effect; results from categorization shows that Whites account for 58.29% of the population, Mixed are 29.19%, followed by Blacks (9.72%) and other races (2.80%). When identification is used, the

share of Whites is approximately 7 percentage points lower, while for Mixed, Blacks, and other races are 4.39, 0.54, and 1.95 percentage points higher, respectively.

Table 2.3.1 – Percent Distribution of Sample across Identified and Categorized Race: São Paulo, 2008

Categorization	Identification			
	White	Mixed	Black	Other
White	47.31%	9.24%	0.25%	1.49%
Mixed	3.55%	21.45%	2.86%	1.32%
Black	0.04%	2.33%	7.06%	0.29%
Other	0.51%	0.55%	0.08%	1.65%

Source: Adapted from (Petruccelli & Saboia, 2013).

Note: (1) Mixed corresponds to the grouping of the *"Morena"* and *"Parda"* categories; (2) Black corresponds to the grouping of the *"Negra"* and *"Preta"* categories; (3) Other, encompasses *"Amarela"*, *"Indígena"*, and *"Outras"*.

In addition, Table 2.3.2 shows different features regarding the use of identification or categorization when assigning race in a population; for each racial category, it is possible to see the agreement level between the two methods. Identification and categorization resulted in the same race for 92% of individuals identifying as White, 63.9% as Mixed, and 68.8% as Black. The extent to which the methods disagree on race is striking for Mixed and Blacks; this pattern of racial mismatch provides relevant insights regarding the implications of using one method or the other: while only 8% of identified whites were placed in a different racial group, this rate was roughly four times higher compared to Mixed (36.1%) and Blacks (31.2%).

The different results obtained by identification and categorization confirm the categorization bias described by Osório. That is, the racial composition of the population resulting from categorization tends to be at the lighter end of the color gradient when compared to the results from categorization. Table 2.3.2 shows additional evidence of that: according to the results presented, it is four times more likely that someone who identifies as Mixed to be categorized as White (27.53%) rather than Black (6.94%) by a third party; this bias works in the same direction for Blacks, with 30.36% assigned to either Mixed (27.91%) or White (2.45%) races.

Table 2.3.2 – Percent Distribution of Persons who Identified as White, Mixed, Black, and of Other Races: São Paulo, 2008

Categorization	Identification			
	White	Mixed	Black	Other
White	92.02%	27.53%	2.45%	31.31%
Mixed	6.90%	63.90%	27.91%	27.80%
Black	0.08%	6.94%	68.82%	6.13%
Other	1.00%	1.63%	0.82%	34.76%

Source: Adapted from (Petruccelli & Saboia, 2013).

Note: (1) Mixed corresponds to the grouping of the *"Morena"* and *"Parda"* categories; (2) Black corresponds to the grouping of the *"Negra"* and *"Preta"* categories; (3) Other, encompasses *"Amarela"*, *"Indígena"*, and *"Outras"*.

The ambiguity of racial classification in Brazil, associated with the whitening effect observed in the data and research on methods of racial classification, shows that research involving data on race requires careful analysis and manipulation. Given the widespread use of the media, government communications, and the discussions led by the Black movement in Brazil, the work presented in this thesis adopts the White/Non-White classification of the population, with Whites on one side and Mixed and Blacks on the other.

3 Data

Two data types are required to produce life expectancy at birth by gender and race: population and death counts. Demographic analysis in Brazil poses challenges due to regional diversity, impacting population and mortality data quality. The South and Southeast regions, with higher socioeconomic development, exhibit more reliable data compared to the less developed North, Northeast, and Midwest regions (Queiroz et al., 2020). Moreover, in Brazil, the COVID-19 pandemic postponed the 2020 census and increased mortality levels (Baptista et al., 2022; Castro et al., 2021; Lima et al., 2021).

Under these circumstances, the analysis presented focuses on the state of São Paulo for the time window between 2010 and 2019. São Paulo is the richest state in Brazil, highly urbanized, being home to the largest share of population among Brazilian states. Moreover, studies investigating completeness of mortality data confirm the data quality (Lima & Queiroz, 2014; Queiroz et al., 2017, 2020). Population counts were obtained from the combination of microdata from the 2010 census and the PNAD-C from the third quarter of 2019; death counts are available through a single source – SIM, a government database where all information about deaths occurring in the country is recorded.

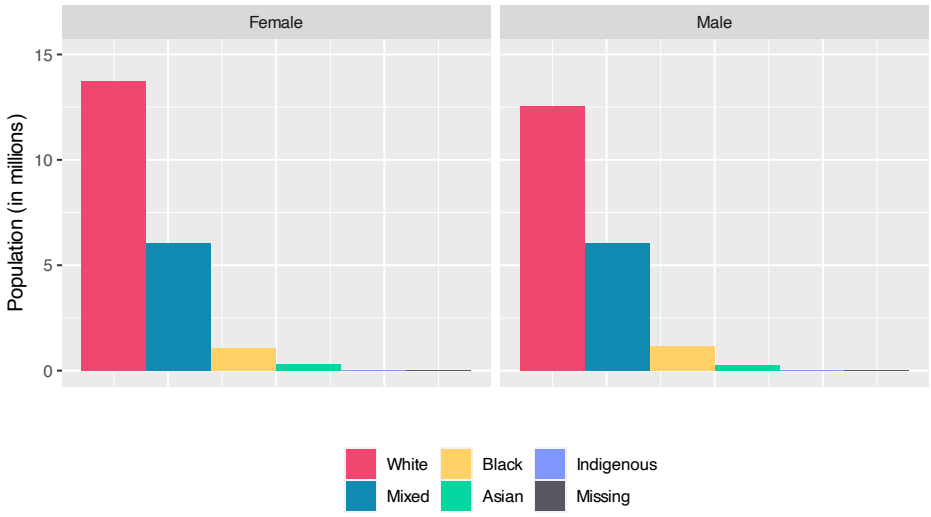
3.1 Population Estimates

Key Features

The 2010 census was the first to include a question about race for every resident in the country (Chiavegatto Filho et al., 2014). Results showed that around 98.5% of São Paulo's female population identified as White, Mixed, or Black. Amongst females, 13.72 million (64.75%) identified themselves as White, 6.06 million (28.60%) as Mixed, and 1.08 million (5.10%) as Black; a similar pattern was observed for males: 12.55 million (62.50%), 6.07 million (30.23%), and 1.16 million (5.78%) of this population identified as White, Mixed, and Black, respectively. The remaining racial categories (Asian and Indigenous) account for approximately 1.50% of males and females.

Figure 3.1.1 shows the distribution of the population in 2010 according to the five racial categories used by IBGE, as well as the total number of individuals for whom this information is missing. As previously presented, the pattern does not differ by gender, with the White category being the largest, followed by Mixed and Black.

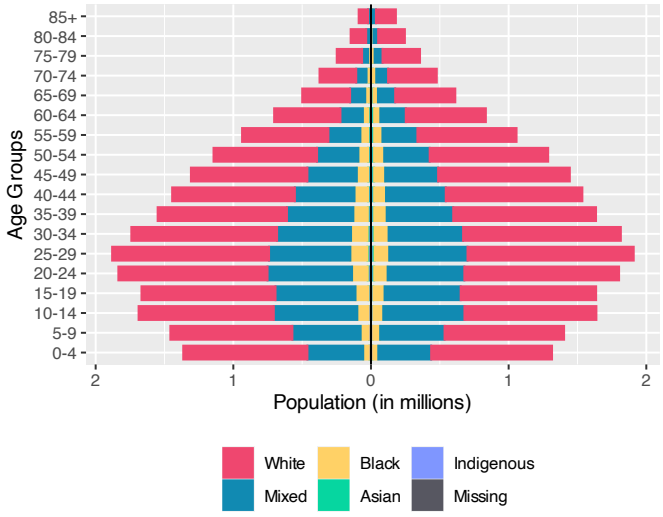
Figure 3.1.1 – Population by Gender and Race: São Paulo, 2010



Source: Own calculations from the 2010 Census.

Figure 3.1.2 introduces the race-age pyramid, which allows for a more complete analysis of the composition of the population by highlighting two dimensions of the data at once: the age and racial distribution considering five-year age groups and identified race.

Figure 3.1.2 – Race-Age Pyramid: São Paulo, 2010

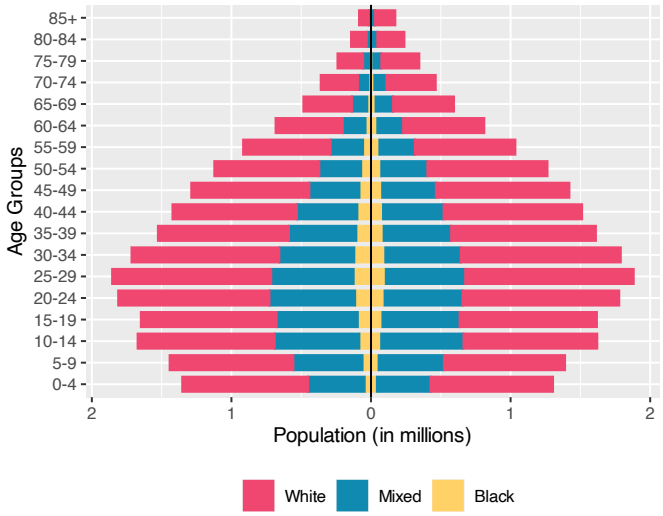


Source: Own calculations from the 2010 Census.

From this pyramid, it becomes clear that the White category is not only the largest in absolute numbers but also in all age groups, followed by Mixed and Black. Asians, Indigenous, and records where race information was missing are considerably small across age groups. Furthermore, the shape of the pyramid indicates a similar pattern for the three most representative races for both genders, varying more significantly in terms of the size of each category.

As the focus of the study is to decompose differences in life expectancy at birth amongst the racial groups, Figure 3.1.3 reproduces Figure 3.1.2 but focuses on the categories used in this study, namely: White, Mixed, and Black.

Figure 3.1.3 – Race-Age Pyramid: São Paulo, 2010
(White, Mixed, and Black)



Source: Own calculations from the 2010 Census.

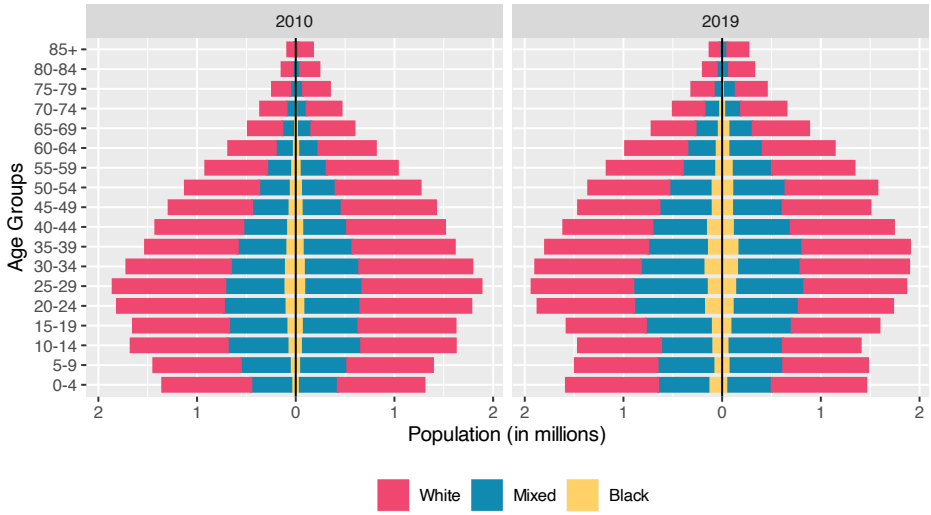
Given the effects the COVID-19 pandemic had in Brazil, PNAD-C was used as a source for the population estimates in 2019. PNAD-C stands for *Pesquisa Nacional por Amostra de Domicílios Contínua*, a comprehensive survey conducted quarterly by IBGE. The central unit of investigation for the survey is the household; on each round, more than 210,000 households are selected, and once selected, a household is visited once per quarter for five consecutive quarters – when it leaves the sample. The survey was permanently implemented in January 2012, covering the entire country. Its sample was designed to produce results at the national level and different aggregations at the subnational level.

In regard to population estimates obtained from this survey, it is essential to highlight that in 2018, IBGE published the revised estimates for the population projections by age and sex for the 2010-2060 period; this revision incorporated the results of the 2010 census and the most recent information on birth records. As a result, in 2019, all PNAD-C estimates were revised to reflect the most up-to-date estimates

of population projections. No further updates have been made since the publication of the initial results from the 2020 census in June 2023.

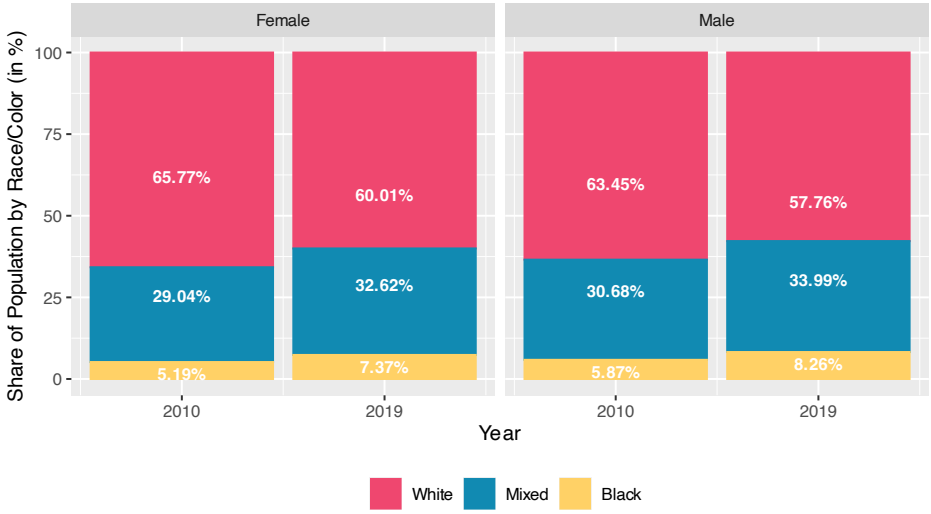
Figure 3.1.4 puts the race-age pyramids for São Paulo in 2010 and 2019 side-by-side; this comparison emphasizes an interesting aspect of the racial composition of the population apart from the expected change in the shape of the pyramid due to demographic forces such as fertility, migration, and mortality: while still being the largest category, the share of whites declined, while mixed and blacks increased; these two groups observed a positive variation between 2010 and 2019 for nearly all age groups.

Figure 3.1.4 – Race-Age Pyramid: São Paulo, 2010 and 2019 (White, Mixed, and Black)



Source: Own calculations from the 2010 Census and the 2019 PNAD-C (third quarter).

Figure 3.1.5 – Population's Racial Composition:
São Paulo, 2010 and 2019 (White, Mixed, and Black)



Source: Own calculations from the 2010 Census and the 2019 PNAD-C (third quarter).

Note: The percentages presented consider only individuals who identified either as White, Mixed, or Black.

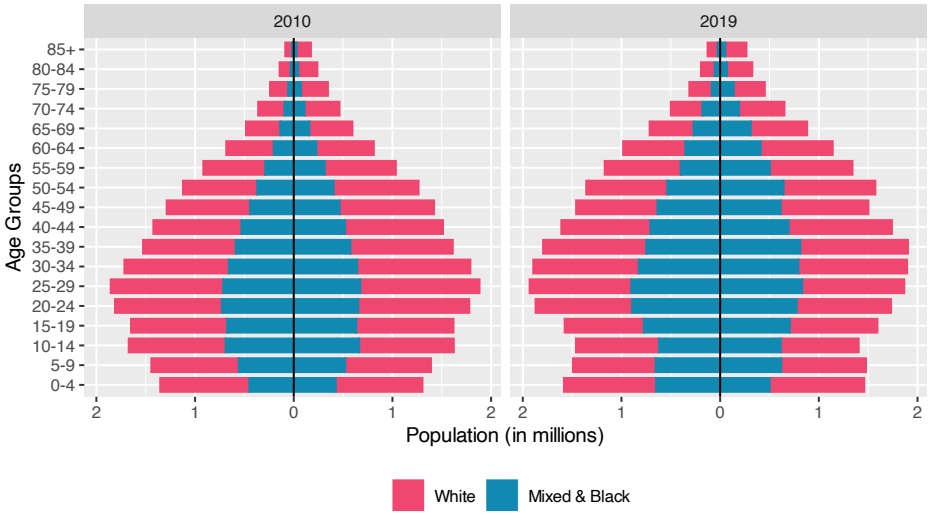
Figure 3.1.5 highlights this feature by showing the relative size of each group in 2010 and 2019 by gender with the top three racial categories considered. for females, the share that identified as White dropped from 65.8% in 2010 to 60% in 2019; for males, it went from 63.4% to 57.8%. Conversely, the share of individuals who claimed to be Mixed or Black increased; if grouped, it rose about six percentage points for males and females.

Previous studies have investigated the drivers of changes in the relative size of racial groups in Brazil (Carvalho et al., 2004; Miranda, 2015; Wood & Carvalho, 1994). The work by Wood & Carvalho (1994) and Carvalho et al. (2004) focused mainly on the period between 1950 and 1991, while Miranda (2015) included data from the 2000 and 2010 censuses to investigate more recent trends. From their results, it is possible to identify two separate historical trends: in the latter half of the 20th century, a significant increase of individuals identifying as mixed in detriment of the White or Black, and in the 2000s, the continued increase of the Mixed category, at a slower pace, accompanied by a growth of the Black category.

As for the causes behind these trends, all authors argue they were not driven by demographic forces such as fertility migration or mortality differentials but rather due to racial reclassification; that is, individuals will change their response when asked to which racial group they pertain (Carvalho et al., 2004; Miranda, 2015; Wood & Carvalho, 1994). This happens for a number of reasons; Miranda (2015) argues that the more recent trends of an increasing number of individuals identifying as Mixed or Black coincide with a period of intensified black political activism, which could be the driver of reclassification.

The trend of an increasing number of individuals identifying with races in the darker end of the spectrum (Mixed or Black) further argues for the approach taken in this work, where the new "Mixed & Black" racial category is used to discuss racial differences in life expectancy at birth. Figure 3.1.6 illustrates the age distribution of the population of São Paulo in 2010 and 2019, considering only these two groups.

Figure 3.1.6 – Race-Age Pyramid: São Paulo, 2010 and 2019 (White, Mixed, and Black)



Source: Own calculations from the 2010 Census and the 2019 PNAD-C (third quarter).

Data Limitations

Aside from discussing some of the most sensitive topics of the data, it is fundamental to address other potential limitations that may affect the interpretation of the results. The presence of missing data can introduce biases; in censuses or sample surveys, it occurs due to various reasons, such as non-response, measurement errors, or data collection limitations. It is crucial to acknowledge and appropriately handle missing data to ensure the reliability of the conclusions drawn from the analysis. By examining these limitations, it is possible to adopt suitable strategies to address them and enhance the overall robustness of the study.

Three main variables were used in this work: (1) age, (2) gender, and (3) race; the presence of missing values for any of these variables requires careful analysis and treatment. If missing values represent a significant dataset share, it would interfere with its use. The examination of the dataset for 2010 indicates that there are no missing values for both age and gender; for race, however, missing values were identified. From the approximately 41.26 million inhabitants of the state of São Paulo at that time, only for about 16 thousand (0.04%) race information was not available; regarding the distribution between genders, no

significant discrepancies were observed – approximately 7 of the 16 thousand were females, and 9 thousand were males, which represent 0.03% and 0.05% of their shares from the population, respectively.

Between 2010 and 2019, São Paulo's population grew by about 11.37%, reaching approximately 46 million people. Regarding data completeness, the scenario remains very similar to that of 2010; there is no missing information for age and gender, and only a small share of the population did not identify as part of one of the five racial categories defined by IBGE. Race is not available for about 6 thousand people (0.01%), of which 1.6 thousand are females and 4 thousand are males; this distribution is considerably different from that observed in 2010 but still accounts for a very small share of the population.

Missing data can yield distorted interpretations of any research's findings; however, as shown, that is not a significant issue regarding the population data used here – the census and PNAD-C have a minimal share of information missing when age, gender, and race are looked at. Individuals without race information were removed and not considered in any estimates; the exclusion of this group is assumed not to distort any estimates.

3.2 Death Counts

Key Features

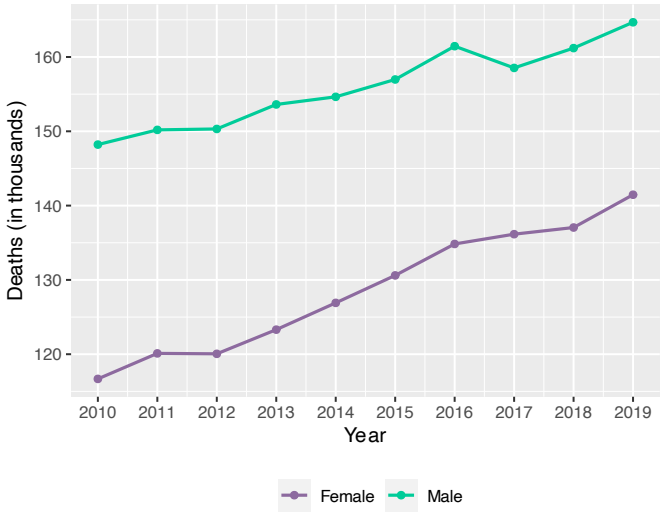
SIM, or *Sistema de Informações de Mortalidade*, is a part of Brazil's health information infrastructure. It is managed by DATASUS, the Department of Informatics of the Unified Health System (SUS). SIM was first established in 1975, and it is the result of the integration of over 40 instruments used to collect mortality data in the country up to that point; in 1979, the system was computerized, which later allowed for decentralized data collection. The data available on SIM is obtained directly from death certificates filled out by the attending physician at the time of the death.

A 2005 study by the World Health Organization, which analyzed mortality systems in various countries, assessed SIM as a medium-quality system (Mathers et al., 2005). The assessment observed six different dimensions: (1) years of mortality data, (2) completeness, (3) years of cause-of-death data, (4) ICD revision used, (5) coverage, and (6) share of deaths coded to ill-defined codes. More recent studies showed a significant improvement in the quality of the information available on SIM (Jorge et al., 2007), particularly regarding the completeness of adult deaths (Lima & Queiroz, 2014; Queiroz et al., 2017, 2020).

These studies have also confirmed a pattern of contrasting completeness levels between regions: while data are considered complete for South and Southeast states, that is not the case for the remaining states. Based on the results from said studies, data for the state of São Paulo are assumed to be complete – that is, the system captures virtually all deaths that happen in the state.

Figure 3.2.1 shows the evolution of the total deaths observed in São Paulo by gander from 2010 to 2019, with the notable exception between 2011 and 2012, when a small stall is observed, deaths have increased in absolute numbers for both genders. The number of deaths decreased between 2016 and 2017 for males (from 161.5 to 158.5 thousand) is also noteworthy. Nevertheless, since 2017, deaths have continued to follow a pattern of growth for both females and males.

Figure 3.2.1 – Total Deaths by Gender: São Paulo, 2010 to 2019



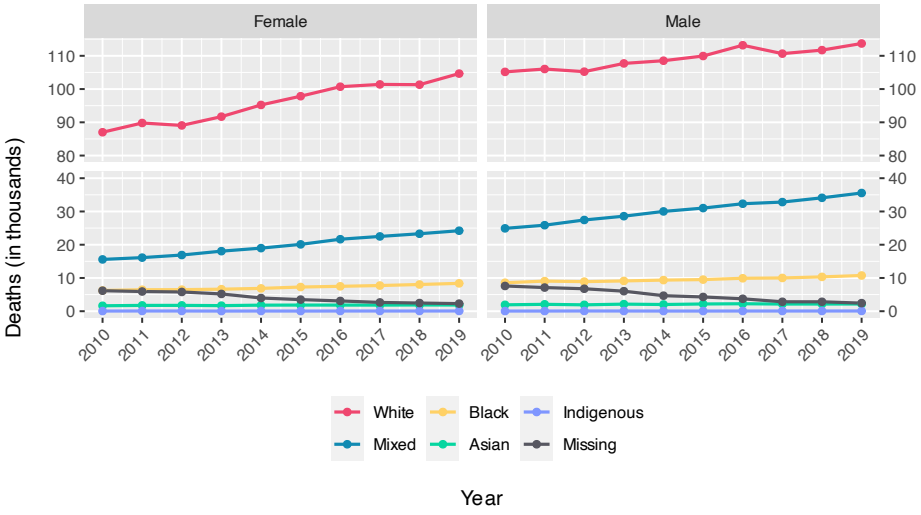
Source: Own calculations from DATASUS.

Note: Does not consider records with missing gender information (0.02% of the total deaths).

Despite its existence for almost half of a century, SIM did not have any information on the race of the deceased until 1996, when the Ministry of Health established the minimum attributes and complementary information to be adopted in all health-related data collection (Osório, 2003). Figure 3.2.2 allows a closer look at the trends in the absolute number of deaths according to gender and race; the y-axis is divided in two sections given the considerably larger amount of deaths occurring amongst whites – the largest share of the population, in contrast with the remaining groups.

The absolute number of deaths for males is higher than that of females for Whites, Mixed, and Blacks, with an overall upward trend for the time interval observed; for the Asian and Indigenous groups, the number of observations is minimal, so no clear pattern is observed. The clear downward trend observed for the category "missing" suggests an improvement in the data quality, with a decreasing number of death certificates where no race information is available.

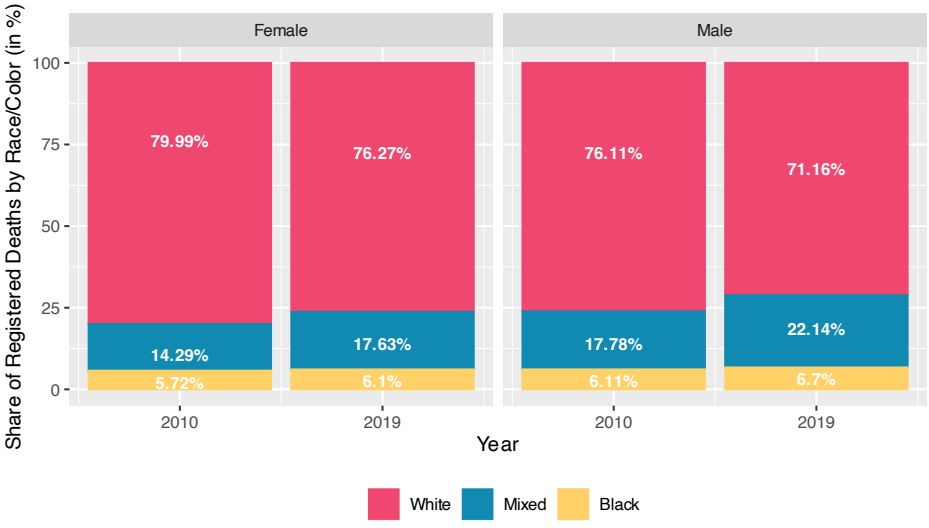
Figure 3.2.2 – Total Deaths by Gender and Race: São Paulo, 2010 to 2019



Source: Own calculations from DATASUS.

Note: Does not consider records with missing gender information (0.02% of the total deaths).

Figure 3.2.3 – Deceased's Racial Composition
São Paulo, 2010 and 2019 (White, Mixed, and Black)



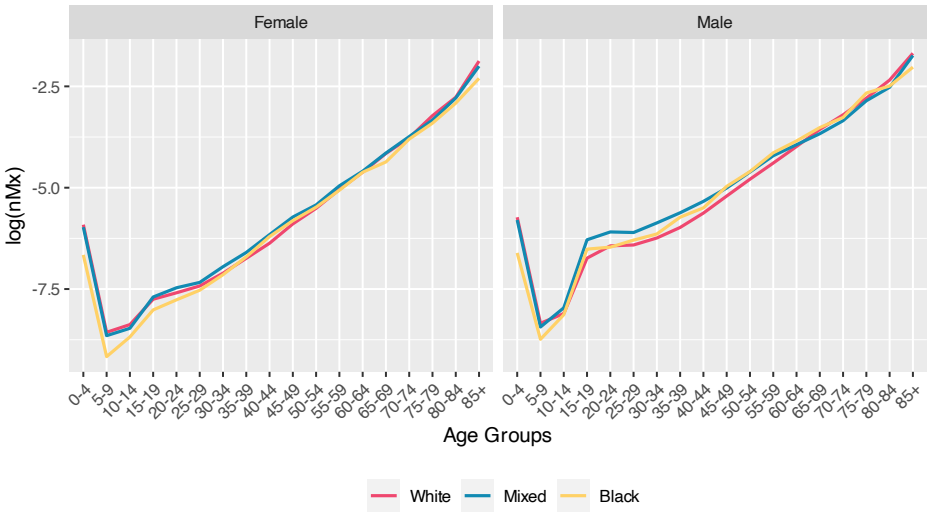
Source: Own calculations from the 2010 Census and the 2019 PNAD-C (third quarter).

Note: The percentages presented consider only individuals categorized either as White, Mixed, or Black.

Figure 3.2.3 shows the relative share of individuals categorized either as White, Mixed, or Black according to the availability of such information on the death certificates; a pattern similar to that observed in the population is evident, with a diminishing share of deaths attributed to Whites and an increase for Mixed and Blacks. Nevertheless, it is crucial to take a step back and consider that any race information available on death records originates from categorization; thus, it could be subject to the same whitening effect observed in the studies that discussed different racial classification methods. No studies have attempted to measure this effect on death records; however, it is a reasonable argument, as this effect is directly associated with the method used.

The previous analyses help highlight general trends in the data; however, they are insufficient in that they do not provide any information about the age distribution of the deaths or allow for more in-depth comparisons between groups. To further advance in the analysis of mortality trends by racial groups, age-specific mortality rates must be used, or in this case, race-age-specific mortality rates.

Figure 3.2.4 – Race-Age-Specific Mortality Rates (Log Scale): São Paulo, 2010 to 2019 (White, Mixed, and Black)



Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Notes: (1) Does not consider records with missing gender information (0.02% of the total deaths); (2) Considers the average number of deaths between 2010 and 2019.

When expressed in terms of the population, the race-age-specific mortality rates indicate a similar pattern and level for Whites, Mixed, and Blacks; in general, death rates for the Mixed category seem to be higher for females, starting at age 10, and for males, starting at age 35. Whites and Blacks show a very similar pattern for most ages, with exceptions made for the age interval 1 to 24 for females and 0 to 14 for males, where mortality for Blacks is higher. Furthermore, the increase in the share of deaths assigned to Mixed and Black populations does not seem to reflect in increased mortality rates for these groups; Figure

3.2.4 shows that the effect of changes in the racial composition of the population seems to have a more significant effect than that of the death records.

Following the previously discussed approach, where the Whites are compared to the group formed by grouping Mixed and Blacks, Figure 3.2.5 presents the race-age-specific mortality rates for these two groups. In this case, the pattern of an increase in mortality for the "Mixed & Black" racial group becomes apparent, with a disadvantage in survivorship for this group at most of the age groups, with an exception made to young-adult males (ages 14 to 29), with similar levels of mortality.

Figure 3.2.5 – Race-Age-Specific Mortality Rates (Log Scale): São Paulo, 2010 to 2019 (White, Mixed & Black)

Sources: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Data Limitations

The findings discussed in the previous subsection do not exist outside of a context where data are imperfect – missing information, underreporting, and biases related to how death certificates are filled out can distort the analysis in more than one way. Unlike what was observed for the data on population estimates, missing values are observed for all variables - age, gender, and race.

Jorge et al. (2007) argue that the data quality from SIM has improved significantly over time, which seems to be the case, as missing values in the age and race variables decreased significantly and steadily between 2010 and 2019, with the share of death records without any race information decreased from 5.21% in 2010 to 1.55% in 2019. Evaluating the complete period, only 0.44% of missing values were observed for age and 0.02% for gender; the share of death certificates with missing race information is

higher than that of the other two variables, around 3.15%. Table 3.2.1 summarizes, by variable and year, the missing values in the database.

Table 3.2.1 – Missing Values by Variable and Year: São Paulo, 2010 to 2019

Year	Age	Gender	Race
2010	1,679 (0.63%)	61 (0.02%)	13.792 (5.21%)
2011	1,583 (0.59%)	65 (0.02%)	13.080 (4.84%)
2012	1,550 (0.57%)	64 (0.02%)	12.583 (4.65%)
2013	1,490 (0.54%)	62 (0.02%)	11.250 (4.06%)
2014	1,520 (0.54%)	75 (0.02%)	8.694 (3.09%)
2015	1,183 (0.41%)	66 (0.02%)	7.813 (2.72%)
2016	1,066 (0.36%)	68 (0.02%)	6.866 (2.32%)
2017	927 (0.31%)	71 (0.02%)	5.522 (1.87%)
2018	800 (0.27%)	66 (0.02%)	5.324 (1.78%)
2019	701 (0.23%)	56 (0.02%)	4.751 (1.55%)
Total	**12,499 (0.44%)**	**654 (0.02%)**	**89.675 (3.15%)**

Source: Own calculations from DATASUS.

Note: "Age" refers to the availability of the deceased's date of birth.

Since race is a fundamental dimension of the analysis presented here, a closer look at the age distribution of the missing racial information sheds light on potential age-specific biases. Figure 3.2.6 shows the distribution of these missing values by 5-year age groups; the first noticeable pattern is that the share of death certificates with racial information available increased for all age groups, in consonance with the results from Table 3.2.1. Furthermore, the age group comprised of individuals aged 0 to 4 is the one where the most information is missing for all years and both genders; the remaining age groups present a constant level of incompleteness, with slight variations for all years.

Figure 3.2.6 – Evolution of Missing Values Regarding Race on Death Records by 5-Year Age Groups, 2010 to 2019

Source: Own calculations from DATASUS.

Note: Results do not consider registers with missing gender information (0.02% of the total deaths).

Unfortunately, it is beyond the scope of this thesis to define a criterion to assign a race to the death certificates that do not contain this information (about 3.15% of the total); thus, these observations were excluded from the analysis. The same procedure was followed for observations where the date of birth (0.44%) or the gender (0.02%) was missing. After accounting for records with multiple variables missing and excluding said elements, 96.46% of the original observations remained, which was assumed to yield consistent results.

Apart from the question of whether all information present in a death certificate is filled out, one should also account for the fact that not every death that happens is recorded. In theory, when a person dies, a new death certificate is filled out; however, it is well-known that despite historical efforts to improve the quality of the data available, deaths still go unregistered in Brazil. Several studies have investigated this issue and concluded that the underreporting of deaths in Brazil reduced significantly over the past decades, and in states such as São Paulo, death registration data is complete (Lima & Queiroz, 2014, 2014; Queiroz et al., 2017, 2020). Given the existence of extensive evidence that death counts are complete for the total population in São Paulo, it would be reasonable to assume this would also be the case when the population is broken down by racial groups.

The use of categorization for race assignment in the data sources is yet another limitation. The race-age-specific mortality rates presented use data from distinct sources: the numerator, namely the average death counts, was obtained from SIM, where race is assigned by categorization – usually performed by the attending physician; the denominator (population estimates) on the other hand, come from the 2010 census and the 2019 PNAD-C, sources that use identification for race assignment. This numerator-denominator bias, as described by Chiavegatto Filho et al. (2014), illustrates the hypothetical situation where the race information for one individual varies when dead (numerator) to the one assigned when alive (denominator), which would systematically distort the race-age-specific mortality rates.

A search of the relevant literature yielded no articles evaluating if the racial information from the death certificates matches that from when this individual was alive, but instead on assessing the convergence or divergence between identification and categorization (Osório, 2003; Petruccelli & Saboia, 2013; Telles, 2002). Nevertheless, it is reasonable to assume the whitening effect observed in studies comparing identification and categorization in the population also plays a role here. Under this hypothesis, individuals who identified as Black while alive would be categorized as Mixed or White on their death certificate, and those who identified as Mixed would be categorized as White; this phenomenon would act in the opposite direction as the movement observed in the population, where more and more individuals move from the White racial group to either Mixed or Black. The choice to unify the Mixed and Black racial groups reduces the possible transitions across racial groups while also minimizing biases from the numerator-denominator bias.

This set of limitations creates an exciting scenario; regardless of the assumption that death counts by gender and race would show high levels of completeness, this evaluation is constructive, as it also adds to the discussion regarding the potential whitening effect or numerator-denominator bias associated with categorization by serving as an indicator of the agreement between the data on race, when assigned by identification, and categorization in a context where an increasing share of the population identifies as either Mixed or Black in detriment of White.

4 Methods

The chapter on the methods applied is divided into three sections. The first section discusses Death Distribution Methods, which are widely used to assess and correct mortality data with completeness issues. Section 4.2 overviews the primary techniques for producing life tables and proposes an indirect life table method. Section 4.3 introduces a method for decomposing differences in life expectancy across races and genders at birth. This straightforward division offers a clear breakdown of the methodologies, addressing completeness issues, proposing alternatives for limited data quality scenarios, and examining differences in life expectancy.

4.1 Death Distribution Methods

Death distribution methods (DDM) are tools used to assess the completeness of adult mortality data collected from CRVS systems. These methods work by comparing two age distributions: the one of the living with the one of the deceased. DDMs serve as demographic indicators of registration systems' death data quality. (Hill et al., 2009). To put it simply, these methods link the age pattern of a population with the age pattern of the observed deaths using a demographic identity (Moultrie et al., 2013). This allows for an assessment of how much of the death counts are accounted for and enables adjustments to mortality estimates if needed. Although DDMs rely on strong assumptions, they are widely used, especially in research examining mortality in countries where CRVS data is suspected to be incomplete (Castanheira & Monteiro da Silva, 2022; Glei et al., 2019, 2021; Queiroz et al., 2020).

The most used methods fit into one of two groups, varying according to how the population's age pattern and the age pattern of the observed deaths are linked. On the one hand, growth balance methods are based on the balancing equation of population change, which states that a population varies between two points in time due to births, deaths, and net migration. On the other hand, synthetic extinct generations methods use the observation that the share of the population alive at a point in time must equal the total deaths from that cohort from that point forward for all ages (Moultrie et al., 2013; United Nations, 1983).

Growth Balance Methods

Initially proposed by Brass (1975), the growth balance method is based on the idea that in a stable population with no migration, constant reporting of deaths across all age groups, and accurate census enumeration, the population only changes due to births and deaths within a given period. This relationship also applies to age groups in the population; in this case, births are the number of individuals who have had a birthday or moved to the next age group. Equation 4.1.1 below describes this exact relationship for closed populations.

$$P_2(x+) = P_1(x+) + B(x+) - D(x+) \tag{4.1.1}$$

This equation can also be expressed as demographic rates; under the assumptions of no migration, constant reporting of deaths across all age groups, and accurate census enumeration, the population's growth rate equals the difference between birth rate (also referred to as birthday rate) and death rates. It is, then, possible to assess to which extent deaths are under-reported. Let $d^o(x+)$ denote the death rate obtained from the observed death counts; assuming these death counts are under-reported to a constant level of c for all age groups, an estimate for the completeness is obtained from the slope of the regression equation fitted to the $d^o(x+), b(x+) - r(x+)$ data points (Moultrie et al., 2013), as shown in Equation 4.1.2.

$$b(x+) - r(x+) = d^o(x+)/c \tag{4.1.2}$$

Brass' method is a simple approach that requires only two pieces of information: the population at a single point in time and the total deaths by age group (Moultrie et al., 2013). However, it comes with a caveat – it assumes a stable population, meaning that age-specific fertility and mortality rates remain constant. Hill (1987) proposed a more flexible approach called the Generalized Growth Balance (GGB) method to address this issue. This method does not require the stable population assumption and can be used when data from two censuses and death counts from the intercensal period are available by age group. With this information, the constant rate is replaced by age-specific growth rates (Moultrie et al., 2013).

It is essential to emphasize the assumptions from Hill's GGB method: the population is closed to migration, and the reporting of deaths across all age groups is constant. The assumption of the accuracy of census enumeration still holds; however, as the method uses age-specific growth rates, which are obtained from populations at two different points in time, enumeration may vary between censuses. Let k_1 and k_2 indicate the accuracy of enumeration between the first and second censuses, respectively; Equations 4.1.3a and 4.1.3b define $P_1(x+)$ and $P_2(x+)$ as a function of the observed population estimates.

$$P_1(x+) = P_1^o(x+)/k_1 \tag{4.1.3a}$$

$$P_2(x+) = P_2^o(x+)/k_2 \tag{4.1.3b}$$

k_1 and k_2 are correction factors for enumeration problems, such as under- or over-counting.

Equation 4.1.4 presents a formalization of the GGB method; it describes a linear relationship between the death rate expressed as the difference between the observed birth and growth rates and the observed death rate obtained from the death counts from the intercensal period; it is then, possible to estimate the completeness of the death counts and completeness of enumeration for each of the two censuses).

$$b(x+) - r(x+) = \left[ln(k_1/k_2)\Big/(t_2 - t_1)\right] + d^o(x+) \times \left[(k_1 \times k_2)^{0.5}/c\right] \tag{4.1.4}$$

When estimating $r(x+)$, the enumeration of census populations used may not be complete. However, the assumption that k_1 and k_2 are the same for all age groups makes it possible to solve for the value of c by calculating the slope of the line fitted to the $b(x+) - r(x+)$, $d^o(x+)$ data points (Moultrie et al., 2013).

Synthetic Extinct Generations Methods

While growth balance methods are described by the balancing equation of population change, synthetic extinct generations methods originate from a different concept. Consider a birth cohort: in the absence of migration, at any given time, the number of individuals from this cohort equals the future number of deaths to those individuals until the last survivor has died. If this expected future death stream is estimated from the observed deaths for a particular year or calendar period, an estimate for the number of individuals alive at that time can be obtained. The ratio between this estimate and the actual number of individuals is a measure of the completeness at which deaths are registered (S. Preston et al., 1980; United Nations, 1983). Furthermore, if the population is stable, an exact relationship can be written for the population size and the observed deaths.

$$\widehat{N}(x) = \sum_{a=x}^{\omega} D^o(a) \times e^{r \times (a-x)} \tag{4.1.5}$$

Equation 4.1.5 summarizes the principle behind the Preston and Coale method (S. Preston et al., 1980). Assume $N(x)$ is the observed number of people aged x and $\widehat{N}(x)$ is the number of people aged x obtained from the observed death counts from age x and higher. If deaths for this population are not fully registered, $\widehat{N}(x) < N(x)$, and the ratio $\widehat{N}(x)/N(x)$ estimates the degree of completeness of death counts for ages x and over (Moultrie et al., 2013).

$$c = \sum_{x=a}^{\omega} \widehat{N}(x) \Big/ \sum_{x=a}^{\omega} N(x) \tag{4.1.6}$$

The Preston and Coale method also relies on the same strong assumptions the Growth Balance method did, namely that the population is stable, migration is negligible or inexistent, and that reporting of deaths across all age groups is constant. Like Hill's version of the Brass' growth balance method, Bennett and Horiuchi (1981) proposed using two different censuses for the population counts to eliminate the stable population assumption. By computing age-specific growth rates, the method can be applied to non-stable populations; however, unlike the GGB method, Bennett & Horiuchi's approach does not provide correction factors for population estimates. This version of the method is known as the synthetic extinct generations (SEG); using this approach, $\widehat{N}(x)$ is obtained by age segments, and the number of people who turned x is obtained as shown in Equation 4.1.7.

$$\widehat{N}(x) = \widehat{N}(x+5) \times e^{5\times {}_5r_x} + {}_5D_x^o \times e^{2.5\times {}_5r_x} \text{ for } x < \omega \quad (4.1.7)$$

for the last age group, the calculations are as shown in Equation 4.1.8.

$$\widehat{N}(\omega) = {}_\infty D_\omega \times \left(e^{{}_\infty r_\omega \times e_\omega} - \left({}_\infty r_\omega \times e_\omega\right)^2/6\right) \quad (4.1.8)$$

The GGB-SEG method

There is no consensus in the literature on the most appropriate approach. However, the GGB and SEG methods are better options than their predecessors because they do not require the assumption of stable populations. Nevertheless, as Agostinho and Queiroz (2016) pointed out, each method has its strengths and weaknesses. On the one hand, GGB provides more accurate results when the census enumeration varies over time. On the other hand, SEG estimates are less affected by errors in age declaration. To combine the most robust features of each method, Hill et al. (2009) proposed a new, mixed method called GGB-SEG. The method's rationale is very straightforward: first, the census enumeration is assessed using GGB; then, the census data is adjusted for problems with enumeration, and then completeness is measured by applying the SEG method.

Methods' Limitations

Hill et al. (2009) evaluated the performance of these methods; the authors used non-stable populations with complete mortality data and simulated a series of data distortions. Their main findings indicate that the methods perform as expected when the data errors correspond to those the methods were designed for and when no assumptions are violated. Furthermore, results are also reliable in the presence of age misreporting, either in the population estimates or death counts. However, all methods yield distorted results if death coverage varies systematically with age or in the presence of significant migration flows (Hill et al., 2009).

Different alternatives deal with distortions caused by migration. On the one hand, Hill et al. (2009) approach this issue by suggesting that age groups more affected by migration flows should not be considered when estimating the completeness of death counts; they claim that this simple adjustment reduces distortions significantly. On the other hand, Hill and Queiroz (2010) propose a sophisticated approach, adjusting the GGB method to account for such flows; the authors found that this adjusted method yields good results for countries with high-quality data and large migration flows.

The DDM R-package (Riffe et al., 2017) was utilized to estimate completeness using the GGB, SEG, and GGB-SEG methods. Hill and Queiroz's methodology is not included in the package. The package's authors' approach was employed to determine the age range for each method, which is the age range with

the smallest root-mean-square error. This approach aims to mitigate distortions that could be observed due to migration, although it does not directly address distortions caused by migration.

4.2 The Life Table

As presented in the previous section, Death Distribution Methods aim to measure and produce correction factors for CRVS mortality data in populations affected by completeness issues. In such a context, careful data quality evaluation is necessary before computing mortality estimates (Queiroz et al., 2020). The most popular demographic tool for measuring mortality is the life table; traditionally, life tables start with an "age" column, which is followed by several age-related functions that serve as different measurements of mortality, survivorship, and life expectation (S. H. Preston et al., 2001; Siegel et al., 2004).

A life table can be described as a set of different pieces that, together, paint a picture of the extinction of a birth cohort (S. H. Preston et al., 2001). This approach calls for what is known as cohort life tables. Although conceptually very intuitive, producing cohort life tables requires high-quality time-series data. Because it mirrors the whole lifetime of the collective of individuals in a birth cohort, its members must be tracked throughout their lifetime, and the life table functions can only be calculated after the last member of that group has died. Furthermore, estimates obtained from a cohort life table do not reflect current mortality and health conditions but rather the conditions that individuals born several years ago were subjected to (Foz, 2021).

Due to the apparent challenges in computing the life table functions for actual birth cohorts, demographers use synthetic cohorts to produce period life tables. Instead of observing the complete mortality experience of a birth cohort to calculate the life table functions, period life tables reflect what would happen to a theoretical birth cohort if it was subjected to the current mortality conditions, defined as the observed age-specific mortality rates in a point in time (S. H. Preston et al., 2001); this assumption still allows for the calculation of all life table functions, and the estimates reflect more recent mortality conditions. However, interpreting these functions requires caution, as period life tables are conditioned to the hypothesis that the current age-specific mortality rates will remain constant in the future (Foz, 2021).

Besides the distinction between cohort and period life tables, many other specifications are used to produce life tables. The length of the age groups presented characterizes complete and abridged life tables. A complete life table's functions are computed for every year of life; in an abridged life table, intervals of 5 or 10 years of age for most of the age range are used. Siegel et al. (2004) argue that, for most purposes, abridged life tables present detailed enough information.

Standard Abridged Period Life Table Method

Standard period life table techniques are widely discussed and presented in textbooks (Foz, 2021; S. H. Preston et al., 2001; Siegel et al., 2004; Wachter, 2014). In such cases, these life tables are derived from

observed age-specific mortality rates, denoted as ${}_nM_x$, which are then converted into age-specific probabilities of death $\left({}_nq_x\right)$. There are different approaches to do such conversion; however, the key aspect is the estimation of the separation factor $\left({}_na_x\right)$ for each age group.

The separation factor expresses the average number of years lived between ages x and $x+n$ by persons who were alive at age x, but died before reaching $x+n$. A frequent approach is to assume that for all age groups above five years of age, deaths happen halfway through the age interval; for mortality under five years of age, the approach proposed by Coale and Demeny (1983, as cited in S. H. Preston et al., 2001) is often adopted. Due to data quality limitations, the chosen approach for this thesis was to have the first age group from age 0 until age 4; in this case, the separation factor, as proposed by Wachter (2014), is 1.5. Table 4.2.1 below summarizes the traditional structure of a period life table with the appropriate definitions and formulas.

Table 4.2.1 – Life Table Functions

${}_na_x$	:	is a factor that expresses the average number of years lived between ages x and $x+n$ by persons who were alive at age x, but died before reaching $x+n$. ${}_na_x = 1.5$ if $0 \leq x < 5$ and ${}_na_x = {}^{n}/_{2}$ if $x \geq 5$.
${}_nq_x$	:	is the probability of a person aged x dying before reaching age $x+n$. ${}_nq_x = {\left(n \times {}_nm_x\right)}/{\left(1 + \left(n - {}_na_x\right) \times {}_nm_x\right)}$
l_x	:	is the size of a birth cohort of initial size l_0 after $x+n$ years. $l_0 = 1$; for $x+n > 0$: $l_{x+n} = l_x \times \left(1 - {}_nq_x\right)$
${}_nd_x$	:	is the number of deaths of members from the birth cohort that happened between ages x and $x+n$. ${}_nd_x = l_x - l_{x+n}$
${}_nL_x$	:	is the number of person-years lived by the members of the birth cohort who are still alive at age x. ${}_nL_x = n \times l_{x+n} + {}_na_x \times {}_nd_x$; for the open-ended interval, ${}_\infty L_x = {l_x}/{{}_\infty m_x}$
T_x	:	is the number of future person-years lived by the members of the birth cohort who are still alive at age x, until the extinction of this cohort. $T_x = \sum_{a}^{\omega} {}_nL_a$

Table 4.2.1 – Life Table Functions

e_x	:	is the average number of years to be lived from age x, more commonly referred to as the life expectancy at age x. $e_x = {}^{T_x}/_{l_x}$

Sources: Foz (2021), Wachter (2014), and S. H. Preston et al. (2001).

However, standard life table techniques can yield distorted results in contexts such as the one described in this thesis, where the data imposes quality, enumeration, and completeness limitations. Alternative approaches to estimating the life table functions can significantly benefit the results. Previous research on similar topics to this thesis has suggested that the Bennett and Horiuchi method (1984) is a more adequate choice (Cunha et al., 2016; Merli, 1998; Paixão et al., 2010; S. H. Preston et al., 1996). This method uses age-specific growth rates and recorded deaths by age to derive the life table functions.

Alternative Approach: The Bennett & Horiuchi Method

This method is an extension of a relationship proposed by S. H. Preston & Coale (1982). The idea behind this relationship is that the age distribution of the observed deaths can be converted into the age distribution of the underlying life table by applying a growth correction factor (Bennett & Horiuchi, 1984; S. H. Preston et al., 1996; S. H. Preston & Coale, 1982). Equation 4.2.1 below, describes said relationship: $d(x)$ and $d(y)$ denote the deaths from the underlying life table at ages x and y ($x < y$), $D(x)$ and $D(y)$ are the observed deaths in the population, and $r(a)$ is the annualized growth rate of the population at exact age *a* (Merli, 1998; S. H. Preston et al., 1996).

$$\frac{d(y)}{d(x)} = \frac{D(y)}{D(x)} \times e^{\int_x^y r(a)da} \tag{4.2.1}$$

From Equation 3.3.1, all life table functions previously described can be obtained; Table 4.2.2 summarizes the main changes in the life table functions under the Bennett and Horiuchi method and their assumptions. It is essential to highlight that the interpretation of these functions is similar to what is summarized in Table 4.2.1.

Table 4.2.2 – Life Table Functions (Bennett & Horiuchi, 1984)

The correction factor is discretely approximated as: $e^{\int_x^y r(a)da} = e^{n\times({}_nr_{x-n} + {}_nr_x)/2}$
${}_nd_x = {}_nd_{x-n} \times {}^{{}_nd_x}/_{{}_nd_{x-n}}$, for the first age group: ${}_nd_0 = {}_nD_0$.

Table 4.2.2 – Life Table Functions (Bennett & Horiuchi, 1984)

${}_nL_x = n \times l_{x+n} + {}_na_x \times {}_nd_x$, for $x < \omega$, and ${}_4a_0 = 1.5$ for both genders and all races.
${}_\infty L_x = l_\omega \times \log_{10}(l_\omega)$

Source: (Cunha et al., 2016).

Similarly to the Death Distribution Methods discussed in the previous section, this approach also requires that migration flows are small enough not to distort the growth rates significantly and that incompleteness – if existing – is constant for all age groups. Because this work deals with life expectancy at birth at a subnational level, further disaggregated by gender and race, possible distortions in the growth rate due to migration flows must be addressed; in this regard, to minimize biases, the growth rates calculated consider only the gender dimension, regardless of which race individuals identify themselves with, as proposed in Paixão et al. (2010).

4.3 Arriaga's Decomposition

This thesis aims to evaluate and highlight differences in life expectancy at birth of subpopulations in São Paulo, namely, racial groups and genders. On the one hand, directly evaluating the gap in life expectancy at birth between the racial groups is very straightforward and indicates potential differences in the mortality conditions. However, the change in age-specific mortality rates might not happen in the same magnitude or direction for all age groups; in this case, the understanding of at which age mortality conditions present the most significant differences between groups and the direction of such changes allows for a more accurate examination of the drivers behind the variation of life expectancy.

Different authors proposed methods to decompose differences in life expectancies into the contributions of different age groups. In this analysis, differences in life expectancy at birth by gender and race are decomposed using the method introduced by Arriaga (1984). The author uses the concept of temporary life expectancies – a measure of the number of years a person is expected to live within a specific age range, to analyze how changes in mortality within specific age groups directly impact the overall life expectancy.

Arriaga argues that the effects of mortality change based on different age groups on life expectancy at birth can be categorized into two types: (1) effects attributed to the change in mortality conditions within a particular age group, the "exclusive effect", and (2) the effect of the interaction between the exclusive effect of each age group and the overall effect, the "interaction effect". It is important to note that the exclusive effect for an age group assumes mortality conditions remain unchanged for the remaining age range.

Moreover, he theorizes that the exclusive effect affects life expectancy directly and indirectly. While the direct effect reflects the result of changes (positive or negative) in life years within a particular age interval, the indirect effect measures the change in life years due to changes in the number of survivors at the end of the age interval from a mortality change within a specific age group (Arriaga, 1984). These effects are additive, that is, when added together, the direct and indirect effects result on the exclusive effect that a given age group has on explaining differences in the life expectancy.

The direct effect relates to the idea that once mortality conditions have changed within the $[x, x+n]$ age group, individuals living through these ages will live a larger (or shorter) number of years than before (Arriaga, 1984). Mathematically, it equals the difference between life expectancy measured for groups *a* and *b* within the $[x, x+n]$ age group, multiplied by the ratio between the number of survivors at exact ages x and 0 for the first group, as shown in Equation 4.3.1 below.

$$ {}_nDE_x = \left(\frac{{}_nL_x^b}{l_x^b} - \frac{{}_nL_x^a}{l_x^b} \right) \times \frac{l_x^a}{l_0^a} \tag{4.3.1} $$

The rationale behind the indirect effect is that the changes in age-specific mortality rates within the age group will also impact the number of survivors that reach the following age group, making this group larger or smaller. The change in the number of survivors, denoted as ${}_nCS_x$ is shown in Equation 4.3.2.

$$ {}_nCS_x = l_x^a \times \frac{l_{x+n}^b}{l_x^b} - l_{x+n}^a \tag{4.3.2} $$

This new group of individuals will continue to live according to the same age-specific mortality rates to which the remainder of the population are subjected to. Equation 3.4.3 shows the formula to calculate the indirect effect.

$$ IE_x = \frac{{}_nCS_x}{l_0^a} \times e_{x+n}^a = \frac{T_{x+n}^a}{l_0^a} \times \left(\frac{l_x^a \times l_{x+n}^b}{l_{x+n}^a \times l_x^b} - 1 \right) \tag{4.3.3} $$

The exclusive effect is defined under the assumption that mortality conditions change exclusively within the age group under analysis; however, this assumption does not allow for capturing the effects due to changes in mortality at all ages. The interaction effect captures the how the change in the age-specific mortality rates affects the years of life to be added (or deducted) to life expectancy at birth due to the new group of survivors (${}_nCS_x$) to reach age $x+n$. Equation 4.3.4 shows the mathematical relationship that defines the interaction effect.

$$ {}_nI_x = \frac{{}_nCS_x}{l_0^a} \times \left(e_{x+n}^b - e_{x+n}^a\right) \tag{4.3.4} $$

The total effect of the age group $[x, x+n]$ on the difference in life expectancy at birth is then equivalent to the sum of the direct, indirect, and interaction effects – as presented in Equations 4.3.1 to 4.3.4. The analysis conducted here uses only the total effect to shed a light in the differences of life expectancy at birth according to racial groups and genders.

5 Results

The chapter discussing the results is divided into three sections. First, the main aspects regarding the evaluation of the completeness of death counts by race are presented. The following section focuses on the life tables produced: life expectancy at birth is presented considering unadjusted and adjusted race-age-specific mortality rates using the three different DDM methods presented and the standard life table method; moreover, life expectancies obtained from the application of the Bennett and Horiuchi method are also presented. Finally, the main findings regarding the decomposition of the differences in life expectancy at birth by race and gender are shown.

5.1 Completeness of Death Counts

To reach the goal of decomposing life expectancy gaps across racial groups and genders, the first step is to use of DDMs to assess completeness of death counts by gender and race. Despite the evidence present in the literature attesting the quality of CRVS mortality data in Brazil, no studies have attempted to include race in their analysis. With the previous knowledge that death counts for the state of São Paulo are complete, a high level of completeness would be expected for all racial groups.

Completeness was evaluated according to the generalized growth balance (GGB), synthetic extinct generations (SEG), and the mixed GGB-SEG methods using the DDM R-package (Riffe et al., 2017). Table 5.1.1 presents a summary of the main results obtained from the application of the methods; in addition to the completeness levels, the last column shows the ratio between parameters k_1 and k_2, as defined in Equations 4.1.3a and 4.1.3b. This ratio indicates differences in the enumeration of the censuses from where population estimates were obtained. If the ratio k_1/k_2 is greater than 1, then enumeration in the first census was better than that of the second, which indicates it worsened between the two censuses; however, a ratio smaller than 1 indicates an improvement in the census enumeration.

Table 5.1.1 – Summary Of Results Death Distribution Methods

White				
Gender	**GGB**	**SEG**	**GGB-SEG**	**Delta (k1/k2)**
Female (1)	103.44%	77.51%	109.29%	1.06
Male (2)	100.13%	85.10%	102.40%	1.04
Mixed & Black				
Gender	**GGB**	**SEG**	**GGB-SEG**	**Delta (k1/k2)**
Female (3)	53.19%	143.45%	80.40%	0.91
Male (4)	76.58%	103.11%	73.62%	0.92

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Notes: (1) age range 15 to 50; (2) age range 15 to 60; (3) age range 15 to 50; (4) age range 25 to 60.

Results from Table 5.1.1 for Whites are presented and discussed first. It is worth noting that the correction factors obtained by the application of GGB indicate a slight "over-reporting" of death counts for females (around 3.4 pp) and complete death registration for males. Conversely, SEG estimates indicate severe under-reporting, 22.49 pp for females and 14.9 pp for males. Results for the mixed GGB-SEG method are closer to the GGB estimates, showing that death counts could be over-reported by 9.29 pp for females and 2.40 pp for males.

The analysis of the results obtained for the Mixed & Black group goes the opposite. Estimates from the application of the GGB method show deaths are under-reported by 46.81 pp for females and 23.42 pp for males. In the opposite direction, SEG estimates show that deaths of Mixed & Black are over-reported for females and males by 43.45 pp and 3.11 pp, respectively. As was the case for White deaths, applying the mixed GGB-SEG method yields similar results to the GGB method; however, completeness for females is about 27.21 pp higher when compared to the GGB estimates, while slight variation is observed for males.

Divergence is evident in the completeness estimates when different methods are considered, particularly when the GGB and GGB-SEG are compared to the SEG method, regardless of the racial group. The "delta" column is helpful to understand why this is happening. As previously discussed, delta (k_1/k_2) measures the variation in census enumeration; while for Whites, the ratio is greater than 1, for Mixed & Blacks, it is smaller. This phenomenon results from racial reclassification, as described in section 3.1 (see Figure 3.1.5); between 2010 and 2019, the share of individuals that identified themselves as White declined while the share of Mixed & Blacks increased, which appears in the results as an enumeration-like problem.

The literature indicates that GGB and GGB-SEG methods perform better than the SEG method when there are variations in enumeration between censuses, which suggests that GGB and GGB-SEG may provide more accurate completeness estimates. for these methods, delta could be interpreted as a measure of changes in the racial composition of the population due to racial reclassification. However, it is essential to note that racial classification is not the same as enumeration; racial classification is a process closely related to intensified Black political activism (Miranda, 2015) while enumeration deals with problems such as under- or over-counting the population. Results from the SEG method are affected to an even greater extent by racial reclassification; as shown by Hill et al. (2009), SEG is quite sensitive to variations related to enumeration. The completeness estimates for this method are particularly distorted for the Mixed & Black racial group; Figures B.2, B.4, B.6, and B.8, available in Appendix B, show the model is considerably ill-fitted for this group. Regardless of the limitations discussed when DDMs are applied to evaluate completeness by racial groups, life expectancy at birth by gender and race is presented using the unadjusted and adjusted death rates.

Robustness Check

As the results presented seem to suffer from potential distortions, in order to assure the methods were correctly applied, the White and Mixed & Black racial groups were grouped together and completeness was reassessed by gender using the GGB, SEG, and GGB-SEG methods. Results are presented in Table 5.1.2. Unlike what was observed in Table 5.1.1, all methods indicate very high completeness estimates, regardless of the gender; completeness levels are above 95% in all cases, in line with the findings presented by Lima & Queiroz (2014), Queiroz et al. (2017, 2020). The measure of variations in census enumeration does not show evidence that combining data from the 2010 census and the 2019 PNAD-C distorted results.

Table 5.1.2 – Robustness Check of Death Distribution Methods

Gender	GGB	SEG	GGB-SEG	Delta (k1/k2)
Female (1)	98.02%	99.45%	95.96%	0.99
Male (2)	100.13%	98.06%	97.69%	1.00

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Notes: (1) age range 30 to 65; (2) age range 30 to 75.

5.2 Life Expectancy at Birth

Following the analysis of completeness of death counts by gender and race, the next step is to calculate the life expectancy at birth for females and males, according to the two racial groups investigated in this work. The calculation of life expectancy considered unadjusted and adjusted death rates. As discussed in

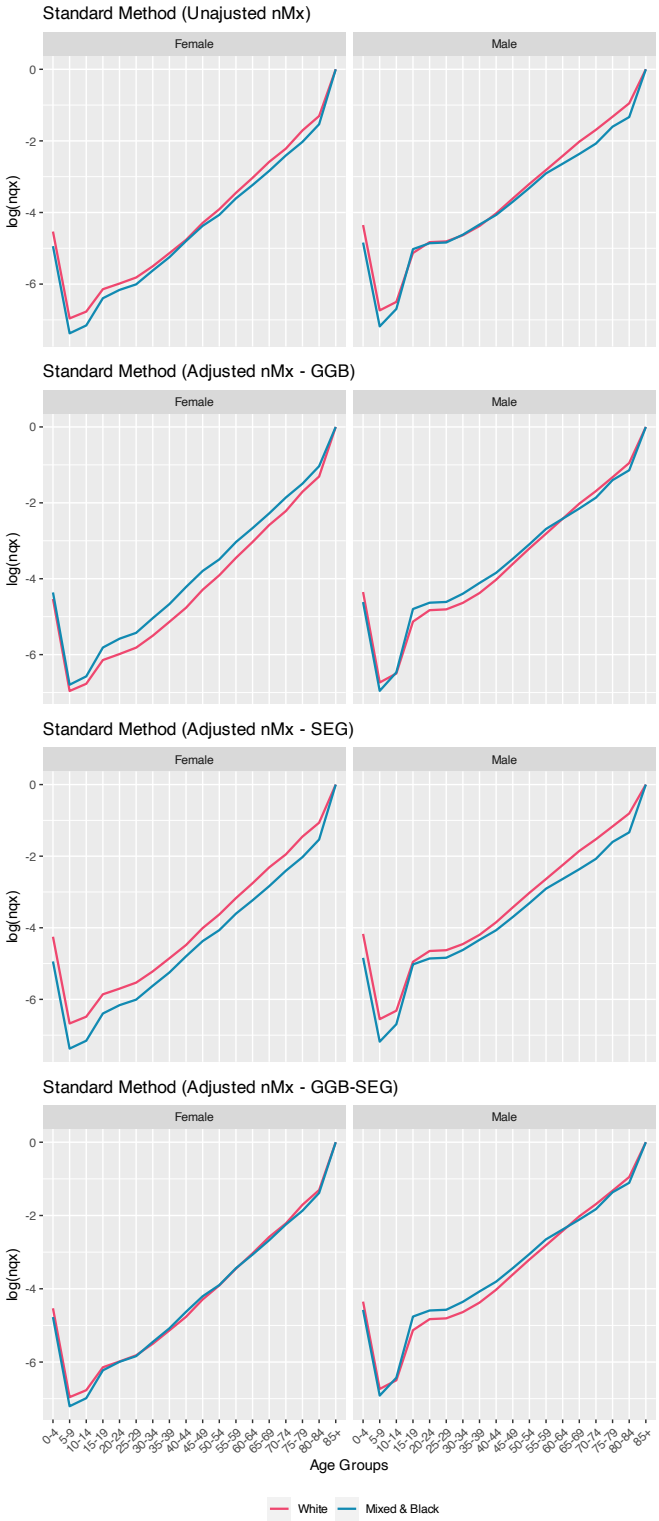
section 4.2, the life tables presented here consider two different approaches: the first one is, arguably, the most frequent version and it is often presented in textbooks – which is usually derived from observed age-specific mortality rates; the second one is based on the work of S. H. Preston & Coale (1982), and the life table is derived from age-specific growth rates and recorded deaths by age.

Standard Life Table Method

The first step to build a life table following the standard method is to compute the race-age-specific mortality rates ($_nM_x$) and from this set of rates, derive all the remaining functions detailed in Table 3.3.1.

Figure 5.2.1 presents the unadjusted and the adjusted race-age-specific probabilities of death for the White and Mixed Black groups by gender. Compared to the unadjusted probabilities, the adjustment by the GGB method emphasizes increased mortality for Mixed & Blacks – across most age groups for females and from age 15 to 59 for males. When adjusted by the SEG method, the opposite is observed: mortality for Whites is higher throughout the whole age range for females and males. Finally, the GGB-SEG approach yields very similar probabilities of death for females across all ages and slight differences for males, with a disadvantage for Mixed & Blacks between ages 15 to 59.

Figure 5.2.1 – Unadjusted and Adjusted Probability of Death (GGB, SEG, and GGB-SEG)



Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Notes: (1) for GGB, SEG, and GGB-SEG, the correction factors were applied to all age groups from 0-4 to 85+; (2) completeness factors were limited to 100%.

Table 5.2.1 shows life expectancy at birth calculated using the standard life table technique under four different scenarios: the first considers unadjusted death rates while the other three observe rates adjusted according to the results from the GGB, SEG, and GGB-SEG methods.

Table 5.2.1 – Life Expectancy at Birth in Years by Gender and Race, Unadjusted and Adjusted Rates

Female				
Race	**Unadjusted**	**GGB**	**SEG**	**GGB-SEG**
White	77.95	77.95 (=)	74.48 (↓)	77.95 (=)
Mixed & Black	81.76	73.91 (↓)	81.76 (=)	79.36 (↓)
Race Gap	**-3.81**	**4.04**	**-7.28**	**-1.41**
MALE				
Race	**Unadjusted**	**GGB**	**SEG**	**GGB-SEG**
White	71.11	71.11 (=)	68.66 (↓)	71.11 (=)
Mixed & Black	74.85	71.38 (↓)	74.85 (=)	70.79 (↓)
Race Gap	**-3.74**	**-0.27**	**-6.19**	**0.32**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: for the GGB, SEG, and GGB-SEG columns, the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Assuming life expectancy at birth estimates from unadjusted race-age-specific mortality rates as the baseline, Mixed & Black seem to live longer lives than their White counterparts regardless of gender. On average, Mixed & Black females live an additional 3.81 years compared to White females; for males, the advantage is slightly smaller, 3.74 years.

Before discussing life expectancy estimates from DDM-adjusted mortality rates, it is necessary to clarify that completeness factors were limited to 100%; thus, for some cases, life expectancy will remain unchanged. To further assist in interpreting the results, the upward and downward arrows and the equal sign in the table indicate how life expectancy at birth changed compared to the estimates from the unadjusted rates.

Results from GGB-adjusted death rates show a different scenario compared to the baseline: while life expectancy for Whites does not show any variation (77.95 years for females and 71.11 years for males), the estimates for the Mixed & Black racial group show a significant decline (-7.85 years for females and -3.45 years for males). This decline shifts the previously observed Mixed & Black advantage in survivorship;

White females have a 4.04-year advantage over Mixed & Blacks, while the difference observed for males is now close to zero.

Moving towards the opposite direction, results from SEG-adjusted death rates show life expectancy for Whites declining (-3.47 years for females and -2.45 for males) while estimates for Mixed & Blacks remain constant (81.76 years for females and 74.85 for males). These results aggravate the race gap in life expectancy observed in the baseline estimates; for females, the gap increased 1.91-fold and 1.65-fold for males. Finally, results from the mixed method show a reduced life expectancy for Mixed & Blacks (-2.40 years for females and -4.06 for males); such reductions paint a picture of convergence in life expectancy across races: for females, the racial gap in life expectancy is of -1.41 years while for males, only 0.32 years.

From the results presented, it becomes clear that racial reclassification is a challenge when calculating race-age-specific mortality rates; the decline in the share of individuals identifying as White, accompanied by an increased share of the population identifying as Mixed or Black leads to higher mortality rates and lower life expectancy for Whites when compared to Mixed & Blacks, as seen for the "unadjusted" estimates for life expectancy. Moreover, the numerator-denominator bias described by Chiavegatto Filho et al. (2014), associated with a potential whitening effect due to the use of categorization for racial classification in death certificates can further distort these rates.

While worthy attempts to estimate and adjust completeness in death counts by racial groups, racial reclassification distorts completeness estimates; furthermore, the application of the census enumeration correction factor (k_1/k_2) to minimize the effects of this process does not seem to achieve the goal, with the GGB and GG-SEG methods resulting in very different life expectancy estimates, especially for females.

In need of more reliable estimates for life expectancy by gender and race, an alternative method is applied – the Bennett & Horiuchi method. This approach to constructing life tables has been used to estimate life expectancy for populations with defective or incomplete data (Merli, 1998) and for different racial groups (Cunha et al., 2016; Paixão et al., 2010; S. H. Preston et al., 1996).

Suggested Approach: Bennett & Horiuchi (1984)

As previously said, instead of starting the life table from the race-age-specific mortality rates, this approach derives all life table functions from the relationship between the observed deaths and the growth rate by age segment. To avoid distortions in the growth rate caused by racial reclassification, the growth rates used in the estimation ignore the racial component, and were calculated only by gender; that is, the same growth rates are used for Whites and Mixed & Black, with the only distinction being according to their gender. It is important to notice that any distortions caused by whitening are not addressed in this approach; if, in fact death records suffer from a racial classification bias where individuals are more likely to be assigned a race in the lighter side of the color spectrum, mortality could be even higher for Mixed & blacks then the estimates produced using the Bennett & Horiuchi method.

This method choice allows for the estimates to reflect mortality differences as they are shown in the death counts. Furthermore, as shown in Equation 4.2.1, because the relationship between the life table deaths and the observed deaths is established based on ratios, reliable life expectancy estimates can be obtained even if death counts are not complete, as long as the assumption that the level of completeness is the same across all age groups. Table 5.2.2 shows life expectancy at birth by gender and race estimated according to: (1) the standard approach without adjustments and (2) the Bennett & Horiuchi method.

In comparison to the life expectancy at birth obtained from the unadjusted race-age-specific mortality rates, the suggested approach paints a very different scenario. Among females, life expectancy for Whites is 1.3 years higher than the one estimated using the standard method; for Mixed & Blacks, the difference is much more significant (-8.68 years). for males, while Whites show a 2.52-year increase, Mixed & Black show a decrease of 8.40 years. Such variations across methods describe a scenario marked by a significant advantage in life expectancy at birth for Whites over Mixed & Blacks (6.17 years for females and 7.18 years for males). These racial differences seem to have such a strong effect that the well-documented advantage of females over males disappears when life expectancy at birth of Mixed & Black females (73.08 years) is compared against that of White males (73.63 years).

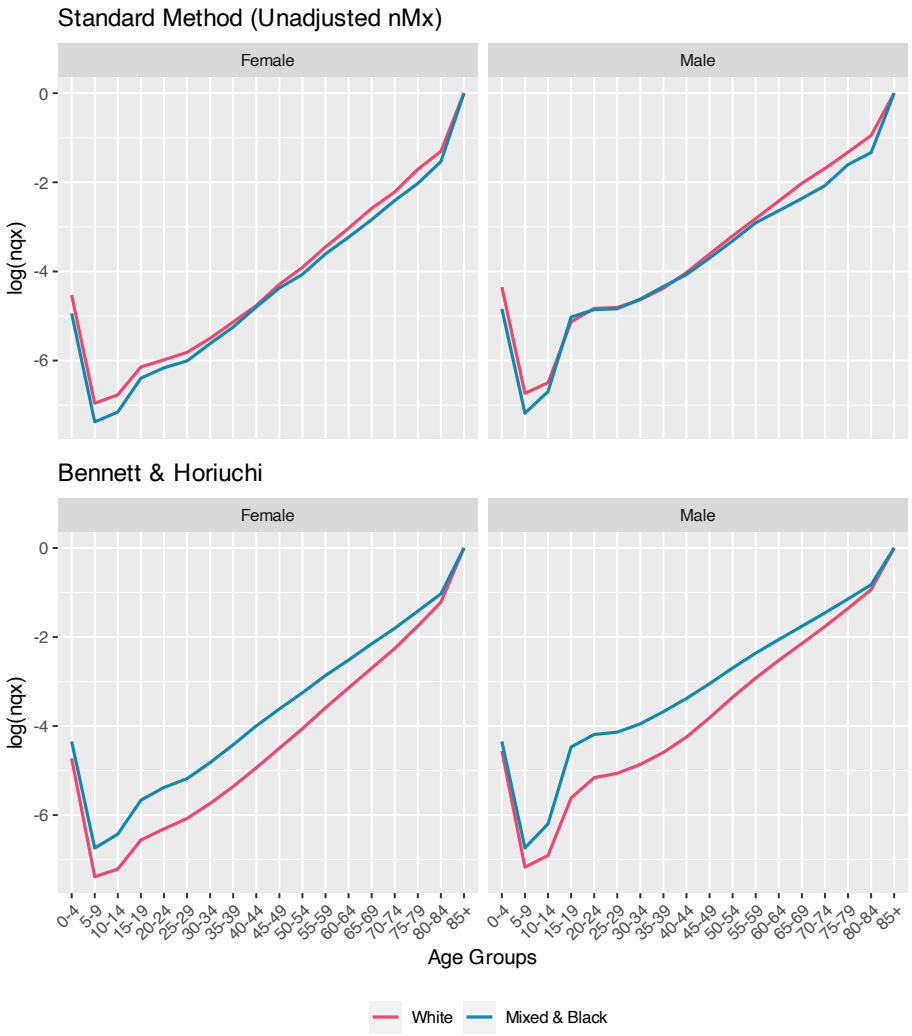
Table 5.2.2 – Life Expectancy at Birth Race Gap:
Standard Method and Bennett & Horiuchi with Unadjusted nMx

Female		
Race	**Standard Method**	**Bennett & Horiuchi**
White	77.95	79.25 (↑)
Mixed & Black	81.76	73.08 (↓)
Race Gap	**-3.81**	**6.17**
Male		
Race	**Standard Method**	**Bennett & Horiuchi**
White	71.11	73.63 (↑)
Mixed & Black	74.85	66.45 (↓)
Race Gap	**-3.74**	**7.18**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure 5.2.2 highlights the gender and race differences in race-age-specific probabilities of death computed using the standard life table method with unadjusted mortality rates and Bennett & Horiuchi method; while for the standard method there is no clear single pattern of differences in nqx, the application of the Bennett & Horiuchi method shows a clear disadvantage for Mixed & Blacks throughout the age range.

Figure 5.2.2 – Probability Of Death:
Standard and Bennett & Horiuchi Life Table Methods



Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Robustness Check

In a similar approach to the check presented for the DDM and aiming to confirm that the Bennett & Horiuchi method, suggested to estimate life expectancy at birth, yields reliable results, Table 5.2.3 presents life expectancy at birth by gender obtained from: (1) *Fundação SEADE*, the statistical office of the state of São Paulo, (2) standard life table method following the specifications presented in Table 4.3.1 with unadjusted rates, and (3) the suggested approach, following the parameters indicated in Table 4.2.2.

Table 5.2.3 – Robustness Check of Life Table Methods

Gender	Official Estimates (1)	Standard Method	Bennett & Horiuchi
Female	79.4	78.53	78.24
Male	73.3	71.76	72.20

Sources: *Fundação SEADE* and own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Notes: (1) Obtained from *Fundação SEADE*; information is only available rounded to one decimal place.

The results here are very consistent: as expected all estimates presented confirm the increased expectation of life women have over men. Official estimates from *Fundação SEADE* for 2019 show higher life expectancy at birth for both females and males when compared to the other two estimates; however, such differences can emerge from methodological differences and a discussion specific to these differences is beyond the scope of this work.

The convergence between estimates from the standard life table method and the approach proposed by Bennett & Horiuchi is striking. for females, the difference is -0.29 years (78.24 - 78.53) and 0.44 years for males (72.20 – 71.76). These results serve as confirmation the method proposed by Bennett & Horiuchi provides consistent life expectancy estimates.

for the purposes of this thesis, life expectancy estimates by race obtained from the application of the Bennett & Horiuchi method are considered consistent and will be used to in the decomposition of life expectancy differences by age groups.

5.3 Age Decomposition of Life Expectancy at Birth Gaps

Using life expectancy estimates obtained using Bennett & Horiuchi's approach, gender and racial differences are decomposed using Arriaga's method; this technique decomposes life expectancy gaps into three additive effects: (1) the direct effect, (2) the indirect effect, and (3) the interaction effect. All three describe the outcomes from changes in mortality conditions – namely, in age-specific mortality rates. The direct and indirect effects describe changes within an age group; while the direct effect measures the

additional (or decreased) life years lived within the age group due to changes in mortality conditions in that age group, the indirect effect measures the impact of changes in mortality within the age group have on the number of survivors that reach the next one and the additional (or decreased) life years to be lived by these individuals. Finally, the interaction effect measures the effects of changes in mortality at all ages. The total effect of a given age group in explaining differences in life expectancy is then obtained by adding the direct, indirect, and interaction effects.

Two different decompositions are presented here: first, the life expectancy racial gap between Whites and Mixed & Blacks is investigated; next, the attention turns to the gender gap in life expectancy. Results are summarized in Tables 5.3.1 and 5.3.2, which show the total effect (in years) and the relative weight (in percentages) of each age group; the direct, indirect, and interaction effects were omitted here, and some age groups collapsed, exceptions made for the first and last groups. The complete results are available in Appendix D.

Racial Differences Decomposed across Genders

To start the analysis, Table 5.3.1 shows the decomposition results by racial groups for both females and males; according to the results presented, White females live, on average, an additional 6.17 years when compared to theirs Mixed & Black counterparts; the advantage among males is even higher, Whites live 7.18 years more than Mixed & Blacks. Moreover, it is essential to emphasize that Whites have an advantage in survivorship in all age groups.

Table 5.3.1 – Racial Differences Decomposed across Genders

Age Group	W - M&B Females		W - M&B Males	
	Years	%	Years	%
0-4	0.32	5.22%	0.18	2.52%
5-19	0.22	3.64%	0.53	7.36%
20-39	0.86	13.92%	1.93	26.90%
40-64	2.77	44.89%	3.24	45.07%
65-84	1.77	28.64%	1.19	16.56%
85+	0.23	3.69%	0.11	1.59%
Total	**6.17**	**100.00%**	**7.18**	**100.00%**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

About the distribution of the relative contribution to the racial gap in female life expectancy, it is imperative to highlight that the age groups 40 to 64 and 65 to 84 are the top two age groups, having the highest shares of the difference in life expectancy between White and Mixed & Black females; these groups account for 73.53% (4.54 years). The age groups 20 to 39 and 0 to 4 are responsible for 19.14% of the difference, followed by the age groups 5 to 19 and the open age group 85+, each one with 3.67%, on average.

For the males, the differences in life expectancy are also centered in the 40 to 64 age group, which accounts for 45.07% (3.24 years) of the 7.18-year gap in like expectancy, followed by the 15 to 39 and 65-84 age groups, adding up to 88.53%, equivalent to 6.36 years. The remaining 0.82 years of the difference are distributed as follows: 0.53 years attributable to differences in mortality conditions in the 5 to 19 age group, 0.18 and 0.11 years attributable to the 0 to 4 and the open age group 85+, respectively.

Furthermore, a comparison between the age distribution of life expectancy differences of females and males is also informative. While differences in mortality conditions in the age group 40 to 64 was responsible for most of the difference for females and males, the female distribution is skewed towards older ages while the male distribution is skewed towards younger ages. The comparison between the 15 to 39 and 65 to 84 age groups across genders is an example of that: on the one hand, while the age group 20 to 39 accounts only for 13.92% of the difference for females, the share is 1.93 times higher for males; on the other hand, while 28.64% of the female racial gap in life expectancy is attributable to the first age group, the share is only 16.56% for males (0.58 times smaller). The similar pattern is observed for the 5 to 19 and 85+ age groups; an important exception is the first age group, differences in mortality in this age group are responsible for twice as much of the racial gap in life expectancy for females than for males.

Gender Differences Decomposed across Racial Groups

Expanding the analysis presented, Table 5.3.2 shows the decomposition results by gender for Whites and Mixed & Blacks; this analysis is informative as it allows for comparison related to the racial pattern of gender differences in life expectancy according. The life expectancy gap between Females and Males is 5.62 years for Whites and 6.63 years for Mixed & Blacks, which indicates that females experience increased survivorship when compared to males regardless of racial groups.

Table 5.3.2 – Gender Differences Decomposed across Racial Groups

Age Group	F - M Whites		F - M Mixed & Blacks	
	Years	%	Years	%
0-4	0.13	2.29%	0.00	0.01%
5-19	0.17	2.95%	0.47	7.11%
20-39	0.87	15.45%	1.93	29.03%
40-64	2.33	41.43%	2.86	43.07%
65-84	2.07	36.90%	1.35	20.43%
85+	0.06	0.98%	0.02	0.34%
Total	**5.62**	**100.00%**	**6.63**	**100.00%**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

For the Whites, the age group 40 to 64 has the highest share of the difference in life expectancy between females and males, accounting for 41.43% (2.33 years) of the difference observed. The remaining 5 groups account for the remaining 58.57% (3.29 years); however, this is not uniformly distributed: 2.07 years are attributed solely to the 65 to 84 age group, which corresponds to approximately 63% of the difference not attributable to the 40 to 64 age group. Other relevant aspects include the increased male mortality associated to the o to 4, 5 to 19, and 20 to 39 age groups.

Mixed & Blacks show a similar pattern regarding the age distribution of the gender gap in life expectancy: most of the difference is related to mortality conditions in the 40 to 64 age group (43.07%, or 2.86 years); however, unlike what was observed for Whites, the age group that comes second is the 20 to 39, with a share of 29.03%, followed by the 65 to 84, 5 to 19, 85+, and 0 to 5 age groups, respectively.

Furthermore, the comparison between the age distribution of gender differences in life expectancy across racial groups is also informative. A similar pattern to what was observed between Whites and Mixed & Blacks is evident here. While differences in mortality conditions in the age group 40 to 64 were responsible for most of the difference for Whites and Mixed & Blacks, the White distribution is skewed towards older ages while the Mixed & Black distribution is skewed towards younger ages. It is striking how the contribution of the age groups 5 to 19 and 20 to 39 is 2.41 and 1.88 times higher for Mixed & Blacks than for Whites, respectively; conversely, the contribution of the 40 to 64 and 65 to 84 age groups is 1.81 and 2.88 times higher for Whites than for Mixed & Blacks.

6 Discussion and Conclusion

Summary and Implications

The primary goal of this thesis was to use Arriaga's decomposition method to evaluate the race gap in life expectancy at birth for females and males. Research focusing on health and mortality differences by race is relatively sparse in Brazil, and life expectancy estimates by gender and race are not readily available. Nevertheless, the existing evidence points towards a systematic disadvantage of Mixed & Blacks (Non-Whites) compared to Whites. The results of this thesis indicate that Whites experience longer lives compared to Mixed & Blacks. Among females, the race gap in life expectancy at birth is 6.17 years, with a right-skewed age distribution of the differences; for males, the race gap is even higher, 7.18 years, with a left-skewed age distribution.

In light of the lack of availability of life expectancy at birth estimates by gender and race, to achieve the main goal of this thesis, it was necessary to produce life tables by gender and race; different methods to assess the completeness of death counts were combined with life table techniques before the final estimates were obtained. In this context, two secondary research question were proposed; the first addressed the evaluation of death counts' completeness and the second focused on which life table technique would yield the most reliable estimates for life expectancy at birth. The introduction of this thesis presents the expected results. At that point, it was expected that life expectancy at birth would differ among races, with Non-White people facing a disadvantage for females and males. Moreover, that completeness would not vary when evaluated for race-based subgroups of the population and that whatever life tables methods were employed would yield similar results. Now, the answers to the research questions will be discussed in light of the relevant literature and the initial hypotheses.

The initial hypothesis that mortality conditions, measured by life expectancy at birth, would differ between racial groups with an advantage for Whites over Mixed & Blacks was confirmed by the results presented and are supported by the literature. White females have the highest life expectancy at birth, 79.25 years, followed by White males with 73.63 years, Mixed & Black females, and males with 73.08 and 66.45 years, respectively. Thus, the race gap in life expectancy at birth is 6.17 years for females and 7.18 years for males. Previous studies on the topic already pointed to this disadvantage for Mixed & Blacks (Cunha et al., 2016; Paixão et al., 2010).

Despite being sparse, literature on racial inequalities in health and mortality differences in Brazil are consistent with these results, as it shows that Non-Whites (Mixed & Blacks) experience increased mortality compared to Whites. Matijasevich et al. (2008) and Wong et al. (2014) showed that infants born to Non-White mothers experience higher mortality rates than those born to White mothers. According to Martins et al. (1991) Non-White women are three times more likely to die due to abortion-related complications than White women; moreover, mortality due to post-abortion complications is partially explained by racial

discrimination in health services (Goes et al., 2020). There is also evidence that Non-White men with low SES are both the primary victims and perpetrators of violence, whereas Non-White women with low SES are the primary victims of domestic violence. (Reichenheim et al., 2011). Finally, only a limited amount of studies investigated differences in mortality by race at older ages; however, Lotufo et al. (2007) presented evidence of racial differences on CVD, with higher mortality rates higher for Non-White women and men compared to their White counterparts.

The analysis of the decomposition's results shows that the age distribution of the race gap for females (6.17 years) and males (7.18 years) has notable differences. For females, the age group that explains most of the difference is 40 to 64, responsible for 44.89% (or 2.77 years), followed by the 65 to 84 age group, which explains 28.64% of the difference (or 1.77 years); thus, the distribution is right skewed, with most of the difference being attributed to older ages. For males, the main age group is also 40 to 64, responsible for 45.07% (or 3.24 years) of the total difference; however, unlike what was observed for females, the age group that follows is 20 to 39, with 26.90% (or 1.93 years), characterizing a left skewed distribution. Thus, the race gap for females seems to be explained by differences in mortality conditions at older ages while differences at younger ages play a more relevant role for males.

To add to the discussion, the comparison between the gender gap across races was also presented. The gender gap for Whites was 5.62 years and 6.63 years for Mixed & Blacks. Among Whites, mortality conditions between females and males differ the most at the 40 to 64 and 65 to 84 age groups, which together account for 78.33% of the total difference; it is important to point out that females experience lower mortality for every age group. On the other hand, 72.10% of the differences in mortality between Mixed & Black females and males results from differences in the 20 to 39 and 40 to 64 age groups. Once again, as observed for the gender differences for Whites, Mixed & Black females also hold the advantage when compared to males.

In summary, the results show that life expectancy varies by race. The decompositions presented indicate that among females, the differences in life expectancy between Whites and Mixed & Blacks result, for the most part, from differences in mortality at older ages, whereas the differences for males occur at younger ages. Furthermore, the decomposition of the gender differences across racial groups highlights a pattern marked by the largest differences in mortality for Whites happening at older ages when compared to Mixed & Blacks.

As there are no life tables by gender and race readily available, a part of this thesis was dedicated to producing these life tables so that it would be possible to decompose the race gap. The two additional research questions of this thesis focus on evaluating the completeness of death counts and discussing how to best produce reliable life expectancy at birth estimates by race.

Extensive research has been conducted on assessing the quality of mortality data in terms of completeness (Diógenes et al., 2022; Lima & Queiroz, 2014; Queiroz et al., 2017, 2020) and the literature

reinforces that completeness improved over the years at the national level, while it varies greatly among regions. Notably, São Paulo shows historically high completeness levels, around 100%. While DDMs have been widely used, no studies employed the methods to assess completeness of subpopulations defined according to racial groups. Based on the results of previous studies, which concluded mortality data for São Paulo is reliable, or more reliable than it is for other regions in Brazil, the initial hypothesis was that death counts would also be complete when the population was broken down by gender and race.

Contrary to the initial hypothesis, the application of DDMs showed contrasting results. On the one hand, results from GGB and GGB-SEG suggested that deaths counts were "over-reported" for Whites and under-reported for Mixed & Blacks. On the other hand, SEG estimates suggested exactly the opposite: White deaths were under-reported and Mixed & Black deaths were over-reported. Moreover, an additional measure provided by the GGB method suggested a significant variation of population enumeration between 2010 and 2019, where estimates for the White share of the population decreased and the share of Mixed & Blacks increased; this finding was particularly interesting as it dialogues with the phenomena observed in Figure 3.1.5 describing the change in the racial composition of the population attributed to racial reclassification.

It is well-known that in the presence of enumeration issues, GGB and GGB-SEG perform better comparatively to SEG (Hill et al., 2009). However, it is important to make a distinction between the situation caused by racial reclassification and what actually is an enumeration problem. Racial reclassification is a process that is closely linked to increased activism by Black communities, as noted by Miranda (2015). On the other hand, enumeration is concerned with issues such as accurately counting the population and avoiding over- or under-counting. Thus, caution is required when evaluating life expectancy estimates obtained from DDM-adjusted race-age-specific mortality rates, as the methods were not designed to deal with situations such as the one described. So, from this thesis' results, it is not possible to draw a definitive conclusion about the completeness of death counts by gender and race in São Paulo between 2010 and 2019 using the GGB, SEG, and GGB-SEG methods.

Finally, the question about which life table method produces the best possible life expectancy at birth estimates dialogues with the idea that a combination of DDMs with standard life table techniques should yield reasonable estimates. However, previous studies have not followed such approach; Cunha et al. (2016) and Paixão et al. (2010) argued that the Bennett & Horiuchi method was the better choice to estimate life expectancy by race. Nevertheless, despite the limitations identified in the DDMs' results, life tables by gender and race were constructed using both standard life table methods and the Bennett & Horiuchi approach; this was done considering a set of different scenarios: (1) standard life table method + unadjusted race-age-specific mortality rates, (2) standard life table method + GGB-adjusted race-age-specific mortality rates, (3) standard life table method + SEG-adjusted race-age-specific mortality rates, (4) standard life table method + GGB-SEG-adjusted race-age-specific mortality rates, and (5) Bennett & Horiuchi method.

Life expectancy at birth calculated using the standard life table method and unadjusted race-age-specific mortality rates shows an advantage for females over males regardless of race, but it suggests that Mixed & Blacks have an advantage over Whites, which contradicts findings from existing research. The use of GGB-adjusted rates reinforces the female over male advantage, but the racial gap in life expectancy at birth is different for females and males; while for the first group, Whites had the advantage over Mixed & Blacks, no significant differences were observed for males. SEG-adjusted rates paint a completely different scenario: Mixed & Black females come in first place, followed by Mixed & Black males, White females, and White males, which again goes against the evidence presented in this study.

Following the path proposed by former studies, the application of the Bennett & Horiuchi method with uniform age-specific growth rates across races yields different results; White females have the highest life expectancy at birth, followed by White males, Mixed & Black females, with Mixed & Black males in last place. These results are the most consistent with previous research and are in line with the hypothesis presented at the beginning of this thesis. Moreover, they are consistent with the literature discussing mortality and health differences by race in Brazil, which shows a systematic disadvantage for Non-White individuals compared to their White counterparts.

These findings contribute to the evolution of research on the topic of racial differences in mortality; so far, no studies attempted to decompose racial gaps in life expectancy by race. Moreover, despite being widely used to assess data quality of death counts, DDMs had not been applied to subpopulations defined by racial membership in Brazil and despite the fact that the results have significant limitations, raising awareness and discussing the reasons why these methods might not be adequate for this type of data is important. Finally, by presenting different methods to estimate life expectancy at birth, this thesis emphasizes how results can be strongly affected by the choice of method; thus, results should always be analyzed in light of relevant literature.

Limitations

Initially, it is fundamental to acknowledge that, ideally, population estimates would all be from censuses; however, for the reasons previously stated, a combination of census and survey data was required; the 2019 estimates from PNAD-C are a result of projections, thus, being subjected to the assumptions made, which can differ from the actual population when counted for a census. Furthermore, death counts also come from a different source – SIM; this could affect race-age-specific mortality rates. Condran et al. (1991) show that age-specific mortality rates are more robust in countries where mortality and population data come from the same source.

As for missing information, no imputations were made to the data, which implies that any records with missing information were excluded during the analysis. This has little effect on variables such as age or gender since the number of records without any information on either variable is significantly small. As for race, death records present a higher share of "missings" – around 3.15%; however, it is beyond the

scope of this thesis to establish a criterion so that race was imputed when missing. The choice to remove death records without this information was based on two facts: first, the share of death records without this information available went down from 5.21% in 2010 to 1.55% in 2019; second, except for the first age group, the share of "missings" is constant throughout the age range, which indicates that excluding these death records would not distort mortality unevenly across ages.

Regarding race, limitations emerge from the observed processes of racial reclassification and the whitening effect. Racial reclassification is the process by which individuals move from one racial group to another over time; this effect was described in detail, including its potential drivers in Chapter 3. This process directly impacts the calculations of race-age-specific mortality rates either by increasing or decreasing the denominator; furthermore, it distorted DDMs' estimates of completeness, in particular those from the SEG method, as shown in Chapter 5

The whitening effect is related to data sources that use categorization, such as SIM, where the death counts were obtained from; previous studies showed that when categorization is applied, individuals who would have identified as Mixed & Blacks are likely to be categorized as White. This effect can also distort race-age-specific mortality rates, by increasing the number of White deaths while decreasing the number of deaths from Mixed & Blacks, which would systematically increase race-age-specific mortality rates for Whites and decrease it for Mixed & Blacks.

Finally, it is important to emphasize that the estimates of life expectancy at birth by race are not evidence of causality; that is, despite the results showing that there are differences in the age distribution of racial gaps in life expectancy at birth, there is no evidence that the gaps themselves are caused by the fact that individuals pertain to a certain group. Race should be treated as a risk marker, not a risk factor itself (Travassos & Williams, 2004).

Conclusion and Opportunities for future research

In conclusion, despite the limitations imposed by the data and the methods, it was possible to produce consistent estimates for life expectancy at birth for the White and Mixed & Black racial groups for the 2010 to 2019 period using data from the state of São Paulo, Brazil. The findings of this study show that there is a race gap in life expectancy at birth, where Whites hold the advantage against Mixed & Blacks; furthermore, the decomposition highlighted the age pattern differences in mortality conditions for Whites and Mixed & Blacks.

Researchers could take many different directions when further exploring this topic. There are numerous avenues to explore, so the possibilities are broad. First, due to data quality limitations, this thesis focused on investigating race-related mortality differentials in São Paulo between 2010 and 2029; further research could expand the study to the national level. This would, naturally, require data quality to be

assessed at the sub-national level due to the country's regional diversity and variability regarding data collection.

One could also focus on estimating age-specific racial reclassification rates; this would be particularly helpful in assessing death counts' completeness, serving as "migration-like" rates, which would allow for applying the DDM method proposed by Hill and Queiroz (2010). Another possibility would be to evaluate death records to observe if there is, in fact, a mismatch between identified and categorized races where individuals who identified as either Mixed or Black are moved towards the whiter end of the color gradient, similarly to what is observed in populations; moreover, producing correction factors for this potential whitening would serve as a means to improve mortality estimates by race.

Additional lines of research also include finding ways around the numerator-denominator bias to avoid potential distortions on race-age-specific mortality rates and using other life table methods different from those used here. Furthermore, a more detailed analysis of said biases could allow estimates to be presented separately for Mixed and Blacks. Finally, despite the limitations associated with cause-of-death data, a decomposition of life expectancy at birth gaps would enrich the discussion on racial inequalities.

References

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., Li, J., & Zand, R. (2021). Racial, Economic, and Health Inequality and COVID-19 Infection in the United States. *Journal of Racial and Ethnic Health Disparities*, *8*(3), 732–742. https://doi.org/10.1007/s40615-020-00833-4

Agostinho, C. S., & Queiroz, B. L. (2016). Estimativas da mortalidade adulta para o Brasil no período 1980/2000: Uma abordagem metodológica comparativa. *Anais*, *0*, Article 0.

Arriaga, E. E. (1984). Measuring and explaining the change in life expectancies. *Demography*, *21*(1), 83–96. https://doi.org/10.2307/2061029

Baptista, E., Queiroz, B., & Lima, E. (2022). Regional COVID-19 mortality in Brazil by age. *Vienna Yearbook of Population Research*, *20*, 1–17. https://doi.org/10.1553/populationyearbook2022.res2.3

Batista, L. E., Escuder, M. M. L., & Pereira, J. C. R. (2004). A cor da morte: Causas de óbito segundo características de raça no Estado de São Paulo, 1999 a 2001. *Revista de Saúde Pública*, *38*, 630–636. https://doi.org/10.1590/S0034-89102004000500003

Bennett, N. G., & Horiuchi, S. (1981). Estimating the Completeness of Death Registration in a Closed Population. *Population Index*, *47*(2), 207–221. https://doi.org/10.2307/2736447

Bennett, N. G., & Horiuchi, S. (1984). Mortality estimation from registered deaths in less developed countries. *Demography*, *21*(2), 217–233. https://doi.org/10.2307/2061041

Brass, W. (1975). Methods for estimating fertility and mortality from limited and defective data. *Methods for Estimating Fertility and Mortality from Limited and Defective Data.*

Braveman, P. A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M., & Posner, S. (2005). Socioeconomic Status in Health ResearchOne Size Does Not Fit All. *JAMA*, *294*(22), 2879–2888. https://doi.org/10.1001/jama.294.22.2879

Carvalho, J. A. M. de, Wood, C. H., & Andrade, F. C. D. (2004). Estimating the stability of census-based racial/ethnic classifications: The case of Brazil. *Population Studies*, *58*(3), 331–343. https://doi.org/10.1080/0032472042000272375

Castanheira, H. C., & Monteiro da Silva, J. H. C. (2022). Examining sex differences in the completeness of Peruvian CRVS data and adult mortality estimates. *Genus*, *78*(1), 3. https://doi.org/10.1186/s41118-021-00151-5

Castro, M. C., Gurzenda, S., Turra, C. M., Kim, S., Andrasfay, T., & Goldman, N. (2021). Reduction in life expectancy in Brazil after COVID-19. *Nature Medicine*, *27*(9), Article 9. https://doi.org/10.1038/s41591-021-01437-z

Chiavegatto Filho, A. D. P., Beltrán-Sánchez, H., & Kawachi, I. (2014). Racial Disparities in Life Expectancy in Brazil: Challenges From a Multiracial Society. *American Journal of Public Health*, *104*(11), 2156–2162. https://doi.org/10.2105/AJPH.2013.301565

Chor, D., & Lima, C. R. de A. (2005). Aspectos epidemiológicos das desigualdades raciais em saúde no Brasil. *Cadernos de Saúde Pública*, *21*(5), 1586–1594. https://doi.org/10.1590/S0102-311X2005000500033

Condran, G. A., Himes, C. L., & Preston, S. H. (1991). Old-age mortality patterns in low-mortality countries: An evaluation of population and death data at advanced ages, 1950 to the present. *Population Bulletin of the United Nations*, *30*, 23–60.

Cunha, E. S. D., Muniz, J. O., Jakob, A. A. E., & Cunha, J. (2016, November 18). *Tábuas de vida por raça/cor: Uma comparação metodológica*. https://www.semanticscholar.org/paper/T%C3%A1buas-de-vida-por-ra%C3%A7a-cor%3A-uma-compara%C3%A7%C3%A3o-Cunha-Muniz/d5297246ca3b12d5cd65cfee136e7ded912e5b3d

da Silva Simões, C. C. (2016). *Relações entre as alterações históricas na dinâmica demográfica brasileira e os impactos decorrentes do processo de envelhecimento da população*. IBGE, Instituto Brasileiro de Geografia e Estatística.

Davis, K. (1945). The World Demographic Transition. *The ANNALS of the American Academy of Political and Social Science*, *237*(1), 1–11. https://doi.org/10.1177/000271624523700102

Diógenes, V. H. D., Júnior, E. P. P., Gonzaga, M. R., Queiroz, B. L., Lima, E. E. C., Costa, L. C. C. da, Rocha, A. S., Ferreira, A. J. F., Teixeira, C. S. S., Alves, F. J. O., Rameh, L., Flores-Ortiz, R., Leyland, A., Dundas, R., Barreto, M. L., & Ichihara, M. Y. T. (2022). Differentials in death count records by databases in Brazil in 2010. *Revista de Saúde Pública*, *56*, 92–92. https://doi.org/10.11606/s1518-8787.2022056004282

Dwyer-Lindgren, L., Kendrick, P., Kelly, Y. O., Sylte, D. O., Schmidt, C., Blacker, B. F., Daoud, F., Abdi, A. A., Baumann, M., & Mouhanna, F. (2022). Life expectancy by county, race, and ethnicity in the USA, 2000–19: A systematic analysis of health disparities. *The Lancet*, *400*(10345), 25–38.

Foz, G. de. (2021). *Métodos demográficos: Uma visão desde os países de língua portuguesa*. Blucher Open Access.

Frenk, J., Frejka, T., Bobadilla, J. L., Stern, C., Lozano, R., Sepúlveda, J., & José, M. (1991). La transición epidemiológica en América Latina. *Boletín de la Oficina Sanitaria Panamericana (OSP)*, 485–496.

Glei, D. A., Barbieri, M., & Santamaría-Ulloa, C. (2019). Costa Rican mortality 1950-2013: An evaluation of data quality and trends compared with other countries. *Demographic Research*, *40*, 835.

Glei, D. A., Paz, A. B., Aburto, J. M., & Barbieri, M. (2021). Mexican mortality 1990–2016. *Demographic Research*, *44*, 719–758.

Goes, E. F., Menezes, G. M. S., Almeida, M.-C. C., Araújo, T. V. B. de, Alves, S. V., Alves, M. T. S. S. B. e, & Aquino, E. M. L. (2020). Racial vulnerability and individual barriers for Brazilian women seeking first care following abortion. *Cadernos de Saúde Pública*, *36*, e00189618.

Hill, K. (1987). Estimating census and death registration completeness. *Asian and Pacific Population Forum/East-West Population Institute, East-West Center*, *1*(3), 8–13.

Hill, K., & Queiroz, B. (2010). Adjusting the general growth balance method for migration. *Revista Brasileira de Estudos de População*, *27*, 7–20. https://doi.org/10.1590/S0102-30982010000100002

Hill, K., You, D., & Choi, Y. (2009). Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data errors. *Demographic Research*, *21*, 235–254.

Horiuchi, S. (1997). Epidemiological transitions in human history. *Health and Mortality Issues of Global Concern*, 54–71.

IBGE. (2022). *Desigualdades sociais por cor ou raça no Brasil*. https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101972

Johnson, C. O., Boon-Dooley, A. S., DeCleene, N. K., Henny, K. F., Blacker, B. F., Anderson, J. A., Afshin, A., Aravkin, A., Cunningham, M. W., Dieleman, J. L., Feldman, R. G., Gakidou, E., Mokdad, A. H., Naghavi, M., Spencer, C. N., Whisnant, J. L., York, H. W., Zende, R. R., Zheng, P., … Roth, G. A. (2022). Life Expectancy for White, Black, and Hispanic Race/Ethnicity in U.S. States: Trends and Disparities, 1990 to 2019. *Annals of Internal Medicine*, *175*(8), 1057–1064. https://doi.org/10.7326/M21-3956

Jorge, M. H. P. de M., Laurenti, R., & Gotlieb, S. L. D. (2007). Análise da qualidade das estatísticas vitais brasileiras: A experiência de implantação do SIM e do SINASC. *Ciência & Saúde Coletiva*, *12*, 643–654. https://doi.org/10.1590/S1413-81232007000300014

Lancet, T. (2022). Advancing racial and ethnic equity in health. *The Lancet*, *400*(10368), 2007. https://doi.org/10.1016/S0140-6736(22)02533-8

Leal, M. do C., Gama, S. G. N. da, & Cunha, C. B. da. (2005). Racial, sociodemographic, and prenatal and childbirth care inequalities in Brazil, 1999-2001. *Revista de Saúde Pública*, *39*, 100–107. https://doi.org/10.1590/S0034-89102005000100013

Lima, E. E. C. de, & Queiroz, B. L. (2014). Evolution of the deaths registry system in Brazil: Associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. *Cadernos de Saúde Pública*, *30*, 1721–1730. https://doi.org/10.1590/0102-311X00131113

Lima, E. E. C., Vilela, E. A., Peralta, A., Rocha, M., Queiroz, B. L., Gonzaga, M. R., Piscoya-Díaz, M., Martinez-Folgar, K., García-Guerrero, V. M., & Freire, F. H. M. A. (2021). Investigating regional excess mortality during 2020 COVID-19 pandemic in selected Latin American countries. *Genus*, *77*(1), 30. https://doi.org/10.1186/s41118-021-00139-1

Lotufo, P. A., Goulart, A. C., & Bensenor, I. M. (2007). Race, gender and stroke subtypes mortality in São Paulo, Brazil. *Arquivos de Neuro-Psiquiatria*, *65*, 752–757. https://doi.org/10.1590/S0004-282X2007000500004

Martins, I. R., Costa, S. H., Freitas, S. R., & Pinto, C. S. (1991). [Induced abortion among low income women: Dimensions of the problem]. *Cadernos de saude publica*, *7*(2), 251–266. https://doi.org/10.1590/s0102-311x1991000200009

Mathers, C. D., Fat, D. M., Inoue, M., Rao, C., & Lopez, A. D. (2005). Counting the dead and what they died from: An assessment of the global status of cause of death data. *Bulletin of the World Health Organization*, *83*(3), 171–177.

Matijasevich, A., Victora, C. G., Barros, A. J. D., Santos, I. S., Marco, P. L., Albernaz, E. P., & Barros, F. C. (2008). Widening Ethnic Disparities in Infant Mortality in Southern Brazil: Comparison of 3 Birth Cohorts. *American Journal of Public Health*, *98*(4), 692–698. https://doi.org/10.2105/AJPH.2006.093492

Merli, M. G. (1998). Mortality in Vietnam, 1979–1989. *Demography*, *35*(3), 345–360. https://doi.org/10.2307/3004042

Miranda, V. (2015). A resurgence of black identity in Brazil? Evidence from an analysis of recent censuses. *Demographic Research*, *32*(59), 1603–1630. https://doi.org/10.4054/DemRes.2015.32.59

Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K., Timæus, I. M., & Zaba, B. (2013). *Tools for demographic estimation*. International Union for the Scientific Study of Population.

Nature. (2023). Why Nature is updating its advice to authors on reporting race or ethnicity. *Nature*, *616*(7956), 219–219. https://doi.org/10.1038/d41586-023-00973-7

Nelson, A. (2002). Unequal treatment: Confronting racial and ethnic disparities in health care. *Journal of the National Medical Association*, *94*(8), 666.

Nogueira, O. (1985). *Tanto preto quanto branco: Estudos de relações raciais*. https://www.semanticscholar.org/paper/Tanto-preto-quanto-branco-%3A-estudos-de-rela%C3%A7%C3%B5es-Nogueira/9b26968ba1b15c963383440df595292481103518

Notestein, F. W. (1945). Population-The long view. *Food for the World.*, 36–57.

Olshansky, S. J., & Ault, A. B. (1986). The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases. *The Milbank Quarterly*, *64*(3), 355–391. https://doi.org/10.2307/3350025

Omran, A. R. (1971). The epidemiologic transition. A theory of the epidemiology of population change. *The Milbank Memorial Fund Quarterly*, *49*(4), 509–538.

Omran, A. R. (1998). *The epidemiologic transition theory revisited thirty years later*.

Osório, R. G. (2003). *O sistema classificatório de cor ou raça do IBGE*.

Paes, N. A. (2005). Avaliação da cobertura dos registros de óbitos dos estados brasileiros em 2000. *Revista de Saúde Pública*, *39*(6), 882–890. https://doi.org/10.1590/S0034-89102005000600003

Paes, N. A., & Albuquerque, M. E. E. (1999). Avaliação da qualidade dos dados populacionais e cobertura dos registros de óbitos para as regiões brasileiras. *Revista de Saúde Pública*, *33*(1), 33–43. https://doi.org/10.1590/S0034-89101999000100006

Paixão, M., Rossetto, I., Montovanele, F., & Carvano, L. M. (2010). Relatório anual das desigualdades raciais no Brasil: 2009-2010. *Rio de Janeiro: Garamond*.

Petruccelli, J. L., & Saboia, A. L. (2013). *Características étnico-raciais da população: Classificações e identidades*. Instituto Brasileiro de Geografia e Estatística–IBGE.

Preston, S., Coale, A. J., Trussell, J., & Weinstein, M. (1980). Estimating the completeness of reporting of adult deaths in populations that are approximately stable. *Population Index*, 179–202.

Preston, S. H., & Coale, A. J. (1982). Age Structure, Growth, Attrition, and Accession: A New Synthesis. *Population Index*, *48*(2), 217–259. https://doi.org/10.2307/2735961

Preston, S. H., Elo, I. T., Rosenwaike, I., & Hill, M. (1996). African-american mortality at older ages: Results of a matching study. *Demography*, *33*(2), 193–209. https://doi.org/10.2307/2061872

Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Blackwell Publishers.

Queiroz, B. L., Freire, F. H. M. de A., Gonzaga, M. R., & Lima, E. E. C. de. (2017). Estimativas do grau de cobertura e da mortalidade adulta (45q15) para as unidades da federação no Brasil entre 1980 e 2010. *Revista Brasileira de Epidemiologia*, *20*, 21–33. https://doi.org/10.1590/1980-5497201700050003

Queiroz, B. L., Gonzaga, M. R., Vasconcelos, A. M. N., Lopes, B. T., & Abreu, D. M. X. (2020). Comparative analysis of completeness of death registration, adult mortality and life expectancy at birth in Brazil at the subnational level. *Population Health Metrics*, *18*(1), 11. https://doi.org/10.1186/s12963-020-00213-4

Reichenheim, M. E., De Souza, E. R., Moraes, C. L., De Mello Jorge, M. H. P., Da Silva, C. M. F. P., & De Souza Minayo, M. C. (2011). Violence and injuries in Brazil: The effect, progress made, and challenges ahead. *The Lancet*, *377*(9781), 1962–1975. https://doi.org/10.1016/S0140-6736(11)60053-6

Riffe, T., Lima, E., & Queiroz, B. (2017). *DDM: Death Registration Coverage Estimation* (1.0-0) [Computer software]. https://cran.r-project.org/web/packages/DDM/index.html

RIVA, P., & AG, D. (2018). *A HUMAN RIGHTS-BASED APPROACH TO DATA*. https://bafybeig72tnsn27gyacxz4wm35gpvr6ithcgtxq3fpyj7aovbbmjiunw4.ipfs.w3s.link/GuidanceNoteonApproachtoData.pdf

Santos, J. A. F. (2011). Desigualdade racial de saúde e contexto de classe no Brasil. *Dados*, *54*(1), 05–40. https://doi.org/10.1590/S0011-52582011000100001

Schramm, J. M. D. A., Oliveira, A. F. D., Leite, I. D. C., Valente, J. G., Gadelha, Â. M. J., Portela, M. C., & Campos, M. R. (2004). Transição epidemiológica e o estudo de carga de doença no Brasil. *Ciência & Saúde Coletiva*, *9*(4), 897–908. https://doi.org/10.1590/S1413-81232004000400011

Siegel, J. S., Swanson, D. A., & Shryock, H. S. (Eds.). (2004). *The methods and materials of demography* (2nd ed). Elsevier/Academic Press.

SlaveVoyages. (2008). Emory University, Rice University. http://www.slavevoyages.org

Telles, E. E. (2002). Racial ambiguity among the Brazilian population. *Ethnic and Racial Studies*, *25*(3), 415–441. https://doi.org/10.1080/01419870252932133

Telles, E. E. (2014). Race in another America. In *Race in another America*. Princeton University Press.

Thompson, W. S. (1929). Population. *American Journal of Sociology*, *34*(6), 959–975. https://doi.org/10.1086/214874

Travassos, C., & Williams, D. R. (2004). The concept and measurement of race and their relationship to public health: A review focused on Brazil and the United States. *Cadernos de Saúde Pública*, *20*, 660–678. https://doi.org/10.1590/S0102-311X2004000300003

United Nations. (1983). *Manual X: Indirect techniques for demographic estimation*.

Vallin, J., & Meslé, F. (2004). Convergences and divergences in mortality: A new approach of health transition. *Demographic Research*, *Special 2*, 11–44. https://doi.org/10.4054/DemRes.2004.S2.2

Wachter, K. W. (2014). *Essential demographic methods*. Harvard University Press.

Williams, D. R., Lavizzo-Mourey, R., & Warren, R. C. (1994). The concept of race and health status in America. *Public Health Reports*, *109*(1), 26–41.

Williams, D. R., Priest, N., & Anderson, N. B. (2016). Understanding associations among race, socioeconomic status, and health: Patterns and prospects. *Health Psychology*, *35*(4), 407–411. https://doi.org/10.1037/hea0000242

Wong, L. L. R., Barros, J. S., Santos, W., Wong, L. L. R., & Sánchez, J. A. (2014). Mortalidad infantil e infantojuvenil en Brasil según sexo y color de la piel. *Wong LLR, Sánchez JA, Organizadores. Situación de La Población Afrodescendiente e Indígena En América Latina–Puntos de Reflexión Para El Debate Sobre Cairo*, *20*, 79–98.

Wood, C. H., & Carvalho, J. A. M. de. (1994). Categorias do censo e classificação subjetiva de cor no Brasil. *Revista Brasileira de Estudos de População*, *11*(1), Article 1.

Appendix A: Abstract in English and German

Abstract

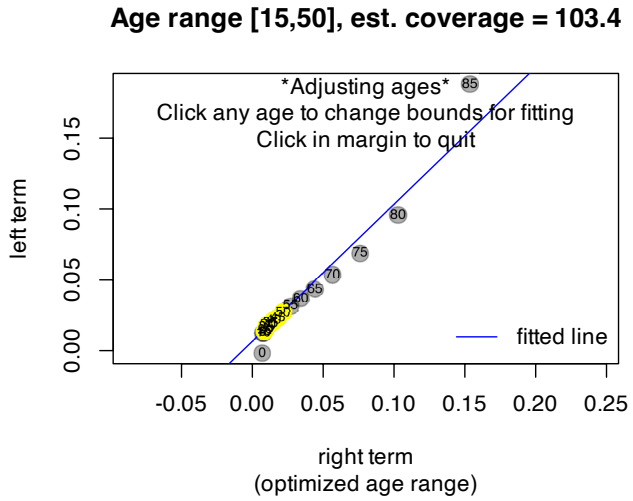
A recent survey revealed that 81% of Brazilians believe the country is racist; moreover, 44% agree that race is the main factor creating inequalities in Brazil. Despite such astonishing numbers, the literature on the demographic implications of racial differences needs to be more extensive. This thesis addresses a gap in the literature by presenting more recent life expectancy estimates by race and by decomposing life expectancy gaps by age groups. A combination of methods was used to estimate life expectancy at birth by race; completeness of death counts was assessed using death Distribution Methods (DDMs), and life expectancy was estimated using two approaches, first, by combining the results from the DDMs with Standard life table techniques and then using the Bennett & Horiuchi method. Finally, gaps in life expectancy at birth were decomposed using Arriaga's decomposition method. The results show that life expectancy at birth varies significantly by race, with the variations being attributed to different age groups when examined separately across genders; moreover, the analysis of the gender gap across racial groups shows variations similar to those observed in the decomposition of the race gap. Additional results indicate that DDMs are not adequate to evaluate the completeness of death counts by racial groups due to racial reclassification. Finally, for this particular context, life expectancy estimates obtained using the Bennett & Horiuchi method are preferred over estimates from Standard life table techniques, even when combined with DDMs. The results confirm previous studies that investigated differences in the cause-of-death profile and mortality and health conditions of Brazilians according to their identified race. Significant disadvantages in life expectancy at birth for Mixed & Blacks when compared to Whites were observed.

Abstrakt

Eine kürzlich durchgeführte Umfrage ergab, dass 81 % der Brasilianer glauben, dass das Land rassistisch ist; außerdem sind 44 % der Meinung, dass die Rasse der Hauptfaktor für die Ungleichheiten in Brasilien ist. Trotz dieser erstaunlichen Zahlen muss die Literatur über die demografischen Auswirkungen von Rassenunterschieden umfangreicher sein. Diese Arbeit schließt eine Lücke in der Literatur, indem sie neuere Schätzungen der Lebenserwartung nach Rasse vorlegt und die Unterschiede in der Lebenserwartung nach Altersgruppen aufschlüsselt. Zur Schätzung der Lebenserwartung bei der Geburt nach Rasse wurde eine Kombination von Methoden verwendet; die Vollständigkeit der Sterbefälle wurde mit Hilfe von Todesverteilungsmethoden (DDM) bewertet, und die Lebenserwartung wurde mit zwei Ansätzen geschätzt, zunächst durch Kombination der Ergebnisse aus den DDM mit Standard-Lebenszeittabellen und dann mit der Methode von Bennett & Horiuchi. Schließlich wurden die Lücken in der Lebenserwartung bei der Geburt mit Hilfe der Arriaga-Zerlegungsmethode zerlegt. Die Ergebnisse zeigen, dass die Lebenserwartung bei der Geburt je nach Rasse erheblich variiert, wobei die Variationen auf verschiedene Altersgruppen zurückzuführen sind, wenn sie nach Geschlechtern getrennt untersucht werden; darüber hinaus zeigt die Analyse des geschlechtsspezifischen Unterschieds zwischen den Rassengruppen ähnliche Variationen, wie sie bei der Zerlegung des rassischen Unterschieds beobachtet wurden. Weitere Ergebnisse deuten darauf hin, dass DDMs nicht geeignet sind, um die Vollständigkeit der Todesfälle nach Rassengruppen zu bewerten, was auf die Neuklassifizierung nach Rassen zurückzuführen ist. Schließlich sind in diesem speziellen Kontext die Schätzungen der Lebenserwartung nach der Methode von Bennett und Horiuchi den Schätzungen nach den Standard-Lebenszeittabellen vorzuziehen, selbst wenn sie mit DDMs kombiniert werden. Die Ergebnisse bestätigen frühere Studien, in denen Unterschiede im Profil der Todesursachen, der Sterblichkeit und des Gesundheitszustands von Brasilianern je nach ihrer identifizierten Rasse untersucht wurden. Es wurden erhebliche Nachteile bei der Lebenserwartung bei der Geburt für Gemischte und Schwarze im Vergleich zu Weißen festgestellt.

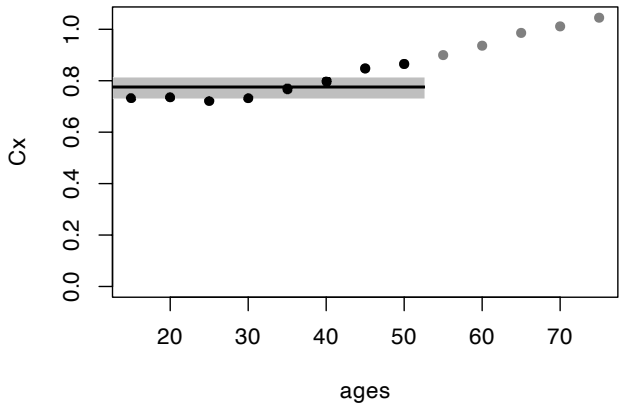
Appendix B: Diagnostics Plots for DDM

Figure B.1 – Diagnostics Plot: GGB applied to White Females



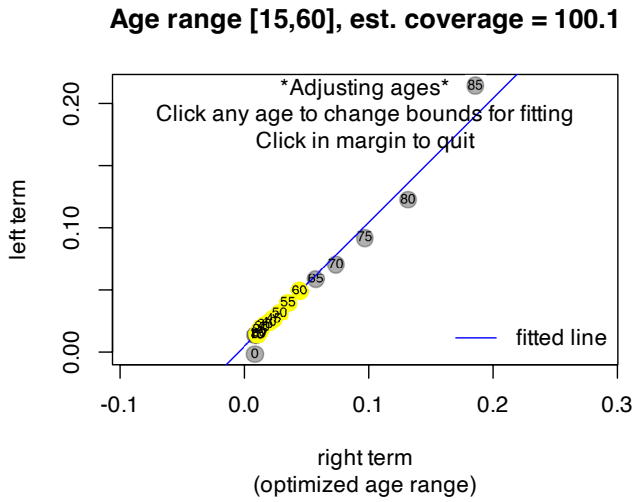
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.2 – Diagnostics Plot: SEG applied to White Females



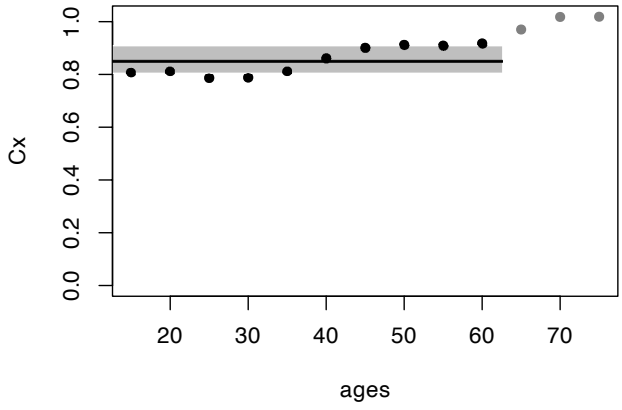
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.3 – Diagnostics Plot: GGB applied to White Males



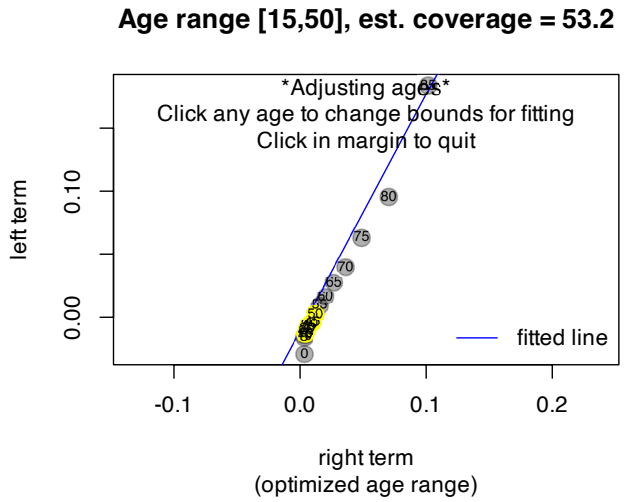
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.4 – Diagnostics Plot: SEG applied to White Males



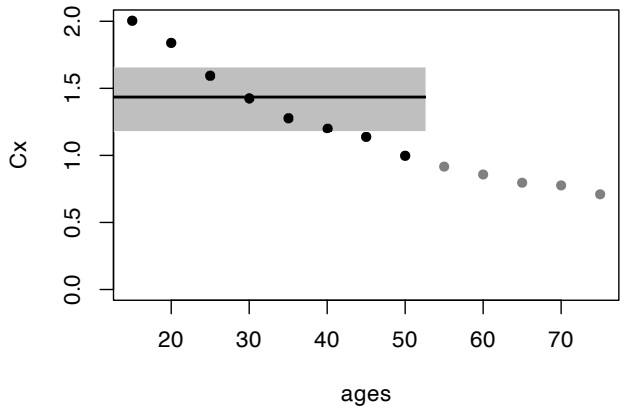
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.5 – Diagnostics Plot: GGB applied to Mixed & Black Females



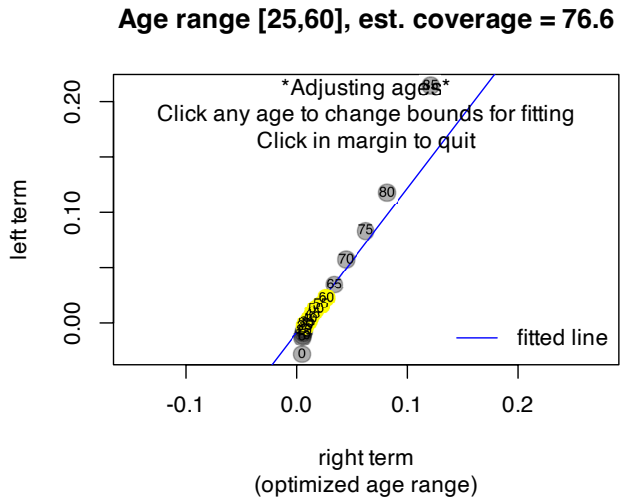
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.6 – Diagnostics Plot: SEG applied to Mixed & Black Females



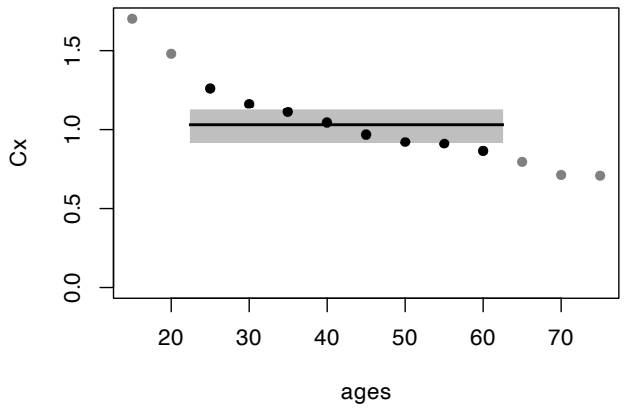
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.7 – Diagnostics Plot: GGB applied to Mixed & Black Males



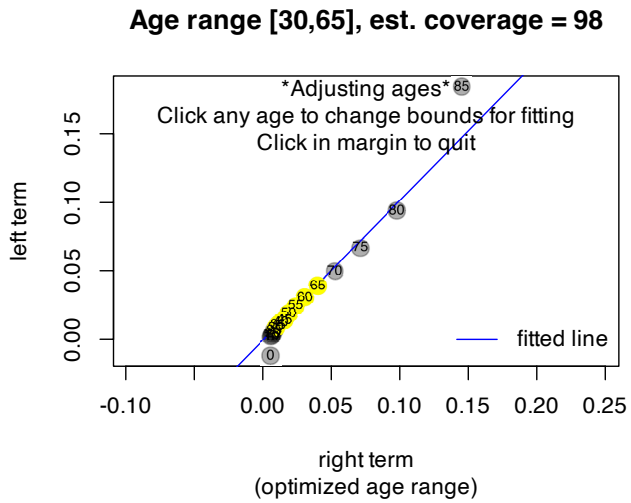
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.8 – Diagnostics Plot: SEG applied to Mixed & Black Males



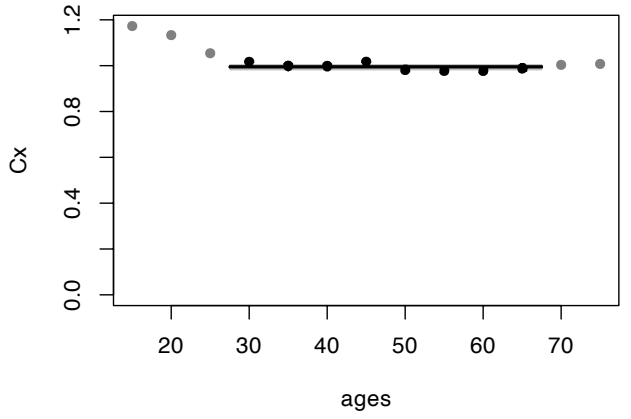
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.9 – Diagnostics Plot: GGB applied to Females



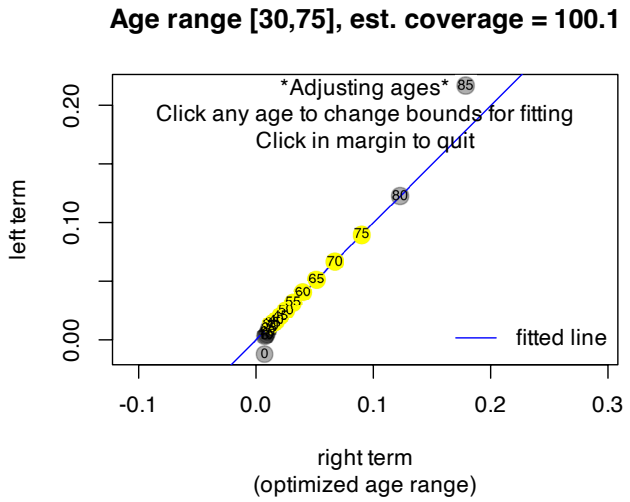
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.10 – Diagnostics Plot: SEG applied to Females



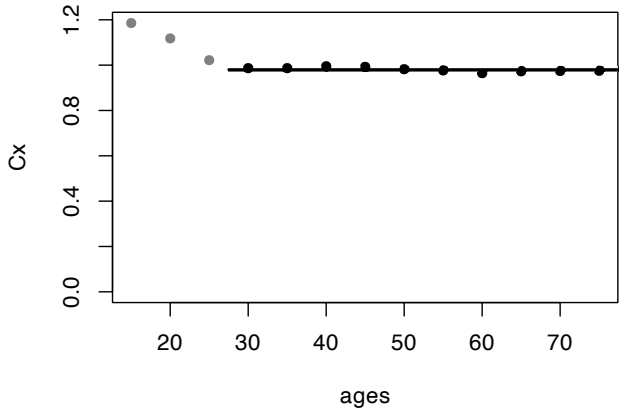
Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.11 – Diagnostics Plot: GGB applied to Males



Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Figure B.12 – Diagnostics Plot: SEG applied to Males



Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Appendix C: Life Tables

Table C.1 – Life Table for White Females

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0027	1.50	0.0107	0.9893	1.0000	0.0107	3.9731	77.9465	77.95
5-9	0.0002	2.50	0.0010	0.9990	0.9893	0.0009	4.9439	73.9734	74.78
10-14	0.0002	2.50	0.0011	0.9989	0.9883	0.0011	4.9387	69.0295	69.85
15-19	0.0004	2.50	0.0021	0.9979	0.9872	0.0021	4.9306	64.0908	64.92
20-24	0.0005	2.50	0.0025	0.9975	0.9851	0.0025	4.9191	59.1602	60.06
25-29	0.0006	2.50	0.0030	0.9970	0.9826	0.0029	4.9055	54.2412	55.20
30-34	0.0008	2.50	0.0041	0.9959	0.9796	0.0040	4.8882	49.3356	50.36
35-39	0.0012	2.50	0.0059	0.9941	0.9756	0.0057	4.8639	44.4474	45.56
40-44	0.0017	2.50	0.0085	0.9915	0.9699	0.0083	4.8289	39.5835	40.81
45-49	0.0028	2.50	0.0137	0.9863	0.9616	0.0132	4.7753	34.7546	36.14
50-54	0.0041	2.50	0.0201	0.9799	0.9485	0.0190	4.6948	29.9794	31.61
55-59	0.0064	2.50	0.0317	0.9683	0.9294	0.0295	4.5735	25.2846	27.20
60-64	0.0099	2.50	0.0482	0.9518	0.9000	0.0434	4.3913	20.7111	23.01
65-69	0.0157	2.50	0.0754	0.9246	0.8566	0.0646	4.1214	16.3198	19.05
70-74	0.0230	2.50	0.1090	0.8910	0.7920	0.0863	3.7442	12.1984	15.40
75-79	0.0399	2.50	0.1812	0.8188	0.7057	0.1279	3.2087	8.4542	11.98
80-84	0.0627	2.50	0.2710	0.7290	0.5778	0.1566	2.4976	5.2455	9.08
85+	0.1533	6.52	1.0000	0.0000	0.4212	0.4212	2.7479	2.7479	6.52

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.1.1 – Life Table for White Females

Standard Method (Adjusted nMx – GGB)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0027	1.50	0.0107	0.9893	1.0000	0.0107	3.9731	77.9465	77.95
5-9	0.0002	2.50	0.0010	0.9990	0.9893	0.0009	4.9439	73.9734	74.78
10-14	0.0002	2.50	0.0011	0.9989	0.9883	0.0011	4.9387	69.0295	69.85
15-19	0.0004	2.50	0.0021	0.9979	0.9872	0.0021	4.9306	64.0908	64.92
20-24	0.0005	2.50	0.0025	0.9975	0.9851	0.0025	4.9191	59.1602	60.06
25-29	0.0006	2.50	0.0030	0.9970	0.9826	0.0029	4.9055	54.2412	55.20
30-34	0.0008	2.50	0.0041	0.9959	0.9796	0.0040	4.8882	49.3356	50.36
35-39	0.0012	2.50	0.0059	0.9941	0.9756	0.0057	4.8639	44.4474	45.56
40-44	0.0017	2.50	0.0085	0.9915	0.9699	0.0083	4.8289	39.5835	40.81
45-49	0.0028	2.50	0.0137	0.9863	0.9616	0.0132	4.7753	34.7546	36.14
50-54	0.0041	2.50	0.0201	0.9799	0.9485	0.0190	4.6948	29.9794	31.61
55-59	0.0064	2.50	0.0317	0.9683	0.9294	0.0295	4.5735	25.2846	27.20
60-64	0.0099	2.50	0.0482	0.9518	0.9000	0.0434	4.3913	20.7111	23.01
65-69	0.0157	2.50	0.0754	0.9246	0.8566	0.0646	4.1214	16.3198	19.05
70-74	0.0230	2.50	0.1090	0.8910	0.7920	0.0863	3.7442	12.1984	15.40
75-79	0.0399	2.50	0.1812	0.8188	0.7057	0.1279	3.2087	8.4542	11.98
80-84	0.0627	2.50	0.2710	0.7290	0.5778	0.1566	2.4976	5.2455	9.08
85+	0.1533	6.52	1.0000	0.0000	0.4212	0.4212	2.7479	2.7479	6.52

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.1.2 – Life Table for White Females

Standard Method (Adjusted nMx – SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0036	1.50	0.0143	0.9857	1.0000	0.0143	3.9644	74.4789	74.48
5-9	0.0003	2.50	0.0013	0.9987	0.9857	0.0012	4.9256	70.5145	71.53
10-14	0.0003	2.50	0.0015	0.9985	0.9845	0.0015	4.9187	65.5889	66.62
15-19	0.0006	2.50	0.0029	0.9971	0.9830	0.0028	4.9080	60.6702	61.72
20-24	0.0007	2.50	0.0033	0.9967	0.9802	0.0033	4.8927	55.7622	56.89
25-29	0.0008	2.50	0.0040	0.9960	0.9769	0.0039	4.8749	50.8695	52.07
30-34	0.0011	2.50	0.0054	0.9946	0.9730	0.0053	4.8520	45.9946	47.27
35-39	0.0016	2.50	0.0078	0.9922	0.9678	0.0076	4.8199	41.1426	42.51
40-44	0.0023	2.50	0.0113	0.9887	0.9602	0.0109	4.7739	36.3226	37.83
45-49	0.0037	2.50	0.0182	0.9818	0.9493	0.0172	4.7036	31.5487	33.23
50-54	0.0054	2.50	0.0266	0.9734	0.9321	0.0248	4.5986	26.8451	28.80
55-59	0.0086	2.50	0.0419	0.9581	0.9073	0.0381	4.4415	22.2465	24.52
60-64	0.0131	2.50	0.0636	0.9364	0.8693	0.0553	4.2082	17.8050	20.48
65-69	0.0208	2.50	0.0990	0.9010	0.8140	0.0806	3.8686	13.5968	16.70
70-74	0.0306	2.50	0.1423	0.8577	0.7334	0.1043	3.4063	9.7282	13.26
75-79	0.0530	2.50	0.2339	0.7661	0.6291	0.1471	2.7776	6.3219	10.05
80-84	0.0833	2.50	0.3448	0.6552	0.4820	0.1662	1.9944	3.5443	7.35
85+	0.2037	4.91	1.0000	0.0000	0.3158	0.3158	1.5499	1.5499	4.91

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.1.3 – Life Table for White Females

Standard Method (Adjusted nMx – GGB-SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0027	1.50	0.0107	0.9893	1.0000	0.0107	3.9731	77.9465	77.95
5-9	0.0002	2.50	0.0010	0.9990	0.9893	0.0009	4.9439	73.9734	74.78
10-14	0.0002	2.50	0.0011	0.9989	0.9883	0.0011	4.9387	69.0295	69.85
15-19	0.0004	2.50	0.0021	0.9979	0.9872	0.0021	4.9306	64.0908	64.92
20-24	0.0005	2.50	0.0025	0.9975	0.9851	0.0025	4.9191	59.1602	60.06
25-29	0.0006	2.50	0.0030	0.9970	0.9826	0.0029	4.9055	54.2412	55.20
30-34	0.0008	2.50	0.0041	0.9959	0.9796	0.0040	4.8882	49.3356	50.36
35-39	0.0012	2.50	0.0059	0.9941	0.9756	0.0057	4.8639	44.4474	45.56
40-44	0.0017	2.50	0.0085	0.9915	0.9699	0.0083	4.8289	39.5835	40.81
45-49	0.0028	2.50	0.0137	0.9863	0.9616	0.0132	4.7753	34.7546	36.14
50-54	0.0041	2.50	0.0201	0.9799	0.9485	0.0190	4.6948	29.9794	31.61
55-59	0.0064	2.50	0.0317	0.9683	0.9294	0.0295	4.5735	25.2846	27.20
60-64	0.0099	2.50	0.0482	0.9518	0.9000	0.0434	4.3913	20.7111	23.01
65-69	0.0157	2.50	0.0754	0.9246	0.8566	0.0646	4.1214	16.3198	19.05
70-74	0.0230	2.50	0.1090	0.8910	0.7920	0.0863	3.7442	12.1984	15.40
75-79	0.0399	2.50	0.1812	0.8188	0.7057	0.1279	3.2087	8.4542	11.98
80-84	0.0627	2.50	0.2710	0.7290	0.5778	0.1566	2.4976	5.2455	9.08
85+	0.1533	6.52	1.0000	0.0000	0.4212	0.4212	2.7479	2.7479	6.52

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.2 – Life Table for White Males

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0033	1.50	0.0129	0.9871	1.0000	0.0129	3.9678	71.1119	71.11
5-9	0.0002	2.50	0.0012	0.9988	0.9871	0.0012	4.9326	67.1441	68.02
10-14	0.0003	2.50	0.0015	0.9985	0.9859	0.0015	4.9259	62.2115	63.10
15-19	0.0012	2.50	0.0059	0.9941	0.9844	0.0058	4.9077	57.2856	58.19
20-24	0.0016	2.50	0.0080	0.9920	0.9786	0.0078	4.8736	52.3779	53.52
25-29	0.0016	2.50	0.0082	0.9918	0.9708	0.0079	4.8342	47.5044	48.93
30-34	0.0019	2.50	0.0097	0.9903	0.9629	0.0093	4.7910	42.6702	44.32
35-39	0.0025	2.50	0.0126	0.9874	0.9535	0.0120	4.7377	37.8792	39.72
40-44	0.0036	2.50	0.0179	0.9821	0.9416	0.0168	4.6658	33.1414	35.20
45-49	0.0055	2.50	0.0270	0.9730	0.9248	0.0250	4.5612	28.4756	30.79
50-54	0.0083	2.50	0.0407	0.9593	0.8997	0.0366	4.4072	23.9144	26.58
55-59	0.0124	2.50	0.0600	0.9400	0.8631	0.0518	4.1861	19.5072	22.60
60-64	0.0186	2.50	0.0889	0.9111	0.8113	0.0721	3.8762	15.3211	18.88
65-69	0.0284	2.50	0.1324	0.8676	0.7392	0.0979	3.4511	11.4449	15.48
70-74	0.0407	2.50	0.1845	0.8155	0.6413	0.1183	2.9106	7.9937	12.47
75-79	0.0614	2.50	0.2660	0.7340	0.5230	0.1391	2.2669	5.0832	9.72
80-84	0.0959	2.50	0.3867	0.6133	0.3838	0.1484	1.5481	2.8162	7.34
85+	0.1856	5.39	1.0000	0.0000	0.2354	0.2354	1.2681	1.2681	5.39

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.2.1 – Life Table for White Males

Standard Method (Adjusted nMx – GGB)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0033	1.50	0.0129	0.9871	1.0000	0.0129	3.9678	71.1119	71.11
5-9	0.0002	2.50	0.0012	0.9988	0.9871	0.0012	4.9326	67.1441	68.02
10-14	0.0003	2.50	0.0015	0.9985	0.9859	0.0015	4.9259	62.2115	63.10
15-19	0.0012	2.50	0.0059	0.9941	0.9844	0.0058	4.9077	57.2856	58.19
20-24	0.0016	2.50	0.0080	0.9920	0.9786	0.0078	4.8736	52.3779	53.52
25-29	0.0016	2.50	0.0082	0.9918	0.9708	0.0079	4.8342	47.5044	48.93
30-34	0.0019	2.50	0.0097	0.9903	0.9629	0.0093	4.7910	42.6702	44.32
35-39	0.0025	2.50	0.0126	0.9874	0.9535	0.0120	4.7377	37.8792	39.72
40-44	0.0036	2.50	0.0179	0.9821	0.9416	0.0168	4.6658	33.1414	35.20
45-49	0.0055	2.50	0.0270	0.9730	0.9248	0.0250	4.5612	28.4756	30.79
50-54	0.0083	2.50	0.0407	0.9593	0.8997	0.0366	4.4072	23.9144	26.58
55-59	0.0124	2.50	0.0600	0.9400	0.8631	0.0518	4.1861	19.5072	22.60
60-64	0.0186	2.50	0.0889	0.9111	0.8113	0.0721	3.8762	15.3211	18.88
65-69	0.0284	2.50	0.1324	0.8676	0.7392	0.0979	3.4511	11.4449	15.48
70-74	0.0407	2.50	0.1845	0.8155	0.6413	0.1183	2.9106	7.9937	12.47
75-79	0.0614	2.50	0.2660	0.7340	0.5230	0.1391	2.2669	5.0832	9.72
80-84	0.0959	2.50	0.3867	0.6133	0.3838	0.1484	1.5481	2.8162	7.34
85+	0.1856	5.39	1.0000	0.0000	0.2354	0.2354	1.2681	1.2681	5.39

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.2.2 – Life Table for White Males

Standard Method (Adjusted nMx – SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0039	1.50	0.0154	0.9846	1.0000	0.0154	3.9614	68.6574	68.66
5-9	0.0003	2.50	0.0014	0.9986	0.9846	0.0014	4.9193	64.6960	65.71
10-14	0.0004	2.50	0.0018	0.9982	0.9832	0.0018	4.9113	59.7767	60.80
15-19	0.0014	2.50	0.0071	0.9929	0.9814	0.0069	4.8895	54.8653	55.91
20-24	0.0019	2.50	0.0096	0.9904	0.9744	0.0093	4.8488	49.9758	51.29
25-29	0.0020	2.50	0.0098	0.9902	0.9651	0.0094	4.8019	45.1270	46.76
30-34	0.0023	2.50	0.0116	0.9884	0.9557	0.0111	4.7505	40.3251	42.20
35-39	0.0030	2.50	0.0150	0.9850	0.9446	0.0142	4.6873	35.5745	37.66
40-44	0.0043	2.50	0.0214	0.9786	0.9304	0.0199	4.6021	30.8872	33.20
45-49	0.0066	2.50	0.0323	0.9677	0.9105	0.0294	4.4788	26.2851	28.87
50-54	0.0100	2.50	0.0486	0.9514	0.8810	0.0428	4.2981	21.8063	24.75
55-59	0.0148	2.50	0.0715	0.9285	0.8382	0.0600	4.0412	17.5082	20.89
60-64	0.0223	2.50	0.1057	0.8943	0.7783	0.0822	3.6857	13.4670	17.30
65-69	0.0340	2.50	0.1567	0.8433	0.6960	0.1091	3.2074	9.7813	14.05
70-74	0.0487	2.50	0.2172	0.7828	0.5869	0.1275	2.6160	6.5739	11.20
75-79	0.0736	2.50	0.3107	0.6893	0.4594	0.1428	1.9403	3.9579	8.61
80-84	0.1149	2.50	0.4464	0.5536	0.3167	0.1414	1.2300	2.0176	6.37
85+	0.2226	4.49	1.0000	0.0000	0.1753	0.1753	0.7877	0.7877	4.49

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.2.3 – Life Table for White Males

Standard Method (Adjusted nMx – GGB-SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0033	1.50	0.0129	0.9871	1.0000	0.0129	3.9678	71.1119	71.11
5-9	0.0002	2.50	0.0012	0.9988	0.9871	0.0012	4.9326	67.1441	68.02
10-14	0.0003	2.50	0.0015	0.9985	0.9859	0.0015	4.9259	62.2115	63.10
15-19	0.0012	2.50	0.0059	0.9941	0.9844	0.0058	4.9077	57.2856	58.19
20-24	0.0016	2.50	0.0080	0.9920	0.9786	0.0078	4.8736	52.3779	53.52
25-29	0.0016	2.50	0.0082	0.9918	0.9708	0.0079	4.8342	47.5044	48.93
30-34	0.0019	2.50	0.0097	0.9903	0.9629	0.0093	4.7910	42.6702	44.32
35-39	0.0025	2.50	0.0126	0.9874	0.9535	0.0120	4.7377	37.8792	39.72
40-44	0.0036	2.50	0.0179	0.9821	0.9416	0.0168	4.6658	33.1414	35.20
45-49	0.0055	2.50	0.0270	0.9730	0.9248	0.0250	4.5612	28.4756	30.79
50-54	0.0083	2.50	0.0407	0.9593	0.8997	0.0366	4.4072	23.9144	26.58
55-59	0.0124	2.50	0.0600	0.9400	0.8631	0.0518	4.1861	19.5072	22.60
60-64	0.0186	2.50	0.0889	0.9111	0.8113	0.0721	3.8762	15.3211	18.88
65-69	0.0284	2.50	0.1324	0.8676	0.7392	0.0979	3.4511	11.4449	15.48
70-74	0.0407	2.50	0.1845	0.8155	0.6413	0.1183	2.9106	7.9937	12.47
75-79	0.0614	2.50	0.2660	0.7340	0.5230	0.1391	2.2669	5.0832	9.72
80-84	0.0959	2.50	0.3867	0.6133	0.3838	0.1484	1.5481	2.8162	7.34
85+	0.1856	5.39	1.0000	0.0000	0.2354	0.2354	1.2681	1.2681	5.39

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.3 – Life Table for Mixed & Black Females

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0018	1.50	0.0072	0.9928	1.0000	0.0072	3.9821	81.7646	81.76
5-9	0.0001	2.50	0.0006	0.9994	0.9928	0.0006	4.9627	77.7824	78.34
10-14	0.0002	2.50	0.0008	0.9992	0.9922	0.0008	4.9592	72.8198	73.39
15-19	0.0003	2.50	0.0017	0.9983	0.9915	0.0017	4.9531	67.8606	68.45
20-24	0.0004	2.50	0.0021	0.9979	0.9898	0.0021	4.9438	62.9074	63.56
25-29	0.0005	2.50	0.0025	0.9975	0.9877	0.0024	4.9325	57.9637	58.68
30-34	0.0007	2.50	0.0036	0.9964	0.9853	0.0036	4.9175	53.0312	53.82
35-39	0.0011	2.50	0.0053	0.9947	0.9817	0.0052	4.8957	48.1137	49.01
40-44	0.0017	2.50	0.0083	0.9917	0.9766	0.0081	4.8626	43.2181	44.26
45-49	0.0025	2.50	0.0126	0.9874	0.9685	0.0122	4.8119	38.3555	39.60
50-54	0.0035	2.50	0.0171	0.9829	0.9563	0.0164	4.7404	33.5436	35.08
55-59	0.0055	2.50	0.0272	0.9728	0.9399	0.0255	4.6355	28.8032	30.65
60-64	0.0081	2.50	0.0395	0.9605	0.9143	0.0361	4.4815	24.1677	26.43
65-69	0.0121	2.50	0.0586	0.9414	0.8783	0.0514	4.2627	19.6862	22.42
70-74	0.0189	2.50	0.0900	0.9100	0.8268	0.0744	3.9479	15.4235	18.65
75-79	0.0282	2.50	0.1318	0.8682	0.7524	0.0991	3.5140	11.4756	15.25
80-84	0.0483	2.50	0.2154	0.7846	0.6532	0.1407	2.9143	7.9616	12.19
85+	0.1015	9.85	1.0000	0.0000	0.5125	0.5125	5.0473	5.0473	9.85

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.3.1 – Life Table for Mixed & Black Females

Standard Method (Adjusted nMx – GGB)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0032	1.50	0.0128	0.9872	1.0000	0.0128	3.9681	73.9149	73.91
5-9	0.0002	2.50	0.0011	0.9989	0.9872	0.0011	4.9334	69.9468	70.85
10-14	0.0003	2.50	0.0014	0.9986	0.9861	0.0014	4.9272	65.0134	65.93
15-19	0.0006	2.50	0.0030	0.9970	0.9848	0.0029	4.9164	60.0862	61.02
20-24	0.0008	2.50	0.0038	0.9962	0.9818	0.0037	4.8998	55.1697	56.19
25-29	0.0009	2.50	0.0044	0.9956	0.9781	0.0043	4.8798	50.2699	51.39
30-34	0.0013	2.50	0.0065	0.9935	0.9738	0.0063	4.8533	45.3901	46.61
35-39	0.0019	2.50	0.0094	0.9906	0.9675	0.0091	4.8148	40.5369	41.90
40-44	0.0030	2.50	0.0147	0.9853	0.9584	0.0141	4.7568	35.7221	37.27
45-49	0.0045	2.50	0.0225	0.9775	0.9443	0.0212	4.6684	30.9653	32.79
50-54	0.0062	2.50	0.0305	0.9695	0.9231	0.0281	4.5450	26.2968	28.49
55-59	0.0099	2.50	0.0481	0.9519	0.8949	0.0431	4.3670	21.7519	24.31
60-64	0.0144	2.50	0.0696	0.9304	0.8519	0.0593	4.1111	17.3849	20.41
65-69	0.0216	2.50	0.1025	0.8975	0.7926	0.0812	3.7598	13.2738	16.75
70-74	0.0338	2.50	0.1557	0.8443	0.7113	0.1107	3.2799	9.5140	13.37
75-79	0.0505	2.50	0.2242	0.7758	0.6006	0.1347	2.6664	6.2340	10.38
80-84	0.0865	2.50	0.3555	0.6445	0.4659	0.1656	1.9157	3.5676	7.66
85+	0.1818	5.50	1.0000	0.0000	0.3003	0.3003	1.6520	1.6520	5.50

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.3.2 – Life Table for Mixed & Black Females

Standard Method (Adjusted nMx – SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0018	1.50	0.0072	0.9928	1.0000	0.0072	3.9821	81.7646	81.76
5-9	0.0001	2.50	0.0006	0.9994	0.9928	0.0006	4.9627	77.7824	78.34
10-14	0.0002	2.50	0.0008	0.9992	0.9922	0.0008	4.9592	72.8198	73.39
15-19	0.0003	2.50	0.0017	0.9983	0.9915	0.0017	4.9531	67.8606	68.45
20-24	0.0004	2.50	0.0021	0.9979	0.9898	0.0021	4.9438	62.9074	63.56
25-29	0.0005	2.50	0.0025	0.9975	0.9877	0.0024	4.9325	57.9637	58.68
30-34	0.0007	2.50	0.0036	0.9964	0.9853	0.0036	4.9175	53.0312	53.82
35-39	0.0011	2.50	0.0053	0.9947	0.9817	0.0052	4.8957	48.1137	49.01
40-44	0.0017	2.50	0.0083	0.9917	0.9766	0.0081	4.8626	43.2181	44.26
45-49	0.0025	2.50	0.0126	0.9874	0.9685	0.0122	4.8119	38.3555	39.60
50-54	0.0035	2.50	0.0171	0.9829	0.9563	0.0164	4.7404	33.5436	35.08
55-59	0.0055	2.50	0.0272	0.9728	0.9399	0.0255	4.6355	28.8032	30.65
60-64	0.0081	2.50	0.0395	0.9605	0.9143	0.0361	4.4815	24.1677	26.43
65-69	0.0121	2.50	0.0586	0.9414	0.8783	0.0514	4.2627	19.6862	22.42
70-74	0.0189	2.50	0.0900	0.9100	0.8268	0.0744	3.9479	15.4235	18.65
75-79	0.0282	2.50	0.1318	0.8682	0.7524	0.0991	3.5140	11.4756	15.25
80-84	0.0483	2.50	0.2154	0.7846	0.6532	0.1407	2.9143	7.9616	12.19
85+	0.1015	9.85	1.0000	0.0000	0.5125	0.5125	5.0473	5.0473	9.85

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.3.3 – Life Table for Mixed & Black Females

Standard Method (Adjusted nMx – GGB-SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0021	1.50	0.0085	0.9915	1.0000	0.0085	3.9788	79.3560	79.36
5-9	0.0001	2.50	0.0007	0.9993	0.9915	0.0007	4.9558	75.3772	76.02
10-14	0.0002	2.50	0.0009	0.9991	0.9908	0.0009	4.9517	70.4213	71.08
15-19	0.0004	2.50	0.0020	0.9980	0.9899	0.0020	4.9445	65.4696	66.14
20-24	0.0005	2.50	0.0025	0.9975	0.9879	0.0025	4.9335	60.5251	61.26
25-29	0.0006	2.50	0.0029	0.9971	0.9855	0.0029	4.9201	55.5916	56.41
30-34	0.0009	2.50	0.0043	0.9957	0.9826	0.0042	4.9024	50.6715	51.57
35-39	0.0012	2.50	0.0062	0.9938	0.9784	0.0061	4.8767	45.7691	46.78
40-44	0.0020	2.50	0.0098	0.9902	0.9723	0.0095	4.8377	40.8924	42.06
45-49	0.0030	2.50	0.0149	0.9851	0.9628	0.0144	4.7780	36.0547	37.45
50-54	0.0041	2.50	0.0203	0.9797	0.9484	0.0192	4.6940	31.2767	32.98
55-59	0.0065	2.50	0.0321	0.9679	0.9292	0.0298	4.5714	26.5827	28.61
60-64	0.0095	2.50	0.0466	0.9534	0.8994	0.0419	4.3921	22.0113	24.47
65-69	0.0143	2.50	0.0690	0.9310	0.8575	0.0592	4.1394	17.6193	20.55
70-74	0.0223	2.50	0.1058	0.8942	0.7983	0.0844	3.7804	13.4799	16.89
75-79	0.0334	2.50	0.1542	0.8458	0.7139	0.1101	3.2941	9.6995	13.59
80-84	0.0572	2.50	0.2502	0.7498	0.6038	0.1511	2.6413	6.4054	10.61
85+	0.1203	8.31	1.0000	0.0000	0.4527	0.4527	3.7641	3.7641	8.31

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.4 – Life Table for Mixed & Black Males

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0020	1.50	0.0079	0.9921	1.0000	0.0079	3.9802	74.8462	74.85
5-9	0.0002	2.50	0.0008	0.9992	0.9921	0.0008	4.9586	70.8660	71.43
10-14	0.0002	2.50	0.0012	0.9988	0.9913	0.0012	4.9536	65.9074	66.48
15-19	0.0013	2.50	0.0066	0.9934	0.9901	0.0065	4.9343	60.9538	61.56
20-24	0.0016	2.50	0.0078	0.9922	0.9836	0.0076	4.8989	56.0195	56.95
25-29	0.0016	2.50	0.0079	0.9921	0.9760	0.0077	4.8605	51.1206	52.38
30-34	0.0020	2.50	0.0099	0.9901	0.9682	0.0095	4.8173	46.2601	47.78
35-39	0.0026	2.50	0.0131	0.9869	0.9587	0.0125	4.7621	41.4429	43.23
40-44	0.0035	2.50	0.0171	0.9829	0.9462	0.0162	4.6903	36.6808	38.77
45-49	0.0050	2.50	0.0247	0.9753	0.9300	0.0230	4.5924	31.9905	34.40
50-54	0.0074	2.50	0.0364	0.9636	0.9070	0.0330	4.4524	27.3981	30.21
55-59	0.0112	2.50	0.0546	0.9454	0.8740	0.0477	4.2506	22.9457	26.25
60-64	0.0148	2.50	0.0716	0.9284	0.8263	0.0591	3.9834	18.6951	22.63
65-69	0.0197	2.50	0.0940	0.9060	0.7671	0.0721	3.6553	14.7117	19.18
70-74	0.0268	2.50	0.1256	0.8744	0.6950	0.0873	3.2567	11.0564	15.91
75-79	0.0448	2.50	0.2014	0.7986	0.6077	0.1224	2.7325	7.7996	12.83
80-84	0.0606	2.50	0.2633	0.7367	0.4853	0.1278	2.1071	5.0671	10.44
85+	0.1208	8.28	1.0000	0.0000	0.3575	0.3575	2.9601	2.9601	8.28

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.4.1 – Life Table for Mixed & Black Males

Standard Method (Adjusted nMx – GGB)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0025	1.50	0.0099	0.9901	1.0000	0.0099	3.9752	71.3776	71.38
5-9	0.0002	2.50	0.0010	0.9990	0.9901	0.0009	4.9481	67.4023	68.08
10-14	0.0003	2.50	0.0016	0.9984	0.9892	0.0015	4.9419	62.4542	63.14
15-19	0.0017	2.50	0.0082	0.9918	0.9876	0.0081	4.9177	57.5123	58.23
20-24	0.0020	2.50	0.0097	0.9903	0.9795	0.0095	4.8735	52.5946	53.70
25-29	0.0020	2.50	0.0099	0.9901	0.9699	0.0096	4.8256	47.7211	49.20
30-34	0.0025	2.50	0.0123	0.9877	0.9603	0.0119	4.7719	42.8955	44.67
35-39	0.0033	2.50	0.0164	0.9836	0.9484	0.0155	4.7034	38.1236	40.20
40-44	0.0043	2.50	0.0214	0.9786	0.9329	0.0200	4.6146	33.4202	35.82
45-49	0.0063	2.50	0.0309	0.9691	0.9129	0.0282	4.4941	28.8056	31.55
50-54	0.0093	2.50	0.0454	0.9546	0.8847	0.0402	4.3230	24.3115	27.48
55-59	0.0141	2.50	0.0680	0.9320	0.8445	0.0575	4.0789	19.9885	23.67
60-64	0.0186	2.50	0.0890	0.9110	0.7871	0.0700	3.7602	15.9096	20.21
65-69	0.0248	2.50	0.1165	0.8835	0.7170	0.0836	3.3762	12.1494	16.94
70-74	0.0336	2.50	0.1552	0.8448	0.6335	0.0983	2.9216	8.7732	13.85
75-79	0.0562	2.50	0.2463	0.7537	0.5352	0.1318	2.3463	5.8516	10.93
80-84	0.0761	2.50	0.3197	0.6803	0.4033	0.1289	1.6944	3.5053	8.69
85+	0.1515	6.60	1.0000	0.0000	0.2744	0.2744	1.8109	1.8109	6.60

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.4.2 – Life Table for Mixed & Black Males

Standard Method (Adjusted nMx – SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0020	1.50	0.0079	0.9921	1.0000	0.0079	3.9802	74.8462	74.85
5-9	0.0002	2.50	0.0008	0.9992	0.9921	0.0008	4.9586	70.8660	71.43
10-14	0.0002	2.50	0.0012	0.9988	0.9913	0.0012	4.9536	65.9074	66.48
15-19	0.0013	2.50	0.0066	0.9934	0.9901	0.0065	4.9343	60.9538	61.56
20-24	0.0016	2.50	0.0078	0.9922	0.9836	0.0076	4.8989	56.0195	56.95
25-29	0.0016	2.50	0.0079	0.9921	0.9760	0.0077	4.8605	51.1206	52.38
30-34	0.0020	2.50	0.0099	0.9901	0.9682	0.0095	4.8173	46.2601	47.78
35-39	0.0026	2.50	0.0131	0.9869	0.9587	0.0125	4.7621	41.4429	43.23
40-44	0.0035	2.50	0.0171	0.9829	0.9462	0.0162	4.6903	36.6808	38.77
45-49	0.0050	2.50	0.0247	0.9753	0.9300	0.0230	4.5924	31.9905	34.40
50-54	0.0074	2.50	0.0364	0.9636	0.9070	0.0330	4.4524	27.3981	30.21
55-59	0.0112	2.50	0.0546	0.9454	0.8740	0.0477	4.2506	22.9457	26.25
60-64	0.0148	2.50	0.0716	0.9284	0.8263	0.0591	3.9834	18.6951	22.63
65-69	0.0197	2.50	0.0940	0.9060	0.7671	0.0721	3.6553	14.7117	19.18
70-74	0.0268	2.50	0.1256	0.8744	0.6950	0.0873	3.2567	11.0564	15.91
75-79	0.0448	2.50	0.2014	0.7986	0.6077	0.1224	2.7325	7.7996	12.83
80-84	0.0606	2.50	0.2633	0.7367	0.4853	0.1278	2.1071	5.0671	10.44
85+	0.1208	8.28	1.0000	0.0000	0.3575	0.3575	2.9601	2.9601	8.28

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.4.3 – Life Table for Mixed & Black Males

Standard Method (Adjusted nMx – GGB-SEG)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0026	1.50	0.0103	0.9897	1.0000	0.0103	3.9742	70.7863	70.79
5-9	0.0002	2.50	0.0010	0.9990	0.9897	0.0010	4.9460	66.8120	67.51
10-14	0.0003	2.50	0.0016	0.9984	0.9887	0.0016	4.9396	61.8660	62.57
15-19	0.0017	2.50	0.0086	0.9914	0.9871	0.0085	4.9144	56.9264	57.67
20-24	0.0020	2.50	0.0101	0.9899	0.9787	0.0099	4.8685	52.0119	53.15
25-29	0.0021	2.50	0.0103	0.9897	0.9687	0.0100	4.8187	47.1434	48.66
30-34	0.0026	2.50	0.0128	0.9872	0.9587	0.0123	4.7629	42.3247	44.15
35-39	0.0034	2.50	0.0170	0.9830	0.9464	0.0161	4.6919	37.5618	39.69
40-44	0.0045	2.50	0.0223	0.9777	0.9303	0.0207	4.5998	32.8700	35.33
45-49	0.0065	2.50	0.0321	0.9679	0.9096	0.0292	4.4749	28.2702	31.08
50-54	0.0097	2.50	0.0472	0.9528	0.8804	0.0416	4.2979	23.7953	27.03
55-59	0.0147	2.50	0.0707	0.9293	0.8388	0.0593	4.0457	19.4974	23.24
60-64	0.0194	2.50	0.0924	0.9076	0.7795	0.0720	3.7175	15.4517	19.82
65-69	0.0257	2.50	0.1209	0.8791	0.7075	0.0856	3.3235	11.7342	16.59
70-74	0.0350	2.50	0.1609	0.8391	0.6219	0.1001	2.8595	8.4106	13.52
75-79	0.0584	2.50	0.2550	0.7450	0.5219	0.1331	2.2766	5.5512	10.64
80-84	0.0792	2.50	0.3304	0.6696	0.3888	0.1285	1.6229	3.2745	8.42
85+	0.1576	6.34	1.0000	0.0000	0.2603	0.2603	1.6517	1.6517	6.34

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: the correction factors were limited to 100%, then applied to all age groups from 0-4 to 85+.

Table C.5 – Life Table for White Females

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	2,481	0.0128	2,481	280,287	0.0089	1,392,749	22,213,982	79.25
5-9	164	0.0065	173	277,805	0.0006	1,388,596	20,821,232	74.95
10-14	200	-0.0165	205	277,633	0.0007	1,387,653	19,432,636	69.99
15-19	402	-0.0023	393	277,428	0.0014	1,386,159	18,044,984	65.04
20-24	522	-0.0031	503	277,036	0.0018	1,383,920	16,658,824	60.13
25-29	667	-0.0012	636	276,532	0.0023	1,381,073	15,274,905	55.24
30-34	919	0.0068	888	275,897	0.0032	1,377,264	13,893,832	50.36
35-39	1,255	0.0193	1,295	275,009	0.0047	1,371,808	12,516,568	45.51
40-44	1,745	0.0154	1,962	273,714	0.0072	1,363,665	11,144,760	40.72
45-49	2,546	0.0061	3,021	271,752	0.0111	1,351,206	9,781,094	35.99
50-54	3,627	0.0239	4,640	268,731	0.0173	1,332,052	8,429,889	31.37
55-59	5,008	0.0284	7,300	264,090	0.0276	1,302,199	7,097,837	26.88
60-64	6,436	0.0380	11,078	256,790	0.0431	1,256,254	5,795,638	22.57
65-69	7,827	0.0434	16,514	245,712	0.0672	1,187,274	4,539,384	18.47
70-74	9,299	0.0372	23,998	229,198	0.1047	1,085,994	3,352,110	14.63
75-79	11,665	0.0299	35,596	205,200	0.1735	937,010	2,266,116	11.04
80-84	13,991	0.0346	50,154	169,604	0.2957	722,635	1,329,106	7.84
85+	27,037	0.0490	119,450	119,450	1.0000	606,470	606,470	5.08

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Table C.6 – Life Table for White Males

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	2,984	0.0174	2,984	284,345	0.0105	1,411,283	20,936,834	73.63
5-9	205	0.0037	216	281,362	0.0008	1,406,267	19,525,551	69.40
10-14	274	-0.0149	281	281,145	0.0010	1,405,023	18,119,283	64.45
15-19	1,048	-0.0048	1,023	280,864	0.0036	1,401,763	16,714,260	59.51
20-24	1,653	0.0032	1,607	279,841	0.0057	1,395,188	15,312,497	54.72
25-29	1,778	0.0045	1,762	278,234	0.0063	1,386,765	13,917,310	50.02
30-34	2,073	0.0114	2,138	276,472	0.0077	1,377,015	12,530,544	45.32
35-39	2,505	0.0180	2,780	274,334	0.0101	1,364,721	11,153,529	40.66
40-44	3,233	0.0140	3,886	271,554	0.0143	1,348,057	9,788,808	36.05
45-49	4,579	0.0138	5,901	267,668	0.0220	1,323,590	8,440,751	31.53
50-54	6,533	0.0207	9,177	261,768	0.0351	1,285,896	7,117,161	27.19
55-59	8,605	0.0268	13,610	252,591	0.0539	1,228,928	5,831,265	23.09
60-64	10,253	0.0400	19,161	238,981	0.0802	1,146,999	4,602,336	19.26
65-69	11,161	0.0432	25,678	219,819	0.1168	1,034,900	3,455,337	15.72
70-74	11,882	0.0356	33,294	194,141	0.1715	887,469	2,420,437	12.47
75-79	12,552	0.0296	41,398	160,847	0.2574	700,738	1,532,968	9.53
80-84	12,067	0.0365	46,951	119,449	0.3931	479,866	832,230	6.97
85+	15,394	0.0399	72,498	72,498	1.0000	352,363	352,363	4.86

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Table C.7 – Life Table for Mixed & Black Females

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	831	0.0128	831	64,097	0.0130	317,577	4,684,199	73.08
5-9	71	0.0065	75	63,266	0.0012	316,144	4,366,622	69.02
10-14	100	-0.0165	102	63,191	0.0016	315,702	4,050,478	64.10
15-19	224	-0.0023	218	63,089	0.0035	314,901	3,734,776	59.20
20-24	301	-0.0031	290	62,871	0.0046	313,630	3,419,876	54.40
25-29	369	-0.0012	352	62,581	0.0056	312,027	3,106,245	49.64
30-34	519	0.0068	502	62,230	0.0081	309,894	2,794,219	44.90
35-39	720	0.0193	742	61,728	0.0120	306,785	2,484,324	40.25
40-44	996	0.0154	1,120	60,986	0.0184	302,130	2,177,539	35.71
45-49	1,355	0.0061	1,608	59,866	0.0269	295,309	1,875,409	31.33
50-54	1,761	0.0239	2,253	58,258	0.0387	285,657	1,580,100	27.12
55-59	2,191	0.0284	3,194	56,005	0.0570	272,040	1,294,443	23.11
60-64	2,475	0.0380	4,261	52,811	0.0807	253,404	1,022,404	19.36
65-69	2,669	0.0434	5,632	48,550	0.1160	228,672	769,000	15.84
70-74	2,738	0.0372	7,065	42,918	0.1646	196,931	540,328	12.59
75-79	2,845	0.0299	8,680	35,854	0.2421	157,569	343,397	9.58
80-84	2,703	0.0346	9,689	27,174	0.3566	111,646	185,828	6.84
85+	3,958	0.0490	17,485	17,485	1.0000	74,182	74,182	4.24

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Table C.8 – Life Table for Mixed & Black Males

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	1,074	0.0174	1,074	82,792	0.0130	410,201	5,501,371	66.45
5-9	92	0.0037	97	81,718	0.0012	408,348	5,091,169	62.30
10-14	162	-0.0149	166	81,621	0.0020	407,692	4,682,821	57.37
15-19	954	-0.0048	931	81,455	0.0114	404,949	4,275,129	52.48
20-24	1,256	0.0032	1,221	80,524	0.0152	399,568	3,870,180	48.06
25-29	1,277	0.0045	1,265	79,303	0.0160	393,351	3,470,612	43.76
30-34	1,458	0.0114	1,503	78,038	0.0193	386,429	3,077,261	39.43
35-39	1,745	0.0180	1,937	76,534	0.0253	377,829	2,690,832	35.16
40-44	2,113	0.0140	2,540	74,597	0.0341	366,635	2,313,003	31.01
45-49	2,653	0.0138	3,419	72,057	0.0474	351,738	1,946,367	27.01
50-54	3,300	0.0207	4,636	68,638	0.0675	331,602	1,594,630	23.23
55-59	3,815	0.0268	6,035	64,003	0.0943	304,926	1,263,028	19.73
60-64	3,964	0.0400	7,407	57,968	0.1278	271,320	958,102	16.53
65-69	3,786	0.0432	8,711	50,560	0.1723	231,024	686,781	13.58
70-74	3,477	0.0356	9,743	41,849	0.2328	184,889	455,757	10.89
75-79	3,095	0.0296	10,206	32,106	0.3179	135,016	270,868	8.44
80-84	2,466	0.0365	9,593	21,900	0.4381	85,516	135,852	6.20
85+	2,613	0.0399	12,307	12,307	1.0000	50,336	50,336	4.09

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Table C.9 – Life Table for Females

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0025	1.50	0.0101	0.9899	1.0000	0.0101	3.9749	78.5331	78.53
5-9	0.0002	2.50	0.0009	0.9991	0.9899	0.0008	4.9476	74.5583	75.32
10-14	0.0002	2.50	0.0010	0.9990	0.9891	0.0010	4.9430	69.6107	70.38
15-19	0.0004	2.50	0.0020	0.9980	0.9881	0.0020	4.9355	64.6677	65.45
20-24	0.0005	2.50	0.0024	0.9976	0.9861	0.0024	4.9247	59.7322	60.57
25-29	0.0006	2.50	0.0028	0.9972	0.9838	0.0028	4.9119	54.8075	55.71
30-34	0.0008	2.50	0.0040	0.9960	0.9810	0.0039	4.8952	49.8956	50.86
35-39	0.0011	2.50	0.0057	0.9943	0.9771	0.0056	4.8715	45.0004	46.06
40-44	0.0017	2.50	0.0086	0.9914	0.9715	0.0083	4.8366	40.1289	41.31
45-49	0.0027	2.50	0.0135	0.9865	0.9631	0.0130	4.7831	35.2923	36.64
50-54	0.0039	2.50	0.0194	0.9806	0.9501	0.0184	4.7045	30.5092	32.11
55-59	0.0062	2.50	0.0307	0.9693	0.9317	0.0286	4.5869	25.8047	27.70
60-64	0.0094	2.50	0.0458	0.9542	0.9031	0.0414	4.4120	21.2177	23.49
65-69	0.0147	2.50	0.0709	0.9291	0.8617	0.0611	4.1558	16.8057	19.50
70-74	0.0223	2.50	0.1057	0.8943	0.8006	0.0846	3.7914	12.6500	15.80
75-79	0.0373	2.50	0.1707	0.8293	0.7160	0.1222	3.2742	8.8586	12.37
80-84	0.0609	2.50	0.2644	0.7356	0.5937	0.1570	2.5761	5.5844	9.41
85+	0.1452	6.89	1.0000	0.0000	0.4367	0.4367	3.0083	3.0083	6.89

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.10 – Life Table for Males

Standard Method (Unadjusted nMx)

x	${}_nM_x$	${}_na_x$	${}_nq_x$	${}_np_x$	l_x	${}_nd_x$	${}_5L_x$	T_x	e_x
0-4	0.0029	1.50	0.0117	0.9883	1.0000	0.0117	3.9708	71.7588	71.76
5-9	0.0002	2.50	0.0010	0.9990	0.9883	0.0010	4.9391	67.7880	68.59
10-14	0.0003	2.50	0.0014	0.9986	0.9873	0.0014	4.9331	62.8488	63.66
15-19	0.0013	2.50	0.0062	0.9938	0.9859	0.0062	4.9141	57.9158	58.74
20-24	0.0016	2.50	0.0080	0.9920	0.9797	0.0078	4.8792	53.0016	54.10
25-29	0.0016	2.50	0.0081	0.9919	0.9719	0.0079	4.8400	48.1224	49.51
30-34	0.0020	2.50	0.0099	0.9901	0.9640	0.0095	4.7964	43.2824	44.90
35-39	0.0026	2.50	0.0129	0.9871	0.9545	0.0123	4.7418	38.4860	40.32
40-44	0.0036	2.50	0.0178	0.9822	0.9422	0.0168	4.6690	33.7442	35.81
45-49	0.0054	2.50	0.0265	0.9735	0.9254	0.0246	4.5656	29.0752	31.42
50-54	0.0081	2.50	0.0398	0.9602	0.9008	0.0358	4.4146	24.5096	27.21
55-59	0.0122	2.50	0.0592	0.9408	0.8650	0.0512	4.1970	20.0950	23.23
60-64	0.0176	2.50	0.0842	0.9158	0.8138	0.0685	3.8978	15.8980	19.54
65-69	0.0257	2.50	0.1209	0.8791	0.7453	0.0901	3.5012	12.0002	16.10
70-74	0.0368	2.50	0.1684	0.8316	0.6552	0.1103	3.0000	8.4991	12.97
75-79	0.0580	2.50	0.2534	0.7466	0.5448	0.1381	2.3790	5.4991	10.09
80-84	0.0881	2.50	0.3612	0.6388	0.4068	0.1469	1.6666	3.1200	7.67
85+	0.1788	5.59	1.0000	0.0000	0.2599	0.2599	1.4534	1.4534	5.59

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table C.11 – Life Table for Females

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	3,526	0.0128	3,526	361,900	0.0097	1,797,156	28,316,397	78.24
5-9	246	0.0065	259	358,373	0.0007	1,791,221	26,519,240	74.00
10-14	311	-0.0165	318	358,115	0.0009	1,789,779	24,728,019	69.05
15-19	646	-0.0023	631	357,797	0.0018	1,787,407	22,938,241	64.11
20-24	848	-0.0031	817	357,166	0.0023	1,783,788	21,150,834	59.22
25-29	1,071	-0.0012	1,020	356,349	0.0029	1,779,194	19,367,046	54.35
30-34	1,493	0.0068	1,442	355,329	0.0041	1,773,038	17,587,852	49.50
35-39	2,047	0.0193	2,111	353,886	0.0060	1,764,153	15,814,814	44.69
40-44	2,841	0.0154	3,195	351,775	0.0091	1,750,888	14,050,661	39.94
45-49	4,050	0.0061	4,806	348,580	0.0138	1,730,884	12,299,773	35.29
50-54	5,608	0.0239	7,174	343,774	0.0209	1,700,934	10,568,888	30.74
55-59	7,497	0.0284	10,930	336,600	0.0325	1,655,674	8,867,954	26.35
60-64	9,284	0.0380	15,980	325,670	0.0491	1,588,400	7,212,280	22.15
65-69	10,955	0.0434	23,113	309,690	0.0746	1,490,668	5,623,881	18.16
70-74	12,613	0.0372	32,549	286,577	0.1136	1,351,513	4,133,213	14.42
75-79	15,214	0.0299	46,426	254,028	0.1828	1,154,076	2,781,699	10.95
80-84	17,542	0.0346	62,884	207,602	0.3029	880,801	1,627,623	7.84
85+	32,756	0.0490	144,718	144,718	1.0000	746,822	746,822	5.16

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Table C.12 – Life Table for Males

Bennett & Horiuchi Approach

x	${}_nD_x$	${}_nr_x$	${}_nd_x$	l_x	${}_nq_x$	${}_5L_x$	T_x	e_x
0-4	4,332	0.0174	4,332	385,986	0.0112	1,914,765	27,869,733	72.20
5-9	307	0.0037	324	381,653	0.0008	1,907,457	25,954,968	68.01
10-14	450	-0.0149	461	381,329	0.0012	1,905,493	24,047,511	63.06
15-19	2,046	-0.0048	1,997	380,868	0.0052	1,899,348	22,142,017	58.14
20-24	2,976	0.0032	2,893	378,871	0.0076	1,887,123	20,242,670	53.43
25-29	3,136	0.0045	3,108	375,978	0.0083	1,872,121	18,355,546	48.82
30-34	3,630	0.0114	3,744	372,870	0.0100	1,854,992	16,483,425	44.21
35-39	4,377	0.0180	4,857	369,126	0.0132	1,833,489	14,628,434	39.63
40-44	5,530	0.0140	6,648	364,269	0.0182	1,804,726	12,794,945	35.12
45-49	7,489	0.0138	9,650	357,621	0.0270	1,763,981	10,990,219	30.73
50-54	10,189	0.0207	14,314	347,971	0.0411	1,704,072	9,226,238	26.51
55-59	12,888	0.0268	20,386	333,658	0.0611	1,617,324	7,522,166	22.54
60-64	14,799	0.0400	27,657	313,272	0.0883	1,497,218	5,904,842	18.85
65-69	15,625	0.0432	35,950	285,615	0.1259	1,338,202	4,407,624	15.43
70-74	16,141	0.0356	45,228	249,665	0.1812	1,135,258	3,069,422	12.29
75-79	16,498	0.0296	54,414	204,438	0.2662	886,153	1,934,164	9.46
80-84	15,359	0.0365	59,761	150,023	0.3983	600,715	1,048,011	6.99
85+	19,166	0.0399	90,262	90,262	1.0000	447,296	447,296	4.96

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Note: ${}_nD_x$ is the average number of observed deaths.

Appendix D: Additional Results of the Decomposition

Table D.1 – Racial Differences Decomposed across Genders

Age Group	W - M&B Females				W - M&B Males			
	Dir.	Ind.	Int.	Total	Dir.	Ind.	Int.	Total
0-4	0.01	0.28	0.02	0.32	0.01	0.15	0.02	0.18
5-9	0.00	0.04	0.00	0.04	0.00	0.02	0.00	0.03
10-14	0.00	0.05	0.01	0.06	0.00	0.05	0.01	0.06
15-19	0.01	0.11	0.01	0.13	0.02	0.37	0.05	0.44
20-24	0.01	0.14	0.02	0.16	0.02	0.40	0.06	0.48
25-29	0.01	0.15	0.02	0.17	0.02	0.36	0.05	0.44
30-34	0.01	0.19	0.02	0.23	0.03	0.38	0.06	0.47
35-39	0.02	0.25	0.04	0.30	0.04	0.43	0.07	0.54
40-44	0.03	0.33	0.05	0.41	0.04	0.48	0.08	0.61
45-49	0.04	0.40	0.06	0.50	0.06	0.51	0.09	0.66
50-54	0.05	0.45	0.07	0.57	0.07	0.53	0.09	0.69
55-59	0.06	0.50	0.08	0.64	0.08	0.52	0.09	0.68
60-64	0.08	0.49	0.08	0.65	0.08	0.45	0.07	0.61
65-69	0.09	0.47	0.08	0.63	0.08	0.37	0.05	0.51
70-74	0.10	0.38	0.06	0.54	0.08	0.26	0.03	0.37
75-79	0.10	0.26	0.04	0.40	0.06	0.15	0.02	0.22
80-84	0.06	0.11	0.02	0.20	0.03	0.05	0.01	0.09
85+	0.23	0.00	0.00	0.23	0.11	0.00	0.00	0.11
Total	**0.90**	**4.59**	**0.68**	**6.17**	**0.83**	**5.50**	**0.85**	**7.18**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.

Table D.2 – Gender Differences Decomposed across Racial Groups

Age Group	F - M Whites				F - M Mixed & Blacks			
	Dir.	Ind.	Int.	Total	Dir.	Ind.	Int.	Total
0-4	0.01	0.11	0.01	0.13	0.00	0.00	0.00	0.00
5-9	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
10-14	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.03
15-19	0.01	0.12	0.01	0.14	0.02	0.38	0.05	0.45
20-24	0.01	0.19	0.02	0.22	0.03	0.45	0.06	0.54
25-29	0.01	0.18	0.02	0.21	0.02	0.39	0.05	0.47
30-34	0.01	0.18	0.02	0.21	0.03	0.37	0.05	0.45
35-39	0.01	0.19	0.02	0.23	0.03	0.38	0.06	0.47
40-44	0.02	0.22	0.03	0.26	0.04	0.38	0.06	0.48
45-49	0.03	0.28	0.04	0.35	0.04	0.42	0.07	0.53
50-54	0.04	0.38	0.06	0.48	0.06	0.47	0.08	0.61
55-59	0.06	0.45	0.08	0.58	0.07	0.48	0.08	0.63
60-64	0.08	0.49	0.09	0.65	0.08	0.45	0.07	0.60
65-69	0.10	0.48	0.08	0.66	0.09	0.37	0.06	0.52
70-74	0.11	0.43	0.07	0.62	0.09	0.29	0.04	0.42
75-79	0.12	0.33	0.04	0.49	0.07	0.18	0.02	0.27
80-84	0.10	0.20	0.01	0.31	0.05	0.09	0.00	0.15
85+	0.06	0.00	0.00	0.06	0.02	0.00	0.00	0.02
Total	**0.76**	**4.25**	**0.61**	**5.62**	**0.74**	**5.12**	**0.77**	**6.63**

Source: Own calculations from the 2010 Census, the 2019 PNAD-C (third quarter), and DATASUS.