Death in New York City:

Gender Disparity in Natural v. Unnatural causes of death.

Sarra Souissi 2630604

Ephraïm Veder 2603454

Ingeborg van der Stuijt 2573043

27 November 2018

Dr. Erika Kuijpers

VU Amsterdam

Digital Humanities & Social Analytics

Table of Contents

Introduction Page 3

Provenance, Promises, and Limitations of the Data Page 5

External Resources Page 6

Methodology Page 7

Results Page 8

Conclusion Page 9

Recommendations for Further Research Page 9

Bibliography Page 11

**Introduction**

Death is inevitable. Death is natural. But natural causes can be prevented. Dying a natural death doesn’t always mean someone’s time has simply come. The time one has to live, can easily be extended or cut short. Diseases, accidents and tragedies cost many lives. But is there a connection between the way someone dies and their gender?

Our research is about the cause of death in New York City between 2007 and 2014. We want to know if there is a gender disparity between men and women in causes of death, specifically in natural causes of death (disease, precondition, and non-accidents) and unnatural causes of death (injury, self-harm, accidents, and overdose). New York City is a very representative city since it has inhabitants of many cultural backgrounds, people of different ages, sexes and income levels. That is why we wanted to test gender disparity for this population because of the representation value it has globally. We can also compare the results of our data analysis to bigger researches, for example data of America as a whole or the world due to the diversity of the New York City Area. We can also compare our data to data of different years, to hopefully see bigger differences. It was also very helpful for our research that this dataset was fully accessible and complete.

Gender disparity in death rates is an interesting topic because of the significant difference in the death rates per cause. In a study done by Susan B. Sorenson in the gender disparities in injury mortality in the US from 1981 to 2007, she concludes that “Gender disparities in injury mortality are consistent and persistent”. Her results show that boys and men are consistently more likely to die from accidental and violent injuries than women are across those years.

Based on this, we can ask a similar question: If men are more likely to die of unnatural causes (injury), then are women more likely to die from natural causes (i.e. disease or preexisting conditions) than men? According to the World Health Organization (WHO),

“Of the 56.9 million deaths worldwide in 2016, more than half (54%) were due to the top 10 causes. Ischaemic heart disease and stroke are the world’s biggest killers, accounting for a combined 15.2 million deaths in 2016. These diseases have remained the leading causes of death globally in the last 15 years” (World Health Organization).

In relation to the claim by the WHO, according to an article from Harvard Health Publishing at Harvard Medical school, risk of death by heart disease, the leading cause of death in America and worldwide, is more likely in women than men. This is because of preexisting conditions in women that aren’t as prevalent or present in men, such as having smaller coronary arteries, and the phenomenon of menopause (Harvard Health Publishing).

Data from a chart provided by Ritchie & Roser called Our World in Data shows that another major cause of death worldwide is Dementia/Alzheimer’s Disease which women are also more likely to die from (Ritchie, H. and Roser, M). According to the Alzheimer’s Association: “there is evidence that biological sex differences may affect mortality in men differently than women” in that women are more likely to be effected by another top-natural-killer than men.

After research into gender disparity between men and women, we are curious if the gender disparity for unnatural causes of death continues past Sorenson’s research time frame, specifically in the New York City area between 2007 and 2014 and if it is true that more women die from natural causes of death, like disease than men do.

Based on the literature of our initial research, we wanted to see if we could get results that hold the same merit in a broader sense. Based on Sorenson’s work, is there still a gender disparity in death by unnatural causes, specifically in New York City from 2007 to 2014. And based on the fact that women are seemingly more likely to die from major diseases such as heart conditions and Dementia/Alzheimer’s Disease, is there a gender disparity in death by natural causes of death between men and women?

Our main research question is: Is there gender disparity in natural causes of death and unnatural causes of death?

We asked ourselves the following sub-questions:

* Is there gender disparity in the deaths by natural causes such as disease and non-accidents? If so, how much of a difference is there?
* Is there gender disparity in the deaths by unnatural causes such as injury, accident, and overdose? If so, how much of a difference is there?
* Can the % differences be compared and if so what does this say about the initial research done?

**Provenance, Promises, and Limitations of the Data**

The data we used for our research is from NYC OpenData. This is a website owned by the City of New York. The data itself is provided by the Department of Health and Mental Hygiene. It is a quantitative dataset that describes all recorded deaths between 2007 and 2014 in New York City per year. There are 1094 rows in total. The dataset is in a CSV excel-format, which is common for large datasets like this one. There are seven variables used:

1. Year of death
2. Leading cause
3. Sex
4. Race Ethnicity
5. Deaths (raw)
6. Death Rate (based on Sex and Racial Ethnicity)
7. Age Adjusted Death Rate (based on Sex and Racial Ethnicity)

To answer the questions of this research, we modified the dataset. The original file holds the information of all the years between 2007 and 2014, so we isolated only the years 2007, 2010 and 2014 to show possible changes over time and provide a reasonable number of data points for our given questions.

To answer our questions regarding gender we had to combine the data in a way that the causes of death are not divided by race, age or any other variables than sex. A limitation of our original dataset was that the column for “Death Rate” was separated by race as well as sex, it was invaluable to our specific research questions. Additionally, the dataset did not provide us with a total population so there was no value to the “Death Rate” column and so we had to use the raw “Deaths” numbers in order to derive our own death rate for easier comparison. Using Microsoft Excel and the SUMIF function, we were able to clean up the data by getting rid of duplicate “Causes of Death” that were a part of the original dataset due to the separation by race. This is explained more thoroughly in the Data Documentation.

Since the data comes from the New York State OpenSource data hub, which is provided by the government in New York City, it is reasonable to argue that the provenance of this data is reputable. But there are possible biases, because the data comes from the department of health there could be more accounts of natural deaths/deaths from disease than unnatural/accidental deaths. The data is most likely from hospital records and mortician records which are more likely to record deaths from natural causes explicitly. Deaths recorded from less reputable sources are more likely to be identified as “unknown cause” or “other cause”.

Another limitation of the dataset is that there is a cause of death labeled as “All Other Causes”. This is problematic because we cannot categorize “All Other Causes” as either natural or unnatural because we simply do not know. This data point must be omitted despite the fact that the weight of this data point may be significant to our findings.

Another limitation of the dataset is the subjectivity of the separation of “natural” versus “unnatural” causes of death in the first place. While we used a dictionary definition for natural causes of death versus unnatural causes of death, the categorization of each of the causes of death is still subjective. For example, one could argue that “Mental and Behavioral Disorders” could be considered a “natural form of death” while others might not. In our case we categorized this cause of death as unnatural.

**External Resources**

Because our dataset does not explicitly divide causes of death as natural or unnatural, we used a definition of “unnatural death” from the medical dictionary section of The Free Dictionary to mediate our categorization for consistency. This source defines “unnatural death” as “A death caused by external causes—e.g., injury or poisoning—which includes death due to intentional injury, such as homicide or suicide, and death caused by unintentional injury in an accidental manner” (“Unnatural Death”). The causes of death derived from the NYC Causes of Death source were categorized as either “natural” or “unnatural” based on this definition.

Because our dataset did not provide a population number for any of the years in which it provided deaths, we were unable to use the death rates given to us by the dataset because they were based on sex and race instead of just sex alone. Because of this, we used the 2010 census population of New York City in order to calculate mortality rate (New York Department of City Planning 12).

**Methodology**

Based on the research question and the sub questions, the causes of death were added based on gender and nature of the death. From there, the death rates based on these factors were found, followed by the average death rate and percent difference in order to visualize a gender disparity, if any. Provided is a summary of the steps taken to query and enrich the data.

1. The data for the years 2007, 2010, and 2014 were isolated respectively from the original dataset.
2. In order to get total deaths for each cause of death by sex only, all of the racial groups per cause of death were grouped together. Each cause of death was also classified as either natural death (“N”) or unnatural death (“U”). The resulting curated data from this step is three charts depicting the cause of death for that year, the total number of deaths by cause per gender, and a column denoting if the cause of death. was natural or unnatural, denoted by a “N” or “U” respectively.
3. In each year, all of the deaths for women denoted by “N” are totaled and the same is done for deaths denoted by “U”. The same is done for men in each year.
4. Death rate is calculated from these totals using the formula for death rate: **(D/P)\*100,000** where D is the total number of deaths given gender and nature of death, P is the total population of New York City (taken from the new York census of 2010), and then multiplying this by 100,000 as a standardized number to record mortality rates.
5. Finally, the results from the three charts were consolidated into two charts to show, the death rates of men and women from natural and unnatural deaths by year.
6. Additionally, the average death rate of the three years by sex were calculated as well as a % Difference in the death rates in order to fully show the disparity between the sexes.

For a more comprehensive overview of the methodology, please refer to the dataset documentation.

**Results**

Using the methods we described previously to curate the original spreadsheet, two bar graphs were created:

This graphs show the following results:

* Women are dying at higher rates than men from natural deaths
* Over the years, the rate of death from natural causes has decreased for both men and women.
* Men are dying at higher rates than women from unnatural death.
* There is a similar trend over the years between both men and women for unnatural deaths in that there was no significant difference.

In addition to the graphs, we found after taking an average of the death rates of women and men from all three years for natural and unnatural deaths that there was a 11.5% difference in the mortality rate of women and men for “natural deaths”, with women having the greater number and a 130.8% difference in the mortality rate of women and men for “unnatural deaths” with men having the greater number.

Our findings basically coincide with our original research. Men have a higher mortality rate than women when it comes to unnatural deaths and women have a higher mortality rate than men when it comes to natural deaths. However when we analyze the differences in percentages between the two sets of results, it is evident that claiming men dying at higher rates than women from unnatural death has much more concrete statistical evidence than that of the claim that women die from natural causes more frequently than men.

**Conclusion**

It is true and clear that men die at higher rates from unnatural causes than women in New York City. This coincides with Sorenson’s piece about death by unnatural causes in the US in general and is further proven with our data given a different set of years and location. Sorenson argues that while men are born with “greater advantages” than women, their likelihood of dying earlier from violence is higher (Sorenson). When looking at our initial data, this holds true in that for every case of unnatural death (i.e. Assault); more men are dying of unnatural causes.

As for the result regarding gender disparity in natural causes of death, while our data proves that women die 11.5% more than men, this is not a concrete fact and calls for further research, especially given the limitations of our dataset. There are also several variables to consider when discussing this claim. For example, women may die from natural causes at a greater rate than men because they generally live longer (Sorenson). The men are the ones dying of accidents! And again, women are predisposed to some of the top causes of natural death, like heart disease and Alzheimer’s disease. So it is important to consider the other surrounding variables such and predispositions. It is interesting to see that this claim holds true in our research, women die at higher rates from some diseases due to their gender attributes.

**Recommendations for further research**

As noted before, our data is limited especially because of the categorization given called “All Other Causes” which had to be omitted for vagueness. With that being said, more specificity may have skewed our data on gender disparity for natural causes of death in the opposite direction. Because the margin for percent difference is so small (especially when compared to the percent difference in unnatural deaths). Therefore, a recommendation for further research would be to investigate what “All Other Causes” entails, categorizing them and redistributing the data to see if results would change.

Another research recommendation would be to expand this topic to other parts of the world, perhaps investigating the death rates in other major cities/regions and examine how the results coincide with our initial research as well as our results. This would be an interesting addition as our research is isolated to the United States. By broadening the range of data we can potentially see world trends, especially regarding if women really do die from natural causes of death at greater rates than men.

Works Cited

Sorenson, Susan B. “Gender disparities in injury mortality: consistent, persistent, and larger than you'd think” *American journal of public health* vol. 101 Suppl 1,Suppl 1 (2011): S353-8.

“The Top 10 Causes of Death.” *World Health Organization*, World Health Organization, 24 May 2018, www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death, 28 Nov 2018.

Harvard Health Publishing. “Gender Matters: Heart Disease Risk in Women.” *Harvard Health Blog*, Harvard Health Publishing, Sep 2006, www.health.harvard.edu/heart-health/gender-matters-heart-disease-risk-in-women, 28 Nov 2018.

Ritchie, H. and Roser, M. “Cause of Death.”, *Our World in Data.* Oxford Martin Programme on Global Development, Feb 2018, ourworldindata.org/causes-of-death, 28 Nov 2018.

“Why Does Alzheimer's Disease Affect More Women Than Men? New Alzheimer's Association Grant Will Help.” *Alzheimer's Association*, Alzheimer's Association, 11 Feb 2016, www.alz.org/blog/alz/february\_2016/why\_does\_alzheimer\_s\_disease\_affect\_more\_women\_tha, 28 Nov 2018.

“Unnatural Death.” *The Free Dictionary*, Farlex, medical-dictionary.thefreedictionary.com/unnatural+death, 28 Nov 2018.

United States. The City of New York. *Population Growth and Race/Hispanic Composition.* New York: Department of City Planning, 2010. Print.