



Leading Causes of Death in NYC 2007-2014

Kathleen Schaffer
Capstone Project
Data 201 - Python for Data Analytics
January 25, 2022



Research area, data source

Area of interest: Wanted to focus on public health issues in NYC

Data source: NYC Open Data Source

Data set: NYC Leading Causes of Death

Inspecting and Cleaning of Data



Columns: Year, Leading Cause (20 in total), Sex, Race Ethnicity, Deaths, Death Rate, Age Adjusted Death Rate (7 in all). All columns were object except for Year (int64)

Rows: 1094

Total number of deaths in data set: 424,998

	Year	Leading Cause	Sex	Race Ethnicity	Deaths
1	2013	Influenza (Flu) and Pneumonia (J09-J18)	F	Hispanic	204
3	2007	Essential Hypertension and Renal Diseases (I10, I12)	F	Not Stated/Unknown	5
4	2014	Cerebrovascular Disease (Stroke: I60-I69)	F	White Non-Hispanic	418
5	2009	Essential Hypertension and Renal Diseases (I10, I12)	M	Asian and Pacific Islander	26
6	2013	Influenza (Flu) and Pneumonia (J09-J18)	M	White Non-Hispanic	618



Cleaning data

Missing data: Only in Deaths, Death Rate, Age Adjusted Death Rate

Removed columns: Death Rate, Age Adjusted Death Rate

Changing data: Changed period in Death column to NaN, shortened text for causes of death.

Removed rows: Created new DF and removed NaN (reduced rows from 1094 to 956 (assumption being that there was no death in that category that year).



Research Question

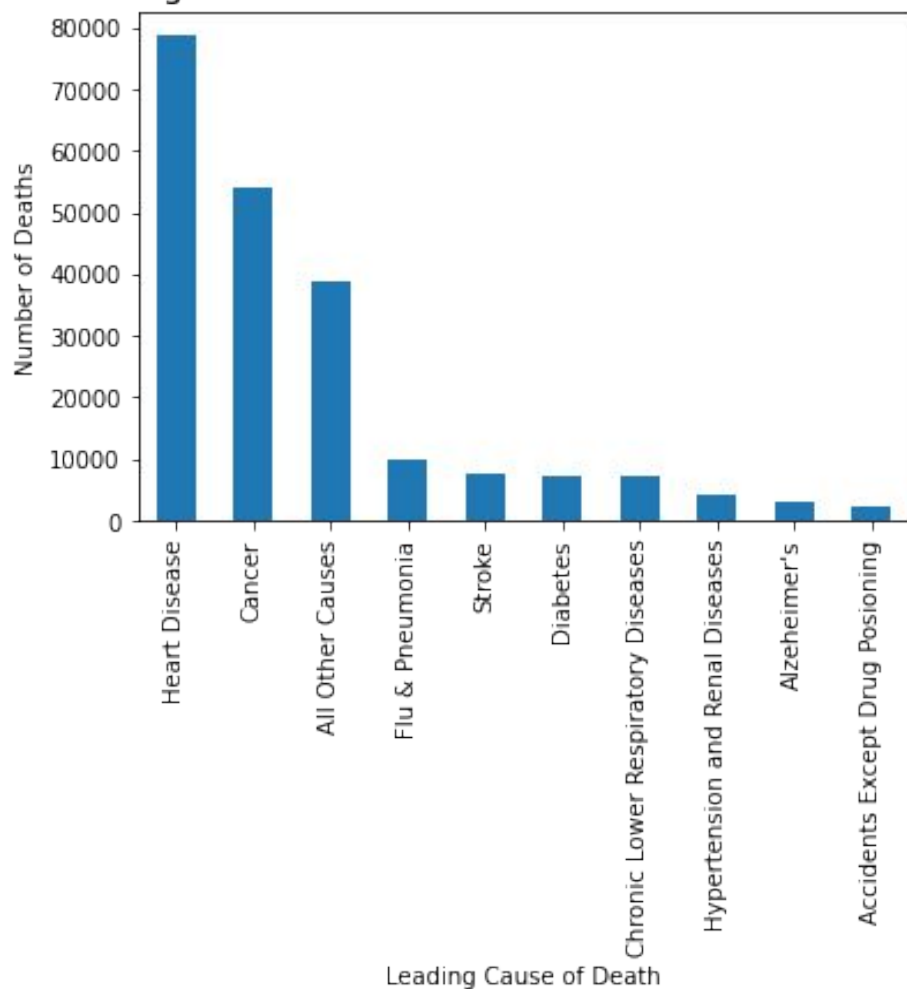
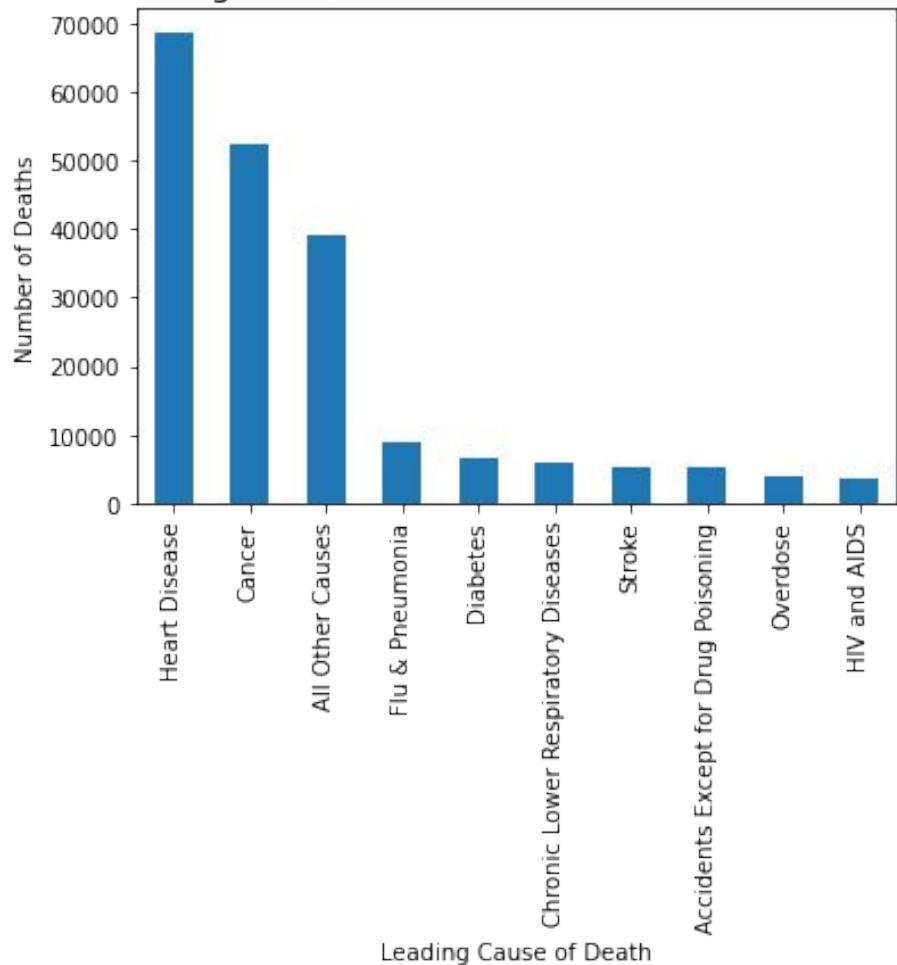
WHAT ARE THE THE TOP 5 LEADING CAUSES OF DEATH IN NYC BETWEEN 2007 AND 2014 AND IS THERE ANY DIFFERENCE BY GENDER?



Top 5 Causes of Death from all years

Diseases of Heart (I00-I09, I11, I13, I20-I51)	147,551
Malignant Neoplasms (Cancer: C00-C97)	106,367
All Other Causes	77,999
Influenza (Flu) and Pneumonia (J09-J18)	18,678
Diabetes Mellitus (E10-E14)	13,794

Leading Cause of Death for Men from 2007-2014 Leading Cause of Death for Women from 2007 to 2014





Conclusion

- 83% of all deaths are from heart disease, cancer, other causes, and the flu (which are the top four causes for both men and women).
- Smaller difference between genders on the top 4-10 causes of death.
 - Similarities: Stroke, Diabetes, Chronic Lower Respiratory Diseases, Accidents Except Drug Poisoning,
 - Differences
 - For men: Overdoses and AIDS related complications
 - For women: Hypertension, Alzheimer's.