Agriculture and combustion, are two major sources of particular matter(PM2.5) that are associated with premature deaths in Canada*

Zhongyu Huang

April 6, 2022

Abstract

The silent killer, PM2.5, was ranked the fifth risk factor for cardiovascular disease and lung cancer. Approximately 4.5 million individuals died due to exposure to PM2.5 globally every year. This report aims to model and predict the trend of premature deaths resulting from PM2.5 in Canada by using the linear regression model. The results indicate that the number of premature deaths can be explained by man-made emissions from agriculture and combustion. Increasing emissions from these two sources are significantly related to increases in deaths.

1 Introduction

In recent years, extreme pollution events bring people's attention to studying the health concerns due to air pollution. The government of Canada reported that an estimated 15,300 premature deaths are linked to air pollution every year (Government of Canada, 2021). That is, approximately 5% of the total number of deaths reported in 2020. Exposure to pollutants such as nitrogen oxides (NOX) or sulfur oxides (SOX) can cause serious health issues including damaging lung function(Government of Canada, 2021). As the main component of smog, fine particulate matter and ground-level ozone (O3) can cause dysfunction of the eye, nose, throat, or lung(Government of Canada, 2021). Fine particulate matter has been identified as one of the risk factors for cardiovascular disease and premature death(Government of Canada, 2021). Other pollutants such as Carbon monoxide (CO) and Ammonia (NH3) are also harmful to human health(Government of Canada, 2021). Most of them are colorless gas to be not noticed. The vulnerable group, children and the elderly, can be harmed more easily at greater risk(Government of Canada, 2021).

This paper aims to predict the premature death toll from the composition of PM2.5 emissions from different sources by building a meaningful linear regression model. The data used here are collected by the organization for Economic Co-operation and Development. The dataset stores information about man-made emissions of PM2.5 by different pollutants and the related death toll in Canada from 2005 to 2019 (OECD,2018). We tried to explain the number of death by related factors. With such a model, it should be clear to determine pollutant sources that are significantly related to premature death. It can also be used to provide directions on how to reduce or avoid air pollution, specifically, reduce the emission of PM2.5.

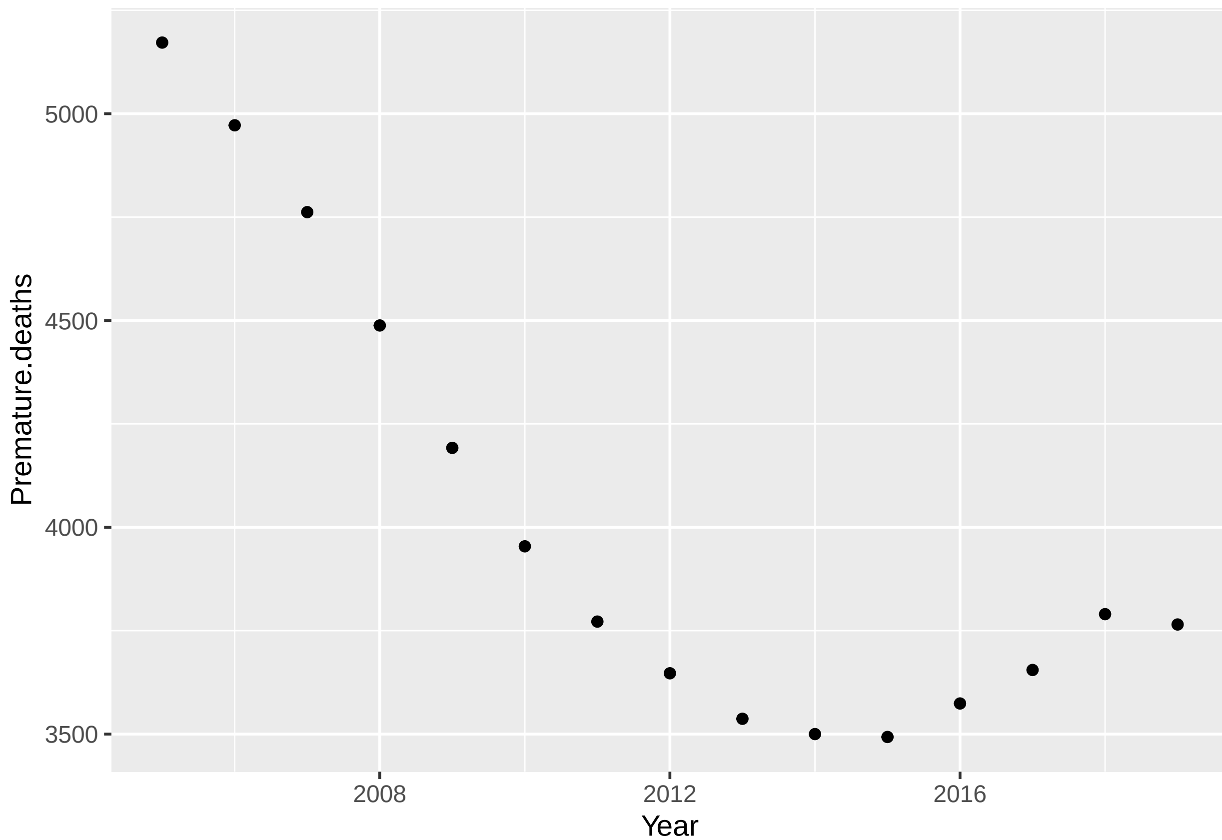
Interestingly, the number of death decreased while the total man-made emissions increased in these 13 years. This can be probably explained by the development of medicine. The results have shown that emissions of PM2.5 from agriculture, combustion, and miscellaneous are statistically significantly linearly related to the death resulting from exposures to PM2.5. More specifically, the model can be expressed as - Premature death $= 7.2419131 + 0.0037454 *$ Emissions from Agriculture $+ (-0.0011323) *$

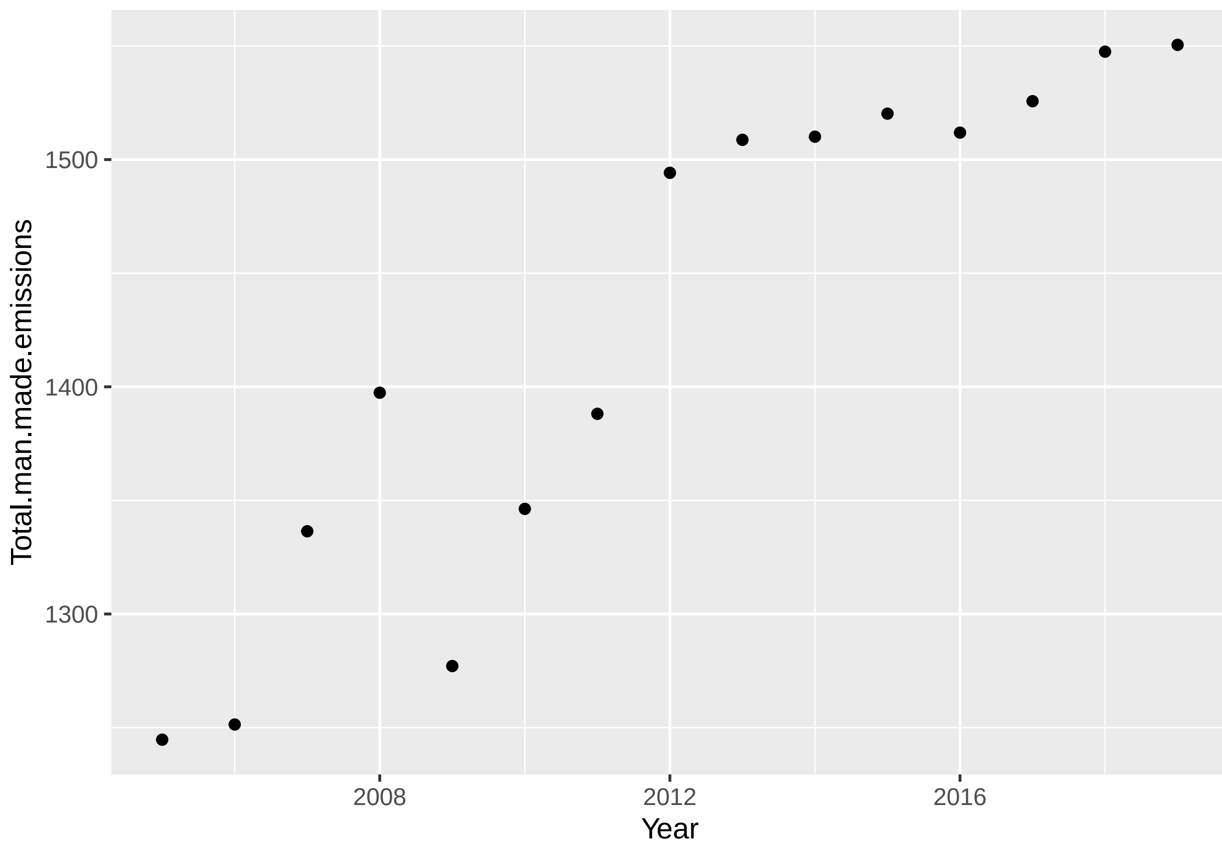
*Code and data are available at: https://github.com/zhongyuhuang/Household-air-pollution-attributable-deaths-Analysis.git

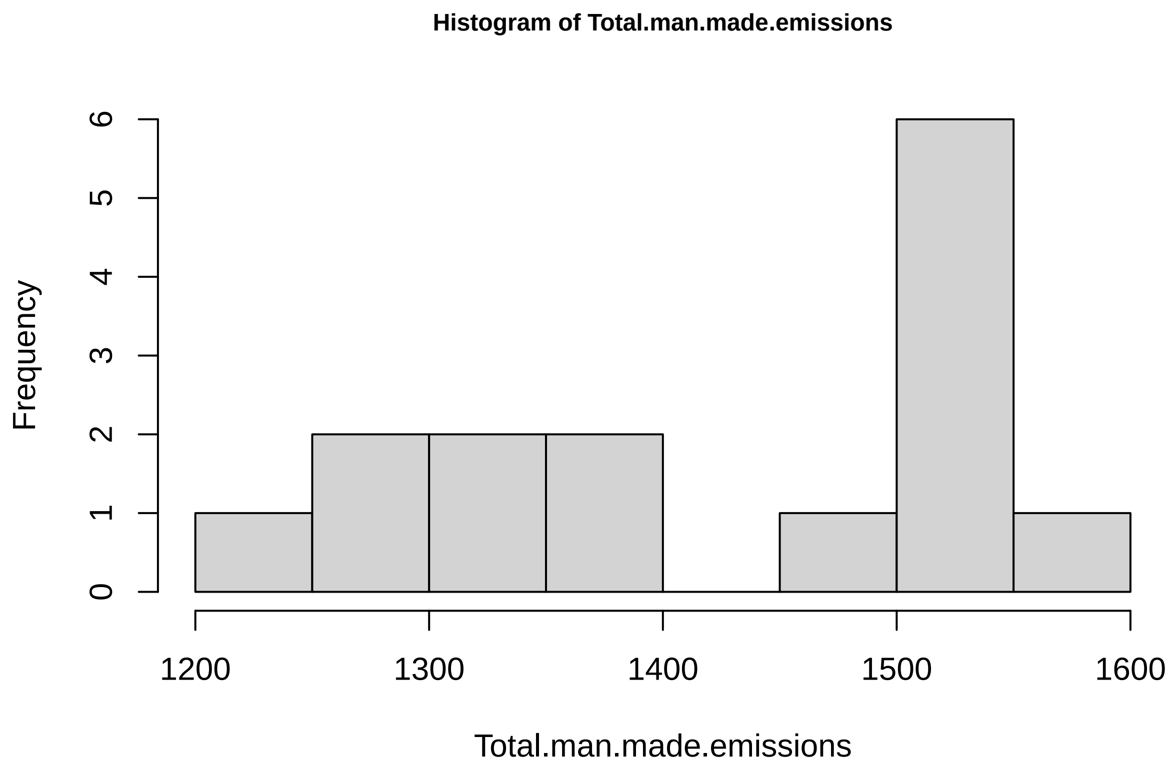
Miscellaneous emissions $+ 0.0003403 *$ Emissions from Combustion $+ \epsilon$. It indicates that a one unit increase in emissions from agriculture would lead to 0.0037454 more death. The unit here for emissions is a thousand tones. It is worth noticing that our model relies on a relatively small dataset which might affect the prediction power of the model. I will introduce the dataset in more detail first, then build and validate the model, and finally discuss the results and limitations.

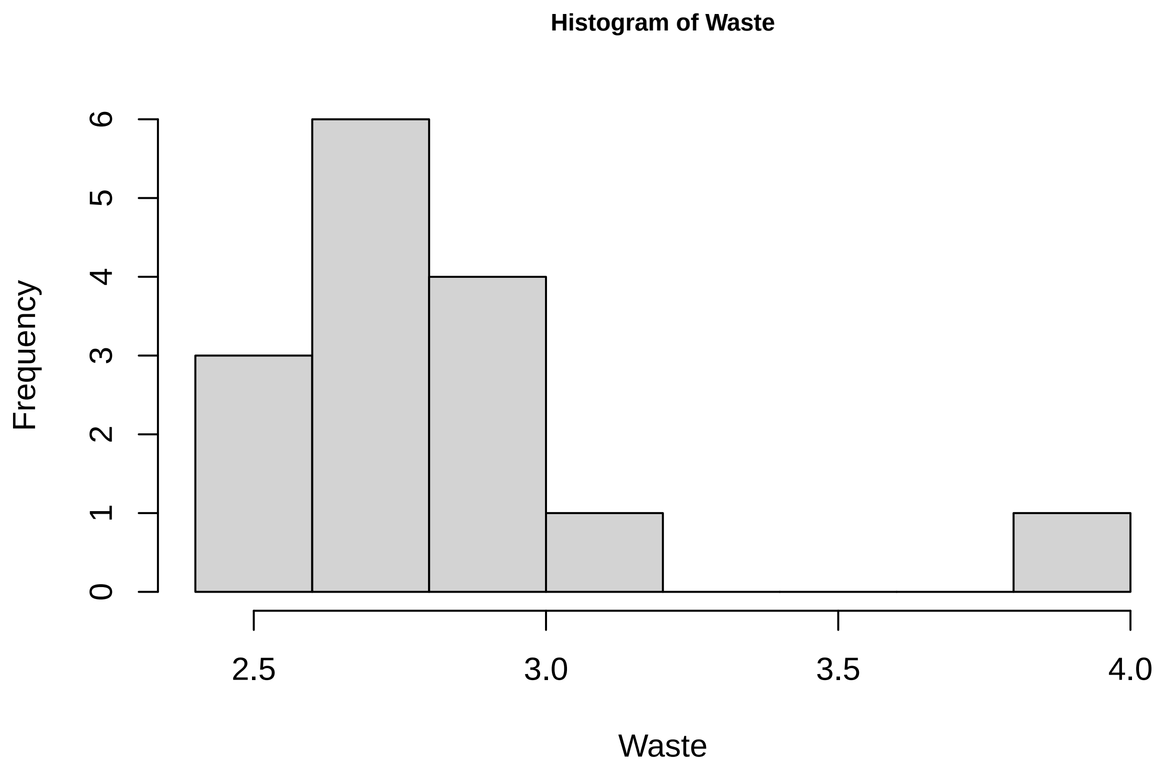
2 Data

Our data is of Air pollution (Figure **??**).

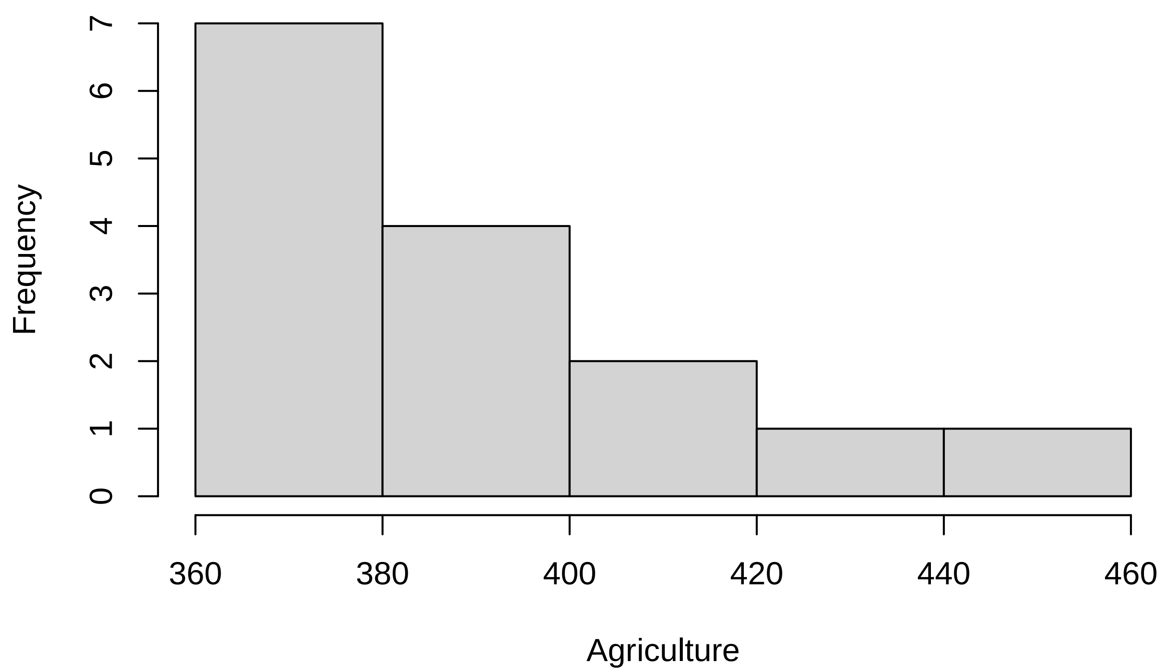


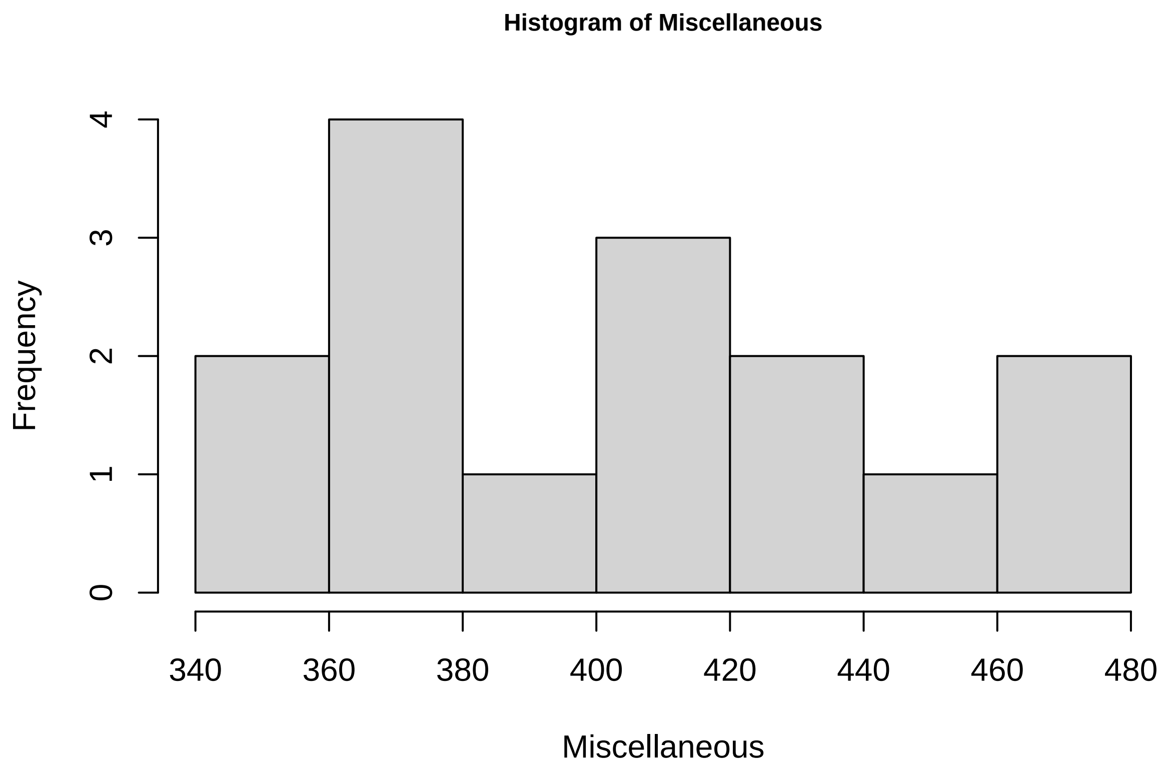

1500
1400
1300
Total.man.made.emissions
2008
2012
2016
Year


Histogram of Total.man.made.emissions
Frequency
0
1
2
3
4
5
6
1200
1300
1400
1500
1600
Total.man.made.emissions

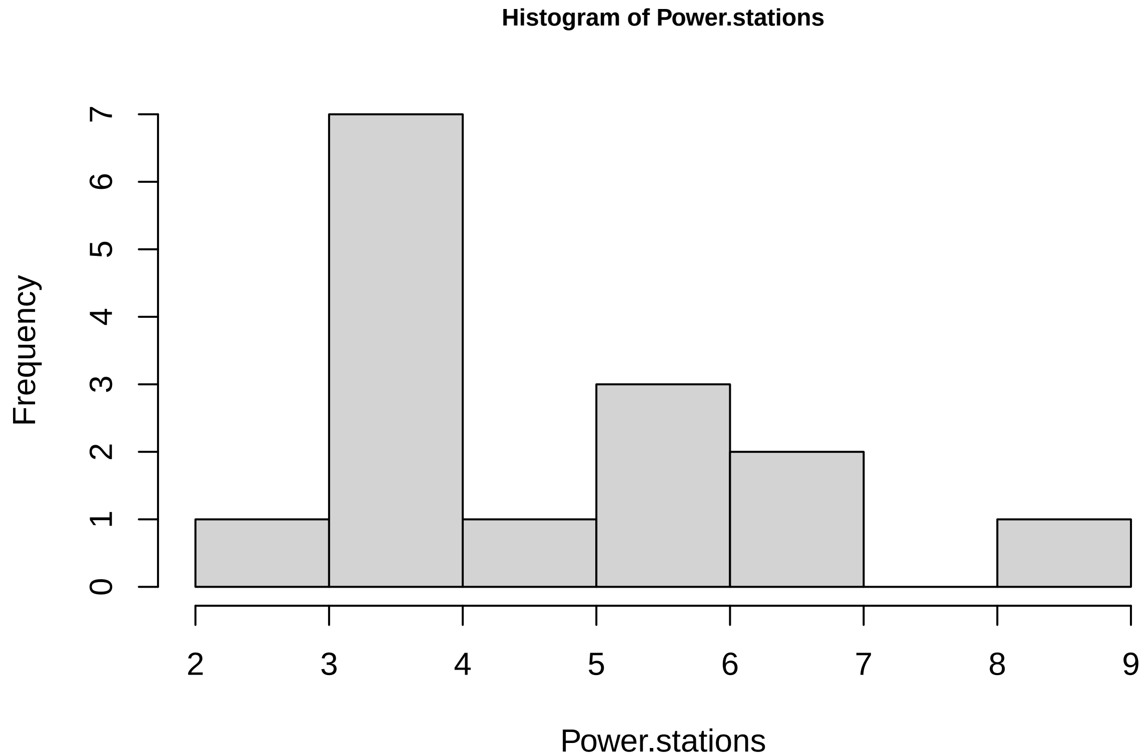

Histogram of Waste
Frequency
0
1
2
3
4
5
6
2.5
3.0
3.5
4.0
Waste

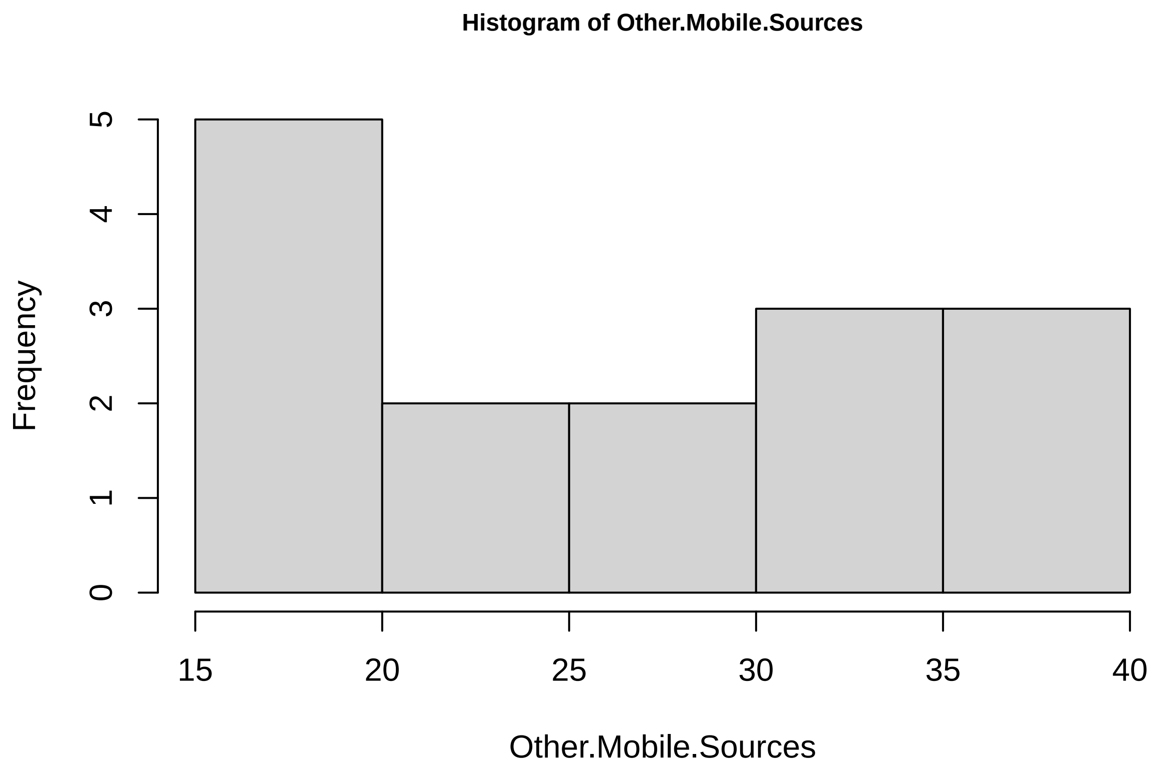
Histogram of Agriculture

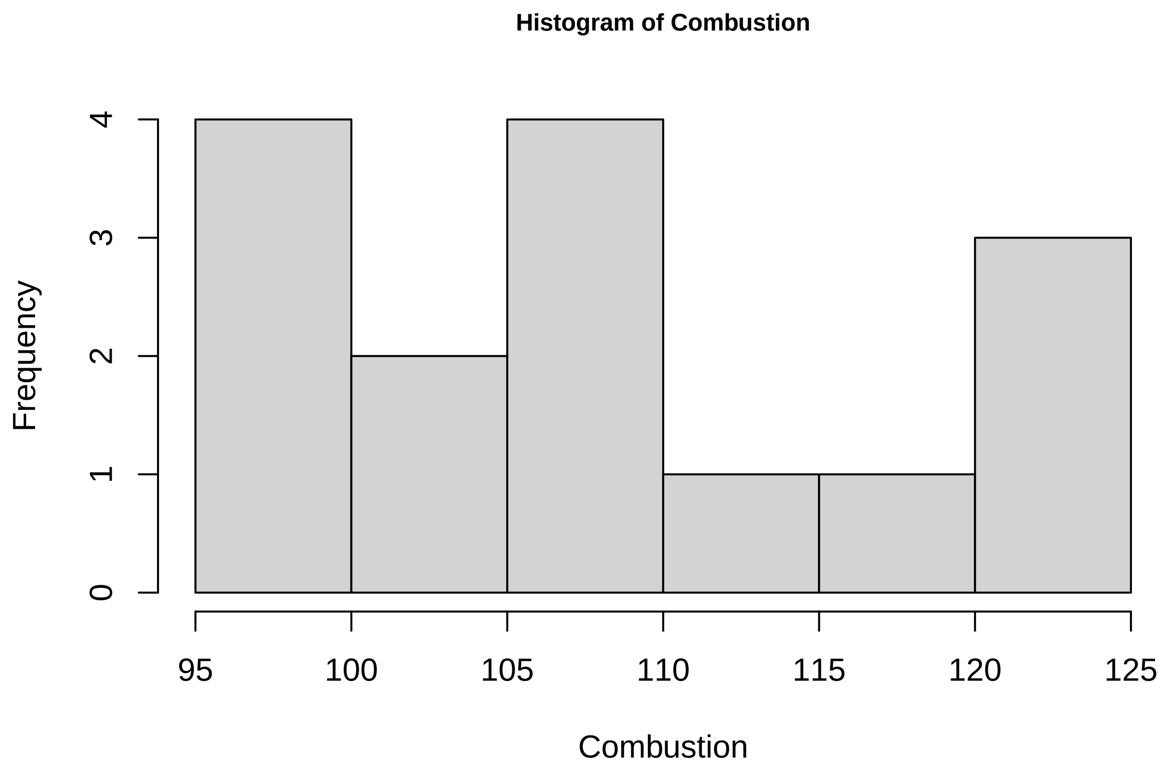



Histogram of Miscellaneous
Frequency
0
1
2
3
4
340
360
380
400
420
440
460
480
Miscellaneous

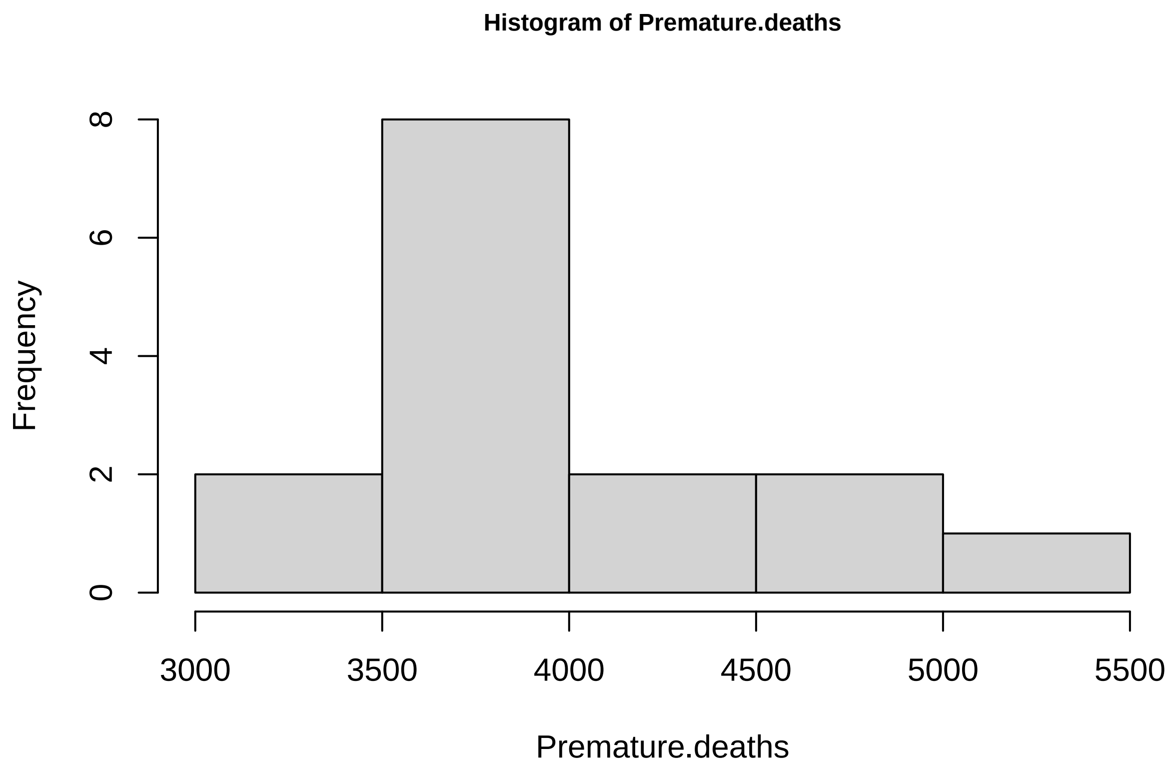

Histogram of Road.Transport
Frequency
0
1
2
3
4
14
16
18
20
22
24
26
28
Road.Transport


Histogram of Power.stations
Frequency
0
1
2
3
4
5
6
7
2
3
4
5
6
7
8
9
Power.stations


Histogram of Other.Mobile.Sources
Frequency
0
1
2
3
4
5
15
20
25
30
35
40
Other.Mobile.Sources


Histogram of Combustion
Frequency
0
1
2
3
4
95
100
105
110
115
120
125
Combustion


Histogram of Industrial.processes.and.product.use
Frequency
0
2
4
6
8
250
300
350
400
450
500
550
600
Industrial.processes.and.product.use



Talk more about it.

Also bills and their average (Figure ??). (Notice how you can change the height and width so they don't take the whole page?)

Talk way more about it.

3 Model

$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)} \tag{1}$$

Equation (1) seems useful, eh?

Here's a dumb example of how to use some references: In paper we run our analysis in `R` (R Core Team 2020). We also use the `tidyverse` which was written by Wickham et al. (2019) If we were interested in baseball data then Friendly et al. (2020) could be useful.

We can use maths by including latex between dollar signs, for instance θ.

4 Results

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional details

References

Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean 'Lahman' Baseball Database*. https://CRAN.R-project.org/package=Lahman.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.