

CA2 – Data Visualisation/Introduction to Statistics

**Alice Corry**

**Your student number is N00211635**

Creative Coding 1

Year 2 2022-23

DL836 BSc (Hons) in Creative Computing

Table of Contents

[1. Introduction](#_eyytwyws84ag)

[2. Description of Dataset](#_mxliscno14xx)

[3. Results](#_qq9cp1ervdwo)

[3.1. Chart 1 - Stacked Bar Chart with Average Line Graph](#_yobgzx579qy4)

[4. Statistical Summary of Dataset](#_brme2cl0ohow)

[5. Discussion](#_edm3u636mf9g)

[6. Conclusions](#_uguwgw2anc3d)

[6.1. What did you learn from the assignment?](#_5pvqihpvyhdf)

[6.2. What would you do differently in the future?](#_8cg46k1zftym)

[7. References](#_w2kwh2y7kqhy)

# 

# Introduction

The aim of this project was to create at least 2 suitable graphs for the chosen dataset. The data visualisation should give a message about your chosen dataset. The project brief was to include several components, including description, results, statistical summary, discussion and conclusion.

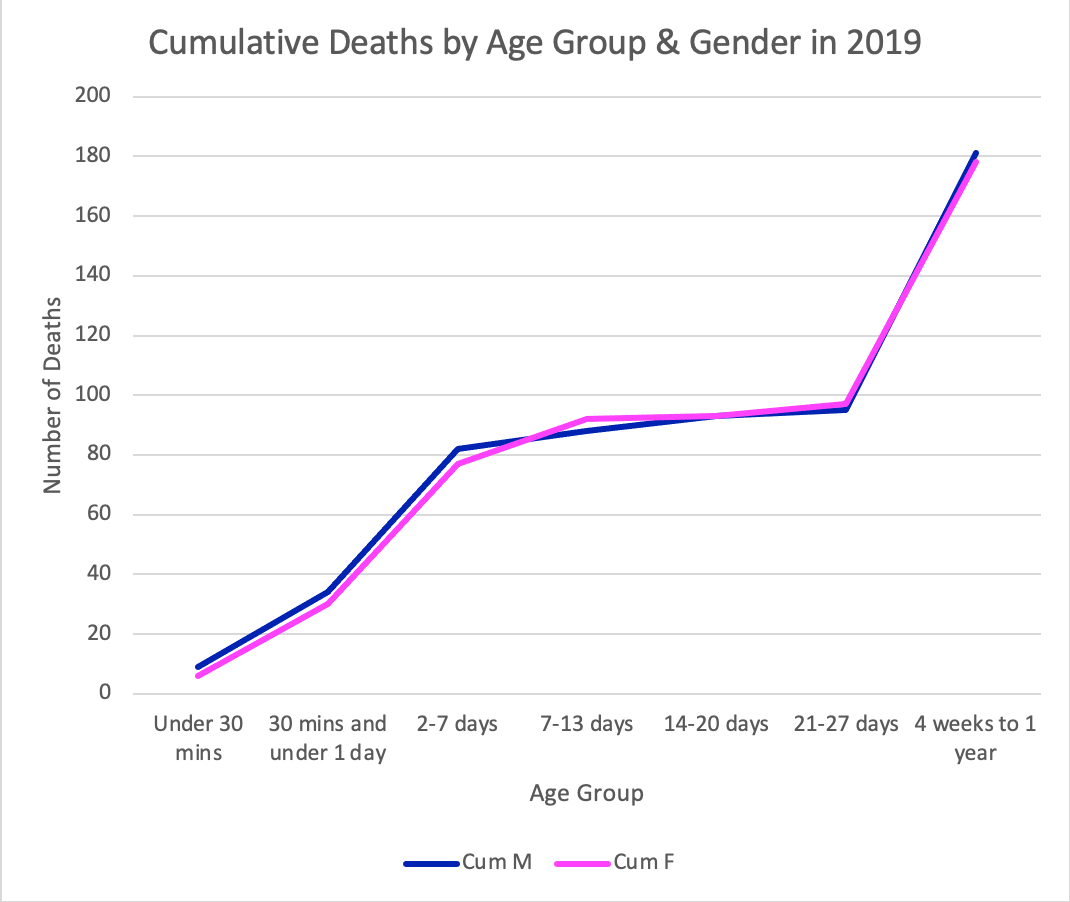
# Description of Dataset

I selected the 2019 data from the ‘Infant Mortality Deaths Under 1 Year’ statistic data set, which gives the number of deaths broken down by gender and age group. This data comes from a reliable source, so I have assumed it is quite accurate and does not need correction

# Results

# Chart 1 - Cumulative Line Graph

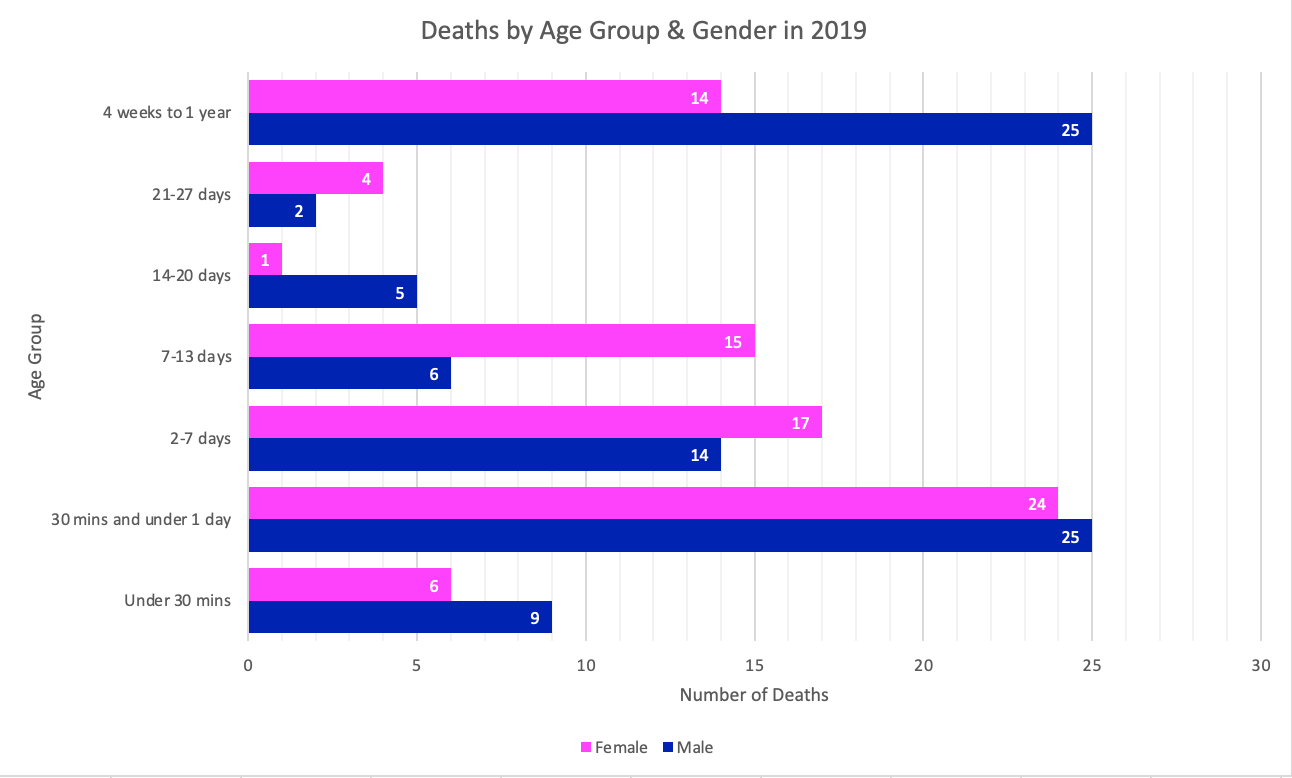
It shows a significant variation between male and female deaths in individual periods, the overall trend is very similar. This indicates a similar level of healthcare for both male and female babies.



# Chart 2 - Clustered Column Chart

The main focus of the graph was comparing deaths between sexes. It allows you to visually compare deaths by gender and age group.

Removed totals because this was redundant data, for example, the under 7 days age group, because they are misleading. Double counting and it appeared like there was a spike in deaths.



# Statistical Summary of Dataset

The information presented is only for infant mortality, or deaths in the first year of life. Also, no information exists on the number of infants who do not die. For this reason, the analysis must be restricted to infants only, and death rates cannot be calculated.

# Discussion

I only used 2019 figures for the graphs/charts in this project since I didn't want to deal with too many categories when visualising my data.

When choosing my graphs/charts, I have decided not to use the following: Stacked Bar Chart, Boxplot and Average Deaths Per Day. A stacked bar chart would not have worked for this data since comparing male and female results is more difficult. A boxplot would also not work since the results are unsuitable for visualisation using this method.

If I had more time or tools, I would have included the following: I would use multi-year data and allow year-based filtering.

Since it lets you see the yearly trend, cumulative is a line graph. A bar chart indicates specific values, whereas a line indicates a trend. In this case, it’s more of a trend than specific values. The colours pink and blue are associated in people's minds with different sex and therefore seem a natural fit for the data items.

A clustered column chart has certain advantages and disadvantages. Easily being able to see the differences for each period is a strength. The weakness is that it is hard to see the total of males and females.

# Conclusions

I think I did well with the project. I had done some Excel & Statistics before, which helped in the early weeks of the project. I had not done some of the harder formulas in Excel before, and although I got the basics and some of the harder statistics fairly easily, I struggled with some of the more complex code.

I found some areas harder than others, but I made sure to ask questions if I needed help with something. Throughout this project, I developed and improved the following skills: project management, time management skills, and the importance of good data visualisation.

Overall, my project went well. There are always things I can improve on, and I can keep practising all the skills I have learned so that if I need to use these skills later on, I will be ready.

# References

*https://data.cso.ie/*. (n.d.). https://data.cso.ie/