Statistical Report on Irish Baby Names From 1964 to 2018

Pradeep Gurunathan, Siddhartha Ranganathan, Harshini Priya Ravi Kumar, Poojaanjali Vasu

10 December, 2019

Introduction

This Statistical report provides a clear overview over the Irish babies names from 1964 to 2018. The various aspects of data such as most popular names for boys and girls throughout 55 years and their popularity trends are emphasized in this analysis. It also includes the Top 10 names across both the genders and also try to estimate the difference in proportions of birth rate of boys and girls.
Note:- Please Run the Rmd file to see the both shiny apps and to check the codes.

Data importing and cleanzing

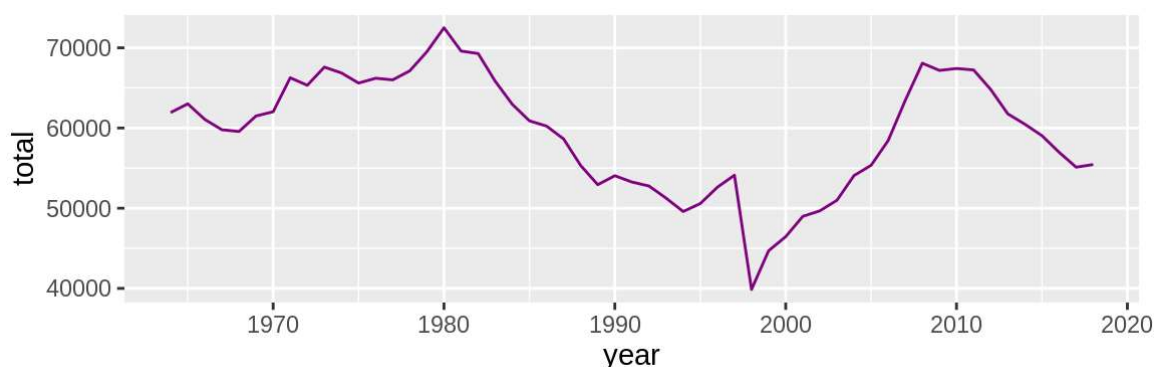| Data importing and cleanzing	Data preview	Packages used

Below code are used to read the individual zip folder which contains separate CSV documents of all Irish Girl and Boy babies details from 1964 to 2018 and storing it in separate dataset's. Secondly, new column fields are created like 'sex' and 'year' in each Girl and Boy dataset's for better understanding of the dataset and for further statistical analysis. Finally, both Boy dataset and Girl dataset are combined row wise to form a new dataset and then sex field is converted as factor.

Observations

Observation 1

Line graph to show the total irish babies born from 1964 to 2018. From the line graph it is observed that there is a steady increase in total births from late 1960's till 1980, then there is steep fall from 1980 till late 1990's.Later, in 1998 total births was least when compared to other years from 1964 to 2018.
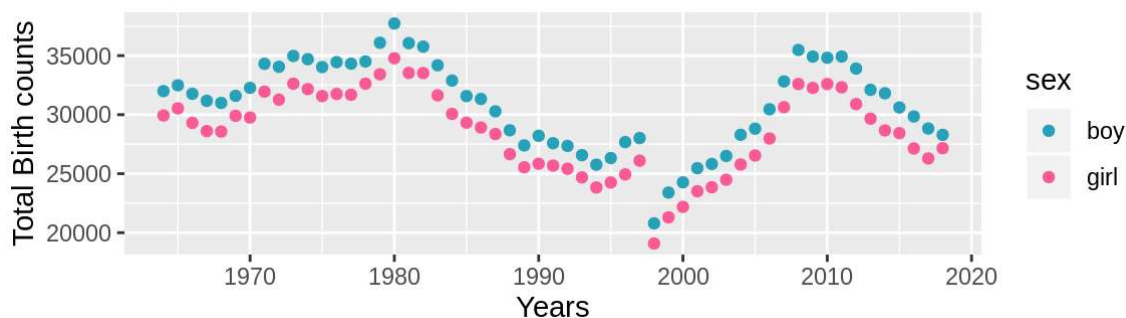

- In year 1980 Total Births was Highest of 72498 births.
- In year 1998 Total Births was least at 39875 births.
- Average number of births from 1964 to 2018 is 59478.6.

Observation 2

Below graph represents total individual Irish Boy-Girl babies birth from 1964 to 2018.
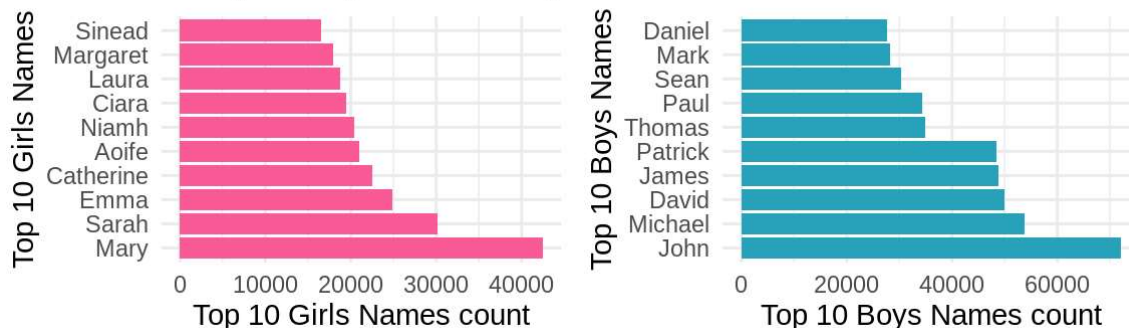
Boy VS Girl Birth Ratio

- Highest Boy-Girl birth difference from 1964 to 2018 is 3149 in the year 2014.
- Lowest Boy-Girl birth difference from 1964 to 2018 is 1135 in the year 2018.
- Average Boy-Girl birth difference is around 2311.

Observation 3

Top 10 Boy and Girl baby name's from 1964 to 2018. It is observed that the name 'Mary' is the most popular Irish Girl baby name that was kept over the years from 1964 to 2018. More than 70000 Irish Boy babies where named as 'John' from 1964 to 2018. It has been the most desired name over the years.
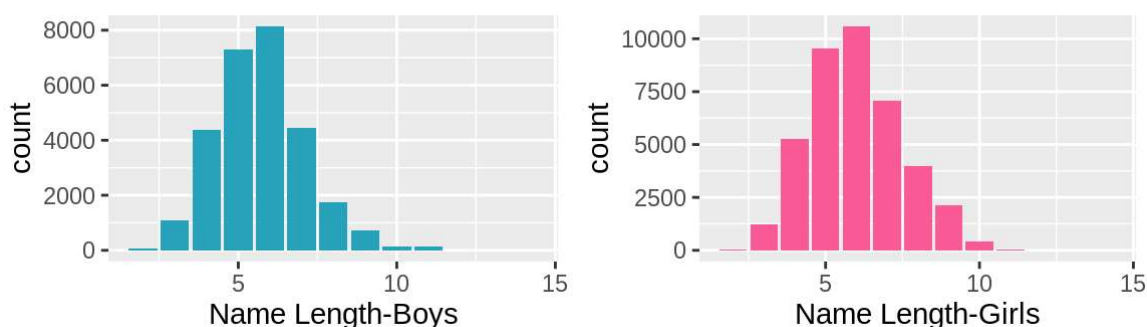
Top 10 Boy and Girl baby name's from 1964 to 2018

Observation 4

Comparing the both graph of Irish Boy-Girl name length we infer that there is not much difference in Boy-Girl name lengths, both are having an average of 5 or 6 as their name lengths. Also, maximum and minimum name lengths for both the Boy-Girl name are 2 and 14 respectively.
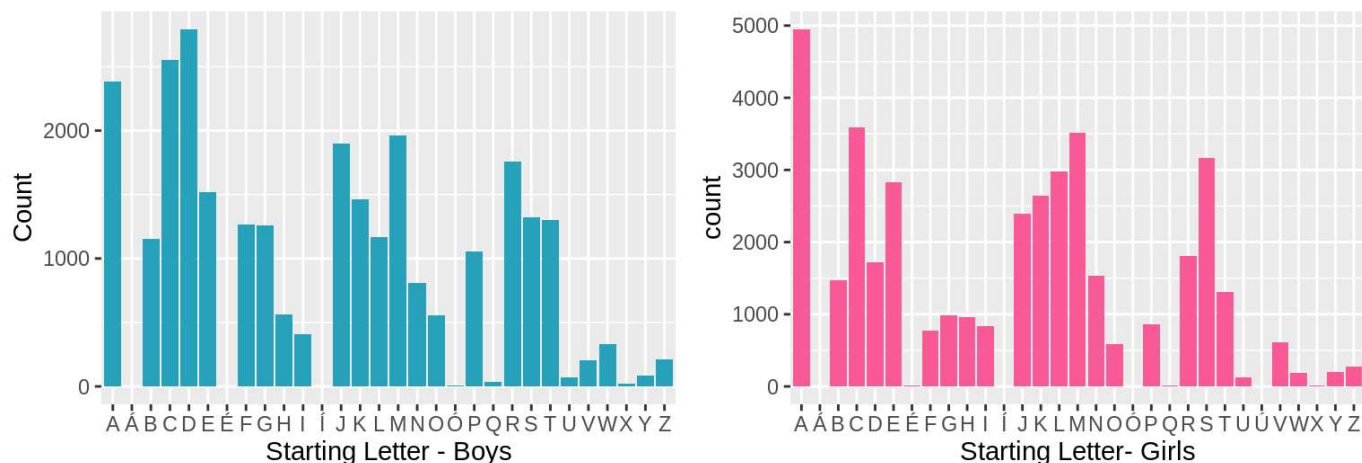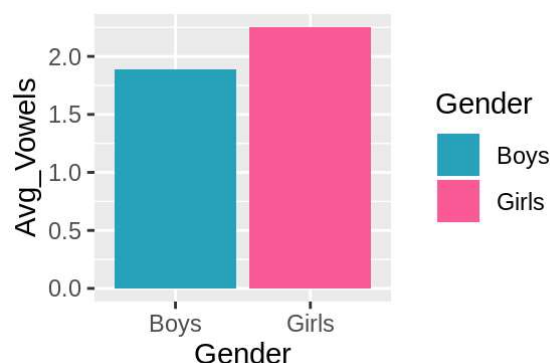
Length of Boy's and Girl's Names

Observation 5

The most frequent Starting letters for the Irish Boy-Girl Baby Names is A, C, D, J, M and A, C, E, L, M respectively. Whereas, the rare occurance Starting letters for both sex is from U to Z.

Starting Letter of the names

Observation 6

From the Vowel's graph, it can be concluded that average vowels count is more in Girl's compared to Boy's names.
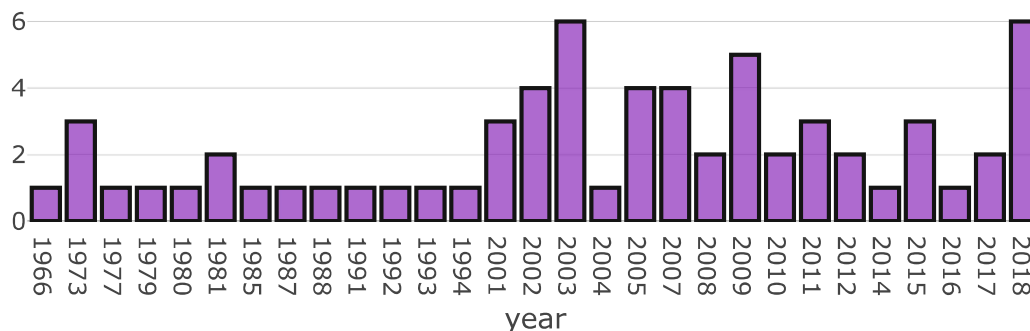

Observation 7

Calculating the highest and lowest popularity of Irish Baby names from 1964 to 2018.

- Name "Emily" increased by 1.63% in popularity since 1964 to 2018 for Irish Girl names.
- Name "Mary" decreased by 11.21% in popularity since 1964 to 2018 for Irish Girl names
- Name "Conor" increased by 1.42% in popularity since 1964 to 2018 for Irish Boy names.
- Name "John" decreased by 10.69% in popularity since 1964 to 2018 for Irish Boy names

Observation 8

Displaying the years that has unique name's from 1964 to 2018. It can be observed that the year 2003 and 2018 has the highest unique name count.

Shiny apps

Top babies	Babies details

For determining the Top 10 Irish Boy and Girl name in each year from 1964 to 2018 below shiny is used.

Note:- Please Run the Rmd file to see the shiny app.

Conclusion

The above evaluation helped us understand the Irish baby names dataset better. Now we have answers to some compelling questions and are in a position to do follow-up and deep dive analysis. The following is the conclusions of the observations:

- It is observed that there is a steady increase in total births from late 1960's till 1980, then there is steep fall from 1980 till late 1990's. Later, in 1998 total births was least when compared to other years from 1964 to 2018. As seen in the line graph.
- It can be inferred that in 1980 there was a highest number of Baby births around 72498 and in 1998 there was a lowest birth rate of 39875.
- In 2014, difference in the Boy-Girl birth rate was highest of 3149. Least birth rate difference was 1135 in the year 2018.
- More than 40000 girls where named after 'Mary' and around 70000 boys where named after 'John' over the 55 years from 1964 to 2018.
- There is not much difference in Boy-Girl name lengths, both are having an average of 5 or 6 as their name lengths. Also, maximum and minimum name lengths for both the Boy-Girl name are 2 and 14 respectively.
- A, C, D, J, M and A, C, E, L, M are the most frequent Starting letters for the Irish Boy-Girl Baby Names respectively.
- Girl name 'Emily' saw 1.63% increase in its popularity over the years from 1964 to 2018. Whereas, there was a 11.21% decrease in its popularity for Girl name 'Mary'.
- 'Conor' name saw 1.42% increase in its popularity over the years from 1964 to 2018 for Boys. 10.69% for boy name 'John' saw decrease in its popularity over the years from 1964 to 2018.
- In year 2003 and 2018 has the highest unique name count of 6.

Roles and responsibilities

Following roles and responsibilities were taken up by the team members :-

- I, **Pradeep Gurunathan (19251698)** had primary responsibility for material in shiny web interactions, report generation and data visualization graphs.
- I, **Siddhartha Ranganathan (19251465)** had primary responsibility for material in plotly interactive graphs, report generation and data visualization graphs.
- I, **Harshini Priya Ravi Kumar (19251251)** had primary responsibility for material in statistical analysis, report generation and data visualization graphs.
- I, **Poojaanjali Vasu (19251512)** had primary responsibility for material in Data importing and cleanzing, report generation and data visualization graphs.