



Statistical Report on Irish Baby Names From 1964 to 2018

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.

Data importing and cleanzing

Data importing and cleanzing

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.

Data preview

```
##           Name Rank births  sex year
## 1         Mary    1   3471 girl 1964
## 2   Catherine    2   1412 girl 1964
## 3   Margaret    3   1392 girl 1964
## 4         Ann    4    866 girl 1964
## 5        Anne    5    834 girl 1964
## 6   Elizabeth    6    722 girl 1964
## 7   Caroline    7    617 girl 1964
## 8   Bridget    8    595 girl 1964
## 9   Geraldine    9    579 girl 1964
## 10  Patricia   10    536 girl 1964
## 11  Jacqueline  11    529 girl 1964
## 12   Pauline   12    460 girl 1964
## 13    Helen   13    437 girl 1964
## 14   Deirdre   14    408 girl 1964
## 15 Bernadette  15    392 girl 1964
## 16    Sandra   16    358 girl 1964
## 17   Eileen   17    344 girl 1964
## 18   Martina   18    339 girl 1964
## 19   Teresa   19    328 girl 1964
## 20    Paula   20    311 girl 1964
## 21   Carmel   21    304 girl 1964
## 22    Susan   22    303 girl 1964
## 23  Kathleen   23    301 girl 1964
## 24   Angela   24    291 girl 1964
## 25  Josephine  25    287 girl 1964
```

Packages used

The analysis and visualization are done using R and we have used below packages from R library.

Packages used:

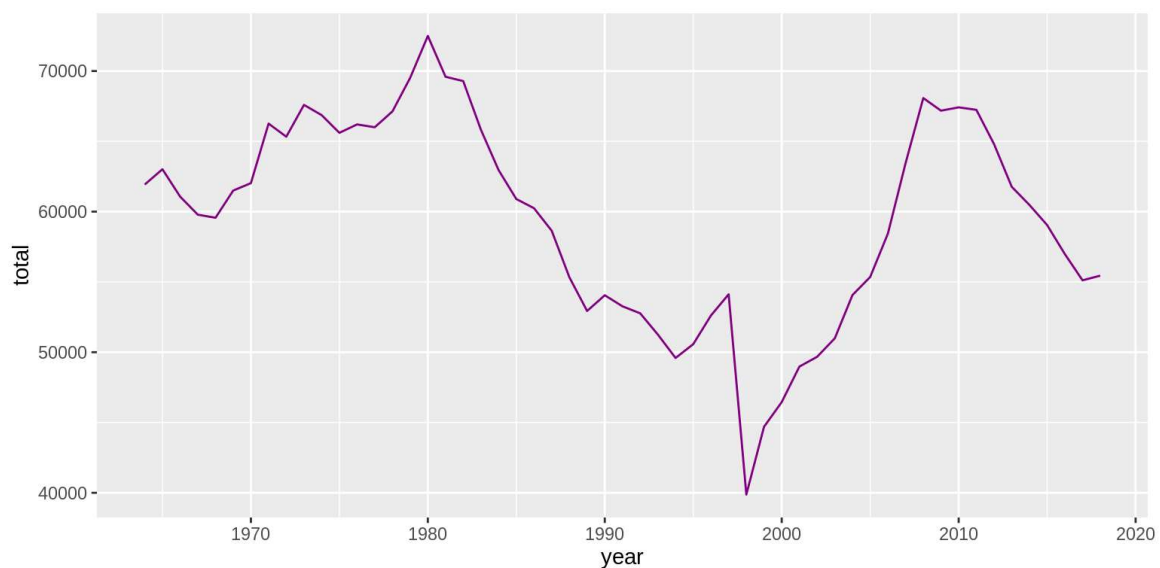
1. tidyverse - Used to reading and manipulation.
2. shiny - Used to plot interactive charts
3. rebus - Used in string operations.
4. grid / gridextra - Used for displaying charts in a grid format.
5. Dplyr: Used for data manipulation.
6. rmdformats - Used for report generation.
7. readr - Used to read the Zip folder.
8. plotly - Used to plot interactive charts.
9. ggplot2 - Used to plot charts.

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.



```
## [1] "Highest number of births from 1964 to 2018?"
```

```
## # A tibble: 1 x 2
##   year `Total Births`
##   <int>         <dbl>
## 1  1980           72498
```

```
## [1] "Lowest number of births from 1964 to 2018?"
```

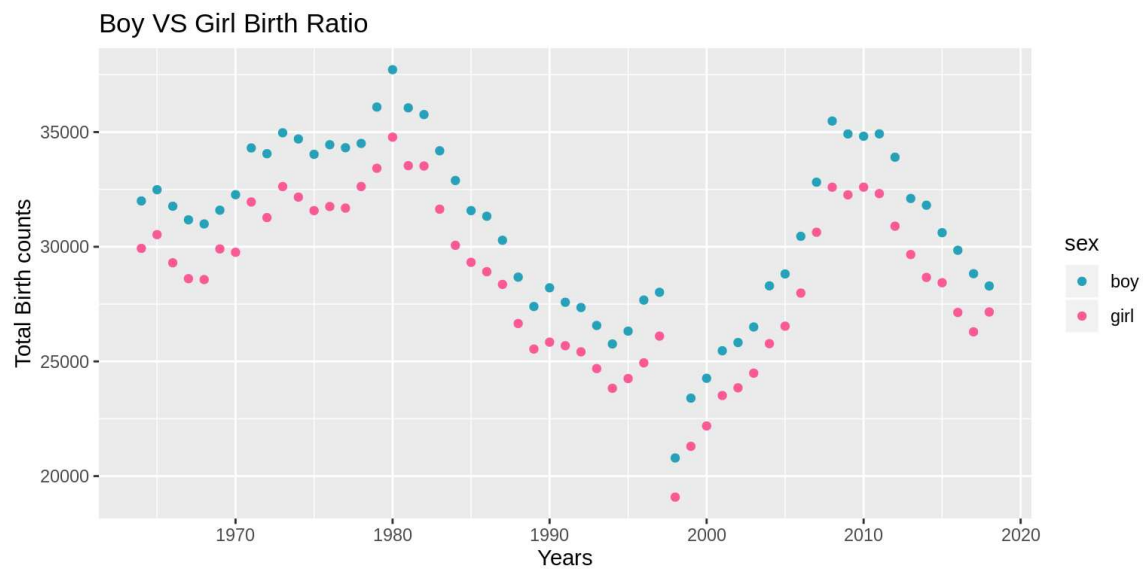
```
## # A tibble: 1 x 2
##   year `Total Births`
##   <int>         <dbl>
## 1  1998           39875
```

```
## [1] "Average number of births from 1964 to 2018?"
```

```
## [1] 59478.6
```

Observation 2

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



```
## [1] "Highest Boy-Girl birth difference from 1964 to 2018?"
```

```
## # A tibble: 1 x 2
##   year `Difference Between Boy-Girl Birth`
##   <int>         <dbl>
## 1  2014           3149
```

```
## [1] "Lowest Boy-Girl birth difference from 1964 to 2018?"
```

```
## # A tibble: 1 x 2
##   year `Difference Between Boy-Girl Birth`
##   <int>         <dbl>
## 1  2018           1135
```

```
## [1] "Average Boy-Girl birth difference from 1964 to 2018?"
```

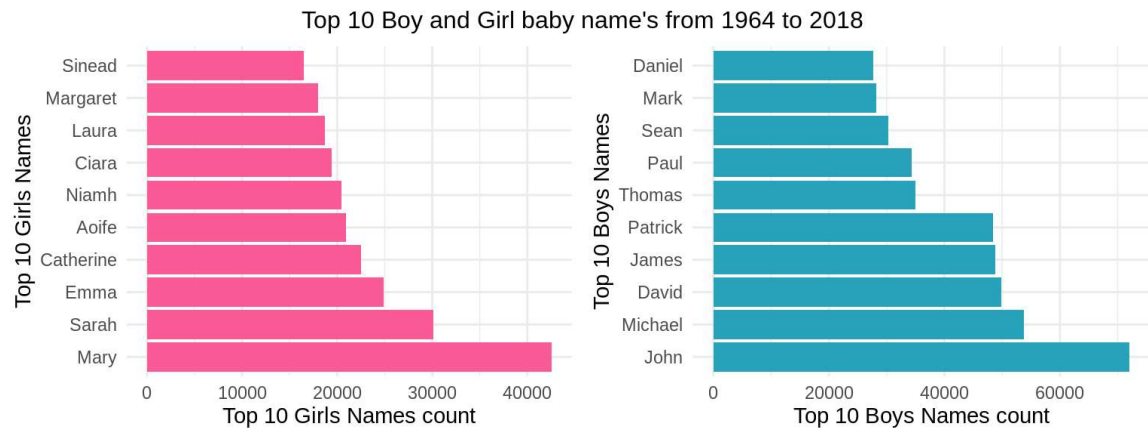
```
## [1] 2311
```

Observation 3

Top 10 Boy and Girl baby name's from 1964 to 2018.

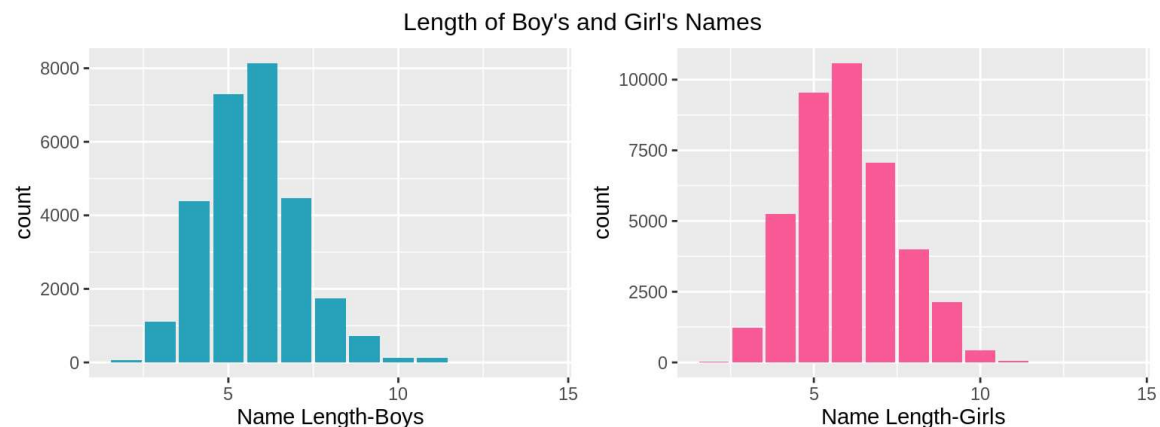
It is observed that the name 'Mary' is the most popular Irish Girl baby name that was kept over the years from 1964 to 2018.

More than 70000 Irish Boy babies where named as 'John' from 1964 to 2018. It has been the most desired name over the years.



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.



```
## [1] "Maximum name length of the Irish Girl Baby name from 1964 to 2018?"
```

```
##           Name length year
## 1 Oluwaseyifunmi      14 2006
```

```
## [1] "Maximum name length of the Irish Boy Baby names from 1964 to 2018?"
```

```
##      Name length year
## 1 Oluwatimileyin    14 2002
## 2 Oluwatimilehin    14 2004
## 3 Michael Junior    14 2010
```

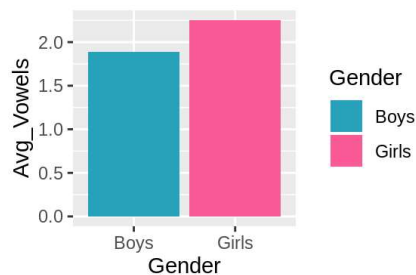
Observation 5

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 occurrence Starting letters for both sex is from U to Z.



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.

```
## [1] "Highest percentage increase in popularity since 1964 to 2018 for Irish Girl names"
```

```
##      Name      Diff
## 1 Emily 1.633839
```

```
## [1] "Highest percentage decrease in popularity since 1964 to 2018 for Irish Girl names"
```

```
##      Name      Diff
## 1 Mary -11.21039
```

```
## [1] "Highest percentage increase in popularity since 1964 to 2018 for Irish Boy names"
```

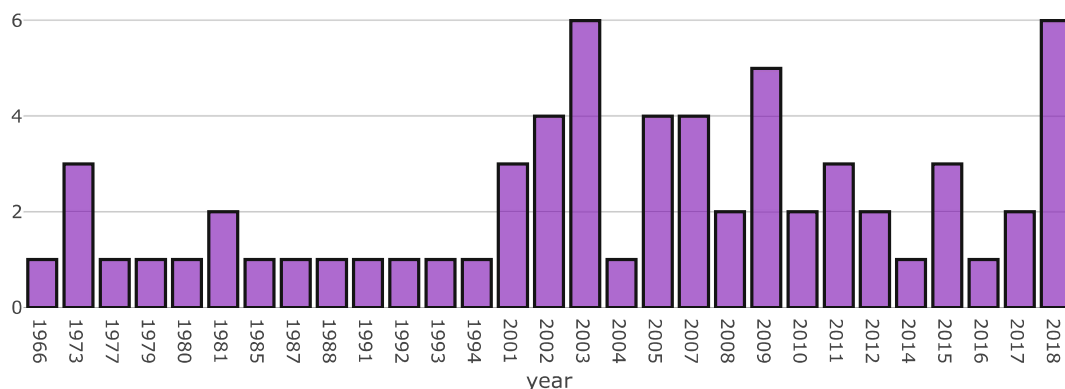
```
##      Name      Diff
## 1 Conor 1.425373
```

```
## [1] "Highest percentage decrease in popularity since 1964 to 2018 for Irish Boy names"
```

```
##      Name      Diff
## 1 John -10.69919
```

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.



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.

About Me

Name:- Pradeep Gurunathan

Email ID- pradeepguru6464@gmail.com

Linkedin: Linkedin profile (<https://www.linkedin.com/in/pradeep-gurunathan>)

Phone Number- +353-894877760