Analysis of US Birth Rates by State Using Google Cloud Platform

Brian Cervantes Alvarez
January 8, 2025

Project Overview

This project looks at birth rates in different U.S. states using data from a public dataset on BigQuery. The goal is to gather, clean, and visualize data about the number of births in each state to understand trends across the country.

Extracting the Data

The birth records from 1969 to 2008 were collected using Google BigQuery, a tool that lets you work with large datasets. The dataset, called bigquery-public-data.samples.natality, was accessed using Python and the google-cloud-bigquery library.

Cleaning the Data

Some parts of the data were incomplete. For example, a few entries were missing state names or birth counts. These incomplete entries were removed so that the final data would be accurate for creating charts.

Exploring the Data

The cleaned data was grouped by state to count the total number of births for each one. A bar chart was used to show the birth counts by state. This helped identify patterns, such as which states had higher or lower birth numbers.

Modeling the Data

No advanced math or predictive tools were used. The project focused on organizing and visualizing the data. The bar chart acts as a simple way to model the information by showing the birth distribution across states.

Conclusion

The bar chart shows the total number of births for each state, with states listed from highest to lowest birth counts. States with tall bars had more births, while those with shorter bars had fewer. This chart helps us see differences in birth rates between states and think about what might be causing them, like population size or economic factors.

Command Line Output

```
$ python main.py
Running Python 3
Connected to the bigquery client.
Found empty value, skipping it
Generating bar chart
2025-01-08 21:37:01.516 python[56361:5313450] +[IMKClient subclass]:
    chose IMKClient_Modern
2025-01-08 21:37:01.516 python[56361:5313450] +[IMKInputSession
    subclass]: chose IMKInputSession_Modern
2025-01-08 21:38:42.685 python[56361:5313450] The class 'NSSavePanel'
    overrides the method identifier.  This method is implemented by
    class 'NSWindow'
```

Data Visualization

Figure 1: Number of Births by State