**SUMMARY**

QUTEZ is a website that collects data on baby names over time. This report presents the objectives, exploration, and insights derived from the analysis of the data collected over the period of 30 years. Explorative questions were posed and answered in order to get an overview of the data, after which analytic questions in form of queries were used to derive interesting insights from the database.

**DATA OVERVIEW**

The analyzed data was stored in a database named “baby\_names\_db” which compromised of two entities – NAMES and REGIONS. Attributes including Names, gender, year, state and births were assigned to the NAMES table, while States and Region belonged to the REGIONS table.

The state column being primary key in the REGIONS table, and a foreign key in the NAMES table established a connection between both of the entities.

Precisely, the data was a collection of baby names, the recipients’ gender, state and region of birth, and the number of babies given each name. The data was collected over a period of three decades.

**QUESTIONS**

DATA EXPLORATION

* How long was the data in the database collected for?
* What is the total number of baby names in the database
* How many male names are in the database?
* How many female names are in the database?
* How many births were recorded in the database?
* How many unique regions are there?
* How many unique states are there?
* How many states are in each region?

KEY QUESTIONS

* Top five (5) most popular male names in the first decade
* Top five (5) most popular female names in the first decade
* Top five (5) most popular male names in the last decade
* Top five (5) most popular female names in the last decade
* What are the single most popular names in each region
* What names overlap across gender (unisex names)?
* Which names have dropped in popularity over the years
* Which year had the highest number of unique names?
* How has the number of babies named “Emma” or “Michael” changed over time?
* What are the top 5 new names introduced in the most recent year that weren’t in use the previous year?

**ANSWERS**

DATA EXPLORATION

* The data stored in the database was collected for a period of thirty (30) years from 1980 to 2009.
* There are a total of **22,240** baby names in the database
* There are a total of **9,730** male names in the database
* There are a total of **14,474** female names in the database
* A total of **98,730,863** births were recorded in the database including **45,856,778** female births and **52,874,085** male births.
* Seven (7) regions are recorded in the database including, South, North, Midwest, New England, Pacific, and Mountain
* Fifty-one (51) unique states are recorded in the database

KEY QUESTIONS

* The most popular male names in the first decade include Micheal, Christopher, Mattew, Joshua and David. This shows that children were mostly given biblical/religious names in the time.
* In the first decade, the most popular female names include, Jessica, Jenifer, Amanda, Ashley and Sarah inferencing that female children had more conventional names at the time.
* The top 5 male names used in the last decade were, Jacob, Micheal, Joshua, Mattew and Christopher which suggests that over the years, many parents still prefer to give their male children religious names, some of which are a repetition of those used in the first decade.
* In the last decade, the most popular female names include, Emily, Maddison, Emma, Hannah and Oliver all of which are conventional names, different but consistent with the trend observed in the first decade.
* Except for the south region where the most popular name used was Christopher, Micheal was found to be the single most popular name used in each of the other regions. This means that the name Micheal was consistently used given to male children in about 80% of the total regions.
* A total of 1,964 names were found to overlap between the male and female genders, of which Michael, Christopher, Mattew, Joshua and Jessica were found to be the top five most common unisex names.

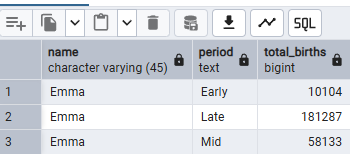
Seeing as the first four names are typical male names. And the last name a typical female name, further analysis to show the years in which these names were used interchangeably was carried out. The result showed that they were used in year 1981, 1980, 1982, 1983 and 1984 respectively.

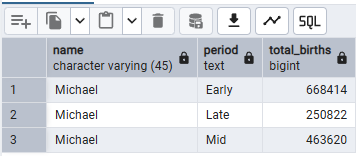
These being early years of the first decade suggests that at the time, there were no outright distinction in terms of gender between the names but came up in the latter years.

* Over the period of 30 years, names such as Micheal, Jessica, Christopher, Amanda, Asley, Robert, David, and James were found to have reduced in popularity this could be due to the invention of new baby names in the latter years.

Even though “Michael” emerged as second most popular name in the last decade, it appears to have undergone the most decline in popularity over the years. This means, it’s decline was very rapid and was in the latter years of the last decade.

* The number of babies named “Emma” and “Michael” was observed to have changed significantly over the years.





As seen in the result table, “Michael” has consistently decreased in use over the years with the least number of occurrences recorded in the last decade, while “Emma” is observed to have consistently increased in use with the highest occurrences recorded in the last decade. This means “Emma” became a more popular name over time, probably due to its conventionality unlike “Michael” which slowly lost its place as a first-choice name.

* Greidys, Chastelyn, Everly, Caylie and Greydis were found to be the top five new names introduced in the year 2009. They being female names shows that in the said year, female names were more invented than male names.

Additionally, analysis revealed that although there were male births than female births in the record, there were more female names than male names. This suggests that male names were more repeated on a ratio of approximately **99:1**, compared to female names with a ratio of **86:1**.

**CONCLUSION**

This analysis of baby naming trends over a 30-year period (1980–2009) has provided significant insights into the evolution of naming preferences across different regions, genders, and time frames. The data reveals that while male names were fewer and more frequently reused, female names were more diverse. Religious and conventional names dominated the early decades—particularly among male children, with names like Michael, Christopher, and Matthew remaining prominent.

Over time, however, certain names such as Michael, Jessica, and David showed a marked decline in popularity, while others like Emma steadily rose to prominence. The emergence of new names in the most recent year, especially among females indicates a shift towards more modern or unique naming preferences.

Regional analysis also revealed a strong consistency in naming trends, with Michael emerging as the most common name in nearly every region except the South. Furthermore, a significant number of names (1,964) were found to overlap across genders, indicating a considerable presence of unisex naming patterns, especially in the early years of the dataset.

**RECOMMENDATIONS**

QUTEZ should integrate a name trend tracking feature that visually shows name popularity over time, automate the detection of new and declining names year-over-year, and apply predictive analytics to forecast future naming trends. This could help expecting parents make informed, trend-conscious choices.

The dataset should be extended beyond 2009 to examine whether the observed trends continued, reversed, or introduced new patterns post-2010. Additionally, explore correlations between name trends and events (e.g., celebrity influence, economic changes, pop culture).