



Application for Big Data Visualization using Google BigQuery and Google Charts

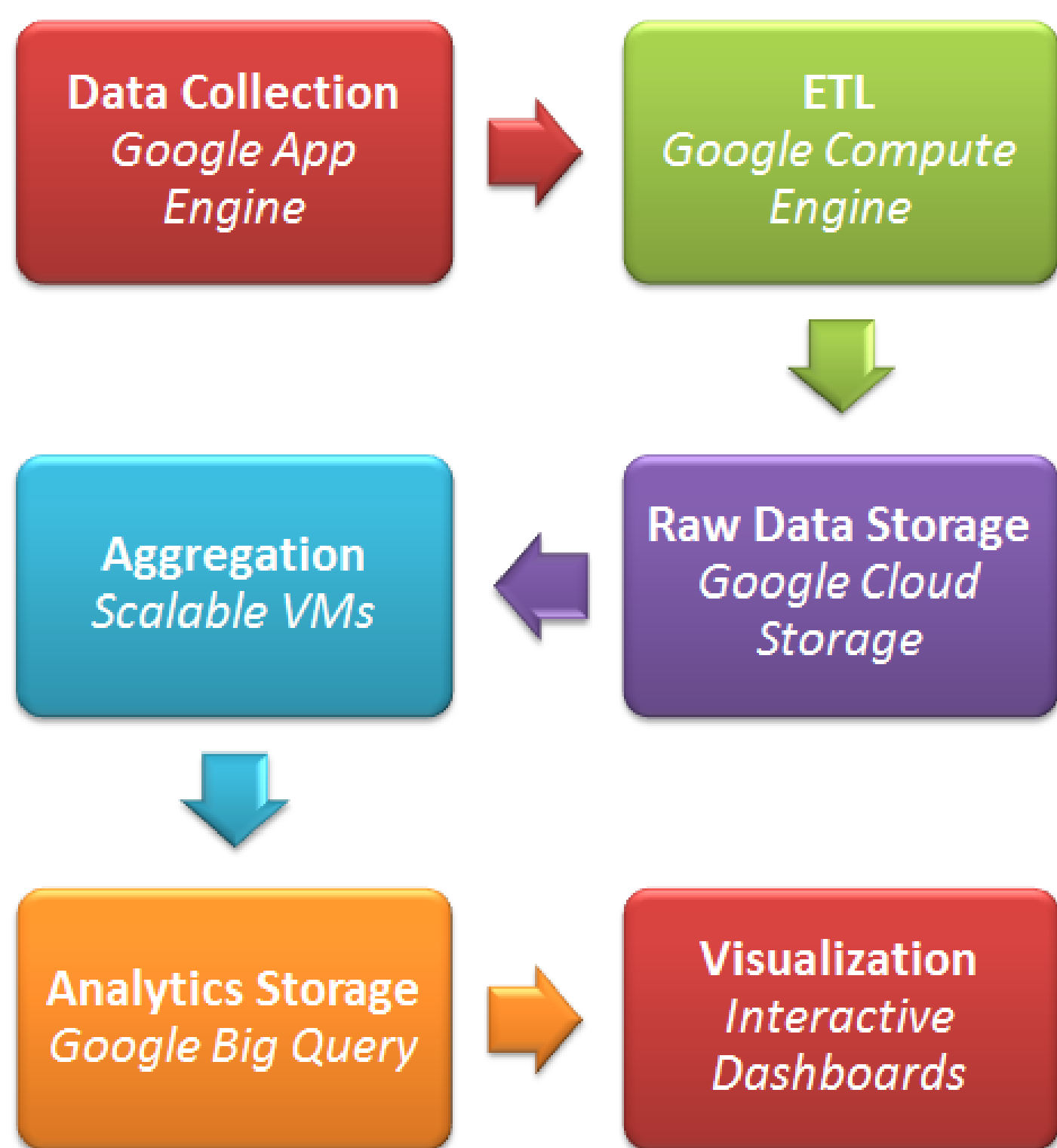
Reem Alattas, Prof. Jeongkyu Lee

University of Bridgeport, Department of Computer Science & Engineering

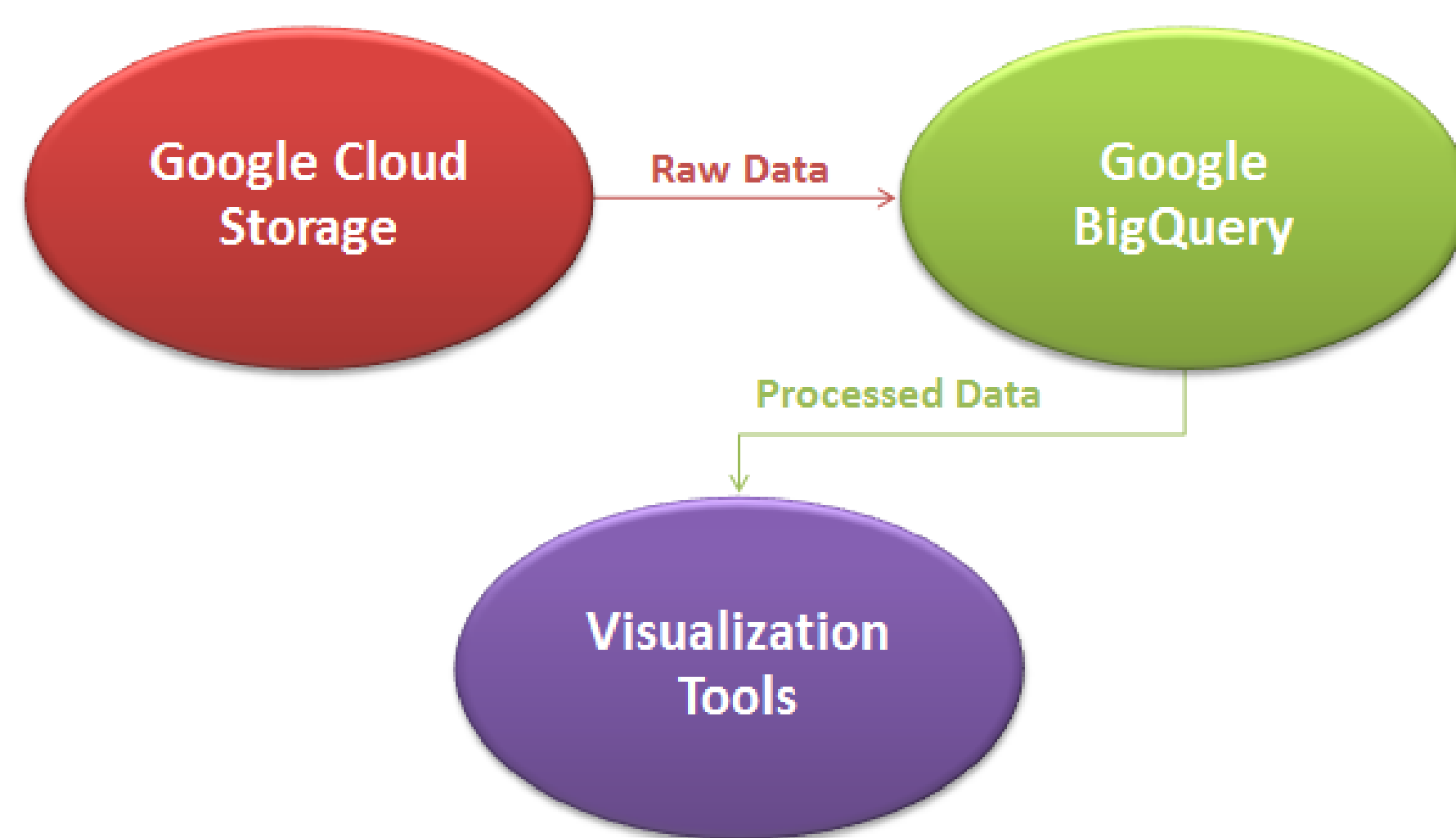
Abstract

Big data visualization conveys information much faster than tables containing numbers and text. Therefore, this poster presents an efficient method for big data visualization on top of Google BigQuery. Google BigQuery is a cloud-based interactive query service for massive datasets built on Dremel. In this project, we used 'natality' dataset that has United States birth data from 1969 to 2008 saved in 137 million rows. For the implementation of dashboard application, we employ Google App Engine using Python and Google Charts with JavaScript.

System Architecture



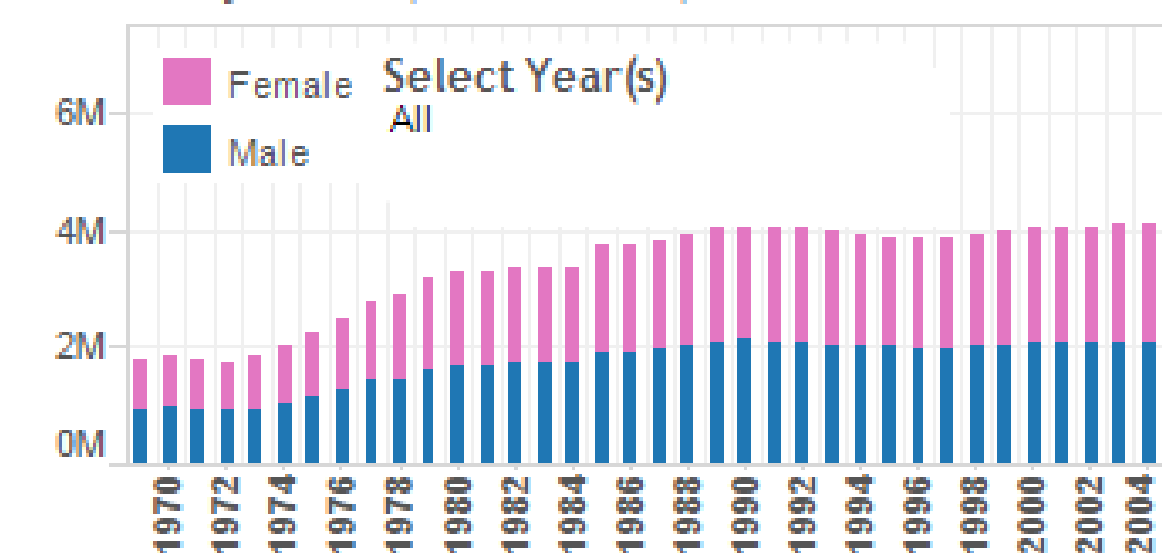
Data Flow Diagram



Results

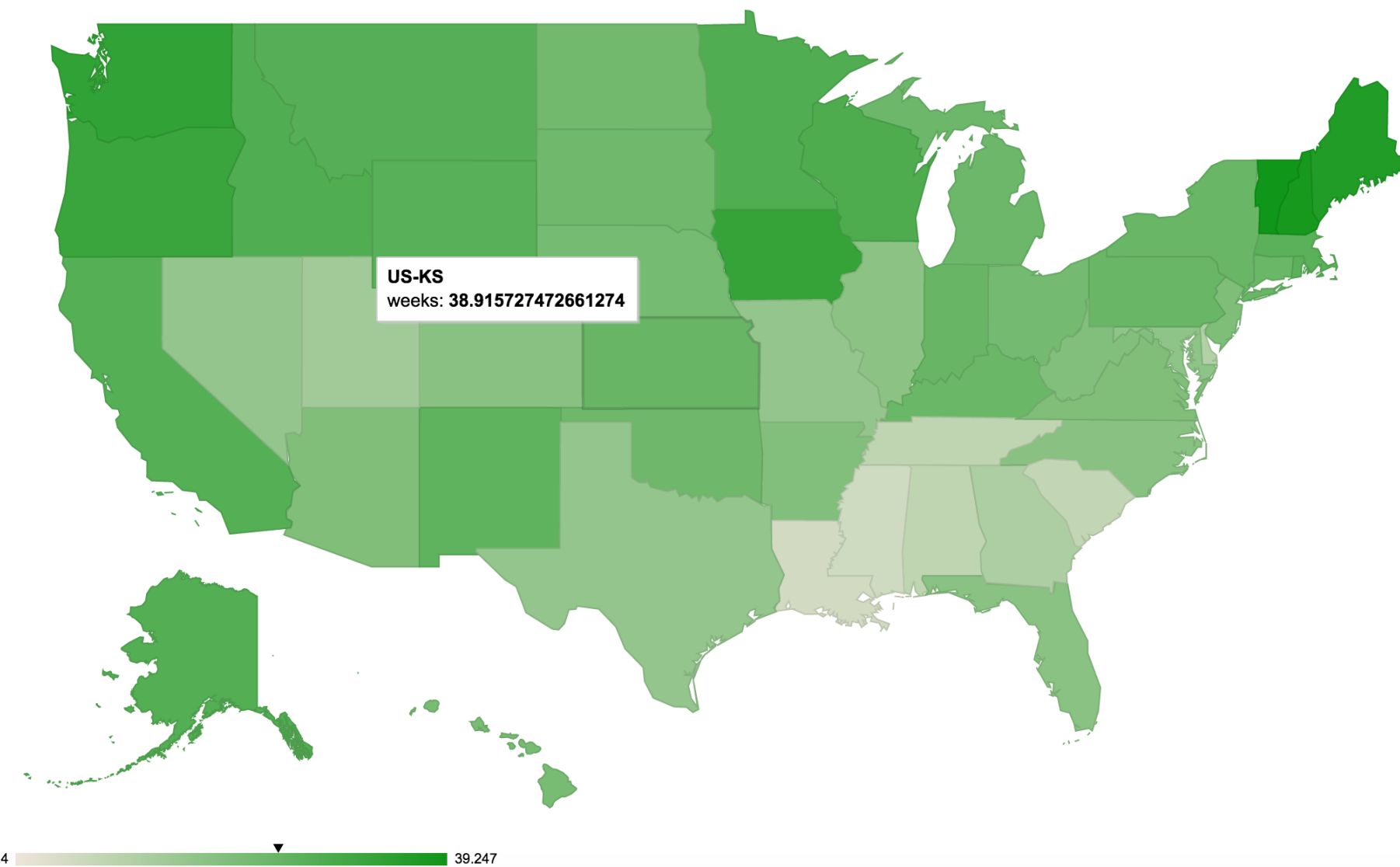
The resulting web application is a graphical dashboard running on development App Engine server and backed by Google BigQuery. It contains a map of the United States with areas identified in colors to answer the question, *Do gestation periods vary by state?*

Births by Year (in millions)



Total Births By Gender

Female	58,784,936
Male	61,719,636



Conclusion

Visualization enables analyzing massive amounts of data in seconds compared to tables. Therefore, we built a dashboard using Python version of Google App Engine to present big data on a map. Creating user credentials was done via OAuth2 to wrap functions that require authentication with decorators. Queries were written in SQL-like syntax to select data from Google BigQuery NoSQL database. Finally, the United States map was made using Google Charts JavaScript classes.

Future Work

The most important direction for future work would be the implementation of additional drag and drop features to enable aggregating data, which in turn simplifies calculating new data items and breaking data by certain factors for better and easier analysis.

References

- [1] Keim, Daniel, Huamin Qu, and Kwan-Liu Ma. "Big-Data Visualization." *Computer Graphics and Applications, IEEE* 33.4 (2013): 20-21.
- [2] LaValle, Steve, et al. "Big data, analytics and the path from insights to value." *MIT Sloan Management Review* 21 (2013).
- [3] Begoli, Edmon. "A short survey on the state of the art in architectures and platforms for large scale data analysis and knowledge discovery from data." *Proceedings of the WICSA/ECSA 2012 Companion Volume*. ACM, 2012.