

CS 329E Elements of Data Visualization: Final Project Report

Interactive Visual Analytics Dashboard using Apple's COVID-19 Mobility Data
Visit the project repository at https://github.com/dmbrannon/datavis_dashboard

Introduction

Our world is currently in a state of uncertainty and turbulence due to the “coronavirus disease 2019”, abbreviated to COVID-19 by the Center for Disease Control (CDC, 2020). The disease was first announced on January 5, 2020, by the World Health Organization as a case of “pneumonia of unknown cause” in Wuhan, China (NPR, 2020). On March 11, just two months later, the WHO declared that COVID-19 was officially characterized as a pandemic (WHO, 2020). As a Civil Engineering major, I am incredibly interested in how the pandemic has affected the transportation of people across the globe. Due to this interest, I decided to focus my final project on Apple’s COVID-19 mobility data, which contains information on the number of requests for directions within the Apple Maps platform from countries all over the world (Apple, 2020). I wanted to see how effective it is for a government entity to issue “stay-at-home” orders as well as how quickly populations respond to such orders.

Methods

In order to understand how the COVID-19 pandemic has affected transportation across the globe, I decided to plot the percentiles given in the Apple mobility data versus the corresponding date. The dataset was fairly clean, other than some sub-regions in various locations (Switzerland, Thailand, Brazil, etc) lacking any values in their respective rows. This did not affect the calculations done in the notebook other than the graphs for those sub-regions being blank. The other notable aspect is that while many locations had data for all three mobility types (driving, walking, and public transit), some locations did not. This affected the analysis only in that the graphs for these locations do not display a plotline for a location’s missing category.

The main technologies used in this project are Jupyter Notebooks, Pandas, Matplotlib, and Bokeh. Each technology has its own unique and useful features. Jupyter Notebooks is an open source web application used to create documents containing live code, visualizations, and narrative text. The beauty of this application is that it brings together both the aspects of human-readability and machine executability. It is currently one of the top tools used for data cleansing, statistical modeling, data visualization, machine learning, and more. For these reasons, it provided the ideal platform to showcase both the project code as well as the resulting visualization dashboard. Pandas is an open source data analysis and manipulation library built for Python. It is mainly used by data scientists who are working with data stored in tabular formats like .csv and .xlsx, making it the optimal candidate to ingest and manipulate Apple’s mobilitytrends.csv file. Matplotlib is a comprehensive open source library for creating static, animated, and interactive visualizations in Python. I decided to start the visualization portion with matplotlib because I have worked with it extensively before this project, making it the

perfect launching point for me to easily explore and visualize the dataset. After getting comfortable with the contents of the dataset, I moved on to Bokeh for the interactive dashboard portion for the project. Bokeh is an interactive visualization library that provides elegant, concise construction of versatile graphics. It is a library I was introduced to during this class and I decided that this project would be a good opportunity to get familiar with its inner workings.

Conclusion

An analysis of the data showed that the world's population did in fact respond to orders from government agencies across the globe to reduce movement. In Figure 1 below, it can be seen that the inflection point for Italy occurred in late February. For a point of reference, the Italian government sealed off 11 towns with police and military checkpoints on February 23, 2020 (*New York Times*, 2020). The Prime Minister then proceeded to halt transportation for most of northern Italy on March 8, 2020. The inflection points for the United States, Germany, and the United Kingdom all appear to happen within the first couple weeks of March. It should be noted that President Trump issued orders to control the spread of the virus in the US on March 16, 2020 (*NPR*, 2020).

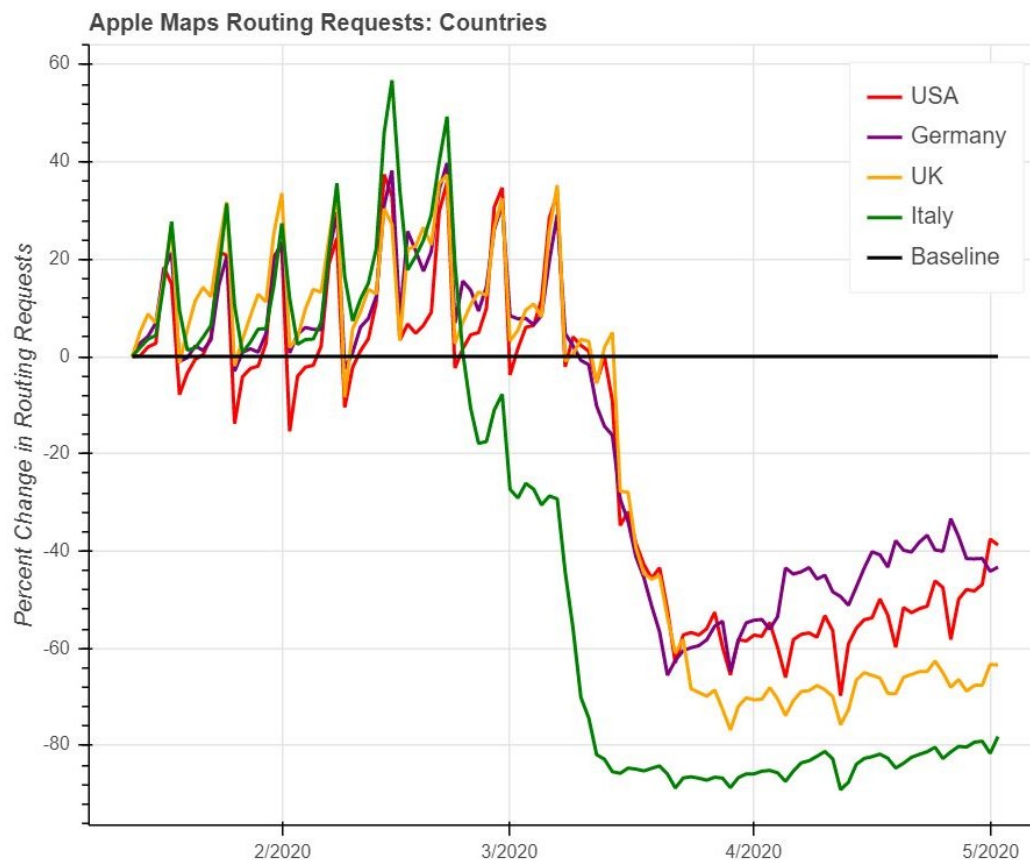


Figure 1: Routing request data for the US, Germany, Italy and the UK

Figure 2 shows the mobility data for Austin, Texas. Governor Greg Abbott issued a disaster proclamation on March 13, stating that the virus posed “an imminent threat of disaster for all counties in the State of Texas” (*Texas.gov*, 2020). It is clear from the figure that the residents of Austin quickly

responded to this proclamation, with Apple Maps requests going negative starting on March 14. Before the inflection point, the figure shows that routing requests had a cyclical nature, with driving and walking requests usually peaking on Fridays or Saturdays. Another notable feature of the graph is that public transit and driving requests usually experienced relatively similar magnitudes of demand before the inflection point. After that point, the data follows the logical conclusion that public transit took the hardest hit among the three categories.

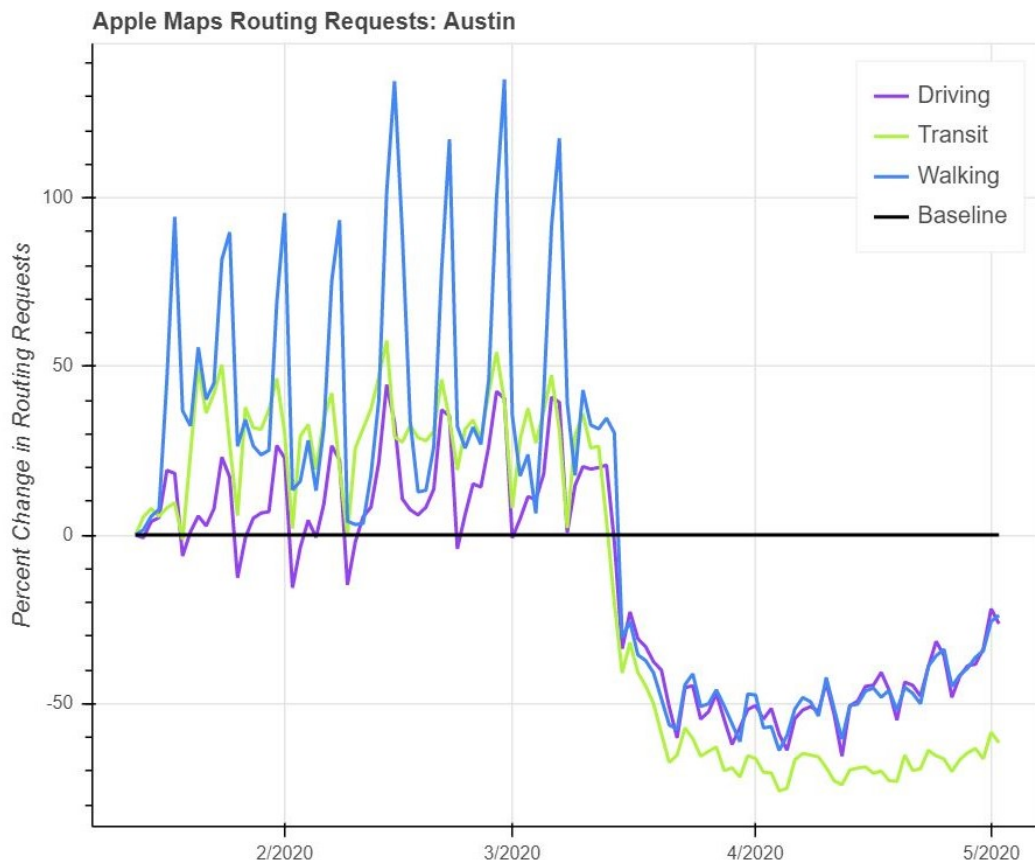


Figure 2: Routing request data for the city of Austin, Texas

On April 27, 2020, Governor Greg Abbott issued Executive Orders to reopen Texas, thus ending the state's quarantine order (*Texas.gov*, 2020). There is fear among the state's population that ending the quarantine too early will lead to a second wave of infections. There are numerous modeling efforts being conducted to view what this second wave could look like (*TACC*, 2020). Only time will tell the full extent of the impact. A continuation of this project would be to create a web application using Plotly's Dash framework and deploy it on a Heroku server in hopes that more people would be able to interact with it. With all the misinformation that is spreading around about the pandemic, it is important to make reliable information easily accessible to the public. For now, I will continue to monitor the dataset as Apple updates it to see if the transportation systems of Austin, Houston, and Dallas begin to return to normal.

Sources

Apple. (2020). *COVID-19 Mobility Data*. Retrieved from <https://www.apple.com/covid19/mobility>

Centers for Disease Control and Prevention. (2020). *Coronavirus Cases Summary*. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/summary.html>

National Public Radio. (2020). *A Timeline of Coronavirus Comments from President Trump and WHO*. Retrieved from <https://www.npr.org/sections/goatsandsoda/2020/04/15/835011346/a-timeline-of-coronavirus-comments-from-president-trump-and-who>

New York Times. (2020). *Italy, Pandemic's New Epicenter, Has Lessons for the World*. Retrieved from <https://www.nytimes.com/2020/03/21/world/europe/italy-coronavirus-center-lessons.html>

State of Texas Governor's Office. (2020). *Governor Abbott Issues Executive Order Mandating 14-Day Quarantine For Travelers Arriving From New York Tri-State Area, New Orleans*. Retrieved from <https://gov.texas.gov/news/post/governor-abbott-issues-executive-order-mandating-14-day-quarantine-for-travelers-arriving-from-new-york-tri-state-area-new-orleans2>

Texas Advanced Computing Center. (2020). *UT Austin COVID-19 Modeling Consortium*. Retrieved from <https://covid-19.tacc.utexas.edu/>

World Health Organization. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19*. Retrieved from <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>