Internet Broadband and Mobile Speeds Across the World

Group One: Emma Abraham, Devin Seay, Melanie Le, Zinat Edjlali

Abstract

This research project was conducted to learn more about broadband and mobile internet speeds across the globe. This interest stemmed from the first dataset that we found that reported broadband and mobile speeds where cursory glances of the dataset revealed that some countries we would not have expected to be in the outlying parameters resided. In particular, we wanted to investigate what factors contribute to the difference in mobile and broadband internet speeds between countries, particularly regarding GDP (gross domestic product). Our results showed that GDP was not necessarily correlated with a country's broadband and mobile internet speeds following data processing and analysis. Further researching the causes behind an area's internet speed leads us to conclude that a country's infrastructure - how well developed the country is internally - is more correlated with faster internet speeds. We present several graphs showing the global differences in broadband and mobile internet speeds as well as a fitted correlation between rurality and internet speed. When interpreting the graph of GDP of countries with the corresponding broadband speeds, it showed a weak correlation between the two, and a definitive conclusion could not be determined about whether GDP had a significant influence on the broadband speed of a country.

Introduction

When we began the project we wanted to learn more about factors contributing to differing broadband internet and mobile speeds for countries around the world. Broadband internet speed is the transmission of data of wide bandwidth. It is a high speed internet connection that is always on and requires a physical connection. Mobile data refers to how fast your mobile connection is and does not require a physical connection. Both broadband internet speed and mobile speed are measured in the unit of Mbps, or megabits per second. After finding the dataset centered around this topic, we were interested to see what attributes of a dataset would be flexible enough to transform information into an understandable visualization. In selecting our data, we wondered whether the dataset would require a lot of cleanup or, to that point, if it was feasible for us to use the dataset for our intended purpose. Because

we all reside in the United States, we thought it would be interesting to use data and conclusions made about the United States to understand the importance of the data. The United States was used to provide context about what the data means compared to all of the other countries included in this analysis.

When we discussed the techniques we wanted to use to present our findings, we thought it would be interesting to learn a technique we had not learned previously. So we used some of the skills we learned from our lab assignments as a starting point. As we went through the project we had some instances where we had to regroup and discuss our findings as they did not match our hypothesis/expectations. Referring to the hypothesis about GDP having an effect on the speed of broadband, after having reformed a regression model on GDP versus global internet speed, we realized that the two were not, in fact, that correlated at all. This prompted us to search for a better and more relevant factor that could impact a country's internet speed. Further research led us to find that the more rural a country, the more likely it is to have slower internet speeds and vice versa.

Data Description

We obtained our primary dataset from Kaggle, as it is a reliable source and chose the topic because it seemed interesting to the group. The dataset covers the different internet speeds of broadband and mobile measured in megabits per second (Mbs) in countries across the world. The dataset reports on the broadband speed, broadband speed ranking, mobile speed, and mobile speed ranking for a total of 179 countries. We thought the data would be significant in reporting the overall accessibility the people within different countries have to a good internet connection and the various activities their connection can support. We noticed that Monaco was ranked the number one fastest for broadband speed and thought it would be interesting to explore, as we expected the countries with higher GDPs like the United States or China to be within the top three or five countries. Our research interests were focused primarily on using the United States as a benchmark for understanding the metrics, as we are familiar with the conditions in the United States. So we thought it significant that the United States was only ranked ninth in broadband speed and twenty-first in mobile speed.

In addition to the primary dataset, we used two secondary data sets to explore relationships between global factors and broadband speed. The first secondary dataset reported on countries' gross domestic product (GDP). The other secondary dataset reported the percentage of a country's population that resided in rural areas.

Data Processing Methodology

      We chose the primary dataset because it appeared to have outliers or standout data points that could reveal a factor that distinguishes countries on a global scale in terms of internet speed. In addition, to see whether internet speeds indicate other signs about the country. We preprocessed the primary dataset to remove the NA values from the mobile speed column which some countries had not reported. We then separated the primary dataset into two separate files, one for broadband internet speed and one for mobile internet speed, to make distinct graphs about both factors. As for the rural population vs. broadband internet speed dataset, we had to manually remove countries that were not listed as the other CSV files in order to make a proper correlation graph between rurality and internet speed. The code was created in Jupyter notebooks, to have access to the Python packages without having to make a new setup locally. The choropleth graphs were made using the plotly package. The correlation graph was made with the seaborn package. Data cleaning was performed using the dropna() function provided by the NumPy package. In order to properly graph the correlation between GDP and Broadband speed (Mbps), we created a function called "Convert" that reduced GDP by a factor of 10^{12} in order to properly graph GDP (2020) versus Broadband (Mbps).

Results

      The first summary barplot shows the ten countries with the highest broadband internet speeds. Monaco, Singapore, and Chile were the countries with the highest values, respectively, and the United States had the ninth highest value of the countries included in this dataset. The second bar plot shows the top ten countries with the slowest broadband internet speeds. The countries with the slowest internet speeds were Afghanistan, Turkmenistan, and Cuba, respectively.

Fig.1 displays a bar graph of the top ten countries with the fastest broadband internet speeds, in units of Mbps
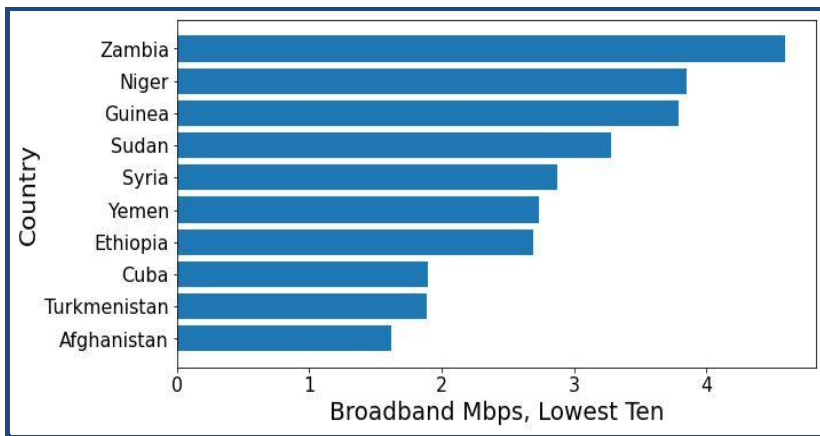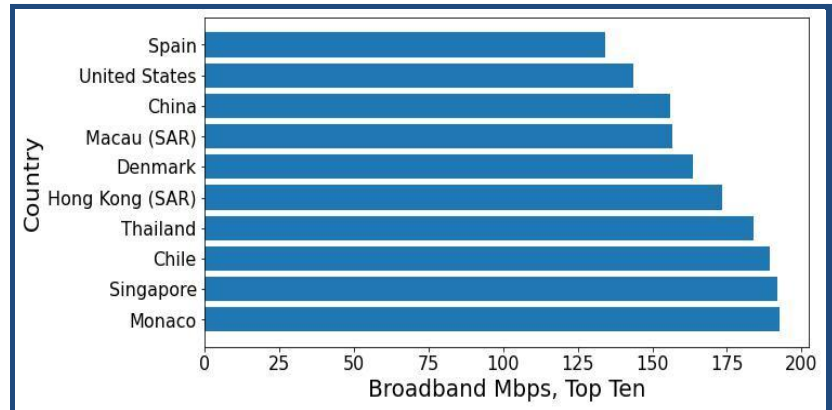

Fig.2 displays a bar graph of the top ten countries with the slowest broadband internet speeds, in units of Mbps
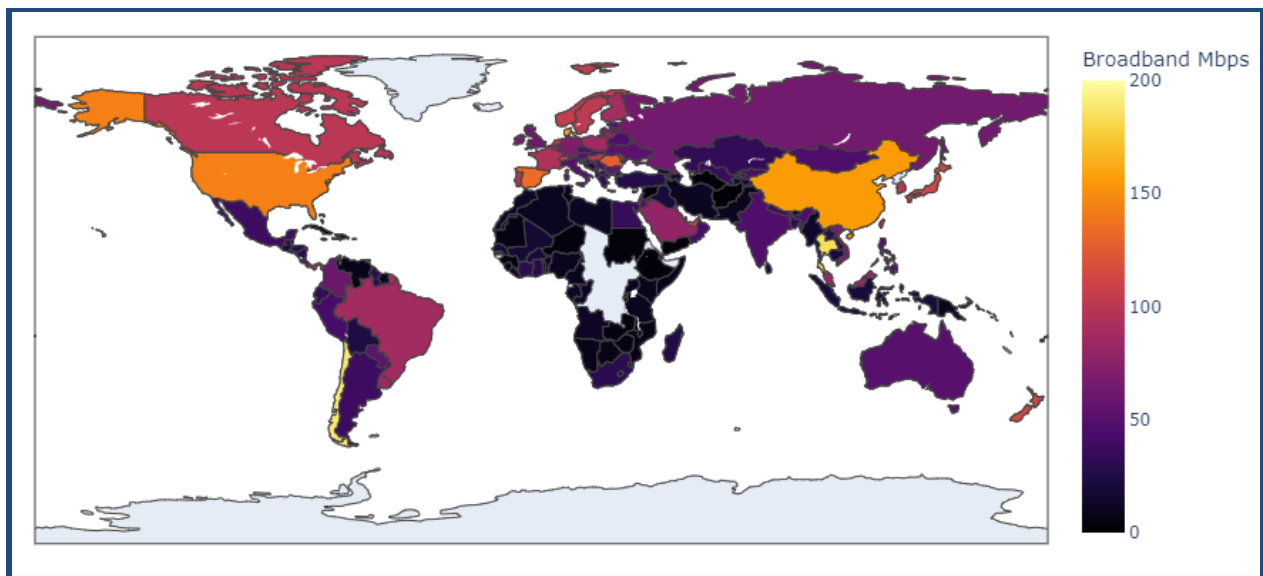

Fig. 3 displays a choropleth graph of global broadband internet speeds in units of Mbps.

A significant thing that we noticed was that, despite the European countries having some of the highest GPDs in the world, they ranked relatively moderate in terms of internet speed. This further contributes to the inference drawn from the data that GDP is not necessarily the best indicator of a country's internet speed. While urban areas are highly developed in most European countries, many towns, villages, and smaller cities are comparably behind in terms of infrastructure, which accounts for the continent's moderate internet speeds.
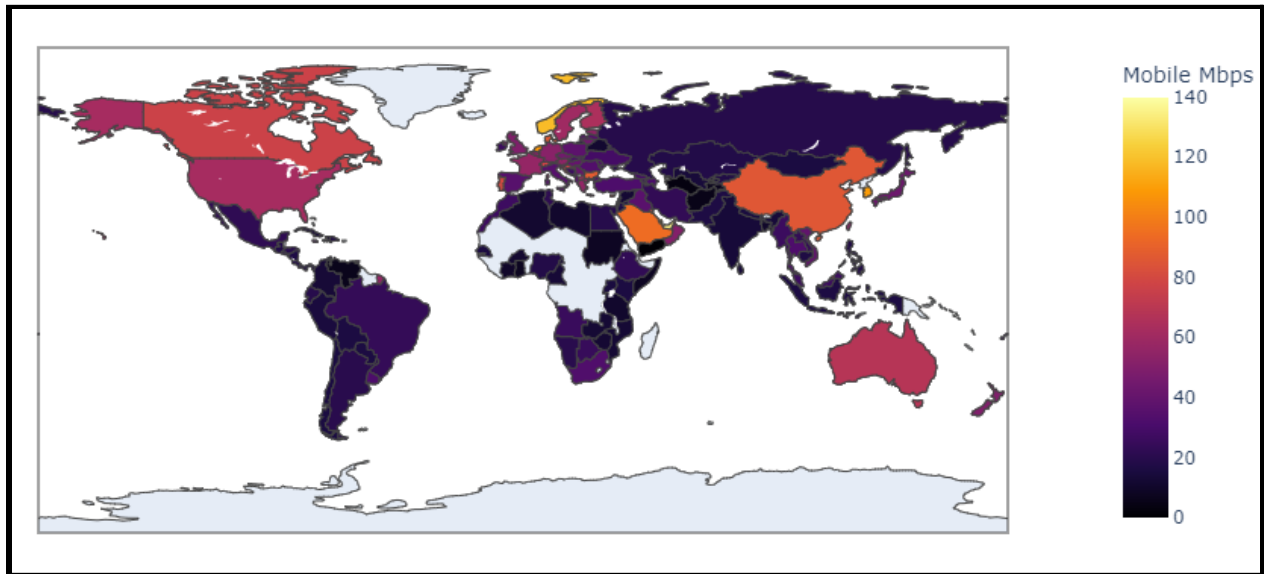

Fig. 4 displays a choropleth graph of global mobile internet speeds in units of Mbps.
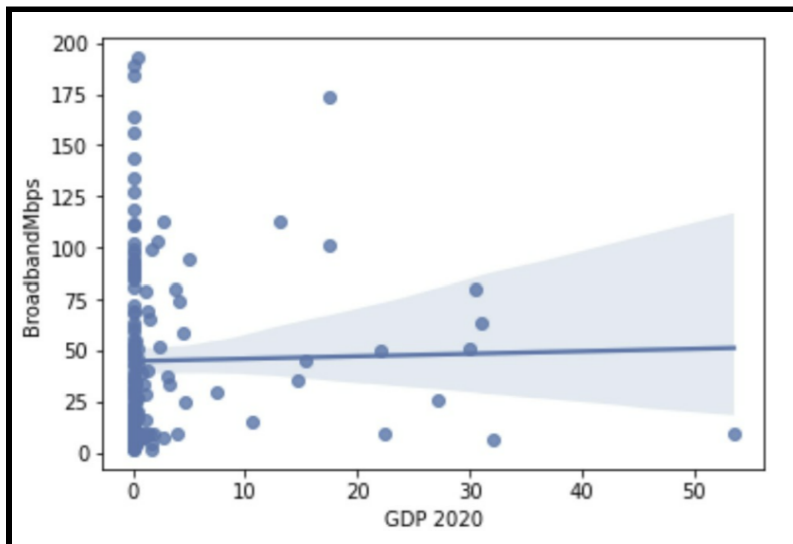

Fig. 5 displays no correlation between global broadband internet speeds (Mbps) of countries and their gross domestic product (GDP) from 2020.

This correlation is best shown in fig. 5, where rural population, or rurality, is plotted alongside

Broadband internet speed (in Mbps). The data for this plot describes a global trend of the correlation between rurality and internet, so that data is not exclusive for a single region but instead supports the conclusion that rurality, rather than GDP, better indicates what a country's internet speeds will be like.

Fig. 6 displays a correlation between global broadband internet speeds (Mbps) of countries and the percentage of the country's population living in rural areas

Testing

```
from unittest import TestCase
max_broadband
tc = TestCase()
test_min = data_broadband.nsmallest(n=10,columns = ['Broadband Mbps'])
print('Test Failures:',tc.assertCountEqual(test_min,min_broadband))

Test Failures: None

test_max = data_broadband.nlargest(n=10,columns = ['Broadband Mbps'])
print('Test Failures:',tc.assertCountEqual(test_max,max_broadband))

Test Failures: None
```

Our group performed two unit tests to confirm that the countries listed in figures 1 and 2 were respectively the countries with the fastest and slowest broadband speeds. After performing unit tests to

test whether or not these graphs were correct, we confirmed that these indeed were, respectively, the top ten countries with the fastest and slowest broadband speeds, as pictured below.

Fig. 7 displays the two unit tests performed and their corresponding results.

<u>Conclusions</u>

From our results above, we concluded that GDP was not necessarily linked with how fast a country's broadband or mobile speeds may be. While countries with high GDP may coincidentally have fast internet speeds, further analysis showed that there was little to no correlation with GDP as a factor in internet speeds. Once it was seen that there was not much correlation between GDP and broadband internet speeds, we looked to other factors and saw that countries with large rural populations had slower broadband internet speeds. This is due to broadband internet's infrastructure requirement, which connects customers with physical cables. In some areas, their connection is through fiber wires which offer faster connections but are more expensive. Some providers use existing telephone wires and connections to allow access to the internet. This option brings slower connections but is less costly as the infrastructure already exists. The US is a high-income nation that can afford the costs of installation and maintenance of rural areas, but there are gaps as sometimes providers refuse to install cables because the cost outweighs what they can reasonably charge. The same cannot be said in lower-income countries that cannot afford the costs of bringing the internet through broadband and instead rely on mobile internet, which is not as fast as broadband internet. These disparities lead to inequalities between rural residents and non-rural residents in the US. This disparity can be seen in rural students left behind because of their poor access to fast internet connections.

<u>References</u>

Antonelli, William. "What's a Good Internet Speed? How to Get the Fastest Internet Speeds and Upgrade
Your Setup." *Business Insider*, Business Insider, 20 Apr. 2020,
https://www.businessinsider.com/what-is-a-good-internet-speed#:~:text=The%20average%20inte
rnet%20download%20speed,internet%22%20by%20the%20FCC).

Early Evidence Suggests Gigabit Broadband Drives GDP - Analysis Group.
https://www.analysisgroup.com/uploadedfiles/content/insights/publishing/gigabit_broadband_sos
a.pdf.

"Early Evidence Suggests Gigabit Broadband Drives GDP." *Analysis Group*,
https://www.analysisgroup.com/uploadedfiles/content/insights/publishing/gigabit_broadband_sos
a.pdf.

"A Guide to Broadband Speed | Moneysupermarket." *Money Super Market*,
https://www.moneysupermarket.com/broadband/speed/.

"How Much Broadband Speed Do Americans Need?" *The Pew Charitable Trusts*,
https://www.pewtrusts.org/en/research-and-analysis/articles/2020/11/30/how-much-broadband-sp
eed-do-americans-need.

"Internet Connection Speed Recommendations." *Help Center*, Netflix,
https://help.netflix.com/en/node/306.

"Mapbox Choropleth Maps." *Mapbox Choropleth Maps in Python*,
https://plotly.com/python/mapbox-county-choropleth/.

"New Study Quantifies the Impact of Broadband Speed on GDP." *Arthur D. Little*,
https://www.adlittle.com/en/insights/press/press-release/new-study-quantifies-impact-broadband-
speed-gdp.

"Ordering in: The Rapid Evolution of Food Delivery - Mckinsey.com." *Mckinsey*,
https://www.mckinsey.com/~/media/mckinsey/industries/technology%20media%20and%20teleco
mmunications/high%20tech/our%20insights/ordering%20in%20the%20rapid%20evolution%20of
%20food%20delivery/ordering-in-the-rapid-evolution-of-food-delivery_vf.pdf.

Holslin , Peter. "The 10 Fastest and Slowest States for Internet Speeds in 2022." Edited by Rebecca Lee
Armstrong, *HighSpeedInternet.com*, 14 Apr. 2022,
https://www.highspeedinternet.com/resources/fastest-slowest-internet#:~:text=The%20national%
20average%20speed%20for,speeds%20have%20gone%20up%2020%25.

Roser, Max, et al. "Internet." *Our World in Data*, 14 July 2015, https://ourworldindata.org/internet.

"Rural Population (% of Total Population)." *Data*,
https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?name_desc=false.

"Visualizing Regression Models¶." *Visualizing Regression Models - Seaborn 0.11.2 Documentation*,
https://seaborn.pydata.org/tutorial/regression.html.

Data Sources:
- Primary Data Set:
  https://www.kaggle.com/prasertk/internet-broadband-and-mobile-speeds-by-country
- Secondary Data Set 1: GDP
  https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?most_recent_value_desc=false
- Secondary Data Set 2 Rural Population Percentage:
  https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?name_desc=false