DATA 201 Group Project Journal

Yazeed, Heran and Chaarvee – Taking an inside look at AirBNBs

24th of September:

Whole Group - We had our first meeting as a group to be in contact with each other, we overlooked the group project information to familiarize ourselves with the task. During this meeting we set up a Trello board which will hopefully help us all keep up to date with what tasks still need to be done for the project and what tasks have been completed. We did still did not produce the final project idea and the datasets for the project during the meeting.

The large gap between the first two dates, was due to us agreeing that we would not start the group project until we all had finished Assignment 2

3rd of October:

Whole Group– Today we communicated with what we would want to scrape for our final project and produced the idea of scraping this website https://www.stats.govt.nz/. We will be comparing population, ethnicity, work and income, education, and housing of the major cities in New Zealand Auckland, Wellington, Christchurch, and Dunedin. Each member of the group will be scraping data for one specific city.

4th of October:

Yazeed – I started on the group assignment however I came across an issue where https://www.stats.govt.nz/. it was not possible to scrape. So, I contacted my group to see if the issue can be resolved or change the website of our project

Whole Group – After Yazeed mentioned the issue of the original website not being scrapable, we had a small meeting in our group chat and decided on changing the website of interest to http://insideairbnb.com/get-the-data/ and https://www.worldometers.info/gdp/gdp-by-country/ . The idea of the project is to compare Airbnb prices from around the world, and we will be using the second site to scrape population numbers and GDP numbers to see if that relates to any of the Airbnb prices around the world

5th of October:

Yazeed – Today I started on the project by reading in the CSV from the Airbnb site for the Airbnb data in New Zealand, I managed to get the scraping and wrangling part done. But I still needed to work on the plotting data which I will work on over the weekend.

Heran Wen - Today I chose data from three countries for my assignment, and I went ahead and did Belgium first. I tried to read its csv link but kept getting errors, then I realized that it was not a csv file but a .gz format, so I changed my approach and got remarkable results.

Chaarvee – I chose the countries whose data I was going to work on. I then read in the csv files into jupyter labs and understood what kind of data I was working with. I could do this for one country today (Spain).

6th of October:

Heran Wen - Today I read all three of my chosen countries successfully on R studio and informed my teammates that my chosen countries are Belgium, Netherlands, and Greece. Then I started to try to graph the read data for these countries.

Chaarvee- I repeated the same process I used for Spain for the data of US.

7th of October:

Yazeed – Today I finished plotting the data for the NZ Airbnb data and then started scraping Airbnb data for Tokyo, which I managed to wrangle and plot the data for Tokyo as well.

Heran Wen- Today I tried to make a price distribution graph, but I ran into difficulties because the price columns in the read data had special symbols, so the code did not run successfully. I tried to remove the special symbols from the price column and successfully made a price distribution graph for three countries.

Chaarvee - I repeated the same process I used for Spain for the data of Australia.

8th of October:

Yazeed – Today I started scraping data for Airbnb's in London and managed to also scrape, wrangle, and plot the data for London then started making some comparison between the other cities/countries I scraped

Chaarvee - I repeated the same process I used for Spain for the data of Thailand.

9th of October:

Yazeed – Today I started scraping the world meters site which has information on world population, GDP and GDP growth. Then plotted the data and made some graphs for each of the countries we had scraped as a group, which helps answer our research question to see if population of the country and GDP per capita had any effect on Airbnb prices.

Heran Wen- Today I tried to make a graph of the number of beds versus the price and wanted to find out if the two were related.

Chaarvee – Today I started on plotting the graphs for US. However, I faced issues with plotting the range of the prices that are present on the x axis. Each bedroom was plotted on the graph which caused the numbers to overlap on the x axis.

10th of October:

Yazeed – Today in the lab I realized that prices for each of the Airbnb prices are in the local currency of the countries scraped which made it a little harder to compare prices in different currencies so I did a little research and found this website https://fixer.io which has an Api that we can hit directly and can convert the prices from different currencies to NZD using the live real time exchange rates which makes it easier for us compare prices of Airbnb's as the prices will be in one currency instead of multiple. I started on creating the price exchange and managed to get some of It working but not all of it.

Heran Wen- Today I tried to make a distribution graph of room types and a graph of bathrooms vs. price, proving that the number of bathrooms is equally related to price

Whole Group – In today's lab, we all kept each other up to date on what we have done and organized an online meeting over zoom for Thursday as we are nearing the end, so it would be good to keep us up to date daily till our presentation day.

Chaarvee – today I tried to research a bit and find out how I can solve the problem with the prices so that they can be plotted on the basis of a range rather than individually.

11th of October:

Heran Wen- Today I have put together data from three countries and created a pair of price distribution graphs. To observe whether GDP and population affect the price distribution of Airbnb.

Chaarvee – Today I finally managed to find a solution for the range problem of the plot. I plotted the bedroom vs prices plot for US. I also plotted the grouped graph for bathrooms vs price and a types of bathroom plot.

12th of October:

Yazeed – Today I continued working on updating the currencies of all AirBNBs me and the group members have scraped to $NZD using the fixer.io API. During the day I managed to get this working finally, which has made it easier for us to compare the prices of all AirBNBs as they are all in NZD now.

Heran Wen- Today I used Yazeed's code and successfully converted currencies from various countries to New Zealand dollars. And successfully ran other price distribution charts

Whole Group – Today we had an online meeting discussing the final parts of the project such as compiling all our graphs together so we can start working on our presentation which we have on Tuesday. We also discussed how we want to present our data and what we want to share in our presentation.

Chaarvee – I realized I had not plotted the price distribution plot. So, I made that plot. I implemented these three plots into all four countries.

13th of October:

<u>Whole Group</u> – Today we had an in-person meeting finalizing the project and started working on our group presentation. We created a PowerPoint so we can all compile our findings and contribute towards the final presentation on Tuesday. We also discussed how we would like to present our findings and what we would like to emphasize during the presentation, such as the main idea of the project, problems we overcame and drawing conclusions. We assigned each other a deadline to finish the presentation which is Monday morning so we can meet during the day and have a run through for the presentation before the main thing on Tuesday. In this meeting we also made a start on writing the final project report.

<u>Yazeed</u> – Today I made a small start to the group presentation by adding some of the graphs I have created and explaining these graphs and what they help us show. I also worked on the introduction of the presentation that gives a concise explanation of our project and what the main goal of it was.

<u>Heran Wen-</u> Today I put the price distribution graph I created and the price box plot for the three countries on a PowerPoint and wrote my conclusions about the relationship between prices and GDP and population.

14th of October:

<u>Yazeed</u> – Today I continued my work for my parts of the presentation. However, I did not get much done.

<u>Heran Wen</u>- Today I merged my two other group members' data with mine and created a box plot of the price distribution across nine countries and managed to conclude that the higher a country's GDP, the higher the price of an Airbnb.

15th of October:

<u>Yazeed</u> – Today I continued to work on my part of the presentation and managed to get most of it done and should have it completed by tomorrow morning before our final group meeting where we will be running through the presentation later in the afternoon.

<u>Chaarvee</u> – Today I added the graphs I had created on jupyter labs to the power point presentation and worked on the presentation.

16th of October:

<u>Yazeed</u> – Today I continued my work on the presentation and managed to get my part done. After this I then created a GitHub repository so we can upload our jupyter notebooks in one place making it easier for the group submission. I also continued to work on the project report by summarizing our project and writing up our ideas, goals, difficulties, what we achieved and what we did not achieve during the project.

<u>Heran Wen</u>- I signed up for a GitHub account today

<u>Whole Group</u> – Today we had our final meeting before our presentation. We had multiple run throughs practicing our presentation. We also continued our work on the final project report.

Chaarvee – Today I prepared for tomorrows presentation using the slides of the presentation.

17th of October:

Whole Group – Today we finally presented our presentation! And continued working on our final project report. We also started uploading all our files onto the GitHub repository

Yazeed – Today I continued to work on our final project report. I managed to write up our goals and ideas for the project, the intended use of our data, what we achieved and what we did not achieve during the project and difficulties we encountered and how we overcame these difficulties.

18th of October:

Whole Group – Today we continued our work on the final project report and uploaded most of our files onto GitHub.

Chaarvee – Today I created a GitHub account.

19th of October:

Yazeed – created a documentation.txt file explaining all our notebooks such as what order they should be run in and what they will perform when being run and uploaded this onto our GitHub repository.

Chaarvee – Today I uploaded all my files onto GitHub.

20th of October:

Yazeed – Today I completed our project report and submitted our presentation into the presentation submission box.

Whole Group – Today we completed our project report and successfully uploaded all our files onto GitHub. We will now do minor checks on our work to make sure we are not missing anything before our final submission.

26th of October:

Whole Group – Today we did minor checks on the project to make sure we are not missing anything and finally submitted our project