Project Team
Iman Hashmi
Rahima Saad
Shahana Ahad
Thays Miran Miller
Ruvimbo Hungwe

Project Overview

We are particularly interested in data regarding Openreach Internet Service Providers, as Openreach is our sponsor. Openreach are rolling out the Ultrafast Fibre Programme across the UK and we aim to build a project closely related to this. We have agreed to investigate different broadband delivery methods, with a focus on Fibre To The Cabinet (FTTC) vs. Fibre To The Premises (FTTP). We aim to use solid data to determine which broadband method is proven to be faster, and how the future may look when everyone is using Full Fibre - will we be better off?

We may also like to consider how the two types of broadband specifically affect Netflix or YouTube usage - for example, does FTTP or FTTC favour faster Netflix download speeds? Which causes longer start up delays?

Other ideas that we have discussed and may consider investigating include:

- Compare the roll out of business fibre to non-commercial
- Is the business broadband any faster? If yes, is the rate of up-take for fibre growing for household/non commercial?
- Cross data for how many people are working from home
- Find data sets for how much data is being used and if fibre is keeping up? Is the rate of growth of speeds keeping up with the demand/consumption?
- Correlate the broadband method to the number of complaints
- Correlate fibre roll out and speed (latency, disconnection, streaming services failure and delay)
- Try to use an api that checks live speed and compare to the historical data
- Datasets for other countries?
- Does economic growth correlate to internet speed?

1. What kind of data research and analysis are you going to take on?

   The data research we are going to carry out will be made up of a combination of real-time observational data from an API and reference data available on the government website. We will store the research data on csv files. In order to summarise our initial findings, we will use descriptive analysis. This process involves gathering and summarising data and will provide an overview of the dataset. We will then use exploratory analysis in order to find relationships in our data to find the answers to our questions. Our project will also include the use of analytical libraries such as Numpy, Pandas and Matplotlib.

2. What industry or areas does it cover?

   We will be covering the telecommunication industry in the UK, looking at the most popular broadband providers.

3. What questions are you planning to answer?
   - FTTC vs FTTP - Does the FTTP upgrade actually make a difference?
     - Compare percentage upload speed, download speed and delays/latency.
     - Compare reliability and speed fttc vs fttp (Could use this from the dataset: Dns failure, jitter, packet loss etc)
   - Comparing streaming services (youtube, bbc iplayer), location, service providers:
     - Is FTTP more reliable in certain locations in the UK?
     - Are certain ISPs superior in terms of their FTTP speeds?
   - Are you better off as a streamer with an FTTP connection?
   - To conclude: In the future when everyone has FTTP are they better off? Why? Why not?

4. What data sources are you planning to use?
   - Gov datasets:
     https://data.gov.uk/dataset/dfe843da-06ca-4680-9ba0-fbb27319e402/uk-fixed-line-broadband-performance
   - API that shows actual speeds

5. Describe the team approach to the project work

   The project will be carried out using the scrum agile methodology. The team will contribute to the product backlog by defining tasks that must be completed for the project to be completed successfully. The team will first create a list of ideas or needs that will be required to accomplish the project. The items on the backlog will then be prioritised based on their benefits/value and projected completion time. Backlog items will also be estimated to determine the amount of effort required to fulfil the requirement. The product backlog will be owned by one team member and will be shared with the rest of the team via google docs.

Example of Product Backlog:

Product Backlog for CFG Telecommunication Project

TASK ID	Story (TASK)	Estimate (Effort)	Priority (1-5)?
1	I want the data to be cleaned to increase productivity and allow for the highest quality information	5	1
2	I want to connect to the API so I can access information I need to answer questions	6	2
3	I want to visualise data to provide further insight into the data	3	3

The team intends to complete the task in iterative 4 day sprints which will require iterative sprint planning. In order to plan the sprint we will first ascertain what outcomes we want to achieve during each sprint to allow us to be in a position to put together the tasks with the goal in mind. The objectives of the sprint will be outlined, tasks will be prioritised, work distributed among the team members, and a time/effort estimate for completion will be provided. After each sprint, we will then hold a sprint retrospective meeting to review the sprint and figure out what went well and what didn't.

6. How are you planning to distribute the workload, how are you managing your code, how are you planning to work on your project?

Initially, we need to meet and break down the project into clear steps. We have already identified everyone's strengths and weaknesses. We can allocate tasks according to these. We will manage our code and share it via GitHub.

We plan to hold a mid-week/ weekly meeting, depending on everyone's preference. This will be used to plan and allocate tasks that need to be completed. A shared google doc will be used to track progress, so everyone can see what needs to be done, and what has been completed.

It may also be worth having a short, once a week meeting if anyone needs help/is struggling, they could attend and we can figure problems out as they arise. Alternatively, if an individual is struggling, they can ask for help in our Slack channel, and a meeting can be arranged.