Data Analyst Technical Challenge

TaxiCab is your neighbourhood friendly Taxi service making your life easier to commute through the busy streets of New York City. Since it began operations in 2019, the taxi service has really picked up and has become the number one service to use in NYC.

Unfortunately, the CEO of TaxiCab, Henry Bentley, does not have complete visibility on how well his company is doing and is not sure if there are potential opportunities that can help his business scale further. Henry has access to all the data but is looking for a Data Analyst with a keen eye to detail to help him with drawing insight from the data. Sometimes the data is very messy and may require additional cleaning before any analysis can be done.

The Data Foundry have won the proposal to help Henry with this problem statement and have chosen you to develop the solution for them. In their requirement gathering phase, these are the areas of concern they want to target:

- Average tip received by a vendor during the day
- Which time of the day is the busiest?
- Clean the Data and fix any anomalies
- Classify trips based on Payment Type (not the number in the excel sheet but the actual payment type. Use metadata table for reference)
- Is there any relationship between congestion surcharge and trip distance?

The outcome needs to be in the form of a report/dashboard/presentation in a visualisation tool of your choice.

Henry Bentley is very excited for this proposal and is keen to provide a huge bonus to the analyst who can provide additional information from the data that he hasn't captured in the problem statement above.

Attached are the dataset from the month of January 2021 and the metadata file to be used as lookup tables. The surcharge data for each tripId is provided in a json file. We are on a tight schedule and our deadline is in exactly one week from today.

Good luck!