Case Study

The purpose of this case study is to demonstrate your data analytics capabilities. You are expected to download, process and analyse the data set in preparation for the interview. In the interview you should run through broader exploration of the data as well as analysis themes focused around the hypotheses given below. The medium of delivery is at your discretion however slides and/or a dashboard are how we would usually report ourselves.

Dataset

The TLC Trip Record Data is a public data set provided by the Taxi & Limousine Commission which collects data for journeys taken in NYC. The necessary data for this case study can be found [here](#).

You should analyse the months Aug-18 and Feb-18 but you can incorporate more months if you wish to do so.

You should look at **yellow cab** journeys only.

Hypotheses

1. There is seasonality in travel patterns and people travel more in Winter than in Summer (if you need to, you can assume Winter is February, and Summer is August)
2. People take more trips in the evenings
3. People tip more on weekends
4. People take longer duration trips on weekends

Framework

- You will receive this case study a week before your interview to allow you to spend as much time on the problem as you see fit. We would expect you to spend not less than 3 hours on preparation.
- There is no restriction on data, research or tools/software you can use.
- The outputs are expected to be delivered in tables and charts within a presentation or a dashboard.
- You should assume the audience has little knowledge of the dataset, so explaining what it is, how it works etc will be important, as well as your conclusions
- You may use any technologies to complete the case study - local/cloud based, from any vendor