Cycling Analysis

The purpose of this assessment is to understand how you approach analysis and the application of data science to a business problem.

You should aim to spend **no more than a day** on this exercise.

You **must** use either Python or R to complete this assessment. You may also use SQL for performing queries and data modelling.

Problem Statement

A company is considering expanding their business into short term rental-cycles. Before investing, they would like to understand how people in London use cycling services in order to develop their strategy. The Data Science team has been asked to look at existing biking usage patterns to help them with this.

From an initial meeting, they suggested that they were interested in what kinds of usage they might expect to see, what profile of customers they might have, as well as operational concerns such as reliability and supply chain management.

The Ask

After the initial meeting, we'd like to get back with an initial response as well as some suggestions for how they can use data to make their expansion a success. Using the TFL cycling dataset (https://cycling.data.tfl.gov.uk/) 'usage-stats'[1] data from 2019 to 2021:

1. Perform exploratory analysis on the dataset
2. Identify possible data science use cases with the data
3. Select a use case and prototype a model to demonstrate the application of data science to solving their problem
4. Make two recommendations for the company

Submission

You must submit the following via email, in a zipped archive:

1. Slide presentation outlining your findings and recommendations (no more than 5 slides)
2. The code you used to perform the analysis

---

[1] For simplicity, assume the data starts with the file:
`143JourneyDataExtract02Jan2019-08Jan2019.csv`