Cycling Analysis

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

Aim to spend around half a day to a day on this exercise.

You **must** use Python to complete this assessment.
You may also use SQL to aid queries and data modelling.
You **must not** use coding assistants (such as Codex) to generate code for your submission.

Problem Statement

One of our clients 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. We have 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 client meeting, we'd like to get back to them 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][2] data from 2021 to 2023:

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 simple model that demonstrates the application of data science to this business
4. Make two recommendations for the client

Submission

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

1. A 3-4 slide presentation outlining key findings and recommendations for the client
2. The code you used to perform the analysis

Do NOT submit any virtual environment modules or other unnecessary artefacts.

---

[1] For simplicity, assume the data starts with the following file: 246JourneyDataExtract30Dec2020-05Jan2021.csv
[2] You may also find it helpful to download the page directly if there are challenges scraping