1. Supervised and unsupervised learning

Both unsupervised and supervised learning are good ways to tackle the given problem.

The main difference between the two is the use of labels. Supervised learning needs to know what each observation represents, what the label for a piece of data is. For example, a 50-year-old rented a standard bike in London could be the observation, and the label could be that they were riding it for 35 minutes.

With enough data and using supervised learning, we would be able to predict that for the next 37-year-old, who rents an electric bike in New York, the time they will ride the bike is 57 minutes.

Given the business problem, supervised learning would be useful to predict customer satisfaction, the cost of the bike (and therefore revenue) or the time and distance the bikes will do – which could be useful for fleet management.

Its main shortcoming is that it cannot predict anything outside the target variables for which it was trained.

Unsupervised learning does not need labelling. With unsupervised learning, the model finds hidden patterns in the data without any specific outcome (or label). In this business problem, unsupervised learning could be useful to see different customers’ segments and how to appeal to them best.

The main issue with unsupervised learning is the lack of a clear target. The patterns it discovers may not be useful for decision making. Also, evaluating the quality is not possible since there is no labeled data to compare these patterns with.

The best option would be to use both supervised and unsupervised learning. However, due to time constraints, only supervised learning will be used. The reason is that supervised learning provides predictive capabilities taht can help with forecasting and can be easily interpreted and evaluated.

1. Supervised and unsupervised learning