A Data Analytics Approach: Leveraging Real-Time Data for Optimizing Bike-Sharing Operations

Yusuf Shehadeh, 7395116
Mohammed Sabaawi, 7381108
Kemal Yazgan, 7396285,
Assia Benmimouni, 7381347
Samuel Rizzo, 7395422
Ivan Kamal Mehieddin, 7396805

May 20, 2023

Contents

Chapter 1

Excecutive Summary

1.1 Business Problem

....

1.2 Data

.....

1.3 The Analytics Solution

....

1.4 Implications

...

1.5 Recommendations

...

Chapter 2

Detailed Report

2.1 Problem Description

Traditional urban mobility heavily relies on internal combustion engine vehicles, which contribute to greenhouse gas emissions, pollution, safety concerns, and inefficiency. To address these issues and meet decarbonization targets, there is a need for a comprehensive transformation of the mobility system. A key trend in this transformation is the rise of shared, fleet-based transportation companies, including bike-sharing platforms.
In this project, we explore how bike-sharing fleet operators can leverage real-time data to monitor and optimize their operations, enhance profitability, and improve service levels. By harnessing data science, we aim to enable fleet operators to achieve success in their operations while benefiting society as a whole.

2.2 Business Goal

The business goal is to assist bike-sharing fleet operators in optimizing their operations, increasing profitability, and improving the service level. This can be achieved through a better understanding of the network of docking stations and an accurate prediction of bike idle time.

2.3 Data Science Goal

The objective of data science is to leverage the available datasets and real-time data to gain insights into the bike-sharing operations. This includes understanding the network of docking stations, analyzing system performance, and developing predictive models to forecast bike idle time. The aim is to provide fleet operators with actionable information to optimize their operations and make informed decisions.

2.4 Data Description

......
We chose to use the idle time at a particular station, because we believe it produces the most advantages.
One of the benefits is that it allows us to identify stations that are in particularly high demand, and also to accurately predict demand at specific stations. This allows us to determine the optimal distribution of bicycles per station, and to reduce the environmental impact by providing only as many bicycles as are actually needed. It also provides the opportunity to improve the infrastructure at a given station to optimize the user experience, for example by expanding bicycle parking or adding charging stations. In addition, monitoring the utilization of a station enables timely replenishment of bikes to avoid bottlenecks and maintain a high level of service. Most importantly, it supports the scheduling of maintenance and repair work at specific stations to ensure the availability and reliability of the bikes. The reasons why we have not opted for the idle time of individual bikes are as follows. First, this method may introduce distortions if certain bikes are inactive for long periods of time due to maintenance or repair work, unnaturally extending the idle time. This could lead to difficulties in distinguishing between actual idle time and technical problems. In addition, privacy concerns could arise, as detailed tracking of individual bikes could reveal users' personal information. Another issue is the neglect of station-specific analysis and utilization, which could lead to inefficient resource allocation.

2.5 Brief Data Preparation Details

......

2.6 Data Analytics

....

2.7 Conclusions

....