PROJECT PROPOSAL

PROPOSED TO:
NYC MTA

PROPOSED BY:
Sultanah Aldossari

Project Proposal

MTA NYC SUBWAY

THE COMPANY

Central Bikes is the largest local bike rental business in New York City and is the official bike rental company of NYC Parks. It serves New York City residents and its visitors through daily bicycle rentals and guided tours of Central Park, the Brooklyn Bridge and more.

PROBLEM STATEMENT

Our planet is undergoing climate change, and transport is one of NYC biggest contributors to carbon emissions. Due to this Central Bikes has contributed to launch “How green is Cycling” campaign for that reason Central Bikes wants to know the top five stations that are crowded in aim to distribute its bike to help in reduce the carbon emissions and to increase customer experience. Thus, it will help in raising the awareness of the impact of cycling on a sustainable future.

QUESTIONS/NEED:

- What are the top five crowded subways in NYC?
- What are the neighbourhoods around the stations that might use the bikes?
- What are the right stations that meets Central Bikes needs?
- will the location of bikes booth at the stations help in increasing the usage of bikes?

DESCRIPTION OF AVAILABLE DATA:

The turnstile usage data of NYC subway provided by the NYC Metropolitan Transportation Authority (MTA). The dataset that will be used is from June 2021 up to August 2021 and it consists of 11 features. Initially, the MTA Dataset will be used and further dataset might be used. The following table shows the fields that will be used during the analysis and its description.

Field Name	Description
STATION	Represents the station name the device is located at
LINENAME	Represents all train lines that can be boarded at this station
DATE	Represents the date (MM-DD-YY)
TIME	Represents the time (hh:mm:ss) for a scheduled audit event
ENTRIES	The comulative entry register value for a device
EXITS	The cumulative exit register value for a device

TOOLS:

- Python will used for coding and data analysis.
- Jupyter Notebook will be used to create and document live code and visualizations.
- SQLite Browser will be used to create and edit database files.
- Different libraries will be used such as:
  -Pandas will be used to create dataframes and for manipulating numerical tables.
  -NumPy.
  -Seaborn and Matplotlib both will be used for visualizations.