Data I/O 2024

Challenge Thesis:

Our client, Cycle, a bike-sharing company situated in Chicago, aims to enhance its grasp of urban mobility by delving into a comprehensive analysis of its bike trip data collected over the past year. By harnessing advanced techniques such as data cleaning, exploratory analysis, and visualization, we seek to unearth actionable insights that will inform strategic decision-making and spur innovation within the company.

Guiding Questions (these are general suggestions not requirements!)

- What are the peak hours for bike usage?
- Which stations experience the highest volume of traffic?
- What are the common destinations of these bike trips?
- How do usage patterns differ between members and casual riders?
- What is the average duration of a bike trip?

Data Set:

🗀 **Data/IO 2024 Dataset**

Data Dictionary:

- ride_id (VARCHAR): Unique identifier for a bike trip.
- rideable_type (VARCHAR): Type of bicycle used for the trip (e.g., electric_bike, classic_bike, docked_bike).
- started_at (DATETIME): Date and time when the trip started.
- ended_at (DATETIME): Date and time when the trip ended.
- start_station_name (VARCHAR): Name of the station where the trip began.
- start_station_id (VARCHAR): Unique identifier for the start station.
- end_station_name (VARCHAR): Name of the station where the trip ended.
- end_station_id (VARCHAR): Unique identifier for the end station.
- start_lat (FLOAT): Latitude coordinate of the start station.
- start_lng (FLOAT): Longitude coordinate of the start station.
- end_lat (FLOAT): Latitude coordinate of the end station.
- end_lng (FLOAT): Longitude coordinate of the end station.
- member_casual (VARCHAR): Indicates if the rider is a member or a casual user