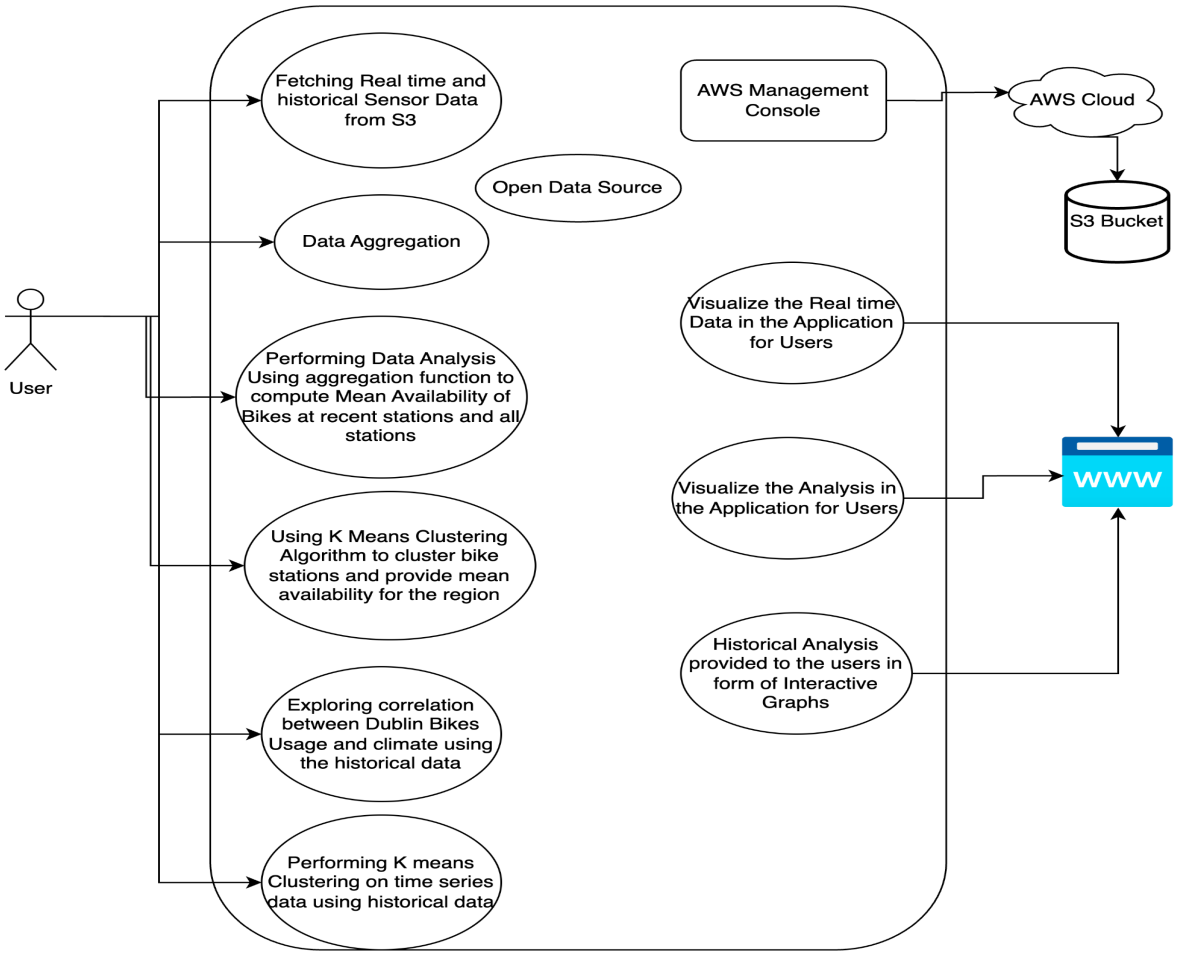
Urban Computing Week 4 Assignment
Urban Computing Application

By: Sanat Paranjape, **Student Id**:23347997
Course Code:CS7NS4 , **Stream**: MSc CS Intelligent Systems

Task 1 - Data Fusion and Analysis

Leading up to this assignment I have successfully stored Dublin Bikes Real Time Data and Historical Data on Amazon Web Services. I was able to access this data and display it on the terminal. Now it's time to perform data fusion and analysis on the DublinBikes data and get insights. The process involves extracting real-time data on bike availability from several stations in order to provide users and urban planners with useful information. The goal of the analysis is to comprehend the general pattern of bike availability over time at different stations.



Steps to perform Analysis on the data :

- Data Retrieval : Obtaining and Preparing data to perform analysis using different machine learning algorithms.
- Data Aggregation : To analyse data using different algorithms it is necessary to aggregate the data to get specific essential values from the data. The dublin bikes data is a huge dataset and contains a lot of columns which provide valuable information. For example, I was planning to get the average availability of Bikes at stations in real time. So I needed to aggregate the data to get the station id,name and availability.
- Data Analysis : I have used the K Means Clustering algorithm on relevant features like location, availability over time, etc. I have divided the regions into 5 different clusters to get the mean availability of bikes across stations within each cluster. For Historical Data, I'm analysing usage patterns throughout the day for all the stations.
- Data Visualization : All this analysis can be visualised by the user in the form of maps and graphs which provides valuable information.

Task 2 - Data Visualization and Actuation

I have utilized Python and Flask Framework to create an interface for Dublin Bikes dynamically showcasing both station locations and bike availability. Additionally, an analysis section was integrated to extract meaningful insights from the showcased data.

Key Points of Implementation:

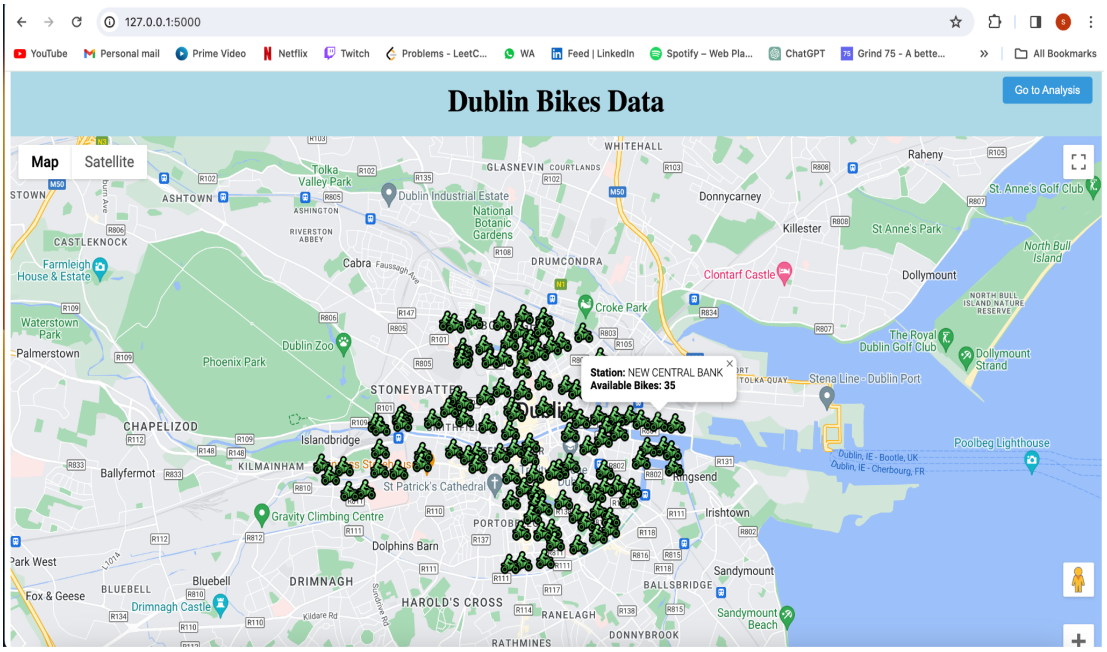
- Employed Python and the Flask framework to craft a web application.
- Integrated live data retrieval from the Dublin Bikes API to capture station locations and bike availability in real-time.

  Real-time Data Presentation:
  - Implemented functionalities to fetch and promptly update station information.
  - Displayed station markers pinpointing their locations while visually representing available bikes in a live manner.

  Analysis Segment:
  - Introduced a specialised section within the web application devoted to data analysis.
  - Utilised Python libraries like Pandas and Matplotlib to process data and conduct statistical analyses.

  Insights Derived from Data:
  - Examined bike availability trends across stations at specific intervals, be it hourly or daily.
  - Generated visual representations in the form of graphs or charts to effectively illustrate trends in average availability or temporal patterns.

Visualisation :



Task 3 - Extra Task

1. Historical Data Analysis:
    - Acquired historical Dublin Bikes datasets covering a substantial time span.
    - Utilised Python's Pandas library to preprocess and analyse historical data.
    - Conducted trend analysis, exploring patterns in bike availability, station usage, and user behaviour over months or years.
    - Derived insights into long-term usage trends, identifying peak periods, seasonality, and overall growth patterns in bike utilisation.

2. Climate Analysis Integration:
    - Retrieved historical weather data, including temperature, precipitation, and other relevant weather metrics for Dublin.
    - Tried connecting the datasets of Dublin Bikes and Climate but wasn't successful.

I did not make use of AI tools in the preparation of this assignment.