Data Analyst / Business Analyst

OMKAR MANKAME

(437) 223-7706 | omkarm001@gmail.com
Linked In Profile / Kaggle Profile
GitHub Projects Repository
Mississauga, L5M 0Y4

Certifications

MICROSOFT AZURE DATABRICKS AND APACHE SPARK – Microsoft — December 2024
BUSINESS INTELLIGENCE WITH DATABRICKS – Edureka — November 2024
DATA ANALYSIS WITH PYTHON – IBM, **Credly Link** — October 2024
GOOGLE DATA ANALYTICS CERTIFICATION – Coursera, **Credly Link** — August 2024
AWS CERTIFIED CLOUD PRACTITIONER - Amazon Web Services(AWS), **Credly Link** — June 2023
FOUNDATIONS OF DATA GOVERNANCE- Collibra — September 2023

Portfolio

PYTHON – Data wrangling and EDA

- Used available datasets of New York Taxi and did exploratory data analysis to answer various dataset related questions.
- Used Pandas and NumPy for EDA. **GitHub Link**

R PROGRAMMING – Dataset Correlation and visualizations using aesthetics, labels and annotations

- Used available datasets of 3 species of penguins. Installed required packages. **Code Link RMDfile**
- Using visualizations and scatterplots, identified relation between body mass and flipper length of penguins. **GitHub Link**
- Created visualizations to demonstrate that there is a strong correlation between the two data attributes. **Kaggle Link**
- Used Aesthetics, ggplot2 and tidyverse packages in R to create better visualizations. Saved project using **RMarkdown**.

CASE STUDY

- Cleansed and aggregated cyclist data (over million values) to create **visualization in Tableau** to understand the usage pattern.
- Demonstrated the story that the data had to tell by achieving specific results to support business objective. For Spreadsheet and Tableau - **GitHub Link**, For SQL and Tableau – **GitHub Link** Found different bike usage based on membership type.

MY SQL – Creating complex result table using Joins, Group By, Aggregate Functions

- Installed MySQL Workbench, worked on its set up.
- Created databases, tables and worked on result table using primary and foreign keys, syntax, aggregate functions.
- Different complex queries were run for different projects. Please review the queries using GitHub Codespace.
- For SQL projects – **GitHub Link**. For SQL and Tableau – **GitHub Link**

TABLEAU – Visualization to demonstrate relations between dog breed and related parameters

- Imported dataset from Google Studio. Data cleansing and data validation using SQL. Data reporting and data profiling in Tableau.
- Created appealing visualization with Dog breed images and data values demonstrating data mapping and dataframe.

GOOGLE STUDIO – Found most utilization of bike rentals on different routes using public dataset

- Google Studio has built in SQL tool – BigQuery. Used the same for data integrity.
- Wide variety of datasets are found on BigQuey data warehouse in Public Datasets, Marketplace as well as on Kaggle.
- The dataset I worked on was found in Public Datasets. It was cleaned and manipulated using BigQuery. (SQL)

BROWSER BASED GENERATIVE AI TOOL FOR DATA VISUALIZATION

- Prepared a clean dataset of different memberships for bike rentals.
- Used generative AI (Gemini) to provide recommendations for better visualization to present to specific teams.
- Based on AI recommendations used Bar Graph-Pie chart visualization for one team and Scattered plot-Heatmap for the other.

EXCEL – Worked on data and made it ready for analysis

- Large Dataset of various engine components of vehicle was available on Kaggle (online data science community).
- Data cleansing using conditional formatting, filtering, various functions – Concatenate, Countif, Dateif, etc.
- Prepared data for analysis. For Spreadsheet and Tableau project - **GitHub Link**. Found direct relation between price and mileage.