Data Analyst/Data Scientist

VALERIYA YUGAY

(+82 10) 6492 6301 | valerya13yu@gmail.com
LinkedIn
Portfolio Website
Busan, South Korea

Summary

Highly organized engineering student with a solid practical background in data management. Proficient in data cleaning, exploratory data analysis and data visualization. Skilled in using various tools and techniques to extract valuable insights from complex datasets. Seeking to leverage data analytical skills to improve corporate performance as a data analyst.

Skills

Technical Skills:

- SQL (SQL Server)
- Python (Pandas, NumPy, Sklearn, MatPlotLib)
- Data Visualization (Tableau, Microsoft Power BI)
- Excel (VLookup, Conditional Formatting, Pivot Tables)

Data Analysis Techniques:

- Web Scraping
- K-means Clustering
- Sentiment Analysis
- Logistic Regression
- Tools (Kaggle, Jupyter Notebook)

Projects

SENTIMENT ANALYSIS ON MY FAVORITE BOOK REVIEWS – Python, Github

- Gathered data using *selenium* and *beautifulSoup* libraries, obtaining over 100 reviews from 50 web pages
- Cleaned and organized the collected reviews with *pandas*, ensuring only relevant data was used
- Conducted sentiment analysis on the reviews, categorizing them into positive, negative, or neutral sentiments
- Aggregated and visualized the data by using **matplotlib** and **wordcloud** to compile a report

MALL CUSTOMERS SEGMENTATION – Python, Kaggle

- Visualized kernel density estimates *(KDE)* to understand the distribution of age, annual income, and spending score among customers using *seaborn*
- Utilized the *Elbow Method* to determine the optimal number of clusters for K-means clustering
- Performed *K-means clustering* with *scikit-learn* to segment customers
- Plotted the cluster centroids to identify distinct groups of customers, enabling the mall to target specific customer segments with tailored marketing strategies

COVID DATA EXPLORATION – Excel, SQL

- Explored Covid-related data using SQL to gain insights into the pandemic's impact on different countries
- Utilized *CTE* (Common Table Expression) and *temporary tables* to efficiently perform calculations and store intermediate results for further analysis
- Created a *view* to store data for later visualizations, enabling easy access and reuse of the calculated metrics

NASHVILLE HOUSING PRICES DATA CLEANING – SQL, Tableau

- Utilized **SQL** to clean data by standardizing format, populating missing data and dealing with duplicated data
- Transformed and filtered data by using *aggregating* and *filtering* function

Work Experience

ESL CONVERSATIONAL TEACHER– English Tochka **–** Remote   February 2023 - August 2023

- Successfully communicated with over 40 diverse individuals from various cultural backgrounds and demonstrated adaptability and creativity in tailoring personalized lesson plans

Education

BACHELOR OF ENGINEERING – Pusan National University – Busan, South Korea   September 2021 - present
Major: Aerospace Engineering