Protsenko Maxim

github.com/Makxxs kaggle.com/makxxxs makcum2901@yandex.ru

EDUCATION

Financial University   2021 - 2025
Bachelor of Applied mathematics and Computer Science

PROFESSIONAL EXPERIENCE

PLC Territorial Development Fund   November 2023 - Now

- Unloading data from the database using SQL and processing it.
- Data cleaning and preprocessing.
- Writing python scripts for parsing and other tasks.
- Building reports on Power BI and presenting it.

SKILLS

Languages: Python, SQL
Tools: Git/GitHub, VS Code, Airflow, Tableau, Power BI, Excel Power Query, ClickHouse, DBeaver, PostgreSQL, Jupyter
Libraries: pandas, NumPy, scipy, Matplotlib, Plotly, Seaborn, Scikit-learn, sqlite3, Scrapy

PROJECTS

Raw Data Cleaning Fifa21 | *Scrapy, Numpy, Pandas*
- Collected a raw dataset of FIFA21 player statistics from sofifa.com.
- Tackled unclean data resulting from inconsistent HTML formatting.
- Showcased the ability to handle and clean messy real-world datasets

Fast Food A/B Testing | *Numpy, Pandas, Seaborn, Matplotlib, Scipy*   3
- Performed a comparison of the sales results of a new product obtained by applying three different marketing strategies.
- Identified the most successful marketing strategy based on data and their analysis.

SuperStore Data Analysis | *Numpy, Pandas, Matplotlib, Plotly*
- Conducted a primary analysis, identifying the main attributes that will influence the analysis and decision-making.

- Analyzed the relationship between sales, profit and various attributes, such as product category, customer segment, region and others..
- Suggested which product categories should be emphasized, which customer segments to attract and in which regions to strengthen efforts to increase profits.

Salary Predict | *Numpy, Pandas, Seaborn, Sklearn*
- Divided the data set into training and test sets using the train test split function. The training set was used to train the model, and the test set was used to evaluate its performance.
- Trained the model on a training dataset using machine learning methods. After evaluating the performance of the model on a test dataset, it was found that the model demonstrates high performance. This is reflected in good metrics, such as a low mean square error (MSE) or a high coefficient of determination (R-squared).

Product Metrics | *SQL*
- Connecting to a database to extract the necessary information.
- Create complex SQL queries for data extraction and aggregation.
- The use of SQL to calculate various metrics and statistical indicators.

Courses

Data Analyst   October 2022 – April 2023
karpov.courses

Data Analytics   December 2022
IBM

Basics of statistics   September 2023
Stepik