

Doniyor Khudoykulov

Data Analyst

Samarkand, Uzbekistan | +998 90 197 72 73 | doniyor@khudoykulov.uz
[linkedin.com/in/doniyor-khudoykulov](https://www.linkedin.com/in/doniyor-khudoykulov) | github.com/doniyor117/portfolio

SUMMARY

Data analyst with a background in business and management. I work end to end: raw source tables, cleaning and modelling, then the analysis and whatever comes out of it, a dashboard or a recommendation. SQL and Python for the work, Power BI for the reporting. Studying Data Science at School 21. Co-founder of AdvoAI, a legal assistant for Uzbek law, live as an MVP.

TECHNICAL SKILLS

SQL: PostgreSQL, MS SQL Server; window functions, CTEs and subqueries, data quality checks and validation

Python: data cleaning, validation and exploratory analysis with pandas and NumPy; charting with Matplotlib, Seaborn and Plotly

Analysis: cohort scoring, segmentation and tiering, percentile and distribution analysis, driver analysis for profit and delivery performance

Statistics: descriptive and inferential statistics, sampling, probability distributions, confidence intervals, hypothesis testing and A/B tests, linear and logistic regression, ANOVA, chi-square

Machine learning: regression, classification, tree ensembles, clustering, dimensionality reduction; scikit-learn, LightGBM, XGBoost, CatBoost, Optuna

Reporting: Power BI with Power Query and DAX, dashboard design, data storytelling

Tools: Git and GitHub, Jupyter, DuckDB, Linux (Fedora, Debian), Bash, PySpark fundamentals

PROJECTS

Olist Seller Quality, SQL

[olist-marketplace-analysis#02---seller-quality](#)

- Scored every seller in SQL on DuckDB from review scores, missed shipping deadlines and cancellations: the worst tier is 5.6% of sellers but only 5.7% of 1-star reviews, while the tier above it carries 33% of orders and 41% of 1-star reviews, so coaching beats removing.

Olist Delivery Promise, Python

[olist-marketplace-analysis#01---delivery-promise-optimization](#)

- Established the cost of lateness first: two thirds of orders arriving 5+ days late were rated 1 star, while review scores flattened past a 7 to 16 day buffer, so extra padding buys nothing.
- Predicted delivery time on ~96k orders with LightGBM quantile regression, targeting a high percentile instead of an average so the late rate is set deliberately: ~18 days at 3.44% late against Olist's ~21 days at 4.5%, and shorter than a seller-state percentile baseline.

DataCo Supply Chain Profitability

[dataco-supply-chain-analysis](#)

- Modelled supply chain orders in Power BI with Power Query and DAX to separate what drives profit from what drives revenue across markets, categories and shipping modes.
- Found faster shipping modes miss their promises more often, not less: First Class delivers in 2 days but runs late 95% of the time against Standard Class at 4 days and 38%, pointing at the promised window rather than carrier speed.

AdvoAI, co-founder

[advoai.uz](#)

- Co-founded and shipped a working MVP of a legal assistant for Uzbek law, built and iterated with AI coding agents; the first prototype placed 8th of 30+ teams in the entrepreneurship track of the National AI Hackathon.
- Designed the retrieval pipeline over Uzbek legal codes: chunking strategy, hybrid dense and sparse search, cross-encoder reranking.

BYD Company Limited, Financial Ratio Analysis

[byd-financial-analysis](#)

- Computed nine profitability, liquidity, efficiency and solvency ratios from BYD's FY2023-2024 statements and wrote the trend analysis linking each movement to a business driver; team assignment, responsible for the calculations and the write-up.

EDUCATION

Higher National Diploma, Data Science

08/2025 – 08/2027

School 21, Samarkand | github.com/doniyor117/machine_learning_school21

- Peer-to-peer, project-based intensive program. Completed the machine learning track through neural networks: supervised learning and validation, decision trees and ensembles, clustering, and dimensionality reduction, implemented in Python.

Bachelor's degree, Management

09/2024 – 05/2028

Silk Road International University of Tourism and Cultural Heritage, Samarkand

CERTIFICATIONS

- [Google Advanced Data Analytics Specialization](#) | [Data Analyst Associate, DataCamp](#) (DA101 and DA501P practical, SQL and PostgreSQL assessed)
- [Linear Algebra and Regression Fundamentals, University of Pittsburgh](#) | [Introduction to Relational Databases, IBM](#)

Languages: Uzbek (native), English (professional working), Russian (reading and listening), German (elementary)