Vasilii Molchanov

Analytics / Data Engineer | **English C1**

wlutanin@outlook.com | Telegram | **molchanov.fyi** | LinkedIn | GitHub

SUMMARY

Four years in data, the last three on incremental ELT on MPP and lakehouse platforms (Iceberg, CDC replicas): data quality checks, step-by-step reconciliations and alerts that stop a load when a check fails.

Domain so far: operational, credit and liquidity risk and provisions in banking.

EXPERIENCE

T-Bank (formerly Tinkoff), Analytics Engineer — Nov 2023 – present

Data warehouse team for integrated risk management

- Build and maintain daily incremental ELT in Greenplum 6 that feeds the bank's loss event database from more than ten source systems. Each load runs its data quality checks as one gate: a failed check stops the write and sends an alert.
- Reconciled the credit fraud pipeline end to end, from about 95,000 source contracts to 2,200 loaded with every drop explained, then backfilled the missed losses.
- Moved pipelines to new sources after upstream retired the anti-fraud tables and split the ledger into hot and archive tables, with outputs compared and records missed since 2023 restored.
- Linked customer compensations to the IT incidents behind them through a ledger posting key. The reload since 2024 added about 1,300 events and cancelled about 3,000 conflicting ones.
- Turned two manual reports into scheduled pipelines: the bank's monthly financial position assessment for the central bank, reconciled with past submissions, and a daily check of interbank lending limits.
- Rebuilt the monthly operational risk report for senior management and an analyst dashboard in Apache Superset 2.1, with custom ECharts visuals, ClickHouse datasets and LLM-drafted event summaries.
- Prepared data for about a dozen central bank and internal audit requests, and onboarded new team members.

Ozon Tech, HR Data Analyst Intern — Jun 2023 – Nov 2023

- Built Power BI dashboards on Jira and HR data to track support team SLAs.
- Designed data loads for a new corporate training system together with its product owner and developers.

HSE University, Research Data Analyst, Institute of Demography — Sep 2022 – Apr 2023

- Automated hypothesis tests on Swiss Household Panel data for a study of single-parent households.

PROJECTS live at molchanov.fyi

- **Workout Calendar**: a training log as a dbt project on DuckDB, with 45 tests on every build in GitHub Actions.
- **Brand Mentions in Russian Lyrics** (master's thesis): 2,967 chart songs from 1991 to 2025, brands extracted by LLMs.

SKILLS

Storage	Greenplum 6 (MPP), Trino and Iceberg (lakehouse), ClickHouse, Vertica, PostgreSQL
Pipelines	SQL, Python (pandas, SQLAlchemy), incremental loads from CDC replicas, star schemas, backfills, MPP query plans, Apache Airflow. dbt, DuckDB and GitHub Actions in personal projects
Data quality	Check gates before writes, reconciliations, row-count and sum alerts, data catalogue, code review
BI	Apache Superset 2.1, ECharts, Handlebars, Jinja, Tableau, Power BI
AI tools	Claude Code and Antigravity in everyday development, LLM APIs for text extraction

EDUCATION

HSE University, MA in Digital Humanities, text analysis and named entity recognition — 2023 – 2026

HSE University, BA in Social Sciences, statistics and data analysis in Python, R and SPSS, GPA 8.52 of 10 — 2019 – 2023