

# Roman Tkachenko

## Data Engineer

Kyiv, Ukraine | +380970757662 | romuchqq@gmail.com

[LinkedIn](#)

## Summary

---

Data Engineer skilled in Python, PySpark and SQL with hands-on experience building end-to-end data pipelines on GCP. Comfortable orchestrating ETL/ELT flows with Apache Airflow and Docker, applying data-quality checks, and collaborating in distributed teams.

## Experience

---

### Python Scraper Developer

OnlineMinds - Remote (Copenhagen, Denmark)

03/2023 - 07/2023

- Worked in a distributed team at a startup to build scalable data scraping pipelines.
- Developed web scraping solutions using Scrapy and Python.
- Automated data collection workflows and structured large volumes of data.
- Collaborated on GitHub, followed version control best practices.
- Contributed to team discussions and agile development cycles.

## Projects

---

### DE-homework-tasks

05/2024 - 06/2024

- Built and orchestrated end-to-end data pipelines as part of a comprehensive data engineering project.
- Developed ETL jobs to fetch, transform, and store sales data (JSON, Avro).
- Set up a PostgreSQL demo DB with Docker, wrote SQL queries.
- Installed and configured Apache Airflow, built DAGs to automate workflows.
- Used GCP (BigQuery, GCS) and PySpark (local) for dual pipeline execution.
- Configured PySpark in Jupyter Notebook for distributed data processing.
- Integrated third-party data (user\_profiles) for enrichment, applied data quality checks.
- Delivered final project aligned with industry-grade data governance and orchestration practices.

### Python-scraper-rental-property

2023

- Designed and implemented a scalable Python data scraping solution using GraphQL and multiprocessing.
- Parsed and structured real estate listings from GraphQL APIs.
- Used multiprocessing for parallel data collection and aggregation.
- Output aggregated dataset into a structured JSON file for analysis.

### Telegram-bot-purchase-discount

2023

- Automated buyer-seller interaction with two Telegram bots to generate and validate promo codes.
- Ensured seamless transaction flow using Telegram Bot API and data validation logic.

## Education

---

BSc – Kyiv National Economics University, 2019-2023

Data Engineering Bootcamp – Robot\_Dreams, 2024

## Certifications

---

[Data Engineering – Robot dreams \(2024-06\)](#)  
[Joining Data in SQL – DataCamp \(2024-03\)](#)  
[Intermediate SQL – DataCamp \(2024-03\)](#)  
[Introduction to SQL – DataCamp \(2024-02\)](#)  
[Foundations of PySpark – DataCamp \(2024-02\)](#)  
[Feature Engineering with PySpark – DataCamp \(2024-02\)](#)  
[Cleaning Data with PySpark – DataCamp \(2024-02\)](#)  
[Big Data Fundamentals with PySpark – DataCamp \(2024-02\)](#)  
[Understanding Data Engineering – DataCamp \(2024-02\)](#)  
[Python Generation \(advanced\) – Stepik](#)  
[Hadoop, Big Data – Stepik \(2023-07\)](#)  
[FastAPI Python – Stepik \(2024-10\)](#)  
[SQL \(Basic\) – HackerRank](#)

## Skills

---

**Core Data Engineering Skills:** Python, PySpark, SQL, NoSQL, Apache Airflow, ETL/ELT, Data Pipelines, Data Cleaning, Data Orchestration, FastAPI, Google BigQuery, Snowflake, Docker, GCP (Cloud Platform), Hadoop, Web Scraping, Multiprocessing.

**Tools & Technologies:** PostgreSQL, MySQL, MS SQL, Cassandra, HDFS, YARN, MapReduce, Apache Spark ML, Anaconda, GitHub, Selenium, Beautiful Soup, Scrapy, Postman, Linux, Bash, Telegram Bot API, Slack, ChatGPT.

**Programming & Other:** Object-Oriented Programming (OOP), Software Design, Automation, APIs.

## Languages

---

- Ukrainian – Native
- English – Intermediate