

Hiring software engineer as easy as calling a taxi.

# Anton

## Data Analyst / BI Engineer

### SUMMARY

Data Analyst / BI Engineer with an extensive background in Computer Science and Software Engineering, bringing over 5 years of hands-on experience in high-level programming languages like Python and SQL to the table. Expertise is anchored in building robust data engineering solutions utilizing technologies such as Apache Spark, Apache Airflow, Databricks, and a variety of cloud platforms including AWS and Azure. Proven track record in data migration, optimization, and visualization with tools like Power BI and Tableau, reinforced by a deep understanding of Data Science principles. Adept at both relational and non-relational databases, showcasing remarkable proficiency in PostgreSQL and experience in MongoDB, MSSQL and Cassandra. This individual has made major contributions to cross-domain projects including blockchain, crowd investing, and securities analysis, evidenced by measurable improvements in data processing efficiency and reliability.

### SKILLS

<b>Main Technical Skills</b>	Python (5 yr.), SQL (5 yr.), Apache HTTP Server (5 yr.), AWS Cloudformation (4 yr.), Databricks (3 yr.)
<b>Python Frameworks and Libraries</b>	Matplotlib (2 yr.), Seaborn (2 yr.)
<b>Data Analysis and Visualization Technologies</b>	Tableau (2 yr.)
<b>Databases &amp; Management Systems / ORM</b>	Cassandra (1 yr.), MongoDB (2 yr.)
<b>Azure Cloud Services</b>	Azure MSSQL (1 yr.)

### WORK EXPERIENCE

#### Data Analyst / BI Engineer, Web3 antivirus extension

**Duration:** 02.2022 - Present

**Summary:** Google Chrome extension focused on providing users with a secure experience in the blockchain world, including risk assessment and ETL pipeline creation.

**Responsibilities:** Migration from Apache stack to Azure Cloud, Automation & consolidation of pipelines, Workflow scheduling with Apache Airflow, Management of repository environment,



Data filtering, anomaly handling, SQL query writing, Data processing with Apache Spark, Unit test writing, Data quality integration, Creating data monitoring systems, Dashboard creation with Power BI & DAX, Database schema creation, Data pipeline reconstruction, Data quality strategy development, Documenting migration and transformation procedures, CI/CD participation, Data visualization with Matplotlib, Seaborn.

**Technologies:** Python, SQL, Apache Spark, Informatica, Databricks, Apache Airflow, Azure, PostgreSQL, MongoDB, Pandas, NumPy, Matplotlib, Seaborn, Power BI, Docker, Azure DevOps

### **Data Analyst / BI Engineer, Analysis platform for securities**

**Duration:** 01.2020 - 01.2022

**Summary:** Migration of business logic and financial data related to mortgage-backed securities to a data warehouse, involving data cleansing and transformation for improved performance and accuracy.

**Responsibilities:** Collaborating with CTO and business analysts, Managing RDS PostgreSQL instances, Documenting professional details, Migrating logic to AWS Redshift, Building DWH on Redshift, Data integration using Informatica services & AWS S3/Redshift, Data modeling with Informatica MDM, Monitoring Redshift health, Creating and optimizing Tableau dashboards, Developing data quality alert system, Data processing with PySpark.

**Technologies:** Python, SQL, Apache Spark, Informatica, AWS, PostgreSQL, Tableau, Docker, Bash scripting, GitHub

### **Data Analyst / BI Engineer, Crowdfunding Platform**

**Duration:** 09.2018 - 12.2019

**Summary:** Creation of a data-driven investment platform connecting entrepreneurs and investors, designed for high availability and security, including data processing pipelines for large volume transactions.

**Responsibilities:** Consolidating diverse data sources, Automating ETL job scheduling with Data Factory, Optimizing database queries, Developing and maintaining ETL/ELT pipelines, Data transformation in Azure Databricks, Creating performance-tuned ETL processes with Azure Data Factory, Developing reports using Power BI, Creating pipelines in Azure DevOps, Optimizing DAX queries, Establishing CI/CD processes.

**Technologies:** Python, SQL, Apache Spark, Databricks, Apache Airflow, Azure, MSSQL, Cassandra, Pandas, NumPy, Power BI, Docker, Bash scripting, Azure DevOps

## **EDUCATION**

- **Computer Science and Software Engineering**

