



# CEYHUN AKAR

## DATA ENGINEER | DATA ANALYST

Data Engineer with experience in building and optimizing data pipelines to support business insights and decision-making. Skilled in using data processing tools, cloud technologies, and distributed computing to create efficient data solutions. Focused on improving data workflows and managing projects with attention to detail and efficiency. Dedicated to learning and adapting to new technologies in data engineering.

### CONTACT

☎ +90 554 475 42 85  
📍 Istanbul - Turkey  
Military Service : Completed  
✉ [akarceyh@gmail.com](mailto:akarceyh@gmail.com)  
🌐 [github.com/akarce](https://github.com/akarce)  
🌐 [akarce.github.io](https://akarce.github.io)  
📧 [medium.com/@akarce](https://medium.com/@akarce)  
🌐 [linkedin.com/in/akarce](https://linkedin.com/in/akarce)  
📺 [youtube.com/@ceysdatahub](https://youtube.com/@ceysdatahub)

### KEY SKILLS

#### PROGRAMMING LANGUAGES

- Python
- Bash Scripting
- SQL

#### TOOLS & TECHNOLOGIES

- Docker
- Confluent
- AWS
- Spark
- Databricks
- NiFi
- Airflow
- Hive
- Kafka
- HDFS

#### DATABASES

- PostgreSQL
- Trino
- MySQL
- IBM DB2
- Cassandra
- CouchDB
- MongoDB
- Elasticsearch

#### BI & VISUALIZATION TOOLS

- Tableau
- IBM Cognos
- Power BI
- AWS Redshift
- Superset
- Kibana

#### LANGUAGES

- Turkish: Native
- English: B2 - 6.0 IELTS Score

### EDUCATION

**Istanbul Technical University**  
Bachelor of Science, Materials  
Science

### PROJECTS

#### Data Pipeline for Real-Time Analysis of Yelp Reviews Sep 2024 [Github Repo](#) / [Youtube Tutorial](#)

- Real-time data pipeline using Kafka, MongoDB, Elasticsearch
- Integrated sentiment analysis with HuggingFace's DistilBERT model
- Created real-time visualizations and dashboards using Kibana

#### Reddit Data Engineering Pipeline Aug 2024 [Github Repo](#) / [Medium Article](#)

- Implemented a data pipeline for extracting and processing data from Reddit.
- Stored raw data in MinIO, set up external tables using hive and managed metadata with Hive Metastore.
- Automated data transformation processes using NiFi, including schema extraction and data manipulation.
- Enabled querying through Trino, Created analysis dashboards using Tableau and Superset.

#### End to End Data Engineering OTP Project July 2024 [Github Repo](#) / [Medium Article](#)

- Engineered a One-Time-Password (OTP) Pipeline, orchestrated by Airflow.
- Utilized Kafka, Zookeeper for broker coordination, and Kafka UI for monitoring.
- Set up alert systems using Slack, Discord, and EmailOperator to push notifications and updates.
- Utilized Airflow XComs for data validation and task communication.

### CERTIFICATIONS & COURSES

#### Elasticsearch v8 - BTK Academy Sep 2024 • Elasticsearch, Kibana, Logstash [Credential](#)

#### NoSQL - Document Databases - BTK Academy Jul 2024 • MongoDB, Couchbase [Credential](#)

#### Apache Kafka - BTK Academy Jun 2024 • Kafka Streams, Kafka Connect, ksqlDB [Credential](#)

#### Data Engineering Nanodegree - Udacity Apr 2024 • AWS, AWS S3, PySpark, NoSQL, Data Lakes

#### Learn SQL Basics for Data Science - Coursera Feb 2024 • Exploratory Data Analysis, Pyspark, A/B Testing

#### IBM Data Engineering Specialization - Coursera Dec 2023 • Data Warehousing, Linux, NoSQL, Shell Scripting

#### Data Engineering Bootcamp - Istanbul Data Lab. Oct 2023 • Big Data, Hadoop, Cloudera, Presto(Trino), HBase