

# Zach Tekker

San Jose, CA | 408-647-5880 | [zachtekker@gmail.com](mailto:zachtekker@gmail.com) | [linkedin.com/in/zachtekker](https://www.linkedin.com/in/zachtekker)

Solutions-driven data engineer with 10+ years of experience in data science, business intelligence, and business analysis across various industries. Possess a strong track record in developing and implementing data pipelines to maximize product value, drive efficiency gains, and foster business growth. Highly skilled in translating data into actionable insights, leveraging business acumen and technical expertise to support product vision and strategy, aligning analytical solutions with wider business goals, and collaborating seamlessly with cross-functional teams to achieve organizational objectives.

## CORE COMPETENCIES

- |                                |                                |                                  |
|--------------------------------|--------------------------------|----------------------------------|
| • Data Modelling & Integration | • Data Engineering             | • Cross-functional Collaboration |
| • ETL, ELT, Data Pipelines     | • Business Performance Metrics | • Project Management             |
| • Business Intelligence        | • Customer/Product Analytics   | • ERWIN                          |
| • Data Analysis & Reporting    | • Stakeholder Engagement       | • Domain Driven Design           |

## TECHNICAL EXPERIENCE

**Cloud Services:** AWS, Azure, Databricks

**Data Management:** Python, PySpark, PostgreSQL, NoSQL, MongoDB, Airflow, Docker, dbt, Apache Kafka, Apache Cassandra, Apache Flink, Apache ZooKeeper, Debezium, Elasticsearch, Tableau, PowerBI, Kibana

**Development:** CI/CD, RestAPI, Agile

## PROFESSIONAL EXPERIENCE

### CloudPeak Analytics

San Jose, CA

#### Analytics Consultant

Jun 2022 – Present

- Executed **ETL automation through AWS Lambda and Databricks' REST API**. Developed a custom function triggered by AWS S3 events to initiate Databricks jobs dynamically. Enhanced data processing flexibility and efficiency by 30%, seamlessly integrating disparate data sources. (Whops Games, Istanbul, Turkey)
- Developed and deployed a real estate **data pipeline on AWS, utilizing EC2, S3, and Lambda, integrating Python with Apache Kafka and Cassandra**. Enhanced data processing with a custom web scraper and OpenAI's GPT, processing 100,000+ listings monthly, reducing retrieval times by 24%, and boosting data reliability for strategic decisions. (NDA Signed, Major Real Estate Firm, Dubai, UAE)
- **Orchestrated the deployment and management of a Spark cluster using Docker Compose on Azure for real-time processing** and visualization of Schengen (EU) visa application data spanning 2012 to 2023. Developed and executed Python scripts utilizing PySpark, Pandas, and Plotly for data ingestion, cleansing, and analytical operations, resulting in the successful transformation and visualization of complex datasets. (IDATA, Ankara, Turkey)
- Led the transition of infrastructure to AWS cloud services, achieving a 38% decrease in costs and a 25% boost in efficiency for a major SME bank. **Developed and implemented a serverless architecture utilizing Lambda and API Gateway for RESTful APIs, and used Step Functions** to coordinate various AWS services, enabling the execution of complex workflows. (NDA Signed, Major SME Bank, London, UK)
- **Developed and deployed a real-time e-commerce data streaming pipeline** using Apache Kafka, Flink, Elasticsearch, and PostgreSQL, enhancing processing efficiency by 30%. Utilized Docker and ZooKeeper for scalable infrastructure, enabling dynamic, data-driven dashboard visualizations with Kibana. (Happy & Authentic LLC, Sacramento, CA)

### Presidency of Turkey, Data Management Department

Ankara, Turkey

#### Project Manager

Mar 2020 – Jun 2022

- Spearheaded the design and implementation of a scalable data engineering pipeline, leveraging Apache Airflow, Kafka, Cassandra, and Docker within an AWS environment to automate the ingestion and processing of data from the Quandl API. **Developed DAGs to orchestrate workflows, configured Kafka for real-time data streaming, and utilized Cassandra for efficient data storage and retrieval**. Successfully containerized the entire architecture using Docker, enhancing system reliability and deployment efficiency, which led to a 32% improvement in data processing speed.
- **Launched the first unified SQL database** for infrastructure investment projects by architecting a comprehensive SQL-based data repository. Integrated disparate data sources to ensure a cohesive data model and streamlined access for analysis.
- Collected and processed export transit time data from 13 logistics companies, creating a predictive model for high-accuracy forecasting global logistics center arrival times. **Leveraged Tableau to implement advanced data visualization capabilities**.
- **Architected and deployed an automated ETL data pipeline**, streamlining the data modeling process for quarterly foreign direct investment forecasts, successfully reducing data processing time by 24% and enhancing decision-making.

## Healthcare Information Management Systems

Istanbul, Turkey

Senior Data Engineer

May 2016 – Feb 2020

- **Designed and implemented a data pipeline** to support anomaly detection in hospital construction bids, utilizing unsupervised learning to identify cost discrepancies per square meter across healthcare facilities. Integrated findings into a Tableau dashboard, enhancing decision-making and reducing potential overpayments by 38%.
- Directed the implementation of a secure Health Information System (HIS), **enforcing role-based access controls (RBAC) and end-to-end data encryption to ensure compliance with the GDPR**. Orchestrated regular security audits and refined data protection policies, enhancing patient data security and supporting timely access for critical healthcare delivery.
- Created a data engineering solution to analyze COVID-19 trends in Turkey versus European nations. Extracted data from the Turkish Ministry of Health and ECDC and integrated it using Azure Data Factory into Azure Data Lake Gen2. Designed mapping data flows, orchestrated pipelines with Azure Databricks activities, and loaded transformed data into a SQL Data Warehouse. **Implemented parameterization, scheduling, monitoring, and CI/CD for the pipelines. Empowered data analysts to generate Power BI** reports comparing Turkey to Europe, enabling insights into the pandemic's impact.
- Developed a real-time monitoring system for the government's Healthcare Investment Projects Database **using Debezium and Apache Kafka**, capturing and logging key events such as status updates, funding adjustments, and regulatory changes. Enhanced transparency and decision-making by providing an immutable audit trail for over 1,000 daily transactions.

## Markafoni (acquired by Naspers)

Istanbul, Turkey

Data Engineer

Jun 2012 – Apr 2016

- **Designed, built, and managed a petabyte-scale Redshift data warehouse** architecture, integrating AWS cloud solutions to support diverse reporting needs. Spearheaded the creation of scalable ETL back ends for Tableau dashboards, empowering 12+ functional units with actionable insights.
- **Engineered and implemented a Python API-based automation framework to update millions of product attributes** in the product catalog within seconds. Leveraged parallel processing and asynchronous operations, resulting in a 25% increase in information management efficiency and facilitating real-time catalog updates.

## ebay (gittigidiyor.com)

Istanbul, Turkey

Senior Data Analyst

Feb 2011 – May 2012

- Prevented monthly losses exceeding \$400K by deploying **tree-based machine learning models** to identify and address fraudulent buyer activities.
- Improved the buyer experience and recovered \$5 million in monthly GMV by updating and **refining legacy blocking rules**.
- Led the Buyer Risk team's efforts in **fraud prevention** for new and guest buyers, enhancing eBay's market reputation and maintaining customer trust.

Data Analyst

Jun 2010 – Jan 2011

- Led weekly reviews of buyer **risk metrics, delivering actionable insights** and consistently communicating key performance indicators (KPIs) to leadership.
- Developed **real-time monitoring and alert dashboards**, facilitating proactive management of site traffic and transaction declines.

## EDUCATION

University of the Pacific

San Francisco, CA

Master's Degree in Computer Science - Data Science

Aug 2022 – May 2024

Middle East Technical University (METU)

Ankara, Turkey

Bachelor of Science in Engineering

Sep 2006 – Jun 2010

## CERTIFICATIONS

AWS Certified Data Engineer - Associate

Apr 2024

AWS Certified Machine Learning - Specialty

Jul 2023