

Sagnik Adusumilli

Data Engineer with 3 years of experience in cloud-based ETL processes, PySpark, SQL and Data Modeling

#15-30 Chichester Pl
Toronto, M1T 3S5
(647) 978-2744
sagnik.adusumilli@gmail.com

EXPERIENCE

TD Bank, Toronto— *Lead Data Engineer*

September 2024 - PRESENT

2025 Q1 Legendary Award Winner for leading critical data ingestion projects, including:

- **Global Anti-Money Laundering (GAML) reports**
- **HR data migration from PeopleSoft to Workday** impacting 95,000+ employees

Led a team of three data engineers across departments to develop an **Azure Data Factory (ADF) and PySpark-based tool** for executing DML statements in PAT and PROD environments. This streamlined hotfixes and schema evolution deployments, saving **1,600+ development hours annually**.

Provide business insights on data formats, load types etc. Help resolve technical challenges.

TD Bank, Toronto— *Data Engineer*

July 2022 - September 2024

Developed an automation tool using **PySpark and Databricks workflows** to monitor ingestion jobs and trigger DEV unit testing. This minimized bugs and reduced development time by **15%**.

Ran ETL jobs using **PySpark and ADF pipelines**, handling data cleaning, transformations and automation.

Extracted and ingested data from various formats (delimited, JSON, API, databases) into the **Standardized Raw Zone (SRZ)** as Delta tables in Parquet format.

Optimized reporting process of application security breaches by migrating Excel logic to SQL using stored procedures. The project implemented **Star Schema** and **SCD Type 2**, minimizing broken reports.

Resolved several critical Ingestion and ETL production issues- some with record turnaround times.

Multiple time “**Above and beyond**” award recipient at TD (2024)

Aviva, Toronto— *Software Developer Intern*

May 2018 - August 2019

Worked as a full-stack developer on an online portal used by underwriters/brokers to create insurance quotes for small businesses. (Tech stack: NodeJS, Angular, MongoDB, Handlebars.js). Presented Data science and machine learning methods during showcase event

SKILLS

Azure Data Factory, Apache spark (PySpark and Spark SQL) on Databricks for Big Data ingestion and ETL

SQL MySQL DBMS, PostgreSQL DBMS and SQL Server Management Studio (SSMS)

Delta Tables

Python, Java, C, SQL

Git, Jira and Bitbucket

PROJECTS (More on [Github](#))

F1 ETL

End to End data engineering project that implements medallion architecture and orchestration to ingest and present F1 data to a dashboard

Bike Toronto ETL

ETL and data analysis project done using PostGreSQL to draw business insights for cyclists in Toronto

Covid Data analysis

A study evaluating the effectiveness of the COVID-19 vaccine across various strains and age groups

EDUCATION

University of Toronto, St George — *B.Sc*

November 2022

Math, Statistics and Computer Science

AWARDS/Certifications

2025 Q1 **Legendary Award Winner** at TD

[Databricks Lakehouse Fundamentals](#)

[Microsoft Certified: Azure Fundamentals](#)

Multiple time “**Above and beyond**” award recipient at TD (2024)