Sagnik Adusumilli

| Sagnik Adusumilli  Big Data engineer with over two years of experience in ETL and Ingestion on Azure with strong skills in PySpark, SQL and data analysis. **CHANGE IT** | #15-30 Chichester Pl  Toronto, M1T 3S5  **(647) 978-2744**  **sagnik.adusumilli@gmail.com** |
| --- | --- |
| EXPERIENCETD Bank, Toronto— *Data Engineer*July 2022 - PRESENT Led a team of offshore/onshore engineers for developing a tool using PySpark and Azure Data Factory (ADF) to automate rollback procedure for production deployment failures. Rollback process involved removing data on cloud, reverting delta table schema and reverting SSMS table changes.  Designed and developed an end-to-end automation tool using PySpark and Databricks workflows to monitor ingestion jobs and trigger DEV unit testing. This saved a significant amount of manual effort and removed human errors.  ETL - Data cleaning, transformations and automations done using PySpark and ADF pipelines  Provided business insights on data such as data formats, load types (incremental, transactional etc.) in, data security and investigated data processing issues.  ETL - Extracted and ingested data in formats such as delimited, Json, API, database, etc., to Standardized Raw Zone (SRZ) as delta tables stored in Parquet file format.   * ETL (On prem) - Developed stored procedures in SQL to extract and load data in star schema format and implemented Slow Changing Dimensions (SCD) type 2. * Used ADF pipelines to run various jobs such as ingestion tasks, set up schedulers, modify SSMS tables using JDBC drivers etc. * Resolved 20+ (highest in the team) critical production issues - some with record turnaround times. Helped team members in resolving their issues.  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   **AWARDS/Certifications**   * [Databricks Lakehouse Fundamentals](https://credentials.databricks.com/b8572f0a-61cd-4ff0-8702-19cd2aad7280) * [Microsoft Certified: Azure Fundamentals](https://www.credly.com/badges/4546297e-cb0f-4873-b592-74bd7408c4f2/linked_in_profile)   Multiple time “**Above and beyond**” award recipient at TD (2024)   * Aviva Hackathon Runner up (2019) | SKILLS  * Apache spark (PySpark and spark sql) * Cloud Services such as: Blob storage, service principals, resource groups etc. * Databricks: notebooks, clusters, workflows, unity catalog * Azure Data Factory * Database Management Systems such as: MySQL, PostgreSQL and SQL Server Management Studio (SSMS) * Delta Tables, Relational tables, semistructured and unstructured data storage formats * Python, Java, C, SQL * Git, Jira and Bitbucket  PROJECTS (More on [Github](https://github.com/SagnikAdusumilli))[Formula 1](https://github.com/SagnikAdusumilli/formula1)End to End data engineering project that implements medallion architecture and orchestration to ingest and present F1 data to a dashboard using Power BI[Bike Toronto ETL](https://github.com/SagnikAdusumilli/Toronto_cycling_analysis)End to End data engineering project that implemented star schema using R and PostgreSQL[Covid Data analysis](https://github.com/SagnikAdusumilli/COVID-Vaccine-Analysis)Data Extraction, Cleaning and analysis using REDUCATIONUniversity of Toronto, St George —*B.Sc*November 2022 Math, Statistics and Computer Science |