

ANURAG SHEKHAR

+1(438) 464-1094 ◇ Montreal, QC, CA ◇ shekhar.anurag4@outlook.com ◇ [LinkedIn](#)

SUMMARY

Microsoft Certified Data Engineer and Data Analyst. Versatile and accomplished Data Engineer with a Master's degree in Computer Science and 4 years of experience in IT, specializing in Data Engineering and Analytics.

EXPERIENCE

AI trainer and Evaluator

Jan 2024 - Current

Freelance

Montreal, Canada

- Evaluated the quality of AI-generated Python & SQL code for Data Analysis and Visualization for correctness.
- Provided expert analysis and feedback to AI models to improve logical chain of thought and accuracy.

Data Engineer

Sep 2020 - Dec 2021

Tata Consultancy Services

Delhi, India

- Developed & optimized ETL pipelines using Azure Data Factory, boosting data processing efficiency by 20%.
- Designed & optimized complex SQL queries to retrieve and manipulate large datasets, improving data accessibility and performance by 40%.
- Automated data validation and cleansing processes using Azure Databricks, reducing data errors by 25% and enhancing the accuracy of business intelligence reports.
- Developed 10+ Microsoft Power BI Reports, Dashboards, and DAX Queries, providing valuable business insights.

ETL Developer

Mar 2018 - Aug 2020

Tata Consultancy Services

Bangalore, India

- Managed & supported 200+ ETL pipelines for 12 projects using SAP BODS and Azure Data Factory.
- Developed and maintained data models to support data warehouse and reporting requirements, resulting in a 20% improvement in data accessibility and usability.
- Conducted data profiling and validation in the data warehouse to identify anomalies and ensure data accuracy, resulting in a 15% decrease in data errors.
- Led a support team of 8 individuals, overseeing 24x7 ETL operations and ensuring SLA target achievement.

SKILLS

Technical Skills	Python (Pandas, NumPy, PySpark), Java, ETL, SQL, Data Warehouse, Spark, Spark SQL, SAP Data Services, Power BI, SAP BW, Github, Excel, DataLake, Data Analytics, GitHub
Cloud Platform	Azure (DataFactory, Databricks, Datalake, Synapse Analytics, Monitoring, Key Vault), AWS(EC2, S3, Lambda, Glue)
Databases	MSSQL, Sybase, Azure SQL, SQL Server, Oracle, Redshift, Teradata
Certifications	Microsoft Azure Data Engineer DP-203, Data Analyst DA-100, Data Fundamentals DP-900

PROJECTS

F1 Races - PySpark, SQL, Azure Data Factory, Data Lake, Databricks, Synapse, Power BI, REST API

- End-to-end implementation of data engineering ETL pipelines including fetching real-time data from API using Databricks, cleaning and transforming the data using Databricks Pyspark notebooks and loading to the presentation layer for analysis using Synapse Analytics and Power BI.

FreeLanceIoT - Java, Play Framework, Akka Streams, HTML, CSS, Javascript, REST, JSON, MVC, JUnit, JaCoCo

- A reactive MVC web application built with Java Play Framework, integrating with the FreeLancer API through Web Sockets and Akka Streams. This enables users to browse and search for freelance job opportunities in real-time.

EDUCATION

Master of Computer Science, Concordia University	Jan 2022 - Aug 2023	GPA - 4.0
Bachelor of Computer Science, University of Delhi	Jun 2013 - May 2017	Avg - 83%