Mr. ADITYA A JOSHI

Data Engineer

Email Id: adityajoshidataengg@gmail.com

Mobile: +91-8624095420

CAREER OBJECTIVE

- Scaling new heights of success and leaving a mark of excellence which involves professional growth. Ability to produce outstanding results individually and as a part of a team.

  I have total 4+ years of experience in IT Industry. Handle Big Data using Google Cloud Platform, Apache Spark, SQL, Python, Hadoop. Good team player as well as able to prioritize own workload and work under own initiative.

OVERVIEW OF TECHNICAL SKILLS

- Having total 4+ years of experience in Data Engineering using **Google Cloud platform, Apache Spark, Python, and SQL.**
- Hands-on experience of Google Cloud Platform components and **Big Data technologies**.
- Having experience in various Google Cloud Products like **Cloud Composer, BigQuery, Cloud Storage, Cloud Compute Instance, Cloud Data Fusion.**
- Experience **Cloud computing – Development, Implementation.**
- Strong experience in **migrating on-premise** Data warehouse applications to Cloud.
- Hands on experience **on Apache Spark.**
- Strong experience in Spark code to **perform ETL**.
- Strong experience in SQL scripts as per the requirement.
- Worked on importing and exporting data from different databases like **Oracle, SQL Server and MySQL into GCP.**
- Capable of processing large datasets (both structured and semi structured) by writing SQL Queries as per business requirements and supporting system application architecture.
- Working experience on **GCP partitioning and clustering.**
- Experience with **RDD, Spark Data frames, Spark SQL.**
- Strong experience in advance **SQL like Window function.**
- Experience in writing **complex SQL queries using joins, functions, sub queries and correlated sub queries.**
- Involved in writing **data transformations, data cleansing using Apache Spark operations.**
- Experience in working on **GCP services - BigQuery, GCS, Airflow Composer, DataProc and.**
- Experienced on cloud Platform services **SaaS, IaaS and PaaS.**
- Event based ingestion of files in cloud's data warehouse solution, on receipt of files at cloud storage (big query in case using Google cloud platform).
- Good knowledge of Software Development Life Cycle and Extensive experience working in an Agile development environment.
- Knowledge on GCP Service Like **Dataflow**.

Knowledge on **Google Looker Studio**.

- Good Experience of core python concepts like **List Comprehensions, decorators, generators, Functions**.
- Coordinate and smooth communication with the team to meet the Business requirements.
- Expertise in implementing end-to-end Cloud computing projects using standard SDLC.
- Good documentation skills and excellent technical, Communication and Interpersonal Skills with strong Client Interfacing Skills.
- Familiarity with delivery methodologies like Agile.

TECHNICAL SKILLS

Big Data (GCP tools)/(GCP Services)	BigQuery, DataProc, Google Cloud Storage , Google Cloud SDK Shell ,Cloud data Fusion ,Compute Engine, SQL
Version Control Software's	Git
Scheduling Tool	Cloud Composer (Airflow)
Programming Languages	Python, PySpark
Tools and Software	PyCharm, Jupyter Notebook
OS	Windows, Linux

WORK EXPERIENCE:

- Working as a GCP Data Engineer at **ASTRA INFOTECH** , from MAY-2020 to till date.

PROJECT DETAILS #3:

Client   : Brussels Airlines.
Domain   : Airline.
Description : Brussels Airlines is a Belgian full-service carrier bases at Brussels Airport. The Airline provides service to destinations across Europe, North America and continue to offer both

Short haul and long-haul services with a diversified tariff structure for the different target group.

Roles and Responsibilities:

- Involved in discussion to automate Google cloud job development process and initializing framework designing.
- Created python code for various Google cloud API clients like BigQuery, DataProc and many more.
- Created Python utilities which will be used to connect source and target systems for the data processing we have used PySpark.
- Updated existing PySpark code and tried to make them more generic across the platform.
- There are various frameworks with their specific use and I am part of developing most of their code.
- Performing pre-processing and cleaning on data using Spark.
- Develop SQL Queries for fetching the data from database.
- Involved in GCP architecture implementation for a complete cloud based big data solution using BigQuery, GCS, DataProc and Airflow.
- Debugging and fixing the issues by sampling the data.
- Responsible for pull, push and committing the code using GITHUB.
- Sorting the extracted data in Oracle database for further querying and analysis.
- Responsible for debugging and troubleshooting during data processing.
- Involvement in analysis of the project requirement.
- Creating database schemas for newly generated data.
- View logging information and start new task or resume task that have failed.
- Attending daily stand-up meetings (Scrum), and Requirement review meetings to analyze requirements for each story card in a sprint.

PROJECT DETAILS #2:

Client   : Ryder supply chain solutions **Domain**:logistics.

Description : Supply Chain Solutions. Ryder is the only supply chain logistics provider to offer an end-to-end portfolio of solutions that includes warehousing & dist*ribution, transportation logistics, e-commerce fulfillment, last mile delivery, reverse logistics, and innovative technology.*

Roles and Responsibilities:

- Created Python utilities which will be used to connect source and target systems for the data processing we have used PySpark.
- Created python code for various Google cloud API clients like BigQuery, DataProc and many more.

- Develop SQL Queries for fetching the data from database.
- Involved in knowledge sharing activities with team.
- Requirements analysis and Change implementation.
- Performing pre-processing and cleaning on data using Spark.
- Involve in the Daily Meetings.

PROJECT DETAILS #1:

Client   : Woolworths   **Domain** : Retail
Description: Woolworths Group is one of Australia and New Zealand leading retail Groups, Supporting Well-Known brands such as Woolworths, Big W and Countdown. Our great team is focused on creating better experiences together, for Our customers, Our communities and for each other.

Roles and Responsibilities:
- Created python code for various Google cloud API clients like BigQuery, Configuration, DataProc and many more.
- Created Python utilities which will be used to connect source and target systems for the data process we have used PySpark.
- Updated existing PySpark code and tried to make them more generic across the platform.
- There are various frameworks with their specific use and I am part of developing most of their code.
- Performing pre-processing and cleaning on data using Spark.
- Develop SQL Queries for fetching the data from database. Involved
- in knowledge sharing activities with team.

ACADEMIC QUALIFICATION:
- **Completed B.E from RTMNU, Nagpur University.**