DHANUSH BABU RAYI
Cloud Data Engineer
dhanushbaburayi@gmail.com
+ 91 - 8106207652

Professional Summary:

- ✓ Having 3+ years of experience in IT as Data Engineer, Hands on experience in migrating on premise ELTs to Google Cloud Platform (GCP) using cloud native tools such as BIG query, DataFlow, DataProc, Google Cloud Storage, Composer, DataFusion.
- ✓ Good Knowledge in Hadoop and its components like **HDFS (Hive) and SPARK.**
- ✓ Exposure to open-source scheduling to **Airflow.**
- ✓ Good exposure to **ELT frameworks and fundamentals.**
- ✓ Good understanding/knowledge of **GCP Architecture and Hadoop Architecture** and its various components.
- ✓ Experience in using different file formats like **Text, Sequence, Parquet, JSON, AVRO, ORC and CSV.**
- ✓ Worked in **Scrum/Agile methodologies**
- ✓ Excellent communication skills.

Carrier profile:

- Currently working at **ClobData Techno Solutions PVT LTD** as **Software Engineer** from **JUNE 2019 to till date.**

Academic Qualifications :

- B. Tech Graduate – 2017 from Dhanekula Institute of Engineering & Technology, College in Gangur, India affiliated to JNTUK.

Technical Skills:

BIGDATA Technologies	Hadoop (HDFS, Hive), Spark Core,Kafka
Languages	Python Scripting, SQL
Cloud Platforms	GCS, Big Query, Dataflow, DataProc, Data Bricks, pub/sub, Airflow, Cloud Fusion
Build Tools	Data Bricks, Jupyter Notebook
Data Formats	CSV, JSON, XML, ORC, Avro, Parquet
Project Execution Tools	**Agile** – Sprint, Scrum, JIRA, Confluence, GitHub
Operating Systems	Windows, Linux

Professional Experience:

Project Profile:

Project #1:

Project: Sales & Repurchase Data Analytics
Duration: November 2021– Till now
Technologies: GCP BigQuery, GCS, DataFlow, Airflow
Role: GCP Data Engineer
Pod Size: 7

Project Description:
Sales & Repurchase - online shopping details of the customers who purchased products on the online site. By using airflow, I can load the data from landing zone to GCS bucket in GCP cloud. After that I can do some transformation in DataFlow and loaded into big query staging layer after that it will be loaded in history layer it is a permanent storage.

Roles & Responsibilities:

- Developed Airflow pipelines for new set of requirements.
- Developed SQL Scripts to load data into Big Query.
- Involved in development of creating data pipeline flow.
- Experience in handling various file formats like CSV, Json, Pipe Delimited files.
- Created stage and history tables as a part of my regular job and created authorized views to the downstream Teams.
- Hand-on experience in SDK creating buckets, loading data and creating Datasets.
- Hand-on experience on GS Util, G-Cloud and BigQuery Commands.

Project #2:

Project: SRP Data Migration description
Duration: September 2020 – August 2021
Technologies: GCS, Big Query, airflow.
Role: GCP Data Engineer
Pod Size: 9

Project Description:
The key benefits delivered as part of this Project will be to enable the business in taking informed strategic decision based on the customer, product and channel indicators. The key business questions that will be addressed as part of the program which will further drive this will give Identification of cost efficiency saves for high cost-to-serve segments (through product, channel, proposition usage) Decision-making tool for sizing / prioritizing segment growth opportunities Product Pricing, staff incentives, cross-selling, RM loading, customer servicing model.

Roles & Responsibilities:

- Responsible for analyzing the requirements and designing the requirements page.
- Developed DDL and DML scripts to Load data into BigQuery from GCS buckets.
- And creating the Datasets, and tables in different layers of BQ projects.
- Used cloud shell in GCP to configure the services Storage, BigQuery.
- Experience in handling CSV, Avro, Parquet file formats.
- Developed Audit tables for reconciliation and metadata tracking purposes.
- Created Authorized views in Big Queryto provide data to Down Streams ML/Reporting and Analytics
- Following Agile Methodology
- Used Source code Management SCM as GitHub and confluence.

Project #3:

Project: Voice Of the Customer
Duration: July 2019 – August 2020
Technologies: GCS, BigQuery, SQL, AirFlow
Role: GCP Data Engineer
Pod Size: 7

Project Description:

Voice of the Customer (VOC) is a global initiative which will bring data together and visualize Customer data fromComplaints, using this data we are analyzing together to find out the Customer IdentificationNumber (OrM). This data will allow the markets and valve streams to see a holistic view ofour customers Movement of customer survey data to cloud and enable the Oilskins dashboards on top of them to enable the customer leans of the markets to take corrective actions.

Roles & Responsibilities:

- Experience in writing spark jobs to process data using DataProc/Data Bricks cluster on Google Cloud Platform.
- Hands-on experience with Airflow to schedule an orchestrated data flow between various internal components in GCP.
- Developed Python script using Google Client libraries to interact with Pub/Sub, GCS and BigQuery.
- Developed SQL based scripts to Load data into BigQuery.
- Create Authorized Views in BigQuery to provide Requires Data to Down Streams
- (ML/Reporting/Stakeholders).

Declaration :

I hear by declare the information as correct and true to my knowledge.