

VENKATESH BANDAPALLE

GCP DATA ENGINEER

Email Id: venkatesh0746@gmail.com

Mobile No: +91 8919649433

Career Objective

To pursue a challenging career and be a part of progressive organization that gives me scope to enhance my knowledge, skill and reach the pinnacle in the field with determination, dedication and hard work.

Summary

- Having 3.2 years of experience in the IT industry as a **Software Engineer** using Google Cloud components.
- from Google cloud platform like **Big Query, GCS, and Composer**.
- Experienced in working with Big Query to load data to Stage, History and Control tables, Develop Authorized views for data analysis and data warehousing.
- Experienced in writing optimized SQL queries.
- Experienced in writing DAGs in AIRFLOW/COMPOSER to automate the Data pipeline.
- Having knowledge on writing Pyspark jobs in Dataproc cluster.
- Awareness in performance tuning, deployment procedures, troubleshooting and solving database issues.
- Used GIT to manage and deploy the code version.
- Having experience in the Software Development Life Cycle (SDLC) and Software Testing Life Cycle (STLC).
- Having knowledge on Linux and Windows environments.
- Good experience in all phases of Agile and participated in all daily scrum meetings with cross teams.

Academic Profile

- Post-Graduation MBA from Sri Ramakrishna (Autonomous) P.G College, Nandyal, Andhra Pradesh.
- Graduation B.Sc. (Computers) from Sri Ramakrishna (Autonomous), Nandyal, Andhra Pradesh.

Technical Skills

- | | |
|----------------------------------|---|
| ➤ Cloud Platforms – GCP | : GCS, Big Query and Airflow/Composer |
| ➤ Languages | : SQL, Python (Scripting) |
| ➤ Big data Technologies | : Hadoop, Pyspark |
| ➤ Build Tools | : Notebook++, Jupyter Notebook |
| ➤ Project Execution Tools | : Agile (Sprint, Scrum), Jira, Confluence, GitHub |
| ➤ Data Formats | : CSV, JSON, Avro, Parquet |
| ➤ Operating System | : Windows, Linux |

Work Experience

- Currently working with **Bluefield Technologies Private Limited**, Bangalore from June 2019 to till date.

Project 1:

PROJECT : CDS Sales and Repurchase Data Analysis
TECHNOLOGIES : Big Query, GCS, Airflow
ROLES : GCP Data Engineer
DURATION : (Feb 2021- Till date)

PROJECT DESCRIPTION: Blue Apron's mission is to make incredible home cooking accessible to everyone. Launched in 2012, Blue Apron is reimagining the way that food is produced, distributed, and consumed. The popularity of meal kit delivery services has surged in recent years as consumer attitudes toward home cooking and grocery shopping have shifted. As a pioneer in the category, Blue Apron helps its customers create incredible home cooking experiences by sending culinary-driven recipes with high-quality ingredients and systematic instructions straight to customers' doors.

ROLES AND RESPONSIBILITIES:

- Extraction, Load and transformation of daily transactional data of client's orders from GCS bucket to Big Query.
- Loading the data from GCS bucket, applying transformations, and cleaning it.
- Creation of Data sets, Tables as per the requirement.
- Writing SQL queries to generate the reports according to the business requirements.
- Developed Audit tables for reconciliation and metadata tracking purposes.
- Creation of DAG files for Airflow to orchestrate the pipelines from creation of the Dataproc cluster performing the jobs and deletion of the cluster every day after the job is done.
- Created Authorized views in Big Query to provide data to down streams (ML/Reporting/Analytics/Dashboards)
- Interacting with the client on daily scrum calls and resolving any blockages/issues to proceed further.

Project2:

Client : Unibail-Rodamco-WestFeild
Project : Clicks Data Migration Analytics
Role : GCP Data Engineer
Technologies : Big Query, GCS, Airflow/Composer
Duration : Oct 2019 – Sept 2020

Description:

Unibail-Rodamco-WestField is Europe's largest B2C Online and Stores Market Place, connecting buyers with suppliers. Improve the customers increasing the open rate of promotional emails by 70% doubling the number of clicks in Europe & US, connect with customers through online too, using several channels including its websites, mobile apps, and emails providing a best product to Customers.

Roles & Responsibilities:

- Gathering requirements from the business team, designing data models, developing design documents and implementing ELT pipeline.
- Responsible for creating the GCS buckets, Datasets, Big Query tables in different layers of BQ projects.
- Developed SQL scripts to Load data into Big Query.
- Involved in Development of creating Data Pipeline to read the Data from GCS and load data to Big Query.
- Used Source Code Management (SCM) as Github.
- Worked on Creating DAG's and tasks in the Airflow scheduling tool.
- Created Authorized views in Big Query to provide data to down streams (ML/Reporting/Analytics/Dashboards)
- Responsible for Ingesting the data from one dataset to different datasets through BQ Queries as a part of ELT and ETL process.
- Responsible for designing the test cases
- Responsible for Unit Testing and SIT (System integration Testing) and capturing the evidence.
- Responsible for SIT walk through to Business to get Sign-off
- Responsible for providing Data Quality Assurance.
- Executing Functional, Regression testing when requires
- Participation in review and client calls