

PROFILE SUMMARY

Google Cloud Platform Data Engineer with 2 years of experience working with GCP stacks including BigQuery, Composer, Cloud Storage, Cloud Dataflow, and migrating legacy systems into GCP technologies. Worked in teams of 10+ team members.

FUNCTIONAL AREAS

- Google Cloud Platform
  - Automation
- Azure
  - Project Management
- Python

TECHNICAL SKILLS

- Python:** NumPy, Pandas
- SQL**
- Cloud:** GCP, Cloud Dataflow, Big Query, Cloud Composer, Google Cloud Storage(GCS).
- Orchestration Tool:** Apache Airflow
- Project Execution Tools:** Agile – JIRA, Git, GitHub
- BI Tool:** Tableau

WORK EXPERIENCE

Torry Harris Business Solutions, Bengaluru, India

- Data Engineer [Google Cloud Platform]** 09/2022 – 04/2024  
**Client:** British Telecom Group  
**Responsibilities:**
    - Designed the Data model as per the client Requirement and prepare the schema in JSON format, add the relevant data for the testing and end process
    - Coordinated the automation of new files in Google Cloud Platform, standardizing column names and data types based on validation files; achieved 60% reduction in manual data entry time and minimized errors in data processing.
    - Implemented data profiling and cleaning techniques to identify and rectify discrepancies in datasets, improving data quality by 20%.
    - Directed the adoption of Terraform practices for IaC to automate BigQuery resource management, achieving a 40% reduction in manual tasks and ensuring efficient version control for enhanced maintainability and scalability.
    - Engineered and optimized scalable data pipelines in Airflow Composer to automate processing, ingestion, and delivery of batch and real-time streaming data through Dataflow, enhancing efficiency by 50%, while establishing reusable and efficient solutions for data management.
    - Streamlined data processing workflow in BigQuery by developing automated scripts to ensure accurate storage of processed data in designated tables, resulting in a 40% reduction in data discrepancies and improving data accessibility for analytics teams.
    - Identified root causes of data pipeline defects affecting data integrity; implemented solutions that decreased data flow interruptions by 40% and increased data accuracy by 25% within a fast-paced tech environment.
    - Spearheaded data validation efforts, partnering with stakeholders to establish and uphold rigorous data quality protocols; these initiatives led to 25% improvement in data accuracy and streamlined decision-making processes.
    - Orchestrated a GitLab to implement CI/CD workflows, which allows teams to automatically test, build and deploy changes; this initiative resulted in a 45% reduction in code conflicts and a 30% increase in developer productivity, saving an estimated 15 hours weekly.
    - Directed the implementation of Jira for project management, increasing team collaboration and efficiency; this led to a 25% reduction in project delays and a 20% improvement in task completion accuracy.

URJA Infra, Aurangabad, India

- Summer Intern** 03/2022 – 09/2022
    - Managed PM Kusum Scheme data collection (surveys, pre-/post-installation). Increased efficiency by 20%.
    - Analyzed AG Pump installation data to optimize workflow. Reduced installation time by 15%
    - Created data visualization dashboards for improved project communication (10% improvement).

CERTIFICATIONS

- Google Data Analytics Professional Certification
- Databricks Lakehouse Fundamentals Certification
- Google BigQuery & PostgreSQL: Big Query for Data Analysis. (Udemy)
- Python Programming through Nareshit Technologies.

EDUCATION

- Bachelor of Technology [B. Tech] [Electrical and Electronics Engineering]** 08/2018 – 08/2022
    - Marathwada Institute of Technology (MIT) [8.63/10 CGPA] [First Class with Distinction]