E PRAVEEN KUMAR
GCP DATA ENGINEERING

Mobilno-9381349593   Email : epraveenkumar0703@gmail.com

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 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, Google Cloud Storage, Composer.

- A qualified Bachelor of Science with a good amount of exposure in cloud data engineering utilizing technologies like GCP, Big Query, SQL and Python(Scripting).

- Experience working with Big Query to load data to Stage, History and Control tables, Developed Authorized views.

- Having knowledge on Linux and Windows environments.

- Good experience in all phases of Agile and participated in all daily scrum meetings with cross teams.

- Awareness in performance tuning, deployment procedures, troubleshooting and solving database issues. Experience on python language scripting.

- Used GIT to manage and deploy the source code version.

ACADEMIC PROFILE

- **B**.Tech from AMMARA INSTITUTE OF ENGINEERING AND TECHNOLOGY under Jawaharlal Nehru Technological University Kakinada.

TECHNICAL SKILLS

✓ Cloud Platforms – GCP	:	GCS, GSutil, BigQuery, Pub/Sub, Airflow
✓ Languages	:	SQL, Python (Scripting)
✓ Big data Technologies	:	Spark, Kafka
✓ Build Tools	:	Notebook++, Jupyter Notebook
✓ Project Execution Tools	:	Agile (Sprint, Scrum), Jira, Confluence, GitHub

✓ Operating System	:	Windows	
✓ Data Formats	:	CSV, JSON, XML, Avro	

WORK EXPERINCE

- Currently working as a Data Engineer on GCP with **BEYOND BLUE IT SOLUTIONS PVT LTD,** Bangalore From June 2019 to till date.

PROJECT-1

✓ Client	: Max Life Insurance
✓ Project	: Customer Data Migration
✓ Role	: GCP Data Engineer
✓ Technologies	: SQL, Big Query, Airflow/Composer
✓ Duration	: Feb 2021 – Mar 2022
✓ Pod Size	: 8

Description:

Max Life Insurance is on a mission to help secure the financial future of its customers by "making life insurance simpler, accessible, and affordable" through its offering of long-term savings, protection, and retirement solutions. The company is committed to improving Indian social and economic development through its corporate social responsibility initiatives in health. In 2019, it was named Disruptive Technology Innovator of the Year by Economic Times India.

ROLES AND RESPONSIBILITIES:

- Developed BQ DDL and DML scripts to Load data into BigQuery from GCS buckets.

- Analyze and organize the raw data as per the business requirements.

- Evaluation of business requirements and objectives.

- Creation of BigQuery Data sets, tables and pipelines to store the processed results.

- Used cloud shell in GCP to configure the services Storage, BigQuery.

- Experience in handling CSV, Avro, Parquet file formats.

- Preparation of data for prescriptive and predictive modeling.

- Developed Audit tables for reconciliation and metadata tracking purposes.

- Hands-on experience with Airflow to schedule an orchestrated work flow between various internal components in GCP.

- Create authorized views in BigQuery to provide required data to downstream applications.

- Using Source code management as a GitHub tool.

PROJECT-2

✓ Clint   : BliBli.com
✓ Project   : Costumer Data Analytics
✓ Role   : GCP Data Engineer
✓ Technologies   : SQL, Big Query, Airflow/Composer
✓ Duration   :Jan 2022 – Nov 2020
✓ Pod Size   : 8

Description:

Blibli.com aims to become a leader in Indonesia's ecommerce sector. Founded in 2010 and launched in 2011, Blibli.com plans to be a one-stop shop for consumers seeking products and services in mobile and tablet devices, sports and outdoor activities, home and living, men's and women's fashion The data platform captures clickstream data as well as data from its backend systems. The business uses the clickstream data to capture intelligence about customer behavior on its website, including trending products and services.

ROLES AND RESPONSIBILITIES:

- Gathering requirements from the business team, designing data models, developing design documents and implementing ETL pipeline.
- Responsible for creating the GCP 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 Kafka and load data to Big Query.
- Used Source Code Management (SCM) as GITHub.
- Worked on Creating DAG's and tasks to deploy and run PySpark jobs 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 data set 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 Generating the Test reports through Python Automation framework
- Responsible for providing Data Quality Assurance.
- Executing Functional, Regression testing when requires
- Participation in review and client calls