

Bhargav Teja J

571-355-3440 | jakku.bhargavteja@gmail.com | [Linkedln](#) | [Portfolio](#)

SUMMARY

Graduated in May 2023 in Data analytical Engineering with 2+ years of experience in building innovative data-intensive applications. Proficient in designing and optimizing scalable data pipelines to solve business problems.

EXPERIENCE

Data Engineer

May 2019 – Jun 2021

Clorder Technology Inc.

Hyderabad, IND

- Utilized Google Analytics to collect raw data and scheduled batch loading into Big Query by designing custom ETL pipelines with Google Cloud Composer (Airflow) to transformed JSON, XML files
- Collaborated cross-functionally with Product, and Marketing stakeholders, to develop 10 KPIs tailored to business objectives, integrated KPI queries into data transformation processes for better analytics
- Connected BigQuery to Tableau to build 24 dashboards and measured performance metrics on marketing campaigns thereby improving global revenue and customer satisfaction
- Streamlined data validation with SQL scripts, solved data inconsistencies, and saved over 10 hours/week
- Improved query efficiency by indices optimization, reduced costs spent by a quarter on allocated resources
- Streamlined CI/CD and version control for Python code using Git and managed workloads using CRON jobs
- Developed on-demand Ad-hoc reports for Leadership on key outcomes, supporting data-driven growth strategies.

PROJECTS

Covid-19 Data Integration | *PostgreSQL, Talend, Snowflake, Big Query, GCP*

Apr 2023

- Defined entities to create EER's and relational models to analyze EMR data (1M records) using PostgreSQL
- Performed ELT operations using Talend and modeled data using snowflake schema and OLAP data cubes
- Built robust pipeline by loading data to GCP buckets using shell, python and transformed data using Data Fusion
- Hosted data on CloudSQL, executed federated queries for rapidly changing data from Pub/Sub stream
- Ingested data to BigQuery using Data stream from local server to maintain data freshness, CDC and SCD's
- Utilized Looker and connected BigQuery to Tableau for visualization and managed access control using IAM

Graph Based Search Engine | *AWS, PySpark, Neo4j*

May 2023

- Engineered ETL pipeline using AWS Glue to extract raw faculty data from the university website to S3 buckets
- Transformed raw data with PySpark for optimal graph modeling and ingested into Neo4j Graph Database.
- Built data models on research activities using COSINE similarity and visualized in Bloom, which saved 50hrs/week
- Designed and developed Web Application to update user's information into the database using Restful APIs.

Market Research Analysis | *Python, Tableau*

Dec 2022

- Conducted data pre-processing in Python on Bike rental data, optimizing Time Series Analysis to predict sales
- Forecasted sales and recommended pricing strategies using 4 Tableau dashboards analyzing seasonality data

School Performance Analysis | *SQL, Python*

Dec 2021

- Collected and Performed EDA on School Card Reports using SQL and Python to improve Education quality.
- Leveraged JOIN, CTEs, and window functions to measure the school's performance for better decision-making
- Visualized the key findings in Python and came up with strategies that led to a 40 percent increase in performance

EDUCATION

George Mason University

Fairfax, VA

Master of Science in Data Analytics

Aug. 2021 – May 2023

Keshav Memorial Institute of Technology

India

Bachelor of Technology in Electronics and Communication Engineering

Aug. 2016 – April 2020

TECHNICAL SKILLS

Languages: Scala, Python, R, SQL

Databases: NoSQL, T-SQL, SQL Server, PostgreSQL, MongoDB

Cloud Services: AWS (EC2, S3), Azure, GCP, Neo4j

Data Warehouse: Snowflake, Big query, Databricks, Redshift

Developer Tools: Git, Docker, Tableau, Kubernetes, Linux, Hive, Kafka

Libraries: Pandas, NumPy, Scikit-learn, TensorFlow, Seaborn, Pyflux