

# Jeevan Reddy Gujjula

• San Jose, CA • 3148979461 • Jeevanreddygujjula7@gmail.com • [Linkedin](#)

*“Dynamic **Data Engineer** with 4+ years of experience designing and managing **data-intensive** applications. Expertise in **Big Data** technologies, **cloud platforms**, **data warehousing**, and **business intelligence**. Proven ability to deliver **actionable insights** and streamline processes, driving measurable **business outcomes**.”*

## WORK EXPERIENCE

---

### T-Mobile | Data Engineer

May 2024 - Current

- **Orchestrated** daily and non-daily data pipelines using **Azure Data Factory (ADF)**, integrating 10+ data streams into **Azure Data Lake Storage** via **Databricks (PySpark)**, improving data processing speed by **20%**.
- Deployed a **Medallion architecture** within a **Data Mesh framework**, creating scalable, governed data access for **enterprise-wide analytics**, reducing time-to-insight by **30%** across teams.
- **Streamlined** data quality by creating **validation rules** and **anomaly detection systems**, reducing reporting errors by **25%**, and bolstering vendor and supply chain performance.
- **Resolved** vendor charge discrepancies via **DBT** and **Snowflake**, renegotiating agreements that cut **operational costs** by **16%**, improving financial forecasting accuracy.

### Johnson & Johnson | Data Engineer II

Jan 2022 - Jan 2023

- Developed an **on-premises Data Warehouse** with **15+ pipelines** using **Python, HiveQL, Spark, and Presto**, improving query execution time by **40%** and reducing stock outs by **18%** through faster inventory insights.
- Migrated **100+ legacy Hive queries** to optimized **Spark transformations**, cutting latency by **40%** and ensuring real-time data access for **pharmaceutical tracking** and **compliance reporting**.
- Designed **Tableau dashboards** for marketing and sales KPIs, boosting campaign ROI by **15%** and optimizing resource allocation across **10+ Johnson & Johnson retail regions** with actionable insights.

### Flipkart | Data Engineer

Dec 2020 - Dec 2021

- Engineered **10+ scalable ETL pipelines**, integrating data sources into **Azure Data Lake, Azure SQL, and Synapse Analytics**, enabling the seamless processing of **5TB+ data daily** to support Flipkart's e-commerce ecosystem.
- Optimized **Customer and Product Analytics, Churn Prediction, and CLTV** models using advanced **SQL queries**, driving a **20% growth in revenue** by identifying high-value customers and optimizing marketing campaigns.
- Automated the **reporting pipeline** for Flipkart's operational metrics, reducing manual effort by **30%** and accelerating decision-making for real-time business needs.

### Hotstar | Data Engineer

Dec 2019 - Nov 2020

- Designed and implemented a **highly scalable AWS-based data architecture** leveraging **S3, Redshift, Glue, and Lambda**, reducing data pipeline latency by **35%**, enabling real-time content delivery for 50M+ active users.
- Ensured **data quality and governance** by developing **50+ validation rules** using **AWS Glue and Athena**, reducing discrepancies in streaming analytics by **40%**, leading to a **10% increase in ad revenue** through accurate campaign targeting.
- Implemented **data quality checks** with custom Python scripts and **CloudWatch alerts**, minimizing operational disruptions and achieving **99.9% pipeline uptime** during peak traffic periods, including the IPL 2019 season.

### Paytm | Data Engineer Intern

Jun 2019 - Nov 2019

- Constructed **batch** and **real-time pipelines** for **user engagement** analytics, improving **retention rates** by **15%** through personalized recommendations.
- Pioneered the development of a **content recommendation engine**, increasing user engagement by **20%**.

## SKILLS

---

**Big Data Frameworks:** Apache Spark, Hadoop, Hive, HBase, MapReduce

**Cloud Platforms:** AWS (S3, Redshift, Glue, EMR, Lambda, SageMaker, CloudWatch), Azure (ADF, Synapse, Data Lake), GCP (BigQuery, Dataflow)

**ETL Tools and Data Orchestration:** Apache Airflow, DBT, Apache Beam

**Data Warehousing:** Snowflake, Google BigQuery, AWS Redshift

**Data Quality & Governance:** AWS Glue Data Catalog, AWS Athena, Data Validation Frameworks

**Programming Languages:** Python(Scikit Learn, Matplotlib, SciPy, Pytorch, TensorFlow), C, Shell Scripting

## EDUCATION

---

**Saint Louis University, St.Louis, MO**

Master of Science, Supply Chain Analytics.