



## JEEVANREDDY KOMMULA

|  jeevanreddykommula99@gmail.com |  Portfolio |

### Professional Experience

#### Johnson and Johnson

May 2022 – Jan 2024

##### Data Engineer

Remote, NJ

- Designed and developed Johnson and Johnson's **Data Warehouse**, implementing **15+ efficient data pipelines** using **HiveQL, Spark, YARN, and Presto**
- Led the creation of a high-performance data processing pipelines on AWS, leveraging **AWS Glue, EMR, S3, Lambda**, and **IAM roles**; achieved a **30%** decrease in processing times, saved **48%** of storage and significantly enhanced system efficiency.
- Engineered comprehensive database management systems utilizing **SQL, Oracle, and MongoDB**, enabling swift and precise manipulation of large datasets, vital for driving data-driven business analysis initiatives and boosting revenue by **25%**.
- Utilized **spark** to develop custom software solutions with employing various algorithms like **MurMur-hashing** for reduced storage, **Round-Robin** for scheduling, aiming **retrieving, parsing**, processing both **structured and unstructured data**, ensuring high **data quality** with less than **0.1% error-rate** and **consistency**.

#### University of Central Missouri

Jan 2022 – Mar 2022

##### Research Data Engineer

Missouri

- Designed and implemented robust data orchestration workflows using **Apache Airflow, Lambda** for trigger alerts, and **CI/CD (Jenkins, EKS, ECS)** pipelines, Managed cloud infrastructure with **Terraform, Docker, K8's**. obtained **99%** reduction in manual intervention.
- Utilized advanced data modeling techniques to streamline complex datasets, enhancing accessibility downstream and gaining insights into **Kinesis** and **Firehose**.

#### Stanford Medicine

Feb 2022 – May 2022

##### Data Engineer

Hyderabad, India

- Collaborated with Missouri Public Health officials to gather COVID-19 data using **API Endpoints** and shaped recurring and non-recurring data pipeline transformations using **DBT** to ingest data in **Snowflake**.
- Moderated in publishing Stanford's Official Interactive and Live Covid-19 Dashboard in **Tableau**.

#### Pike Solutions

Jun 2020 – Jun 2021

##### Data Engineer

Hyderabad, India

- Directed the implementation of a sophisticated Data Warehouse for an E-fitness client, managing 100+ tables and coordinating 25+ ETL Pipelines using **Azure Data Factory**. Extracted data from **MongoDB, MySQL, App Store Connect API**, and **WooCommerce**.
- Optimized ETL logic, orchestrated pipelines using **Airflow**, and analyzed event logs to integrate data with **Power BI** Dashboards for 3+ client teams; Secured **99%** accuracy, surpassing tight deadline with minimal team.

#### Pike Solutions

May 2019 – May 2020

##### Data Analyst

Hyderabad, India

- Developed an optimized **Supply Chain** procurement for the E-fitness company and presented results on a **PowerBI** dashboard.
- Implemented **CRISP-DM** methodology, diagnosed **Exploratory Data Analysis (EDA)**, and helped reduce orders leadtime by **15%** by integrating production and sales data in **Power BI** dashboards.

### Technical Skills

**Programming Languages:** Python (Pandas, NumPy, Seaborn, Streamlit), SQL, Scala, R.

**Databases:** Oracle, MS SQL Server, PL/SQL, MongoDB, HBase, Hive, DBT

**Data Engineering:** Informatica, Hadoop, Airflow, Apache Spark, Docker, Kafka

**Cloud Technologies:** Snowflake, Bitbucket, Azure, Databricks, Redshift

**Visualization Tools:** Tableau, PowerBI, Google Data Studio, Excel

### Education

#### University of Central Missouri

Aug 2021- Dec 2022

Master of Science in Computer Science

Missouri

#### LOVELY PROFESSIONAL UNIVERSITY

Aug 2016-May 2020

Bachelors in technology in Computer Science and Engineering

Punjab, India