

# Janapareddy Naveen

---

📍 Hyderabad ✉ naveenj15597@gmail.com 📞 9989413810 🌐 <https://www.linkedin.com/in/naveen-janapareddy15597/>

## Summary

Data Specialist with 4+ years of experience in Python, Hadoop and Spark having good knowledge in Hadoop and Spark Architectures. Experienced in coding in Python for pre-processing and having hands on experience in AWS S3, EMR and EC2 with exposure to Agile environment in an ETL project. Currently gaining Knowledge on other AWS services such as Redshift and Athena.

## Experience

**Associate, Data | MassMutual India | Hyderabad | September 2021 - September 2022**

Project: DNA\_RULES\_ENGINE

- Built and Maintained ETL pipelines for large datasets to load data into a Data Warehouse.
- Designed a STAR Schema in Data Warehousing Model with a centralised Fact Table indexing multiple dimensional tables.
- Expertise in writing Pyspark SQLs and handling Historic data using SCD Type 2 logic.
- Performed Query tuning for optimizing the query thereby reducing the runtime by 30%.
- Crafted the Automation of the ETL Process using Apache Airflow DAGs reducing 80% of manual effort.

Project: GET\_DAG\_STATUS - POC

- Built a low-cost object-oriented design solution in a fast-paced environment for monitoring the status of crucial DAGs thereby maintaining Business as Usual and increasing the surveillance by 90%.
- It saved around 4 hours/ day of manual effort by team members to troubleshoot and prepare a report.

**Senior Systems Engineer | Infosys Ltd | Mysore | June 2018 - September 2021**

Project: FP\_Data\_Analysis

- Developed an end-to-end solution for reporting the success rate for a batch of around 1000 intakes which helps around 200+ intakes.
- Worked Closely with the Client to get a good understanding of the Business Requirements.
- Implemented Trifacta scripts to provide clean and aggregate data and crafted complex Spark SQL queries.

## Education

**Degree of Master of Science | Liverpool John Moores University | Liverpool, UK | 2023 |**

**Bachelor of Technology in ECE | Anil Neerukonda Institute Of Technology and Sciences | 8.04 | Visakhapatnam | 2018 |**

## Certifications

**HackerRank Certification on Python (Basic) | HackerRank | 2022**

**HackerRank Certification on SQL(Basic) | HackerRank | 2022**

## Skills

- Big Data Tools: Hadoop, Spark-Python/Scala, Hive, Pig, Sqoop
- Cloud: AWS – S3, EMR and EC2
- Databases: Vertica, Postgresql, Oracle, MySQL
- Languages: Python, Basics of Java and Scala