

# Aashritha Venepally

## DATA ANALYTICS ENGINEER

I'm a highly skilled Data Analytics Engineer with 5+ years of experience in analyzing complex datasets, developing actionable insights, and driving business optimization. Proficient in Python, SQL, Tableau, Power BI, and cloud platforms, leveraging advanced analytics and machine learning to deliver impactful, data-driven solutions. Recognized for adaptability, collaboration, and results-oriented problem-solving.

✉ bigdata.aashritha21@gmail.com

📞 3028876590

in [www.linkedin.com/in/aashritha-a-336333106](https://www.linkedin.com/in/aashritha-a-336333106)

## SKILLS

- Big Data Technologies, Statistical and Data Analysis
- SQL (MySQL, SQL Server, Oracle)
- Python, Java Script, C#, .Net Programming
- Hadoop Eco System
- AWS Cloud Services : EC2, Lambda, ECS, S3, RDS, DynamoDB, Redshift, Athena, EMR, Kinesis, Glue, Lake Formation, Data Pipeline, Quick Sight, Sage Maker, CloudWatch, IAM.
- Azure Cloud Services: Functions, Blob, Data Lake, SQL & Cosmos DB, Synapse Analytics, Data Factory, Stream Analytics, Databricks, Monitor, App Insights, Event Hub, Kusto.
- Databricks Platform
- Snowflake
- Apache Airflow
- Data Visualization/ BI: Tableau & Power BI
- Generative AI, AI/ML , LLM
- Google Analytics (Sheets, Looker Studio, Sheets, App Scripts, PLX Dashboards)

## CERTIFICATIONS

[AWS Certified Data Engineer](#)

[Azure Certified Data Engineer](#)

[Databricks Certified Data Engineer](#)

[Forge Data Analytics and Visualization](#)

[Hacker Rank - SQL Intermediate](#)

[Python for Data Engineering](#)

[JIRA Fundamentals Badge](#)

## EXPERIENCE

### Data Lead

Global Logic, USA Sep 2024 - Present

- Designed and maintained dynamic dashboards in Google Sheets and Looker to analyze submitted KPIs for Google's LLM responses, providing actionable insights that contributed to enhanced model training and optimization.
- Automate the extraction and reporting of KPI data, including daily low productivity metrics, ensuring timely delivery of insights.
- Streamline the generation of daily low productivity reports to identify trends and areas for improvement.

- Analyze KPI's and productivity data to provide insights into machine learning model performance and enhance prompt generation.
- Utilize Google Apps Script and advanced formulas to automate various reporting and data management processes, improving efficiency.
- Analyze complex datasets and create automated workflows to streamline reporting and enhance data insights.
- Work with cross-functional teams, including PMs, PGMs, and other stakeholders, to gather requirements and deliver customized reporting solutions.
- Ensure the accuracy and consistency of data across WBR / MBR reports by implementing quality control measures. Continuously assess and improve reporting processes, integrating automation tools to enhance efficiency.
- Maintain documentation for reporting processes, data sources, and automation workflows for transparency.

## Python Engineer

**Speridian Technologies, USA** *Apr 2024 - Sep 2024*

- Actively worked and lead a project that centers on writing, evaluating, and correcting a series of API calls and text/code responses generated by one of the Google's (LLM) Gemini.
- Adhering to Google guidelines and documentation on Gemini AI/LLM principles to ensure up to 100% of LLM responses reflect these standards.
- Ensuring the highest quality final product for Google Gemini by curating, reviewing the LLM responses evaluated by curators and final review of Python API's, code comments, data, and documentation.
- Led a multidisciplinary team of software engineers and content analysts in delivering Python-based solutions for LLM projects, ensuring alignment with client requirements, timely execution, and optimal performance outcomes.
- Evaluating the effectiveness, accuracy, and quality of code responses produced by LLMs.
- Utilize Python to develop scripts for automated evaluation metrics.

## Data Engineer

**Skill Voice Inc, USA** *Sep 2022 - Apr 2024*

- Conducted exploratory data analysis (EDA) and feature engineering to identify key trends and insights using **Python** and **SQL**, and communicated findings to stakeholders in a clear and concise manner.
- Integrated **Azure Databricks** with other Azure services such as **Azure Data Lake Storage**, **Azure SQL Database**, **Azure Cosmos**, **Azure Blob** and Azure Synapse Analytics for end-to-end data solutions.
- Implemented and maintained **Azure Synapse Analytics** (formerly SQL Data Warehouse) for efficient data warehousing and analytics purposes.
- Integrated diverse data sources by leveraging **Azure Data Factory** linked services, enabling a unified view of the data for analysis and reporting.
- Developed real-time data processing solutions using **Spark Streaming** and Databricks Structured Streaming, integrating with Azure services like **Event Hubs** for continuous data ingestion.
- Implemented data security measures, including Azure Data Lake Storage security policies and role-based access control (RBAC), to ensure data privacy and compliance.
- Integrated data engineering solutions with **Azure DevOps** for version control, continuous integration, and continuous deployment (CI/CD) processes.
- Established monitoring solutions using **Azure Monitor** and **Splunk** Log Analytics, proactively identifying and resolving performance bottlenecks in data pipelines.
- Collaborated with cross-functional teams, including data scientists and business analysts, to understand data requirements and deliver actionable insights.
- Provided technical support and troubleshooting assistance for data engineering solutions, ensuring minimal downtime and optimal system performance.

## Big Data Developer

**Perfexion Information Technologies, India** *Apr 2020 - Jun 2021*

- Implementing data ingestion processes to bring in data from various sources into the **Hadoop ecosystem - HDFS**.
- Utilizing tools like **Apache Sqoop** to transfer data efficiently.
- Developing and optimizing **MapReduce** programs or utilizing higher-level abstractions like **Apache Spark** for data processing.
- Writing and debugging **HiveQL** queries for data transformations.
- Utilized **Python** to integrate and interact with various components of the **Hadoop ecosystem**, including **HDFS**, **Hive**, and **HBase**.
- Proven ability to **optimize** and **fine-tune** Hadoop clusters for enhanced performance and scalability.
- Managing and optimizing **data storage** strategies for efficiency and performance.

- Extensive experience designing, developing, and orchestrating complex data workflows using **Apache Airflow**.
- Identifying and resolving performance bottlenecks in Hadoop clusters.
- Collaborated closely with cross-functional teams, including data scientists, analysts, and system administrators, to ensure seamless integration of Hadoop solutions.

### Assistant System Engineer (Data Analytics Engineer)

**Tata Consultancy Services, India** *Dec 2018 - Mar 2020*

- Proficient with **ETL** (Extract, Transform, Load) processes.
- Experience working with **Java** and **Python** programming languages using libraries (**NumPy**, **Pandas**, **Matplotlib**, **boto3**), frameworks (**Flask**, **Django**) for Data Engineering tasks.
- Expertise in working with various databases like **SQL** databases (PostgreSQL, MySQL) and **NoSQL** databases (MongoDB, Cassandra).
- Proficiency in creating and maintaining **database schemas** and **data models**.
- Proficient in **AWS** services such as **EC2**, **Lambda**, **S3**, **Glue**, **Redshift**, **Athena**, **EMR** and **Quick Sight**.
- Designed and developed visually appealing and intuitive dashboards to display **KPIs** using tools such as **Tableau**, **Power BI**, or custom solutions.
- Proficiency with version control systems, such as **Git**.
- A strong understanding of the software development life cycle.
- Ability to work independently and collaboratively in a fast-paced environment
- Ability to troubleshoot and debug issues.

### Data Science Using R Intern

**Brain O Vision Solutions Pvt Ltd, India** *Jun 2018 - Nov 2018*

- Analyzed and interpreted complex data sets to provide actionable insights.
- Proficient in identifying and handling missing data and outliers in datasets using **R** and **Python**.
- Leveraged **scripting** and **programming** skills in **R** and **Python** to automate data cleaning processes, improving efficiency and reproducibility.
- Mastered the use of tidyverse packages such as **dplyr** and **tidyr** for **data manipulation**.
- Demonstrated expertise in identifying **outliers** through visualization tools like box plots and histograms in **R studio** using packages like **ggplot2** for creating informative plots.
- Explored working with **Tableau** and **Power BI** for creating insightful and interactive data visualizations.

## EDUCATION

---

### Masters in Information Systems Technology

**Wilmington University, Delaware, USA** *Aug 2021 - Dec 2022*

4.0 GPA

### B Tech Computer Science and Engineering

**Jawaharlal Nehru Technological University Hyderabad** *Sep 2014 - May 2018*

3.5 GPA