

# HARSHINI YERRAMSHETTY



+19452893920

DATA ANALYST



harshiniyerramshetty04@gmail.com

## PROFESSIONAL SUMMARY

- **4+** years of experience in data analysis, reporting, and business intelligence using **Power BI, SQL, and Tableau**.
- Skilled in data wrangling, ETL, and big data processing using **Python, R, Spark, Hive, and NoSQL databases**.
- Experienced in machine learning, statistical analysis, regression, and predictive modelling to generate actionable insights.
- Developed and automated dashboards, ETL pipelines, and analytics solutions on cloud platforms including **AWS and Google Cloud**.

## TECHNICAL SKILLS

Languages	Python, R, SQL, Bash
Packages	NumPy, Pandas, Scikit-learn, Tensor Flow, Matplotlib, Seaborn, Plotly, NLTK
Cloud Tools	AWS (EC2, S3, RDS Redshift, EMR), Google Cloud (Big Query, Kubernetes)
Databases	MySQL, PostgreSQL, MS SQL Server, MongoDB, Hive, Presto, AWS RDS, AWS Redshift, AWS Redis, Big Query
Machine Learning:	Linear Regression, Logistic Regression, Naive Bayes, Decision Trees, Random Forest, Support Vector Machines (SVM), K-Means Clustering, K-Nearest Neighbors (KNN), Random Forest, Gradient Boosting Trees, Ada Boosting, PCA, LDA, Natural Language Processing
Technologies & Primary Tools:	Artificial Intelligence, Machine Learning, ML, PMML, AI, Big Data, Hadoop, Spark, Cloudera, Storm, Kafka, Hive, Impala, Flume, Sqoop, C, C++, Python, Graph Lab, AWS, R, Spark, scikit-learn
ETL Tool	SQL Server Integration service (SSIS)

## EDUCATION

- Master's in **Business Analytics** from Webster University.

## WORK HISTORY

**Senior Data Analyst**

➤ **2024 - Present**

**RABO BANK**

**Chesterfield, MO, USA**

### Key Responsibilities:

- Led the collection, cleaning, and transformation of large-scale datasets from multiple sources using **Python (Pandas, NumPy) and SQL**.
- Designed and optimized relational database schemas, queries, and stored procedures to support complex reporting and analytics.

- Conducted advanced **exploratory data analysis (EDA)** to identify trends, patterns, and anomalies in business and customer data.
- Developed and maintained interactive dashboards and reports in **Power BI and Tableau** for monitoring KPIs, business performance, and operational metrics.
- Performed statistical analysis including **hypothesis testing, regression analysis, confidence intervals, and A/B testing** to support business strategies.
- Automated reporting processes and data pipelines using **Python and SQL**, improving efficiency and reducing manual effort.
- Integrated structured and unstructured data from multiple sources, including **APIs, Excel, and databases**, for comprehensive insights.

**Environment:** Python (Pandas, NumPy, Scikit-learn), SQL, Power BI, Tableau, Excel, Machine Learning, Statistical Analysis, Data Visualization, APIs.

## Data Analyst

Argo Group

11 / 2023 - 8 / 2024

Spring Field, MO, USA

### Key Responsibilities:

- Collected, cleaned, and transformed data from multiple sources using **Python (Pandas, NumPy)** and **SQL** for analysis.
- Developed and optimized **SQL** queries, stored procedures, and views to extract meaningful insights from large datasets.
- Performed **exploratory data analysis (EDA)** to identify patterns, trends, and anomalies using Python and visualization tools.
- Built interactive dashboards and reports in **Power BI and Tableau** to support business decision-making.
- Applied feature engineering and data pre-processing techniques to prepare datasets for analysis and predictive modelling.
- Automated repetitive data extraction and reporting tasks using **Python and SQL scripts**, improving efficiency.
- Integrated structured and unstructured datasets from multiple sources including **APIs, Excel, and relational databases**.
- Created data visualizations including **bar charts, line charts, heat maps, scatter plots, and geographic maps** to communicate insights effectively.

**Environment:** Python (Pandas, NumPy, Matplotlib, sea born), SQL, Power BI, Tableau, Excel, Data Visualization, Statistical Analysis.

## Data Analyst

Viasat

05 / 2021 - 07 / 2023

Chennai, India

### Key Responsibilities:

- Designed **MySQL table schemas** and implemented stored procedures to extract and store customer usage, session, and subscription data.
- Queried and validated large-scale datasets from **MySQL**, detecting inconsistencies and ensuring data quality for reporting.

- Processed raw telemetry and operational data using **Hadoop HDFS and Spark RDD** transformations to enable large-scale analytics.
- Integrated data from multiple sources including **MongoDB, APIs**, and internal systems for comprehensive network and customer insights.
- Developed **Power BI dashboards** and reports to visualize network reliability, customer satisfaction, and subscription trends.
- Built predictive models for churn analysis, bandwidth demand forecasting, and customer lifetime value using **Python and machine learning techniques**.
- Collaborated with engineering and operations teams to convert analytical insights into actionable strategies for improving satellite services.

**Environment:** MySQL, Python (Pandas, NumPy, Scikit-learn), Hadoop, HDFS, Spark (RDD), MongoDB, APIs, Power BI, Tableau, Machine Learning.