

# Satya Sai Vamshi Krishna Arza

Chicago, IL | arzavamshikrishna@gmail.com | (773)-835-4154 | [LinkedIn](#) | [GitHub](#) | [Tableau](#) | [Portfolio](#)

## SUMMARY

Data Analyst with 3.5 + years of experience and hands-on experience in Python, SQL, and data visualization, with a proven ability to deliver actionable insights and optimize business processes. Proficient in machine learning, big data technologies, building ETL pipelines and cloud platforms like AWS and Azure. Passionate about transforming complex data into meaningful insights.

## EDUCATION

Master's in Data Analytics (CGPA 3.8)  
University of Illinois at Springfield

August 2022 - December 2023

## PROFESSIONAL EXPERIENCE

### Data Analyst | InfoPro Ltd | Chicago, Illinois

March 2024 - Present

- Streamlined data processing workflows, leading to a 25% increase in data accuracy and efficiency.
- Led **data analysis, ingestion, and integration** initiatives, ensuring seamless acquisition and consolidation of diverse data sources.
- Performed complex data analysis to identify trends and insights, which contributed to strategic business decisions.
- Demonstrated expertise in writing, analyzing, and **debugging SQL** queries for data manipulation and reporting purposes.
- Created interactive dashboards using **Tableau** and **Power BI**, improving data visibility and decision-making.
- Worked closely with cross-functional teams to understand data needs and deliver actionable insights.

### Associate Software Engineer | Aviro Global Solutions Ltd | Bengaluru, India

August 2019 - July 2022

- Leveraged Big Data technologies such as **Hadoop, MapReduce (M/R), Hive, and Spark** to process and analyze vast data efficiently.
- Orchestrated **ETL pipelines** for data extraction, transformation, and loading, streamlining data workflows and ensuring data quality.
- Developed and implemented **machine learning algorithms** to extract insights from large structured and unstructured datasets, improving predictive accuracy by 25%.
- Led the integration of **AWS** and **Azure** services, reducing data processing time by 40%.
- Utilized **Apache Spark** and **PySpark**, cutting processing time by 50%.
- Collaborated with teams to deliver key software features, resulting in a 20% increase in user satisfaction.
- Optimized database performance, achieving a 15% reduction in query times.
- Implemented **Tableau data visualizations**, resulting in a 15% improvement in data interpretation efficiency.

### Business Development Intern | Flipspaces Ltd | Remote

December 2020 – May 2021

- Conducted market research and analyzed economic data using various tools, including IMPLAN, to support strategic decisions.
- Performed complex data analysis using **SQL** and **Python libraries like pandas** to identify trends and patterns.
- Developed **predictive models** to identify high value customer segments, resulting in a 25% increase in targeted sales.
- Designed SQL database for efficient customer data management, leading to 20% faster resolution time.
- Created interactive dashboards using Tableau and Power BI, improving data visibility and decision-making.
- Communicated with departmental leaders on reports aligned with strategy and ensured **data accuracy**

## SKILLS

• R	• Hadoop	• MySQL	• Data Integration
• PYTHON	• Hive	• CSS	• Data Ingestion
• SQL	• MapReduce	• Tableau	• Data Modelling
• Java	• Big Data	• ETL Development	• Statistical Analysis
• Amazon Web Services (AWS)	• GitHub	• API Integration	• Data Streaming
• Snowflake	• PostgreSQL	• Data Mining	• Data Analysis
• Azure	• Git	• Microsoft SQL Server	• Data Processing
• JavaScript	• Business Intelligence	• Data Cleaning	• Data Quality
• Spark	• Amazon Redshift	• Data Pipeline	• Shell Scripting
• UNIX	• Docker	• Data warehousing	• IMPLAN
		• Data Visualization	

## HIGHLIGHTED PROJECTS – [\[GitHub\]](#)

- Graduate Admission Data Analysis** [LINK](#)  
A project predicting graduate admission chances using machine learning algorithms, emphasizing GRE scores and CGPA.  
**Tools Used:** R, R Studio, Caret, Keras, TensorFlow
- Cricket Sport Data Analysis** [LINK](#)  
Analyzed IPL cricket data to provide insights on match performance and player statistics.  
**Tools Used:** Amazon S3, Databricks, Python, SQL, PySpark, Apache Spark
- Superstore Data Analysis** [LINK](#)  
Processed and Visualized store data using AWS Services  
**Tools Used:** Amazon S3, Databricks, Python, SQL, PySpark, Apache Spark
- HR Workforce Data Analysis** [LINK](#)  
Analyzed HR data and created a dashboard to inform decision-making on hiring, retention, and workforce management.  
**Tool Used:** MySQL, PowerBI
- Tableau Projects** [LINK](#)  
Created interactive and impactful visualizations using Tableau Public, including trends, distributions and comparisons.

## CERTIFICATIONS

- Google Data Analytics Professional Certification by Google
- Excel Data Analysis by LinkedIn
- Python Data Structures by University of Michigan