

# VISHAL GIRIJAPURAM

Albany, New York

☎ +1-518-606-7963

✉ [girijapuramvishal@outlook.com](mailto:girijapuramvishal@outlook.com)

🌐 [VishalGirijapuram](#)

## CAREER OBJECTIVE

---

Aspiring Data Engineer with a strong foundation in building and optimizing data pipelines using Azure services such as Azure Data Factory, Azure Databricks, Azure Synapse Analytics, and Azure Data Lake Storage. Proficient in managing large-scale datasets and implementing ETL processes to streamline data integration and transformation. Eager to contribute to a dynamic organization by applying expertise in cloud technologies, data analytics, and Azure DevOps to drive data-driven decision-making and enhance business operations. Committed to continuous learning and delivering high-quality solutions in a fast-paced, collaborative environment.

## Professional Summary

---

- Experienced in utilizing **Azure Data Factory (ADF)** to design and orchestrate scalable **ETL/ELT pipelines**, facilitating seamless integration of diverse data sources and automating workflows.
- Proficient in building optimized ETL/ELT processes, ensuring data quality and performance using platforms like **Azure Data Factory** and **Azure Databricks**.
- Extensive hands-on experience with **Azure Databricks** for processing large datasets, performing advanced analytics, and implementing machine learning models.
- Expertise in leveraging **Azure Data Lake** to manage and organize large-scale data storage, ensuring secure and efficient data access.
- Knowledgeable in **Azure Synapse Analytics** for advanced data warehousing solutions and large-scale data analysis.
- Experienced in utilizing **Snowflake** to design and manage scalable cloud-based data warehousing solutions, optimizing data storage, query performance, and enabling seamless integration with diverse data sources for advanced analytics and reporting.
- Proficient in creating interactive dashboards and visualizations using **Power BI** and **Tableau**, effectively translating complex data into clear, actionable insights for stakeholders.
- Skilled in setting up **CI/CD pipelines** with tools like **Azure DevOps**, automating deployment and ensuring continuous integration for data engineering processes.
- Skilled in operating within **Agile** and **Scrum** methodologies, participating in daily **stand-up meetings** to ensure smooth communication, iterative development, and timely project delivery.

## Skills

---

**Programming Languages:** SQL, Python, Java, JavaScript, R.

**Scripting Languages:** HTML, CSS.

**Operating Systems:** Windows, macOS, Linux.

**Databases:** SQL Server, MySQL, MongoDB.

**Networking:** Cisco Networking.

**Analytical Tools:** Numpy, Pandas, Seaborn, Matplotlib, Excel, Tableau, Power BI.

**Cloud Platforms:** Azure Cloud Services, Azure Data Factory, Synapse Analytics, Azure Data Lake, Azure Databricks, Azure DevOps, Azure cloud Power platform, Snowflake, AWS Glue, Amazon Redshift, and Amazon S3, Amazon EC2, AWS Cloud Services, Google Cloud, Compute Engine.

**Dev Ops and CI/CD:** Azure DevOps, AWS CodePipelines, Git Hub, Bitbucket, Git Lab, Jenkins.

**Machine Learning:** Clustering, Linear and Non-Linear Classifications, Regression, Natural Language Processing(NLP)

**Application Softwares:** Matlab, Google Colab, Jupyter, Anaconda, VS Code Data Visualization, PyCharm, AnSys.

**Libraries:** Scikit learn, statsmodel, tensorflow, keras, nltk

**Additional Skills:** Prompt Engineering, Analytical and Program Solving Skills, Project Management, Agile, Numerical Reasoning.

## Work Experience

---

### Data Engineer - Intern

**JUL 2023 – DEC 2023**

*New York State Energy Research and Development Authority (NYSERDA)*

*Albany, New York*

- **Collected** and **analyzed** data related to diversity, equity and inclusion within the organization i.e, Evaluate demographic data, employee surveys, and other relevant information to identify patterns and trends.
- Work closely with **cross-functional teams** to understand business requirements and provide data-driven insights.

- Processed and cleaned large datasets using few **data scaling techniques** to ensure data accuracy and integrity.
- Created **visualizations** and **dashboards** to communicate employee-related trends and insights.
- Utilized **Azure Data Factory (ADF)** to orchestrate and automate data workflows, integrating data from multiple sources.
- Managed large datasets in **Azure Data Lake** for secure and scalable storage, enabling efficient data processing.
- Leveraged **Azure Databricks** for advanced **data analytics** and **machine learning** model development on large datasets.
- Used tools like **Tableau** or other visualization platforms to present data in a clear and compelling manner.

## Assistant System Engineer - Data Engineer

NOV 2020 – JUL 2022

*Tata Consultancy Services (TCS)*

*Chennai, India*

- Designed and developed ETL pipelines using **Azure Data Factory (ADF)** to automate data workflows and integrate multiple data sources.
- Managed and stored large datasets in **Azure Data Lake**, ensuring efficient data retrieval and security.
- Processed and transformed large datasets using **Azure Databricks** to enable advanced analytics and machine learning models.
- Collaborated with stakeholders to design and optimize data models in **Azure Synapse Analytics**, improving data accessibility and performance for business intelligence.
- Utilized **Snowflake** to design and manage scalable cloud-based data warehousing solutions, optimizing data storage, query performance, and enabling seamless integration with multiple data sources for advanced analytics and reporting.
- Developed interactive and insightful dashboards using **Power BI** to visualize key business metrics, enabling stakeholders to make data-driven decisions and monitor performance in real-time.
- Implemented **CI/CD pipelines** for seamless deployment and integration of data engineering solutions using **Azure DevOps**.
- Monitored and optimized data pipelines for performance and cost-efficiency on the Azure platform.
- Ensured data accuracy and integrity by conducting regular audits and implementing data governance practices.
- Supported data security and compliance through role-based access control and encryption practices within Azure services.

## Certifications

- Snowflake - The Complete Masterclass.
- Azure Data Factory - Azure Data Engineering on Azure and Fabric.
- End to End Azure Data Engineering Real Time Project.
- Certified in Microsoft Power BI: The Complete Guide.
- Certified by HackerRank - SQL (Intermediate).
- Certified by HackerRank - SQL (Basics).
- Certified in Machine Learning A-Z™: Hands-On Python R In Data Science.
- Certified in Programming for Everybody (Getting Started with Python).
- Certified in SQL - MySQL for Data Analytics and Business Intelligence.

## PROJECTS

### F1 Racing Database Analytics | Python, Azure Data Factory, Snowflake, Power BI

SEP 2024

- Migrated historical F1 racing data from an on-premise SQL Server database to Snowflake using **Azure Data Factory (ADF)**, ensuring secure, scalable, and efficient data transfer to the cloud.
- Leveraged **Snowflake** for processing and transforming large datasets (e.g., lap times, race results, driver stats) using **SQL** and Snowflake's built-in functions to calculate key performance metrics like average lap speed and race standings.
- Designed an optimized data model within **Snowflake** that integrated multiple data sources (e.g., races, drivers, teams). Partitioned data to improve query performance and ensure fast retrieval of critical race information.
- Developed dynamic Power BI dashboards connected to Snowflake, allowing users to filter and slice data by season, team, driver, and apply slicers for in-depth visualizations of race trends, pit stops, and team standings.
- Generated comprehensive race performance reports with key metrics such as pole positions, fastest laps, and podium finishes. Enabled granular analysis with filters, providing season-long trends and driver comparisons.
- Published the interactive dashboard on **Power BI** Service, enabling real-time data access for stakeholders. Integrated role-based access control (RBAC) within Snowflake to ensure data security, and facilitated collaboration across teams.

## Predictive Maintenance for IoT Devices | ADF, Databricks, Datalake, Power BI, Python APR 2024

- Developed an end-to-end data engineering solution to predict maintenance needs for IoT devices in a home automation system using historical and real-time data.
- Utilized **Azure Data Factory** to orchestrate ETL pipelines for data ingestion and transformation, ensuring data from multiple sources was accurately processed.
- Employed **Azure Data Lake** for scalable data storage and Azure Databricks for advanced data processing and model training.
- Created predictive models with **Python** to forecast maintenance needs, achieving high accuracy and uncovering critical patterns in device performance data.
- Designed interactive dashboards in **Power BI** to visualize maintenance predictions, trends, and alerts for actionable insights.
- Implemented **CI/CD pipelines** to automate the deployment and updates of data processing workflows and predictive models.
- Demonstrated expertise in integrating cloud services and data engineering tools to enhance system reliability and reduce downtime.

## Sales Performance Dashboard | ADF, Azure SQL Database, Azure Synapse, Power BI JAN 2024

- Integrated and processed sales data from multiple sources using **Azure Data Factory**, orchestrating ETL workflows to ensure data accuracy and consistency.
- Stored and managed the data in **Azure SQL Database**, performing data transformation and aggregation for detailed sales analysis.
- Leveraged **Azure Synapse Analytics** to perform advanced queries and create a data model for in-depth sales performance analysis.
- Developed an interactive **Power BI** dashboard to visualize key sales metrics, trends, and performance indicators, enhancing decision-making capabilities.
- Implemented dynamic filters, slicers, and drill-down features in Power BI to enable users to explore sales data in a detailed and interactive manner.
- Conducted data validation and testing to ensure the dashboard accurately reflects sales performance and provides actionable insights for stakeholders.

## Home Automation using IoT | Arduino IDE, Cloud service (Adafruit IO and ThingSpeak) FEB 2020

- Implemented IoT communication through **Adafruit IO**, enabling real-time monitoring and control of connected devices over the internet.
- Integrated **ThingSpeak** as a cloud service to enhance data logging, analytics, and remote monitoring capabilities within the home automation system.
- Created intelligent interactions between devices, allowing for coordinated actions such as turning off lights when the TV is switched off.
- Developed a user-friendly interface for remote control, allowing homeowners to manage and monitor devices through a web or mobile application.

## EDUCATION

**State University of New York at Albany** **Albany, New York**  
*Master's of Science in Data Science - CGPA-3.48/4.00* *May 2024*

Related Coursework : Machine Learning, Python programming, DataBase Systems, Topological Data Analysis, Advanced Statistics.

**Anurag Group of Institutions** **Hyderabad, India**  
*Bachelor of Technology in Electronics and Communication - CGPA - 8.78/10.0* *May 2020*

Related Coursework : C Programming, C++, JAVA, Analog and Digital Signal, Probability and Statistics, DataBase Management, Numerical and Logical Reasoning, Image Processing, Mathematics, Data Structures.

## Extracurricular Activities

- Participated in a workshop: CODEBEAT 2018 – A workshop on IOT by BITS Palani.
- Participated in a workshop: Real time Applications using MATLAB with Raspberry PI - by Cranes Software International Limited.
- Participated in a 36 Hour HACKATHON: 'Drowning detection device' - using Node MCU, GPS Tracker, moisture sensor and IFTTT application.

## Achievements

---

- Rewarded as one of the top 5 employees for the dedication, delivery and attitude towards work.
- Achieved second in place in MSP-430, an online contest organized by Texas instruments.
- Consistently recognized for academic achievement, ranking among the top students in my Bachelor of Electronics program.