

Vinay Kumar Camarushi

Data Analyst, Jersey City, NJ

vinaycvk01@gmail.com • 9137300497

[GitHub](#) [Portfolio](#) [LinkedIn](#)

SUMMARY

- Results-driven Data Analyst with 1.5+ years of experience in leveraging Python (Pandas, NumPy, Matplotlib, Keras, TensorFlow) along with SQL, Java, and C for data-driven decision-making.
- Proficient in developing interactive dashboards and data visualizations using Power BI and Tableau to provide actionable insights.
- Strong expertise in data modeling (Star Schema, Dimensional Modeling) and ETL development for optimized data pipeline workflows.
- Expertise in SQL to query and manipulate large datasets from diverse data sources for insightful business intelligence reporting.
- Skilled in machine learning techniques, including linear regression, logistic regression, decision trees, and random forests for predictive analytics.
- Experience in cloud-based data solutions including Azure Data Lake, Databricks, Data Factory, and AWS Redshift, improving scalability and performance.
- Proven ability to automate data workflows, reducing manual efforts and improving operational efficiency by implementing Python-based automation.
- Adept at data governance, ensuring data quality, security, and compliance with industry best practices.
- Strong communicator with experience in collaborating with cross-functional teams, stakeholders, and executives to translate business requirements into technical solutions.
- Passionate about continuous learning and innovation, staying up to date with the latest trends in data science, analytics, and cloud computing.

EDUCATION

Master of Computer Science

Aug 2022 – May 2024

University of Central Missouri, Lee's Summit, MO

B. Tech in Electronics & Communication Engineering

Jul 2016 – Sep 2020

SRKR Engineering College, Andhra Pradesh, India

TECHNICAL SKILLS

- **Languages:** Python, SQL
- **Data Analytics & ML:** NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, TensorFlow, Keras, Seaborn
- **Data Visualization:** Power BI, Advanced Excel (Pivot Tables, Lookups, Dates), Tableau
- **ETL & Cloud:** SSIS, Informatica, Azure (Data Lake, Databricks, Data Storage, Data Factory, SQL Database, Blob Storage)
- **Databases:** MySQL, SQL Server, Redshift, Oracle
- **Operating System:** Windows, Linux, MacOS
- **Version Control:** Git

CERTIFICATIONS

- Microsoft Certified Data Engineer Associate (**DP - 203**)
- Microsoft Certified Data Analyst Associate (**PL-300**)
- Microsoft Certified Azure Fundamentals (**AZ-900**)
- Microsoft Certified Data Fundamentals (**DP-900**)

PROFESSIONAL EXPERIENCE

TATA Consultancy Services, India

Jan 2021 – Jul 2022

Data Analyst

- Single handedly automated data ingestion pipelines by processing Excel-based datasets with Python (NumPy, Pandas), reducing manual effort by 95% and streamlining data flow into Azure Data Lake.
- Developed and maintained Power BI dashboards for Commercial and Performance metrics for Diageo Great Britain Pvt. LTD.
- Designed SQL Server Management Studio Views to improve data accessibility and alignment with business objectives.
- Debugged and improved Azure Data Factory Pipelines, increasing data processing accuracy.
- Migrated legacy data solutions to AWS, utilizing S3 (data storage), Redshift (data warehousing), and EMR (data processing), cutting operational costs.
- Integrated Tableau dashboards with Amazon Redshift, MongoDB, and Oracle, streamlining enterprise-wide data analytics.
- Managed data infrastructure across cloud and on-premises environments, ensuring high availability, security, and reliability.
- Collaborated cross-functionally to translate business requirements into scalable technical solutions.
- Optimized Tabular Model in Visual Studio, database queries, and data pipelines to enhance performance and efficiency.

ACADEMIC PROJECTS

Online Grocery ecommerce Site

Aug 2023 – Dec2023

- Developed a fully functional e-commerce platform leveraging Flask (backend), MongoDB (database), and JavaScript (frontend UI).
- Deployed the application on AWS to ensure scalability, high availability, and reliability.

Empirical Evaluation of Machine Learning Techniques for Text Classification

Jan 2023 – May 2023

- This project assesses the performance of different machine learning algorithms in classifying text data into predefined categories.
- Analyzed the performance of Naïve Bayes, Support Vector Machines, Decision Trees, and Neural Networks for text classification.
- Assessed models' accuracy, efficiency, and scalability in real-world text mining applications.

Large Scale Matrix Multiplication using Hadoop MapReduce

Jan2023 – May 2023

- Executed high-performance parallel matrix multiplication leveraging Hadoop MapReduce for large-scale data processing.
- Utilized distributed computing to optimize performance for massive datasets.

Airline Ticket Management System

Aug 2022 – Dec 2022

- Designed a web-based ticket booking system using Flask, Python, MongoDB, HTML, CSS, and JavaScript.
- Built an admin panel for airline staff to manage flight schedules, bookings, and generate reports.