Vinay Kumar Camarushi
Data Analyst, Overland Park, KS
vinay.kumar@mycvtalent.com • +1 913-227-9343
https://vinaykumar-dataanalyst.github.io/

SUMMARY

- Around 5 years of experience in Data Analyst, Report Development, and writing SQL scripts, Proficiency in gathering requirements, collecting data, and designing/developing Reports for decision-making.
- Crafted impactful dashboards in Tableau, Power BI, and Looker Studio, utilizing compelling data visualizations to communicate insights and tell data-driven stories.
- Practical understanding of Data modeling (Dimensional and relational) concepts like Star-Schema, Snowflake Schema Modeling, and Fact and Dimension tables.
- Optimizing data storage and processing costs on Azure by utilizing managed services like Azure Blob Storage, Azure SQL Database, and Azure Data Lake Storage to reduce cloud expenses.
- Skilled in developing predictive models using machine learning algorithms (linear regression, logistic regression, decision trees, random forests) to forecast future trends and outcomes.
- Utilizing SQL to extract and analyze large datasets from various sources, identifying multiple key insights that informed data-driven business decisions.

EDUCATION

Master of Computer Science

University of Central Missouri, Lee's Summit, MO

B. Tech in Electronics & Communication Engineering

SRKR Engineering College, Andhra Pradesh, India

TECHNICAL SKILLS

- **Languages:** Python, R, SQL, SAS
- **Packages:** NumPy, Pandas, Matplotlib, SciPy, ggplot2, Scikit-learn, TensorFlow, Keras, seaborn
- **Software/IDEs:** Visual Studio Code, Jupyter Notebook, Talend
- **Visualization Tools:** Tableau, Power BI, Advanced Excel (Pivot Tables, Power Query, Macros & VBA, Lookups, Dates), Alteryx, Google Analytics, Looker Studio
- **ETL & Cloud:** SSIS, Informatica, Azure (Data Lake, Databricks, Data Storage, Data Factory, Azure App Service, Azure SQL Database, Azure Blob Storage), AWS
- **Data Mining and Exploration:** Knime, Web Scraping (Scrapy)
- **Database:** MySQL, SQL Server, MongoDB
- **Other Skills:** Data Cleaning, Data Wrangling, Data Warehousing, Data Mapping, Data Modeling
- **Operating System:** Windows, Linux, MacOS

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

Walmart, Kansas, USA   *Jan 2024 – Present*
Data Analyst

- Identified a previously unknown customer segment with high online grocery order value through Seaborn analysis, leading to a targeted marketing campaign that increased online grocery sales by 30%.
- Utilized Python libraries like scikit-learn to develop and train a churn prediction model, achieving an accuracy of 10%. This model helped identify customers at risk of churning, allowing for targeted interventions and improved retention strategies.
- Managed the performance of the Snowflake platform and executed performance tuning strategies, optimizing SQL queries and troubleshooting performance issues, improving query performance and enhancing user experience by 50%.
- Developed ETL pipelines using Azure Data Factory to extract, transform, and load data from various sources into Azure data warehouses, streamlining data integration processes and reducing errors by 40%.
- Created 50+ DBT models and macros to automate data transformation and loading processes, reducing manual effort by 60%.
- Leveraged Azure SQL Database's integration with cloud-based analytics tools (Azure Synapse Analytics, Azure Machine Learning) to perform advanced analytics.
- Used SSIS to develop comprehensive ETL packages, facilitating data validation, extraction, transformation, and loading into Data Warehouse and Data Mart Databases.
- Wrote and optimized complex SQL queries to extract and transform large datasets from relational databases (PostgreSQL), retrieving 1 million records for analysis.

TATA Consultancy Services, India   *Jan 2021 – Jul 2022*
Data Analyst II

- Extracted and preprocessed nearly 2 TB of data from diverse sources (databases, APIs) using Python. This rigorous data cleaning process, including outlier detection and removal (15%), resulted in a 10% improvement in predictive model accuracy.
- Leveraged data-driven strategies to streamline the recruitment process, achieving a 15% reduction in time-to-fill positions and enhancing candidate selection efficiency.
- Performed Data Cleaning, Data Screening, Data Exploration, Data visualization, Feature Selection and Engineering using Python libraries such as Pandas, NumPy, Scikit-learn (Random Forests), and Matplotlib.
- Applied machine learning techniques using scikit-learn to build predictive models, identify patterns and extract insights from complex datasets.
- Increased data-driven decision-making by 25% using Looker's powerful visualization and analytics capabilities.
- Designed and implemented SSIS packages to automate data extraction, transformation, and loading (ETL) processes for data warehouses, resulting in a reduction in manual effort and a 20% improvement in data processing speed.
- Accomplished interactive Power BI reports leveraging advanced DAX and Power Query modeling techniques to unlock deeper data insights and facilitate informed decision-making.
- Operated Azure Databricks to perform advanced data analysis and machine learning tasks on large datasets, improving data-driven decision-making by 25%.

Macro Software Solutions, India   *Aug 2018 – Dec 2020*
Data Analyst

- Crafted interactive Power BI and Tableau dashboards that mapped customer journeys, pinpointing key touchpoints and gauging customer preferences throughout their lifecycle. These data visualizations unlocked actionable insights, improving customer engagement, experience, and satisfaction.
- Created multiple informative visualizations (histograms, scatter plots, boxplots) using Matplotlib and Seaborn to communicate complex data insights to stakeholders, leading to a 30% improvement in decision-making efficiency.
- Utilized Azure Data Factory to automate data ingestion and transformation processes, reducing manual effort by 50% and ensuring data consistency.
- Extracted and cleaned terabytes of data from relational databases (MySQL) and cloud data warehouses (Snowflake) to prepare for analysis, ensuring data accuracy and reducing analysis time by 20%.
- Built data cleaning routines using scripting languages (Python, SQL) to automate the cleaning process for large datasets, improving data analysis efficiency by 20%
- Improved data quality of supplier delivery schedules by data cleaning techniques, leading to more accurate forecasts and reduced delays in product deliveries.
- Collaborated with marketing to design and execute A/B tests, measuring the impact of marketing campaigns on ROI.