VINAY REDDY VERAMAREDDY

Vinayreddyveramareddy@gmail.com
+18172052908
Montana, U.S
Vinay|LinkedIn

Results-driven Data Analyst with 3 years of experience in data analysis, visualization, and modeling using tools like SQL, Python, Tableau, and Power BI. Proficient in managing big data frameworks (Hadoop, Hive, Cassandra) and cloud technologies (AWS, Azure) to deliver scalable solutions. Skilled in statistical analysis, ETL pipelines, and creating impactful dashboards to enhance business intelligence and support decision-making. Expertise in healthcare analytics with Medicare and Medicaid claims analysis, leveraging SAS and ArcGIS for trend analysis and resource optimization. Strong knowledge of database management (Snowflake, PostgreSQL, MongoDB) and certified in Power BI, Big Data Hadoop, and Python for Data Science. Adept at strategic planning, stakeholder communication, and working in Agile environments, driving data-driven outcomes efficiently.

EXPERIENCE

Data Analyst
State of Montana | January 2024 – Present

- Created and delivered 20+ **Tableau dashboards** and **analytical reports**, implementing secure **role-based access controls** for data integrity and **remote collaboration**.
- Designed advanced **Excel reports** with **pivot tables**, **charts**, and **graphs**, enabling stakeholders to visualize **key performance metrics** and streamline **data communication**.
- Integrated datasets from diverse sources, including **BCBS** and **MMIS**, into Snowflake, ensuring efficient **data consolidation** for **big data analytics**.
- Designed and implemented **data-driven projection models** to optimize **decision-making**, supporting strategic planning and resource allocation for a $1.3B healthcare budget.
- Developed **forecasting models** to predict trends in **Medicare** and **Medicaid reimbursement** and enrollment for a population of over 1.3M, improving **resource management** and **healthcare delivery**.
- Conducted advanced **data analysis** on large datasets using **SAS**, identifying key insights from over 14M records to drive **policy development** and **evidence-based decisions**.
- Built **statistical models** to uncover actionable insights, leveraging **ETL pipelines** and tools like **Snowflake**, **Tableau**, and **SQL** for seamless data integration and analysis.
- Utilized **ArcGIS** to generate **geospatial insights**, mapping trends and identifying regional disparities to enhance **targeted healthcare initiatives**.
- Automated **data visualization** processes, leveraging **SQL** and **Tableau** to identify correlations and provide actionable insights, enhancing **business intelligence** and **decision support systems**.
- Collaborated with cross-functional teams to develop **health metrics dashboards**, consolidating KPIs and improving organizational alignment with strategic goals.

Data Engineer
Volantsoft | June 2022 – January 2024

- Utilized **Azure Data Lake Storage** and **Azure Data Factory** for efficient **data integration**, lowering infrastructure costs by 25% and improving system performance.
- Conducted **exploratory data analysis** on large datasets using Python and **Jupyter Notebook**, generating valuable insights for business strategy.
- Engineered complex SQL queries to extract, manipulate, and analyze data from remote databases, improving retrieval performance.
- Wrote a **Pre-Requisite document** for migration to JIRA using Tasktop, enhancing the software development process.
- Employed **Git and GitHub** for version control, improving team collaboration, reducing deployment errors by 20%, and enhancing code quality.
- Designed and implemented **ETL modules** using tools like **SSIS**, streamlining data transformation processes, and improving data processing efficiency.
- Created **dynamic Tableau dashboards** with calculated fields, parameters, and hierarchies, providing advanced **data exploration capabilities**, and increasing user engagement by 25%.
- Designed and implemented **Power BI scorecards** and **dashboards** with advanced visualization elements such as stacked bars, scatter plots, and geographical maps, significantly enhancing **data visualization** and improving **decision-making processes**.
- Utilized **Pandas libraries** in Python to import and analyze customer data, uncovering actionable insights that drove key business decisions and contributed to a 15% revenue increase.
- Developed **Azure Data Migration** and **Cleansing rules** for Integration Architecture, ensuring accurate and consistent data flow across systems, reducing data errors by 20%.
- Automated reporting functionalities using **Power BI tools** and MySQL, streamlining report generation for supply chain analysts and reducing reporting time.
- Leveraged **Cassandra** for managing large-scale **NoSQL databases**, improving **data availability** and retrieval efficiency for high-volume datasets.
- Applied **Scala** to query and analyze datasets stored in **HDFS**, optimizing query performance and reducing processing time.
- Developed comprehensive data mapping rules and guidelines, ensuring accurate data integration across systems and improving data consistency.
- Demonstrated expertise in **MongoDB** for storing and retrieving structured and unstructured data, enhancing accessibility and utilization.

Data Engineer
Slesha | March 2022 – June 2022

- Developed and optimized **ETL workflows** using **Informatica** and **Talend** for seamless **data acquisition** and **transformation**.
- Implemented **CI/CD pipelines** and automated code deployment using **Visual Studio Team Services (VSTS)**.
- Validated data fields from downstream systems to ensure **data consistency** and **integrity**.
- Built and managed **pipelines** in **Azure Data Factory (ADF)** leveraging **Linked Services**, **Datasets**, and **Pipelines** to **extract, transform, and load (ETL) data** from diverse sources like **Azure SQL**, **Blob Storage**, and **Azure SQL Data Warehouse**.
- Resolved **L3 issues**, installed new components, and streamlined operations through **automation** and the integration of a **CI/CD model**.
- Optimized **big data processing** with **Spark Streaming** and **Spark SQL**, leveraging advanced techniques such as **accumulators**, **broadcast variables**, and **caching strategies** to enhance **job performance**.
- Contributed to the design and implementation of a **Data Lake** using the **AWS Big Data ecosystem** and **Hadoop** frameworks.
- Hands-on with **AWS services** including **EMR**, **S3**, **EC2**, **RDS**, **ELB**, **DynamoDB**, **Glue**, **SNS**, **SQS**, and **CloudFormation**, with experience in Redshift Spectrum and **AWS Athena** for querying data stored in **S3**.

Software Engineer (Intern)
TLC | January 2019 – March 2019

- Designed and implemented **ETL pipelines** using tools like SSIS, improving data processing efficiency by 40% and ensuring smooth data flow between systems.
- Automated supply chain reporting using **Power BI** and **MySQL**, reducing report generation time by 50%.
- Utilized **Cassandra NoSQL databases** to manage large-scale data storage and retrieval, enhancing data availability and accessibility.
- Queried and analyzed datasets stored in **HDFS** using **Hive**, optimizing query performance and reducing processing time.
- Leveraged **Azure cloud technologies** like Data Lake Storage and Data Factory to streamline data integration, achieving a 25% reduction in infrastructure costs.
- Developed and optimized **Power BI dashboards** and scorecards, incorporating advanced visualizations like stacked bars, scatter plots, and Gantt charts, leading to improved **decision-making capabilities**.
- Analyzed customer data using **Python** and **Pandas libraries**, delivering actionable insights that contributed to a 15% increase in revenue.
- Engineered **data migration** and **cleansing rules** to ensure seamless integration across systems, reducing data errors.
- Used **Git and GitHub** for version control, improving team collaboration and reducing deployment errors.

Additional EXPERIENCE

Graduate Research Assistant
TAMUC | August 2020 – September 2021

- Integrated the **Capon algorithm** with the ROOT MUSIC algorithm, resulting in a 75% reduction in complexity and a 45% improvement in signal processing precision.
- Authored and presented research titled *"Capon Root-MUSIC-like Direction of Arrival Estimation Based on Real Data,"* accepted for publication by IEEE at VTC Fall 2021.
- Enhanced **signal retrieval processes** under uniform and non-uniform noise conditions, improving accuracy by 45% through innovative algorithm.
- Designed and implemented new methodologies for complex signal matrices using **Python** and **MATLAB**, optimizing computational efficiency and execution time.
- Applied the **Real Hermitian Symmetric Toeplitz algorithm** with No Decomposition to significantly reduce computational load by 36% while increasing accuracy from 74% to 89%.
- Achieved a 57% improvement in the effective utilization of retrieved source datasets, even in highly correlated environments.
- Conducted advanced exploratory data analysis on signal datasets, uncovering patterns and correlations that informed enhancements to signal processing models.

CERTIFICATIONS

- **Microsoft Certified: Power BI Data Analyst Associate**
  Microsoft | Issued Aug 2023 | Expires Aug 2025
- **Problem Solving (Intermediate)**
  HackerRank | Issued Jul 2023
- **Rest API (Intermediate)**
  HackerRank | Issued Jul 2023
- **SQL (Advanced)**
  HackerRank | Issued Jul 2023
- **Tableau 2022**
  Udemy | Issued Jan 2023
- **Big Data Hadoop and Spark Developer**
  SkillUp Online | Issued Feb 2022
- **Data Engineering Nanodegree**
  Udacity | Issued Feb 2022
- **Tableau Essential Training**
  National Association of State Boards of Accountancy (NASBA) | Issued Feb 2022
- **Hands-On Essentials - Data Warehouse**
  Snowflake | Issued Jan 2022
- **Introduction to Big Data Tools for Beginners**
  Simplilearn | Issued Jan 2022
- **Introduction to Indexing and Aggregation in MongoDB**
  Simplilearn | Issued Jan 2022
- **Python for Data Science and AI**
  Coursera | Issued Mar 2021
- **PG Diploma on the Embedded System and IoT**
  National Skill Development Corporation | Issued Jan 2020

SKILLS SUMMARY

Data Analyst and Data Scientist Skills: Machine Learning, Scikit-learn, Statistical Modeling, Predictive Analytics, Data Mining, Data Wrangling, Data Transformation
Languages: SQL, Python, Spark, Scala, SAS, C++
Visualization: Tableau, Power BI, ArcGIS
Databases: Oracle, PostgreSQL, MongoDB, Snowflake, Cassandra, SQL Server
Big Data: HDFS, Hive, Airflow, PySpark, Kafka
Cloud Technologies: AWS (S3, EC2, Athena, Crawler), Azure (Data Factory, Data Lakes, Azure DevOps)
ETL and Data Warehousing: SSIS, Data Marts, Business Intelligence Tools: Git, GitHub, Terraform, Jenkins, Tasktop
Methodologies: Agile, SDLC
Technical Skills: MS Excel, MS Word, Data Migration, Data Cleansing
Healthcare Analytics: Medicare, Medicaid Claims Analysis
Soft Skills: Data Accuracy, Problem-Solving, Cross-Functional Collaboration

EDUCATION

- **Master of Science in Computer Science**
  Texas A&M University | December 2021
- **Bachelor of Engineering**
  Sri Indu College of Engineering & Technology | June 2019

PUBLICATIONS

- **"Propagator Rooting Method Direction of Arrival Estimation Based on Real Data"**
  2021 IEEE Military Communications Conference (MILCOM) | December 30, 2021
  Presented a computationally efficient DOA estimation method utilizing the Propagator Method and Root-MUSIC algorithm, significantly reducing computational complexity by transforming data matrices into real-valued forms.
- **"Capon Root-MUSIC-like Direction of Arrival Estimation Based on Real Data"**
  VTC2021 | September 28, 2021
  Introduced a hybrid approach for DOA estimation combining Capon and Root-MUSIC methods, achieving a 75% reduction in complexity for matrix operations with coherent narrowband signals.