

KEERTHI KRISHNA AIYAPPAN

☎ 315-374-9628 ✉ keerthikrishna.official@gmail.com  [linkedin.com/in/keerthi-krishnaa](https://www.linkedin.com/in/keerthi-krishnaa)

EXPERIENCE

GSK (ViiV Healthcare) - via Atrium

Jun 2024 – Present

Data Strategy and Operations Associate

Durham, NC, USA

- Deployed a Hypothesis Testing algorithm using R in an Azure Databricks environment, measuring the impact of a parallel data strategy on patient data stewardship on 3+ metrics, validating a 5% reduction in data duplicates
- Automated the generation of 150+ slides of HIV claims data using Plotly, reducing manual slide creation time by 80%
- Analyzed weekly sales data for two products, using DAX in Power BI, Window Functions, Views and CTEs in SQL within Azure Databricks, delivering trend and forecast reports to 5+ internal business partners
- Led UAT testing, data validation, and error/bug documentation, contributing to future dashboard enhancement development and ensuring platform migration of 400+ users from an internal app to Power BI
- Collaborated as a team of two to build an anomaly detection and data monitoring dashboard using Power BI, implementing logics to identify irregular data changes and improve data quality

NA Sportz Interactive

Nov 2021 – Jun 2023

Data Science and Analytics Associate

Mumbai, MH, India

- Produced over 300 analytical reports using SQL and MS Excel, assisting broadcasters and experts with insights on Indian Super League teams and players
- Developed a multivariate A/B Testing model to optimize the MatchCentre icon placement on the home page of a client website using SciPy and Statsmodels, resulting in a 9% increase in page clicks
- Identified and automated 25+ KPIs to track soccer player and team performance metrics, leveraging tidyverse and ggplot in R and CTEs, Window Functions, and Joins in SQL, delivering daily reports to clients and senior management
- Streamlined the creation of sport highlight packages Scipy Optimizers in Python, cutting video editing time by 80%
- Devised initial ratings for 200+ players for the official UEFA Champions League fantasy game using SciPy, pandas and NumPy, eliminating errors caused by human biases

BilledRight Healthcare Solutions

Jun 2020 – May 2021

Data Analyst

Chennai, TN, India

- Crafted over 10 dashboards in Tableau identifying 10+ KPIs, enhancing the delivery time of weekly reporting processes
- Managed the Extract, Transform and Load of 10+ business processes using Glue and Athena within AWS, transforming business processes into data tables for reporting and dashboards
- Handled 100+ ad-hoc data queries, utilizing Window Functions and CTEs in SQL, ensuring 100% accuracy in responses
- Developed 5+ reports in MS Excel utilizing Pivot Tables and VLOOKUPS, enhancing client-level channel data visibility

PROJECTS

Data Warehousing Solution for NY Metro Authority | SQL, Snowflake, dbt, Power BI

Sep 2024 – May 2024

- Implemented a data warehousing solution covering 5 business processes for the NY Metro Authority, adopting Enterprise Bus Architecture techniques, using Snowflake for data storage, dbt for ETL processes, and Power BI for reporting
- Authored end-to-end documentation for the data warehousing solution, ensuring data quality with 10+ automated tests in dbt and enabling scalability for future data workflows

What Drives Shipping and Transport of Pharmaceuticals Drugs? | Scipy, Re, Statsmodels

Jan 2024 – May 2024

- Analyzed 10,000+ HIV drug shipment records from the USAID data portal leveraging SciPy and Statsmodels in Python using S3 and Sagemaker in AWS, identifying global shipping cost drivers, devising a solution to cut shipping costs by 2%

EDUCATION

Syracuse University

Aug 2023 – May 2025

Master of Science in Applied Data Science

Syracuse, NY, USA

Bharathiar University

Aug 2016 – May 2019

Bachelor of Science in Visual Communication and Electronic Media

Coimbatore, TN, India

SKILLS

Programming Languages and Data Visualization: Python, R, SQL, Spark, Tableau, MS Excel

Cloud and ETL: Snowflake, Docker, dbt, Hadoop, Airflow, Kubernetes, BigQuery, GCP, Kafka, Minio, Cassandra, Hive

Libraries : Scikit-Learn, Pandas, Numpy, Matplotlib, Seaborn, Re, Plotly, Streamlit, Statsmodels, SciPy, Shiny, Caret

Modeling: Regression, Classification, Anomaly Detection, Forecasting, ANOVA, Hypothesis Testing, A/B Testing

AWS Services: Lambda, Athena, S3, Sagemaker, Redshift, QuickSight, CloudWatch, Glue, RDS, EC2

Azure Services: Databricks, Data Factory, Data Lake, Synapse, Power BI

LEADERSHIP

- Served as the Lead Data Consultant for the 2022-23 season of the Indian Super League, acting as the data science liaison between broadcasters, organizers and my employers