Nikhil Thota
San Jose, CA 95126 • nikhil.thota@sjsu.edu • + 1 (669)-677-1551 • LinkedIn • Github • Medium

---

Objective

Data Analyst with 4 years of experience and a Master's degree in Data Analytics from SJSU, specializing in Report generation, ETL/ELT pipelines, data migration and growth mindset. Built 20+ pipelines using SQL, Python, business intelligence (BI) tools, with a real touch of ML applications. Looking out for Summer Internships.

EDUCATION

- **MASTER OF SCIENCE IN DATA ANALYTICS, SAN JOSE STATE UNIVERSITY**   **San Jose, CA**
  Jan 2023 - Expected. December 2024
- **B.TECH IN COMPUTER SCIENCE AND ENGINEERING, IIIT VADODARA**   **Gandhinagar, India**
  Aug 2015- Sept 2019

TECHNICAL SKILLS

- **Databases**: PostGreSQL, RDBMS, Neo4j, Redshift, BigQuery, MySQL
- **Programming languages:** R, Python, SQL, DAX, Hive SQL, Java (Object Oriented Programing Programming), PySpark
- **Data Visualization Tools:** Data Studio, Matlplotlib, Plotly, Power BI, Tableau, MicroStrategy, Grow (BI), Looker
- **Workflow Tools:** Airflow DAGS, DBT Cloud (ETL/ELT), Excel VBA
- **Cloud:** AWS S3, Redshift, DynamoDB, BigQuery, Google Cloud Buckets, Snowflake, Zeppelin, Databricks
- **Mathematics:** Statistics, Probability, Regression, Optimizations, Eigenvalues, eigenvectors, p-value, hypothesis testing
- **Machine Learning:** Keras, Tensorflow, Pandas, Numpy, Seaborn, scikit-learn, Requests, Data Analysis, Data Science

EXPERIENCE

Data Analyst | PHARMEASY   **DEC 2021 - Jan 2023**

- Orchestrated a **4%** upswing in revenue by optimizing product listings accuracy, utilizing SQL,NER (NLP), CNNs for data driven rules on Product Catalog and implementing automation through Apache Airflow and SageMaker.
- Pioneered a transformative shift through cross-functional team collaborating, resulting in a remarkable **40%** reduction in ad-hocs by building data pipelines, models, and dashboards with proficiency in Airflow, Presto, Redshift, and MicroStrategy.
- Delivered an impressive **8%** boost in user conversion rates by analyzing real-time data seamlessly using Connectors (API), BigQuery, Google Analytics, DAX, and crafted compelling Power BI visualizations.
- Developed **8+** MicroStrategy Dashboards, & emailers in collaboration with demand business teams to aid in transformation of the firm's supply chain model from state based model to zone based model, thereby improving the demand by **15%.**

Data Business Analyst | Saras Analytics   **DEC 2020 - Dec 2021**

- Executed Agile SDLC while Optimizing marketing strategies resulting in **10%** sales uplift, by analyzing customer segments and leveraging SQL, DBT, and Looker for precision targeting across Google ads, Meta Ads, and other omni-channel platforms.
- Optimized & Refined BigData data models through Airflow, utilizing tools like DBT Cloud, S3, Python and SQL for ETL/ELT processes, enhancing efficiency and scalability across hive/presto systems reducing costs by **23%.**
- Built Data Lakes on BigQuery, Snowflake, and Redshift, reducing data redundancy by an impressive **20%** and elevating data transparency and accuracy from **40%** to **> 95%** by employing data cleaning & transformation.
- Spearheaded the development of P/L dashboards for Amazon Sellers, resulting in a **20%** cost reduction by integrating factors such as budget, forecasted sales, demand and COGS, providing key business insights using Compelling Tablea

Business Analyst | Capillary Technologies Pvt. Ltd   **JAN 2019 - Dec 2020**

- Delivered a **4x** ROI as an intern by orchestrating data-driven campaign plans, collaborating cross-functionally with teams such as brand marketing, campaign delivery, finance, and legal, to execute successful campaigns in the retail industry.
- Mitigated fraud transactions by **17%** in offline retail environments through the implementation of real-time, data-driven rules using SQL, Association rule mining, K-means clustering and predictive modeling.
- Drove a **20%** increase in CRM growth by deploying advanced models incorporating RFM, Demand Forecasting, and Market Basket Analysis for customer retention, and cross-sell product analysis through real time campaigns using Ads.
- Retained **6+** clients by computing new metrics like CSI, NPS by leveraging **20+** SQL logics by incorporating inputs from Economics, Finance stakeholders and clients feedback to understand the fundamentals of the B2B and B2C systems.

PROJECTS

- **BART, VTA DATA MODELLING:** Employed a suite of tools including MySQL Workbench, Python, DBT, Neo4j, MongoDB, Mongo Atlas, BigQuery, GCP, AWS Glue, Kubernetes, and Flask in a comprehensive data modeling project for SJSU.   **Jan 2023 - May 2023**
- **Urban Audio Classification:** Lead a team of four in a research project focusing on comparative analysis of urban audio classifiers. Aimed to minimize use of labeled audio data. Vision Transformers (ViT) vs CNN variants, for SOTA results.   **Sept 2023 - Dec 2023**
- **Polypharmacy effects:** Developed a GCN using Node2Vec features for dr-dr interactions, predicting side-effects.   **Sept 2023 - Dec 2023**