

# VATSAL JIGAR KADAKIA

📍 Bayonne, NJ (Open to relocation) 📞 (813) 843-9863 ✉ vatsalkadokia20@gmail.com 🌐 Vatsal Kadokia 📁 Portfolio

## Education

**Stevens Institute of Technology - Charles V Schaefer School of Eng, Hoboken** Aug 2022 - May 2024

*Masters of Science in Data Science*

GPA: 3.6/4.0

**University of Mumbai, India**

Aug 2018 - May 2022

*Bachelor of Engineering in Computer Engineering*

GPA: 3.8/4.0

## Experience

**Data Science Intern, Changing The Present**

Aug 2024 - Present

- Developed ETL pipelines using SQL (AWS Redshift) for data pipelines and QuickSight for visualization, enabling daily tracking of metrics across various business domains, resulting in 3 FTE saves and reducing manual intervention by 25%
- Collaborated on data transformation project for 3800+ student organizations using Google Sheets. Extracted, cleaned and analyzed data using Matrixify(Shopify app), VLOOKUP and Pivot Tables leading to better data management
- Collaborated with data science team to develop ML models for student data by providing clean and processed data, saving weekly 15+ hours of manual data cleaning work, accelerating forecast accuracy and decision-making by 25%
- Improvised ARIMA and GARCH (Time Series Analysis) models using R to forecast student's spending power on private courses, significantly improving accuracy score by 46%, achieved p-value<0.05 by performing EDA
- Performed statistical analysis on student survey data using R, aiding 2,750 people and saving \$1.5M in private courses

**Data Analysis IT Intern, Minerals Technologies Inc**

Jun 2023 - Dec 2023

- Established predictive ML models, Pandas and NumPy for data processing and compared accuracy of Regression and Classification models; analyzed real-time customer sentiment, increasing actionable insight by 40%
- Engineered geospatial data visualization in Oracle Analytics Cloud (OAC) to visualize Pet Care product distribution of \$629 millions, tool is now utilized by 4+ analysts for strategic resource allocation streamlining product management
- Revamped data project with OAC's ML module, slashing duplicate data by 30%, facilitating dashboard creation
- Collaborated with cross-functional teams to migrate client data from Oracle EBS to Oracle Analytics Cloud(OAC)
- Designed and published comprehensive sales dashboard using Tableau resulting in 13% increase in sales team efficiency
- Automated incident and request handling of company's ServiceNow Chatbot by enhancing pre-existing JavaScripts in Natural Language Unit(NLU), reducing employee manual work by 35% and saving company's \$7,200 product fees
- Presented business reports using SQL and Excel VBA on 11 chatbot KPIs, improving employee engagement by 28%

**Data Analyst, Edvicer Private Limited**

Jan 2022 - May 2022

- Created different KPI's using Table calculations, Level-of-Detail (LOD) expressions, Actions, Sets, Parameters, ad hoc reports, calculated key figures manipulating strings, arithmetic calculations, aggregation options, and logical functions
- Visualized data by creating Charts and Graphs (bar graphs, line charts, pie charts, Tree maps, Bubble Charts, Waterfall Charts,Bump Charts, Funnel Charts, and Bollinger Charts) based on client's need
- Used ETL tools to extract data into Tableau Desktop and data wrangling using Alteryx in data preparation phase

## Project Experience

**For the Love of Alcohol! | Python, Unsupervised ML**

Jan 2024 - May 2024

- Evaluated 3,500 alcohol transactions using the Apriori algorithm to identify product relationships, offering insights for improved placement such as linking gift products and bundling strategies to enhance customer experience

**Weather API Data Streaming & Analysis | PySpark, PyML, PySQL**

Aug 2023 - Dec 2023

- Architected a Weather Analytics ETL pipeline using PySpark Streaming, automating the processing of around 25 hourly weather parameters for 5+ locations, reducing manual intervention for comparative analytics by 70%

**"You Are What You Eat" Customer Segmentation | Python, A/B Testing, Regression**

Jan 2023 - May 2023

- Constructed customer segmentation with 3 clusters(73.6%, 14.6%,11.8%) using WCSS based on dietary preferences within customer database. Applied K-means clustering aiding business to manage inventory, product promotions

**Product Sentiment Analysis of iPhone 15 | Python, NLP, Web Mining, API**

Aug 2022 - Dec 2022

- Analyzed sentiments from 20K tweets with VADER/TextBlob, scoring 0.09/0.12, used LDA for topic modeling (perplexity 1066), used Twitter API to create Semantic Network of +ve and -ve words boosting marketing insights

## Skills

**Programming Languages:** Python, R, SQL

**Data Visualization:** Tableau, Power BI, Ms Excel, Matplotlib

**Databases:** MySQL, PostgreSQL, SQL Server, NoSQL, MongoDB, Hadoop, Hive, Databricks, Snowflake

**Cloud Services:** Amazon Web Services (AWS), Google Cloud Platform (GCP)

**Methodologies:** Agile, Jira, Confluence