

VATSAL JIGAR KADAKIA - DATA ANALYST

✉ vatsalkadokia20@gmail.com [in LinkedIn](#) [Portfolio](#)

Experience

Changing The Present, NY

Jul 2024 – Present

Data Science Intern

- Established a robust **ETL workflow** that automated data extraction, transformation, and loading processes, saving the data team over 15 hours per week previously spent on manual data handling tasks.
- Improvise **Time Series Analysis** model using **R** to forecast student organizations participation rate, significantly improved the accuracy score by 38%.
- Collaborated with a colleague in Data Transformation project of 3800+ student organizations using **Google sheets**. Extracted data using **Matrixify** app(Shopify), cleaned and analyzed data using **VLOOKUP** and **Pivot tables**.
- Designed 3 **Excel dashboards** after analyzing data using **SQL** queries and shared with business user.

Minerals Technologies Inc, Bethlehem, PA

Jun 2023 – Dec 2023

IT Intern - MTI Oracle

- Integrated **Oracle Analytics Cloud** with **Oracle EBS** and APIs to exchange data and automate data transfer processes, ensuring seamless integration and interoperability between different systems.
- Engineered an advance geospatial model in Oracle Analytics Cloud(OAC) to visualize and assess mineral product distribution, tool is now utilized by 6+ analysts for strategic resource allocation decisions streamlining **product management**.
- Revamped data project with OAC's **Machine Learning module**, slashing duplicate data by 40% and facilitating timely report and **dashboard creation**.
- Automated incident and request handling of company's **ServiceNow Chatbot**, slashing employee manual work by 35% and improving ServiceNow accessibility by 50%.
- Designed and published 2 dashboards with top-notch security by leveraging **Tableau User filter** and other complex Row level security methods.

Education

Stevens Institute of Technology, New Jersey

Aug 2022 – May 2024

Masters of Science in Data Science

Coursework: Time Series Analysis, Deep Learning, Applied Machine Learning, Statistics Methods, Optimization Methods

University of Mumbai, India

Aug 2018 – May 2022

Bachelor of Engineering in Computer Engineering

Coursework: DBMS, Data Structures, Data Analytics, Software Engineering, Object-Oriented Programming

Technical Skills

Programming language: Python, R, Java Script, HTML, CSS, Bootstrap

IDE's: Visual Studio Code, Jupyter Notebook, Spyder, R Studio

Visualization tools: Tableau Desktop, Power BI, Looker, Advanced Excel(Pivot Tables, VLOOKUP)

Database: MySQL, SQL, SQL Workbench/J, DBeaver

Cloud Technologies: Amazon Web Services(Lambda, S3, EC2, DynamoDB, SageMaker)

Tracking Tools/Methodologies: Jira, ServiceNow, Agile, Waterfall, SDLC, Docker

Competencies: Data Mining, Forecasting, Critical Thinking, Problem Solving, Communication Skills, Presentation Skills

Projects

Pyspark weather API data streaming and analysis | *Numpy, Pandas, SQL*

Prof.U Prasad

- Engineered a **PySpark** streaming application integrating SQL and machine learning algorithms to dynamically fetch weather data from custom-selected locations and dates via a **weather API**(Processed over 25 fields).
- Enhanced functionality with PyML operations to compare weather data from two locations simultaneously, aiding event scheduling teams and skincare brands in strategic decision-making.

Grocery Store Optimization | *Matplotlib, A/B testing, data modeling*

Prof.Hadi Safari

- Conducted **Z-test, Chi-Square** test to analyze significant difference between sign up rate for cheaper expensive mailer of delivery club. Performed **ANOVA** to find p-values of all columns(response variable as Sales column).
- Obtained 96% loyalty score accuracy, 93% behavior classification with Regression algorithms, significantly transforming **store layout and marketing**. Applied causal impact to determine relationship between customer loyalty and delivery campaign.

Product Sentiment Analysis of iPhone 15 | *Scikit-learn, NLP, predictive models*

Prof. Jingyi Sun

- Analyzed sentiments from 20K tweets with **VADER/TextBlob**, scoring 0.09/0.12, and used LDA for topic modeling (perplexity 1066), boosting marketing insights. Utilized **Twitter API** to scrape tweets and created a **Semantic Network** to identify the most prominent positive & negative words associated with product.

Publications and Courses

Tableau Desktop Specialist : Data Visualization Certificate

Automated Healthcare System using AI based Chatbot: Developed a project focused on addressing health issues through one-on-one interactions between patients and a bot. Key functionalities included disease prediction, medicine prescription, and appointment scheduling with clinics and hospitals.

Springer, Singapore · Mar 22, 2023

De Gruyter, Berlin, Germany · Feb 20, 2023