SARISH CHACHARE

Data Analyst

+91 - 7517748084
chacharesarish@gmail.com
github.com/chacharesarish
linkedin.com/in/sarish-chachare

PROFILE

Data Analyst with a B.Tech in Computer Engineering and hands-on experience in Python, SQL, Excel, Power BI, and Tableau. Skilled in data cleaning, exploratory analysis, visualization, and predictive modeling, with a track record of delivering actionable insights through end-to-end analytics projects. Demonstrated ability to analyze large datasets, identify trends, and support business decisions. Currently enhancing expertise through the Newton School Data Science Professional Certification Program while completing multiple practical projects.

EDUCATION

B.Tech, Computer Engineering
Vishwakarma University, Pune
2020 – 2024
CGPA: 7.61 / 10

TECHNICAL SKILLS

- Excel
- SQL (BigQuery)
- Tableau
- PowerBI
- Python
- GitHub

EXPERTISE

- Dashboard Development
- Data Analysis & Modeling
- Surfacing Actionable Insights
- Data-Driven Decision Support
- Predictive Modelling
- Reporting & Storytelling
- GitHub

EXPERIENCE

Credit Risk Modeling
Created a machine learning–driven credit risk prediction system using Python and Streamlit to evaluate borrower default probability. Performed feature engineering and resampling on an imbalanced dataset (~10% defaults). Trained Logistic Regression, Random Forest, and XGBoost models; optimized XGBoost with Optuna, achieving AUC 0.98. Applied SHAP and LIME for model interpretability and built an interactive Streamlit dashboard for real-time risk visualization and decision support.

Vendor Performance Analysis
Built a complete analytics pipeline using SQL for ETL, Python (Pandas, Matplotlib, Seaborn) for data processing and statistical analysis, and Power BI for interactive vendor performance dashboards. Cleaned and merged multiple retail sales and inventory datasets to evaluate vendor contributions, inventory turnover, and cost efficiencies. Identified that top 10 vendors accounted for ~65% of purchases and uncovered 72% cost savings from bulk purchasing, guiding strategic vendor diversification and inventory optimization decisions.

Post-Pandemic E-commerce Sales Analysis
Analyzed 100,000+ order and customer records from 2019–2022 to uncover e-commerce trends before, during, and after the pandemic. Used SQL for data extraction and Python for cleaning, analysis, and visualization, revealing that total sales grew 162% in 2020, then declined 46% in 2022 relative to the previous year. Identified changes in average order value (AOV), order volume, regional sales distribution, and product performance, providing actionable insights for targeted marketing and inventory strategy. Maintained reproducible code and analysis on GitHub.

TechSphere E-commerce Analytics
Analyzed ~108,000 transactions totaling $28M+ revenue using SQL and Excel for data extraction, cleaning, and sales analysis. Identified trends across regions, product categories, and quarters, revealing that North America contributed ~52% of sales and flagship products like 27'' 4K monitors and AirPods drove the highest revenue. Insights guided strategic marketing, inventory optimization, and regional prioritization. Maintained clean documentation and reproducible reports on GitHub.

LA Web Traffic Analysis
Analyzed 8M+ web traffic records for lacity.gov using SQL to extract and transform sessions, bounce rate, device, and browser data from 2014–2019. Built Tableau dashboards to visualize usage patterns, revealing that desktop users accounted for ~66% of sessions and mobile had the highest bounce rate. Identified weekday traffic peaks and critical spikes during events like Dec 2017 wildfires, and provided UX and content strategy recommendations to improve engagement and communication during high-demand periods.

Medicare Part D Analysis
Analyzed 1M+ Medicare Part D prescription records using SQL to uncover cost and utilization patterns in national claims data. Identified that just 10 drugs accounted for ~30% of total spending, revealing extreme cost concentration and price-volume imbalances (e.g., premium drugs vs cheaper alternatives). Quantified a potential $543M+ annual savings opportunity by modeling targeted utilization adjustments and generic substitutions. Built data-driven recommendations for utilization management and prescriber-behavior interventions, with reproducible SQL scripts and documentation on GitHub.

Data Science Professional Certification (Ongoing) – Newton School
Comprehensive training in Python, SQL, Tableau, Power BI, Machine Learning, and Data Analytics workflows.