SAGARIKA SARDESAI

| LinkedIn: /sagarikasardesai | Tableau: /sagarikasardesai | GitHub: /sagarikasardesai | sagarikasardesai13@gmail.com | 6199537007 | Open to Relocation | Visa Sponsorship Required

EDUCATION

MS, Data Science — **San Diego, CA**
UC San Diego, CGPA 3.89 — *Sep 2022 - Apr 2024*
Btech, Computer Science and Engineering — **Vellore, India**
Vellore Institute of Technology, CGPA 8.61 — *Jul 2016 - Jun 2020*

TECHNICAL SKILLS

Languages: **SQL, Python**, R, Cypher, PySpark, SparkSQL
Libraries: **Pandas, NumPy, Matplotlib, Seaborn, SciPy, Statsmodels, Scikit-learn**
Data Vis and BI Tools: **Tableau, Qlik Sense, Advanced Excel (Advanced Formulas, XLOOKUP, VLOOKUP, Pivot Tables, Stat tools, Macros, Charts)**
Cloud Technologies: AWS (Redshift, Glue, S3)
Statistical Methods: **ANOVA, Hypothesis Testing (A/B testing), T-Statistics, Regression Analysis**
ML Algorithms: **Regression (Linear & Logistic), Classification, Clustering (KMeans & Hierarchal), XGBoost, Random Forest, Decision Trees, SVM, PCA, Neural Networks**
Other: Jupyter Notebook, Google Colab, Anaconda, **HP ALM, Git, Jira**, Confluence

EXPERIENCE

Data Science Research Assistant — **San Diego, CA**
UC San Diego — *May 2023 - Nov 2023*

- Derived **valuable insights into consumer attitudes and decision-making** regarding End-of-life choices.
- Ensured **data accuracy by 5% through cleaning, exploration, and visualization** in the Data Quality Report, revealing significant patterns and trends.
- Utilized **exploratory data analysis, statistical analysis, and NLP techniques such as VADER and RoBERTa to measure sentiments** and derive meaningful conclusions about consumer behavior.

Business Analyst — **Pune, India**
Credit Suisse — *May 2021 - Jun 2022*

- **Identified root causes, patterns, and trends** by conducting 300+ Test Trades, resulting in a 70% decrease in defects and unveiling new feature development areas.
- **Designed Key Performance Indicators (KPIs)** by analyzing over 80,000 production and test data records using **advanced SQL queries** sourced from company databases and software.
- Created 300+ functional and user acceptance test (UAT) scenarios for **B2B applications utilizing KPIs and Advanced Excel tools**.
- **Performed ad-hoc analysis** of test results data, **delivering actionable insights to cross-functional teams**.
- Led defect management processes to streamline the trade lifecycle.

PERSONAL PROJECTS

Predictive Analysis for Detecting Credit Card Transaction Fraud

- Utilized **supervised machine learning** to detect credit card fraud, resulting in estimated savings of 21M USD with a 3% False Discovery Rate.
- Analyzed 90,000+ credit card transactions, conducting **data exploration, cleaning, and visualization to uncover insights**.
- **Engineered over 1000 variables** from existing dataset fields for comprehensive analysis.
- Employed the Kolmogorov-Smirnov test to **feature select 20-25 relevant variables reducing training time by 30%**, improving model accuracy and fraud detection capabilities.

Covid19 Data exploration and Visualization

- Utilized **SQL to analyze** the OWID Covid-19 Dataset of 85,000+ rows, **identifying patterns in global infection and mortality counts, country-specific statistics, and regional peak counts.**
- **Forecasted infection numbers** for highly affected countries and **presented insights visually with Tableau** to aid decision-making.

NY Property Fraud Detection (Unsupervised Anomaly Detection)

- Detected potential property tax fraud in 10,000+ NY property records using **statistical analysis and unsupervised ML**.
- **Uncovered insights through data exploration, cleaning, and visualization.**
- **Feature engineered 59 z-scaled variables** using statistical analysis to enhance fraud detection accuracy.
- Employed heatmaps to **visualize variables influencing high fraud scores**, gaining deeper insights into fraud patterns.

Investment Scope of SBIR Awarded Companies

- Conducted in-depth **analysis of the SBIR Awarded Companies containing 165,000+ records using Python and SQL**, revealing patterns in industry sectors and year-wise award amounts across US agencies and branches.
- Improved **data accuracy by 20%** by **cleaning methods including data type adjustments, keyword extraction, and imputation.**
- **Established relationships** between US departments and agencies using **Neo4j graphs**, enhancing investment analysis context.
- **Web scraped recent news** about relevant companies, stored in **MongoDB** for further reference.