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: **Supervised Learning, Unsupervised Learning, Ensemble Learning, Deep Learning**
Other: Jupyter Notebook, Google Colab, Anaconda, **HP ALM, Git, Jira**, Confluence

EXPERIENCE

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

- **Derived valuable insights about consumer judgments, preferences, and decisions** regarding End-of-life decisions.
- Ensured data accuracy by **cleaning, exploring, and visualizing datasets in the Data Quality Report, identifying significant patterns and trends.**
- **Applied exploratory data analysis, statistical analysis, and NLP, including VADER and RoBERTa, to measure sentiments and draw meaningful conclusions about the consumers.**

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

- **Identified root causes, patterns, and trends** by booking 300+ Test Trades, leading to a 70% reduction in defects and uncovering new feature development areas.
- **Designed Key Performance Indicators (KPIs) by the data-driven approach of analyzing over 80,000 production and test data records, using advanced SQL queries**, extracted from company-related databases and software.
- **Created 300+ functional and user acceptance test scenarios, with the use of KPIs and Advanced Excel (Data cleaning, Pivot tables, etc)**, for B2B use cases.
- **Conducted ad-hoc analysis of test results data**, reporting actionable insights effectively to cross-functional teams.
- Spearheaded defect raising, tracking, and resolution to streamline the trade lifecycle.

PERSONAL PROJECTS

Predictive Analysis for Detecting Credit Card Transaction Fraud

- **Detected fraud in credit card transactions using supervised machine learning models, resulting in estimated savings of 21M USD with a 3% False Discovery Rate (FDR)**.
- **Analyzed 90,000+ credit card transaction records, exploring, cleaning, and visualizing data** in the Data Quality Report to identify root causes, patterns, and trends, uncovering valuable insights.
- **Engineered 1000+ variables** from existing dataset fields, ensuring comprehensive analysis.
- Selected 20-25 pertinent variables using the Kolmogorov–Smirnov (KS) test as a Filter and Forward Selection Wrapper, enhancing model accuracy and fraud detection capabilities.

Covid19 Data exploration and Visualization

- **Analyzed the OWID Covid-19 Dataset comprising 85,000+ rows using SQL, identifying patterns and trends related to global infection and mortality counts, country-specific statistics, and peak counts in various regions.**
- **Forecasted infection counts** for countries with the highest infection percentages, for informed decision making.
- Effectively **visualized extracted insights using Tableau** for comprehensive understanding and decision-making.

NY Property Fraud Detection (Unsupervised Anomaly Detection)

- **Identified potential tax fraud instances in 10,000+ NY property records**, by applying statistical analysis and unsupervised machine learning techniques.
- **Discovered valuable insights by exploring, cleaning, and visualizing the data distribution in the Data Quality Report for the NY properties dataset.**
- Utilized statistical analysis to engineer 59 z-scaled variables, identifying principal components for comprehensive analysis and improving fraud detection accuracy.
- **Utilized heatmaps to visualize key variables contributing to high fraud scores, for a deeper understanding of fraud patterns.**