

David Jian

San Jose, CA | (408) - 816 - 3728 | davidjian00@gmail.com | linkedin.com/in/davidjian00 | github.com/LunarkoiFish

Education

- **Languages:** SQL, Python (Pandas, NumPy, Matplotlib), HTML VBA, HTML/CSS
- **Tools:** Excel, Tableau, Visual Studio Code, MongoDB, Jupyter Notebook, MongoDB, Github, AWS, Big Data
- **Skills:** Spreadsheets, Business Intelligence, Data cleaning, API extraction, Web scraping, ETL, Data visualization, Databases (SQL & noSQL), Statistical analysis, Machine learning (regression, clustering), Scikit-learn, Communication

Projects

Diabetes Risk Predictor Model

- Developed a Logistic Regression model to predict diabetes risk based on patient health indicators before optimizing the model through hyperparameter tuning (C-values for regularization) and feature scaling with StandardScaler.
 - Built an interactive Flask web application for real-time risk assessment.
- Visualized key insights in Tableau, highlighting significant diabetes risk factors.

Sales Performance Analysis

- Developed interactive Tableau dashboards to visualize sales trends, KPI performance, and product subcategory comparisons, enabling stakeholders to monitor year-over-year metrics.
- Performed end-to-end data preparation including cleaning, standardizing, and transforming sales datasets using Python, Excel, and SQL for accurate visualization and analysis.
- Implemented dynamic filtering and drill-down capabilities for product and location, allowing users to explore historical sales data and derive actionable insights.

CSM Hospital General Information Analysis

- Cleaned and structured hospital data using Python and PostgreSQL, writing SQL queries to analyze ratings, services, and ownership types.
- Designed and published an interactive Tableau dashboard to visualize service availability and rating distributions across U.S. regions.

Telco Customer Churn Prediction Project

- Developed a predictive model to identify high-risk churn customers for a telecom company.
- Conducted exploratory data analysis and feature engineering on customer demographic and usage data.
- Trained logistic regression and tree-based models to optimize performance, achieving high accuracy and AUC scores.

Work Experience

Medical Records Assistant

March 2024 - Present

Canyon Springs Post Acute – Skilled Nursing Facility (SNF)

- Maintained and updated patient records using PointClickCare (PCC), ensuring accuracy, compliance, and confidentiality in alignment with HIPAA standards.
- Utilized Excel to create and manage spreadsheets for tracking patient admissions, discharges, and documentation workflows.
- Prepared reports and logs in Microsoft Word to support compliance audits and cross-departmental communication.

Education

UCI Data Analyst Bootcamp

September 2024 - March 2025

Certificate in Data Analytics

- Training in Excel, VBA, Python, SQL/NoSQL, data collection, ETL, and advanced analytics, including machine learning, neural networks, and big data. Gained hands-on experience in data visualization (Tableau, Python), interactive dashboards, and applied projects covering data ethics and real-world business cases.

San Jose State University

2018 - 2023

Bachelors of Science Chemistry, Concentration in Biochemistry

San Jose, CA

- Coursework: Data Analysis & Interpretation, Quantitative & Qualitative Analysis, Experimental Design, Data Collection & Cleaning, ChemDraw