Johnny Yu

Johnnyyu14@gmail.com | 415-312-2772 | johnnyyu14 | johnnyyu14.github.io

Education

University of California Santa Barbara — Santa Barbara, CA
Bachelor of Science in Statistics and Data Science — *September 2018 – June 2022*

Awards: Dean's Honors L&S, Awarded $5000 to participate in UCSB's NSF-funded Central Coast Data Science Fellowship for the 2021-2022 academic year

Skills

Programming Languages: Python, SQL, R, SAS, Excel

Visualizations/Libraries/Analysis: Tableau, Power BI, NumPy, Pandas, DaX

Work Experience

Live Proxies — Remote
Business Analyst — *March 2020 – Present*

- Conducted data analysis using SQL, Power BI, and Excel to support business decisions and strategies, including trend identification, revenue forecasting, and marketing campaign optimization, resulting in a 25% increase in sales
- Optimized data warehousing solutions for high-volume transaction processing (OLTP) and analytics (OLAP) to increase efficiency and accuracy in reporting and forecasting, while also conducting data quality checks and reconciliations, reducing time spent on data cleaning by 50%
- Developed and maintained data models, databases, and warehouses, supporting data-driven decision making, governance, security, and compliance, resulting in improved business performance metrics
- Collaborated with stakeholders and IT teams to prioritize data requirements and implement restructuring and optimization strategies, resulting in increased margins and revenue streams
- Utilized ETL processes and tools such as SSIS to extract, transform, and load data, and created visualizations to communicate findings and drive new business opportunities, leading to increased revenue and market share

ENVENT Labs — Santa Barbara, CA
Data Analyst — *September 2021 – June 2022*

- Led analysis of 2500+ complex data sets using advanced statistical techniques and machine learning to uncover crucial demographic predictors of climate opinions, achieving a 10% increase in operational efficiency
- Spearheaded the implementation of highly efficient SQL query processes and cutting-edge supervised/unsupervised ML models that accurately predicted individual climate scores with an impressive 85% accuracy rate
- Collaborated with cross-functional teams to efficiently collect, clean, and model large amounts of data from various sources, resulting in successful presentations of complex findings to technical and non-technical stakeholders
- Communicated data-driven insights with exceptional presentation skills, resulting in a significant 20% increase in project funding and cementing the company's reputation as a pioneering leader in the cutting-edge field of climate research

Projects

Forecasting Short-Term Future COVID-19 Cases — *September 2021 - December 2021*

- Developed and implemented time-series forecasting models to predict daily COVID cases, utilizing estimated percentage of COVID-related outpatient doctor visits as features
- Compared and evaluated the performance of Decision Tree Regressor and SVR models, identifying limitations and proposing potential solutions for improving forecasting accuracy
- Communicated findings and insights to stakeholders, providing recommendations for utilizing the models in decision-making processes

2016 Presidential Election Prediction Model — *January 2021 - June 2021*

- Built and evaluated multiple predictive models, including Random Forest and Logistic Regression, to predict county-level election results
- Communicated results and limitations to a team of analysts, contributing to ongoing research on election forecasting
- Collaborated with team members to improve model performance, resulting in an increase in accuracy by 0.8%
- Conducted extensive exploratory data analysis to identify key demographic variables that impact election outcomes