JACOB LINAO

jacobnlinao@gmail.com
linkedin.com/in/jacobnicholasl
Portfolio: jacoblinao.netlify.app
714.353.6338

EDUCATION AND HONORS

University of California, Irvine
Master of Science – Business Analytics (MSBA), Concentration in Data Science   June 2022
- Merage Scholarship Recipient (Top 20%), GPA: 3.84/4.00
- Courses: Data Science Statistics, Adv. Machine Leaning, Adv. Predictive Analytics and Forecasting, Marketing Analytics, Big Data and Cloud Computing, Natural Language Processing, Customer and Social Analytics, Management Analytics

California State Polytechnic University, Pomona
Bachelor of Science -Information Systems & Business Intelligence (B.S.)   May 2021
- Director of Data for Digital Analytics Club, Head Teaching Assistant for College of Business, VP of Academics for Zeta Phi Rho

SKILLS

Programming: Python, SQL, PL/SQL, R, Git

Data Tools: Pyspark, AWS EMR + S3, SSMS, Oracle DB, PostgreSQL, Tableau, MS Office, Power BI, SPSS, WEKA, Alteryx

PROFESSIONAL EXPERIENCE

Graduate Data Scientist   Irvine, CA
The Guitar Center Company   Jan 2022 - Present
- Building models that predict high churn probability customers, segment customers for personalized transactions, and find propensity to target high ROI customers
- Communicated insights to VP of Data by devising plans to facilitate marketing campaigns and encourage trade-ins
- Raise business questions by developing and tracking KPIs to support digital platform and in-store experience

Data Scientist Intern   Diamond Bar, CA
Niagara Bottling   May 2020 - Aug 2021
- Led production of an **NLP** Random Forrest model in **Python** as a user facing web application to resolve an ITSM issue, reducing false positive rate by 20%, saving est. $75K/YR in resource allocation, and benefitting 500+ users nationwide
- Performed EDA in **Excel** (pivot tables, macros, vlookups) and created weekly **PowerBI** reports to spearhead weekly team meetings for 10 analysts over data driven suggestions, increasing operational efficiency by 80%
- Compiled 40+ **SQL** objects to achieve results such as automated data processing and 100% reduction in cycle time
- Designed ad-hoc **SQL** queries to eliminate data transformation errors for manufacturing lines and produced 10+ client reports

Data Analyst Intern   Burbank, CA
NBCUniversal   Sep 2020 - Dec 2020
- Led data migration in extracting and aggregating data and liaising insights to an international team of directors and stakeholders
- Extracted and cleaned over 10M+ rows of SAP records using **Python**, distilling down to 5 **KPIs** and designed 20+ **Tableau** visualizations to summarize operational trends and drill down analysis to stakeholders to help inform strategies
- Built department dashboard for 25+ users to monitor vendor trends by pulling and consolidating data from 3 **SQL** DB sources

PROJECTS AND RESEARCH

Data Science Research Lead (*bit.ly/3CBlD0L*)   Remote
Hawaii International Conference on System Sciences   May 2021 – Aug 2021
- Won Best Project - Created a TFIDF **Logistic Regression** model for segmenting customer sentiments and utilized **LSA** for topic modeling to find 3 positive/negative areas of business focus to combat the ill effects of Covid on LA restaurants

Data Analytics Research Assistant (*bit.ly/3Bykl5f*)   Pomona, CA
Center for Customer Insights and Digital Marketing   Jan 2020 - Mar 2020
- Led a team to the finals at GFK's data science hackathon by building a product and presenting insights to VP of Marketing
- Conducted a successful **A/B test** to measure school website efficacy, resulting in 120% traffic increase

Predicting Mental Health – *Class Project (bit.ly/3JK6LRE)*
- Used **Python** to implement KNN, Decision Trees, Naïve Bayes to understand predictors that affect mental health in the workplace
- Achieved highest recall score (86%) on Kaggle by using Recursive Feature Elimination and Grid Search optimization

Decreasing Florida Carbon Footprint – *Hackathon at IBM (proprietary)*
- Utilized time-series algorithms including **ARIMA** and Holt's Method to assess seasonality in household energy usage trends
- Built end-to-end data pipeline of data collection via web scraping, feature engineering, model building, and model ensembling