Cole Mak

mrcolemak@gmail.com | https://colemak.me | San Francisco, CA

EDUCATION

University of California - Santa Barbara   **Expected Grad: June 2025**
B.S. - Statistics and Data Science | B.A. - Film and Media Studies

Relevant Coursework: Probability & Statistics, Linear Regression Analysis, Data Structures & Algorithms, Statistical Machine Learning, Principles of Data Science in R and SQL, Statistical Design & Analysis, Real Analysis, Linear Algebra, Stochastic Processes

PROJECTS

Music Genre Classifier Application (Python, AWS, FastAPI, Docker, Heroku)   **Sep 2024**
- Built a song genre classification application using FastAPI with an ETL and model training pipeline, achieving 86% testing accuracy.
- Leveraged Docker to containerize and deploy API on AWS ECS for scalable public access.
- Created a Flask frontend user interface and deployed the application on Heroku for enhanced UI/UX.

Spotify Playlist Recommendation App (Python, Heroku)   **Jun 2024 - Aug 2024**
- Launched a Heroku web app leveraging Spotify API to compare playlists to recommend similar tracks using cosine similarity.
- Implemented a data pipeline utilizing the Spotify API with PKCE authorization to fetch, preprocess, and vectorize song data using TF-IDF, one-hot encoding, and normalization.
- Designed a Flask backend and a frontend with HTML, CSS, and JavaScript

Sales and Customer Dashboard (Tableau)   **Jul 2024 - Aug 2024**
- Followed a structured project timeline, including user demand analysis and mockup creation before designing two dashboards (Sales and Customer) to personalize functionality.
- Designed interactive dashboards featuring BANs of KPIs with percentage increases, sparklines, visualizations like bar-in-bar charts for year-over-year sales comparison, and histograms for customer order distribution.
- Implemented dynamic calculated fields and customizable filters, enhancing UI/UX and enabling detailed analysis of sales and customer data.

Diamond Price Prediction Model (R)   **May 2024**
- Conducted Exploratory Data Analysis (EDA) on a dataset of diamond features and prices to understand data distributions, correlations, and key statistics, facilitating informed feature selection for modeling.
- Applied simple linear regression models to validate assumptions including linearity, normality, and homoscedasticity of error terms.
- Utilized AIC, BIC, and R squared adjusted for model selection, achieving 95% testing accuracy.

WORK EXPERIENCE

Kumospace   **October 2023 - May 2024**
Marketing Intern   New York
- Leveraged data-driven insights to refine influencer selection and optimize marketing strategies, contributing to over 60% increase in revenue in 2 months.
- Analyzed datasets of content creators on TikTok, Instagram, and YouTube to identify those aligning with company values.

SKILLS

Technical: Python (NumPy, Pandas, Scikit-Learn, tensorflow/keras, pytorch, seaborn, Matplotlib, Flask, FastAPI), SQL, R, Tableau, Heroku, Docker, AWS, HTML/CSS, Git
Machine Learning: Linear Regression, Logistic regression, KNN, CNN, SVM, Decision Tree, random forest, naive bayes, NLP, XGBoost, LDA, QDA
Statistics/Math: Hypothesis testing, Regression Analysis, Bayesian statistics, Linear Algebra, Real Analysis