# ANGEL XU

San Francisco, CA | 510-424-3509 | anqi_xu@berkeley.edu | Linkedin | Github | Website

## Education

**University of California, Berkeley** | **Master of Engineering** | GPA: 3.8 /4.0                    08/2022 – 05/2023
   ◆   Industrial Engineering & Operations Research
   ◆   Coursework: Principles and Techniques of Data Science, Optimization Analytics, Applications in Data Analysis.

**Dalian University of Technology** | **Bachelor of Science** | GPA: 3.9 /4.0                    09/2018 – 07/2022
   ◆   Information Management and Information Systems

## Technical Skills

**Core skills:** Machine Learning, Statistical Analysis, Data visualization, Predictive Modeling, Geospatial Analysis, Database Desgin.

**Programming Languages:** SQL, Python (Pandas, Scikit-learn, Matplotlib, TensorFlow), R, SAS, C, C++, Java.

**Software:** MySQL, Power BI, Tableau, Excel (Pivot Table, VLOOKUP), AWS, QGIS, Google Cloud, Google Analytics.

## Work Experience

**Data Analyst** | *Predictive Modeling, Timeseries Forecasting, Python* | University of Victoria | Github        07/2021 – 10/2021
   ◆   Predicted COVID-19 cases, deaths, and recoveries in **Python** using **Linear Regression, ARIMA**, etc, compared forecast performance using RMSE, and visualized confidence intervals of results with **Python Plotly**.
   ◆   Generated interactive data visualizations predicting Canada's cumulative cases with ipywidgets in a **Jupyter Notebook**.

**Monetization Data Analyst** | *A/B testing, Statistical Analysis, SQL* | Bytedance (TikTok's parent company)        02/2022 – 06/2022
   ◆   Built data pipelines, designed **5 dashboards** to track ad metrics, and analyzed **300k+ sales data** to boost ad revenue.
   ◆   Conducted **A/B tests** using **SQL** to test the hypothesis that extending the duration of opening screen ads from 3s to 5s will improve the performance of ads metrics on Toutiao, resulting in **a 28.8% increase** in click-through rate.
   ◆   Partnered with the product team and design team to promote the **go-to-market process** of 3D Layer News Feed Ads.
   ◆   Collected feedback from salespeople and advertisers to provide actionable insights for feature development.

**Commercial Data Analyst** | *Data Analysis, Product Management, Tableau* | Meituan.        06/2021 – 09/2021
   ◆   Monitored daily sales of ads and produced **20+ dashboards** in **Tableau** for weekly team meetings.
   ◆   Performed abnormal data analysis using **Excel Pivot table** and **identify root causes** with the advertising platform department.

## Project Experience

**Data-Driven Assessment of Wildfire Risk in California** | *QGIS, Python, Tableau* | Visuals        11/2022 – 04/2023
   ◆   Developed a wildfire risk assessment framework using Python, QGIS and used machine learning models including **Random Forest, Logistic Regression, SVM, and Decision Tree** to predict damage levels of structures with over 80% recall rate.
   ◆   Utilized **Generalized Linear Model** to quantify the relationship between intensity measures, structural characteristics, and wildfire risk, providing actionable insights for homeowners, firefighters, etc. on prioritizing wildfire risk mitigation efforts.

**Analysis of TMDB Movie Dataset and Revenue Prediction** | *Data Mining, Regression, Clustering* | Github        10/2022 – 12/2022
   ◆   Crawled **57k+** movie data through TMDB Open API, performed exploratory data analysis using **SQL queries** in **Python** in terms of the release year, genre, ratings, runtime, etc., and built comparative visualizations in a **Jupyter Notebook**.
   ◆   Conducted **feature engineering** and preprocessed text data to build **Bag-of-Words** using NLTK in **Python** for modeling.
   ◆   Applied multiple **machine learning models** to predict movie revenue and validated the result through **the bootstrap method**.

**Research on User Download Behavior based on Google Play App Data** | *ML Classification, Text Mining*        11/2020 – 12/2020
   ◆   Created a data preprocessing pipeline on Google app data, utilized multiple **machine learning classifiers** (Logistic Regression, SVM, Random Forest) in **Python** to predict application ratings, and **achieved an accuracy of 89.3%** with SVM.
   ◆   Performed polarity analysis on product review in **Python** and recommended targeted suggestions to app developers