

Angel Xu

510-424-3509 | anqi_xu@berkeley.edu | [Linkedin](#) | [Github](#) | [Website](#) | San Francisco, CA

Education

- University of California, Berkeley | Master of Engineering** | GPA: 3.9 /4.0 08/2022 – 08/2023
- ◆ Industrial Engineering & Operations Research (STEM)
 - ◆ Coursework: Principles and Techniques of Data Science, Introduction to Financial Engineering, 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

Work Experience

- Data Scientist Intern** | Omnium (CPG Consulting Firm) | San Francisco, CA 05/2023 – 08/2023
- ◆ **Data Analysis, SQL, SQL Server:** Utilized SQL queries on SQL Server to pull key metrics for Oatly's weekly sales report.
 - ◆ **Data Warehousing, SQL:** Created a data loading and transforming pipeline to connect syndicated databases using SQL, resulting in a runtime decrease of **200%** for the production of company-wide clients' Power BI dashboards.
 - ◆ **Machine Learning, Customer Analysis, Python:** Applied clustering algorithms to Oatly's customer segmentation based on dollar sales and growth index, optimizing promotion efficiency, and ROI for different customer segments.
 - ◆ **Statistical Modeling, Python, Excel:** Collaborated with Oatly sales team to develop a sales rates regression model to analyze P&L and pricing impact, resulting in informed pricing strategies and increasing Oatly's revenue by 26% in frozen products.
 - ◆ **Data Visualization, Power BI:** Designed, and implemented multiple Power BI dashboards to track client's sales performance.
 - ◆ **Agile Practices, SQL, AWS:** Led the company's Data Building Tool implementation and presented to the leadership team.
- Data Analyst Intern** | Bytedance (TikTok's parent company) | Beijing, China. 02/2022 – 06/2022
- ◆ **Data Visualization, Tableau:** Designed 5 **Tableau** dashboards to track metrics, analyzed 300k+ sales data to boost revenue.
 - ◆ **A/B Testing, SQL, SSIS:** 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.
 - ◆ **Product Testing, Market Research, SAP:** Partnered with the product team and design team to test a new product.
- Data Scientist Intern** | University of Victoria | Victoria, Canada | [Github](#) 07/2021 – 10/2021
- ◆ **Time-series Forecasting, Python:** Built a rolling forecasting pipeline using time-series models to predict COVID-19 cases, deaths, and recoveries, compared forecast performance, and visualized confidence intervals of results with Plotly.
 - ◆ **Data Visualization, Tableau:** Generated interactive line charts in Tableau to predict Canada's cumulative cases.
 - ◆ **Version Control, Git, MS Azure:** Managed and updated tasks in projects on Git version control system and Microsoft Azure.
- Business Data Analyst Intern** | Meituan | Shanghai, China 06/2021 – 09/2021
- ◆ **Data Visualization, Tableau:** Produced 20+ dashboards in Tableau to monitor product daily sales for weekly team meetings.
 - ◆ **Root Cause Analysis, Excel:** Used Excel to perform data analysis, identified root causes and reported to the engineering team.
 - ◆ **Leadership:** Led the launch of the new product, including pricing, market research, ad testing, and promotion plans.

Project Experience

- Data-Driven Assessment of Wildfire Risk in California** | *ArcGIS, Python, Tableau* | Berkeley, CA
- ◆ Developed a wildfire risk assessment framework using Python & ArcGIS and used **machine learning models** including LSTM, SVM, Linear Regression, Decision Trees to predict damage levels of structures with over 80% recall rate.
- Analysis of TMDB Movie Dataset and Revenue Prediction** | *Python* | [Github](#) | Berkeley, CA
- ◆ 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 **regression models** to predict revenue in a Jupyter Notebook.

Technical Skills

Core skills: Data Analysis, Data Visualization, Machine Learning, Statistical Modeling, ETL, Cloud Computing, Version Control.

Programming Languages: SQL, Python (Pandas, Numpy, Scikit-learn, Tensorflow, NLTK), SAS, C, R (Tidyverse), Git, Linux.

Software: MySQL, Power BI, Tableau, Excel (Pivot Table, VLOOKUP, VBA), SQL Server, AWS, Google Analytics, Snowflake.