

Xiaoge Zhang

77 N Almaden Ave
San Jose, CA, 95110

Linkedin
zhang.xiaoge@northeastern.edu 4085924660

Education

Northeastern University Sep 2022 – Present
Master's Degree, Analytics(STEM) | GPA:3.9/4.0
San Jose, CA
Coursework: Probability, Database Management systems, R, Python, Advanced Data Analysis

Guangdong University of Technology Sep 2016 – Jul 2020
Bachelor's Degree, Human Resources Management | GPA:3.5/4.0
Guangdong, China

Specialized Skills

- **Programming::** Python (Scikit-Learn, Pandas, NumPy, Seaborn), R(tidyverse, ggplot2), SQL
- **Tools:** MySQL, Tableau, Git, Figma
- **Methodologies:** Exploratory Data Analysis, Machine Learning (Regression, Classification, Clustering, Correlation), Statistical Inference, Probability, Hypothesis testing

Work Experience

Luxury Appraiser November, 2020 – May , 2022
Liujiu Suiyue Watch Maintenance Co., Ltd. Guangzhou, Guangdong, China

- Reviewed and analyzed over 7,000 watches to ensure quality appraisal and streamlining the process.
- Communicated with both technical and non-technical stakeholders through all phases of the appraisal development cycle; Gathered data from appraisal and customer service department and analyzed it to identify customer needs
- Online and Offline customer service: Communicated with customers online, guided customers to operate the identification process, check the completeness of the information filled out by customers; answered follow-up questions, cumulatively tracking 15000+ customers.

Project Experience

Udemy Project: Hands-On Apache Airflow Practice December , 2022 –January , 2023

- Utilized Docker to manage Apache Airflow and different executors; Mastered core functionalities of Apache Airflow, such as DAGs, Operators, Tasks, and Workflows
- Integrated Apache Airflow into a Big Data ecosystem, including Hive, PostgreSQL, and Elasticsearch
- Enforced Apache Airflow to solve real-world data processing problems

Bike Sharing Prediction (Code in R/Python) October , 2022 –December , 2022

- Conducted data cleaning to remove irrelevant attributes. Utilized Python, Numpy, and Pandas in imputing null values and data errors, corrected data types, and removed unreasonable values and outliers
- Identified trends and patterns in bike sharing using statistical analysis and data visualization. Performed statistics and data visualization to understand the distribution and relationships of the data
- Trained and constructed regression model to predict bike rentals and evaluated their performance
- Used linear regression to explore the relationship between bike rentals and various factors
- Implemented feature engineering to create new features and improve the predictive power of the model

Credit Card Data Management System October , 2022 – December, 2022

- Created an Entity Relationship Diagram to show the data architecture and design the database schema
- Imported data into the database and created 5 tables to classify and organize the data
- Wrote SQL queries and stored procedures to manipulate and retrieve data from the database