

EDUCATION

The University of Texas at Austin, M.S. in Data Science (Aug. 2021 – Present)

Columbia University in the City of New York, B.S. in Mechanical Engineering (Sep. 2013 – May 2017)

- Departmental GPA: 4.05 / Cumulative GPA: 3.87

EXPERIENCE

Senior Consultant, Design Laboratory - Data Engineer, Microsoft (Aug. 2021 – Present)

- Working in a data engineer role with Microsoft as a Design Laboratory consultant

Teaching (Oct. 2018 – Apr. 2020)

- Established a sole proprietorship business providing in-person and online lessons in the Seattle area
- Taught Python and math to middle and high school students via a personalized curriculum

Mechanical Engineer, ASML (May 2017 – Feb. 2018)

- Developed mechanical designs that improved the performance of ASML's photolithography machines
- Designed and conducted statistical experiments and analyzed experimental data to guide design decisions

Researcher, Columbia University Creative Machines Lab (Sep. 2015 – May 2017)

- Designed and manufactured Eva, a robotic face capable of making facial expressions and head movements
- Eva is an open-source platform designed to facilitate artificial intelligence research in Python
- Video demonstration and published, peer-reviewed paper: <https://www.zanwarfaraj.com/eva.html>

Data Analyst Intern, Microsoft (June 2013 – Aug. 2013)

- Improved Bing's search results by using SQL, regex, and NLP to find and normalize business name synonyms
- Normalized 38,910 distinct businesses and propagated website information between 49,145 distinct businesses
- Project summary: https://www.zanwarfaraj.com/pdf/Data_Mining_Bing.pdf

DATA SCIENCE PORTFOLIO

View project reports using the links below

Predicting Hotel Rates and Cancellations and Segmenting Customers (<https://zanwarfaraj.github.io/machine-learning/>)

- Trained regressors and classifiers via scikit-learn to predict booking rates and cancellations for two hotels in Portugal
- Performed customer segmentation with k-means clustering to identify customers for a possible rewards program

Using Consumer Price Data to Analyze Global Cost of Living (<https://zanwarfaraj.github.io/data-analysis/>)

- Compared the cost of living in cities around the world and analyzed the price of common goods and services
- Created data visualizations using pandas, Matplotlib, and Tableau to communicate findings

Designing and Implementing a MySQL Database on AWS for Financial Data (<https://zanwarfaraj.github.io/database/>)

- Designed a relational database for data about US stocks by creating an ER diagram and performing normalization
- Created a Python script to clean and transfer financial data from Interactive Brokers into the MySQL database

SKILLS / TOOLS

- Data engineering
- Database design
- Python (scikit-learn, pandas, Matplotlib)
- Java, C#, R
- Software development
- Machine learning
- SQL, MySQL
- Jupyter Notebook
- Data cleaning
- Data visualization
- Tableau, Excel