San Diego, CA
630.397.1159

BERNARD WONG

bew030@ucsd.edu
github.com/bew030

EDUCATION

University of California, San Diego **August 2017 - June 2021**

- B.S. in Data Science (3.4/4.0) and B.S. in Probability and Statistics (3.0/4.0). Cumulative GPA: 3.27/4.0.
- Relevant Coursework: Java 1/2; Data Structures; Statistical Approaches to A.I; Spatial Data Science and Applications; Data Analysis and Inference; Recommender Systems and Web Mining; Data Management; System for Scalable Analytics; Data Visualization; Signal Processing; Stochastic Processes; Computational Stochastics; Probabilistic Modeling and ML.

EXPERIENCE

Data Science Intern **Scripps Research** **June 2019 – October 2019**

- Created a professional research report outlining the relationships between sleep stages, heart rate, and motion.
- Used Pandas and Numpy to clean out large data sets created by wearable technology and built a well documented package containing methods that use MatPlotLib and Seaborn to create useful visualizations.
- Applied Markov Chains to predict heart rate and used Scikit-Learn to detect time series anomalies.

PROJECTS

COVID-19 Airborne Transmission Modeling **2021**

- Implemented an agent-based model using statistics, GeoMesa, and Matplotlib that simulated airborne transmission of COVID due to interactions between passengers in a bus under various user set attributes.
- Utilized school bus route datasets to create accurate route simulations using ArcGIS and GeoPandas.
- Used Comet to create a web application displaying the model allowing for easy user input and visualization.

Toilet Paper Production and Distribution Optimization **2021**

- Modeled the production and consumption of toilet paper both before and during COVID-19 using a stock-adjustment model and linear equations.
- Visualized supply and demand using Matplotlib & calculated optimum toilet paper limit using gradient descent.

Lyft Driver Valuation **2019**

- Performed a case study on Lyft drivers analyzing lifetime value and projected employment time and gave suggestions on how to increase each value using clustering and statistical analysis.
- Used Pandas and Numpy to extrapolate profit and Matplotlib, SciPy and Numpy to visualize classifications.

StockX Analytics **2019**

- Developed methods for StockX's open source API that arranged pricing and sales data into dataframes.
- Created an automated system that organized and scraped email receipts decreasing time spent accounting.

AWARDS

UCSD DS3 Lyft Datathon **1st/200** **2019**

- Collaborated with a team to write a report on the relationships between traffic, repairs, and collisions in San Diego which was selected by Lyft as the winning analysis.
- Cleaned and interpolated datasets using Pandas and Numpy, created choropleths and map timelapses using GeoPandas, and conducted hypothesis testing to determine if an area had a significantly higher death rate.

LANGUAGES AND TECHNOLOGIES

Languages and Tools: Python; SQL; Java; JavaScript; R; Matlab; Swift.
Python Libraries: Pandas; Numpy; SciPy; Matplotlib; Seaborn; Scikit-Learn; NLTK; ArcGIS; GeoPandas; Pillow; BeautifulSoup4; Regex; Selenium.
Machine Learning: Statistical Analysis; Nearest Neighbor Classifiers; Decision Trees; Linear Classification; PCA; NLP; Multi-Class Classification; Dimensionality Reduction; Boosting Algorithms; Recommender Systems; Graph Theory.