Ederlyn M. Tanangco

Data Scientist | Data Analyst
Arlington, VA | 703.409.4141 | Ltanangco8@gmail.com | LinkedIn | GitHub | Portfolio

Data Scientist with a psychology background who tells meaningful stories with data. Inspired to help translate large data into actionable results to be able to connect consumers with products and services in a team setting.

SKILLS

- **Programming Languages:** Python, SQL
- **Machine Learning & Libraries:** Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Natural Language Processing, Sentiment Analysis, Statistical Analysis, Predictive Modeling, Predictive Analytics, Cluster Analysis
- **Analytical Tools:** Data Visualization, Tableau, Flask, Jupyter Notebook, Microsoft Excel, VS Code, Git, GitHub

EXPERIENCE

Data Science Immersive | General Assembly | Washington, D.C. | Mar 2020–June 2020

Applied skills in computer programming and statistics in a 500+ hour immersive program. Developed a portfolio of client projects:

- **Beer Recommender Application:** Produced an application, deployed on Flask, that returned similar beers to users using the Scikit-Learn library's Pairwise Distances & Cosine Similarity. Dataset was taken from Kaggle and analyzed, cleaned and performed Exploratory Data Analysis. Used clustering methods such as KMeans and DBSCAN. Evaluated the model using K-Nearest Neighbors. Presented the project to a technical and non-technical audience.
- **Sentiment Analysis on Covid-19:** Completed a sentiment analysis of Covid-19 Twitter posts in the United States to see the general public sentiment in response to state policies. Our team pulled 60k posts with relevant keywords to "coronavirus" and "executive orders". Results concluded that the current public sentiment in response to state policies and Executive Orders were relatively neutral.
- **Web API & NLP:** Created a conceptual application called "Health Today". Based on user engagement, the team predicted and analyzed two separate categories ("food" versus "exercise") in order to find which subreddit category would be most beneficial when marketing and deploying a health application. Completed Natural Language Processing (NLP). Multinomial Naive Bayes pipelined with CountVectorizer received an accuracy score of 84%. TfidfVectorizer with logistic regression was used as a comparison model. Results concluded that "exercise" was a better predictor for our health application.

All Plumbing, Inc. | Arlington, VA | Aug 2007–Feb 2020

Operations Manager

- Led an office team of 6 and a field team of 25, overseeing day-to-day operations. Organized and facilitated multiple company-wide training, which entailed coaching, mentoring and troubleshooting.
- Integrated and trained associates on a new field management system which cut invoice payment turnaround time ~40% and increased revenue 5-7% in 6 months.
- Developed tailored sales techniques to sell and upsell residential and commercial projects to existing and prospective clients.

Marketing Manager

- Developed and maintained marketing campaigns that included digital, social media analysis, email, and print, resulting in a 20% increase in sales leads per quarter.

EDUCATION

General Assembly, Data Science Immersive | Washington, D.C | March 2020–June 2020

George Mason University | B.S. Psychology | Fairfax, VA | Jan 2010–May 2014

Undergrad Research Asst. | George Mason Human Factors (Human-Computer Interaction)

- Conducted over 200 participant simulations for studies in spatial attention, multiple modalities, auditory and visual perception, and cognitive workload
- Collected hundreds of participant data, assisted with data analysis using Excel and MATLAB