BENYIR PACHECO

Wyomissing, PA | 347-777-7970 | benjijosepacheco@gmail.com| [Linkedin](#) | [Github](#)

EDUCATION

HUNTER COLLEGE
New York, NY

Bachelor of Arts
Feb 2022

- Major in Mathematics; Minor in Computer Science; Cumulative GPA: 3.44
- Relevant Coursework: Python Programming, Data Structures and Algorithms, Relational Databases & SQL, Numerical Methods, Real Analysis, Partial/Ordinary Differential Equations, Probability Theory, Linear Algebra

Certifications**:** Data Analysis Nanodegree – Udacity

SKILLS

Languages: Python, SQL, English (Native), Spanish (Native)
Data Visualization: Matplotlib, Seaborn, Streamlit
Data Wrangling: Pandas, NumPy, BeautifulSoup, Requests
Machine Learning: Scikit-Learn, SciPy, Statsmodels, XGBoost
Tools: Jupyter Notebooks, PyCharm, Git, Tableau, Microsoft Power BI, pgAdmin4

PROJECTS

PROSPER LOANS DATA VIZ
Jan 2022 – Mar 2022

https://github.com/benjpacheco/ProsperLoans_Data_Visualization

- Created univariate-multivariate data visualizations using **Python**, **Pandas**, **Matplotlib**, **Seaborn**, and **NumPy** for loan amounts along with several predictor variables from Prosper data to find positive/negative correlations.
- Exploration findings show Prosper prefers to lend at 0.32% interest rate, interest rate vs. APR have pearson correlation of 0.990, and loan amounts vs. APR have correlation of -0.320. Moderate relationship between loan amounts and the rates (interest rate/APR) with several different Credit Grades/Prosper Ratings given to Prosper customers.

WERATEDOGS TWITTER WRANGLING
Oct 2021 – Dec 2021

https://github.com/benjpacheco/WeRateDogs_Wrangling

- Leveraged **Python**, **Pandas**, **NumPy**, twitter API (**tweepy**), **Requests**, to gather tweets from twitter account WeRateDogs in order to assess, wrangle the data from dirty to clean/tidy, and then analyze.
- Top breeds predicted by the neural network data were the Golden Retriever, Pembroke, and Labrador Retriever, along with most favorited tweet "Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad)" with 146369 likes.

E-COMMERCE SITE A/B TEST
Aug 2021 – Sep 2021

https://github.com/benjpacheco/Ecommerce_AB_Test

- Assessed whether an e-commerce site needed to implement the new web-page, keep the old web-page, or run the experiment longer to make a decision utilizing **Python**, **Probability**, **Null Hypothesis** testing, and **Logistical Regression**.
- Computed p-value of 0.8874, z-score of -1.213 for right-tailed test with Type I error rate 5%. Regression approach yielded p-value of 0.1899 which implies our results are statistically insignificant. Results along with observed actual differences show that the new and old pages have equal chance of converting users. We fail to reject the null hypothesis. Recommended e-commerce site to keep the old page.

TMDB EDA
May 2021 – June 2021

https://github.com/benjpacheco/TMDB_EDA

- Performed exploratory data analysis on The Movie Database (TMDB) data set containing information of about 10,000 movies measuring key values such as revenue and user ratings.
- Identified and implemented data wrangling solutions before conducting further analysis including but not limited to: pivoting, melting/exploding dataset, and handling missing/NaN values with **Python**, **Pandas**, **NumPy**.
- Western was the most popular genre over time, Woody Allen & Steven Spielberg tied at 30 films for most amount of films produced, and Net Profit vs. Revenue for all movie genres had pearson correlation of 0.860 without inflation and 0.877 with inflation which implies an increase of revenue for higher budget investment regardless of accounting for inflation.