


BETHANY THOMPSON

DATA SCIENTIST

thompson.bethany.01@gmail.com 

(903) 780 - 1486 

github.com/ThompsonBethany01 

in/bethany-thompson-01 

ABOUT ME

I bring a truly creative approach to every challenge both in life and in work. My diverse data science portfolio demonstrates my ability to apply this creativity, whether I'm discovering insights in data or skillfully telling its story through professional quality reports and presentations. My task-oriented nature allows me to consistently complete projects from conception to completion in a timely manner.

EDUCATION

Codeup, Certificate of Completion

San Antonio, TX | July 2020 - January 2021

Fully-immersive, project-based 22-week Data Science career accelerator that provides students with 670+ hours of expert instruction in applied data science. Students develop expertise across the full data science pipeline (planning, acquisition, preparation, exploration, modeling, delivery), and become comfortable working with real, messy data to deliver actionable insights to diverse stakeholders.

Tyler Junior College, General Studies Associates of the Arts

Tyler, TX | August 2015 - May 2016

Magna Cum Laude

SKILLS

Tools

- Tableau
- Python
- SQL
- Git
- Jupyter

Libraries

- Pandas
- Numpy
- Skikit-learn
- Matplotlib
- Seaborn

Concepts

- Data Storytelling
- Applied Statistics
- Communication
- Project Planning
- Presenting

PORTFOLIO

Development Projects

Capstone Project - SpotiScry

January 2021

Our team used Spotify's API to acquire data on thousands of songs. To discover the science behind the art of music, we explored and statistically tested what features influenced a song's popularity. We then created a model to predict how popular a song is based on these key insights found.

Predicting a Repo Language

November 2020

Using python to web scrape our own data set, my partner and I collected information on the top trending GitHub repositories. We cleaned the data with Natural Language Processing to further explore and model the data. We found key insights by controlling for each coding language and then visualizing with word clouds and various seaborn plots.

Analyzing Web Logs

November 2020

Given a web log of users accessing the curriculum site, I was asked to find insight on fraudulent activity and general trends with various student groups. I determined further action needed on suspicious activity. In addition, I optimized the deliverable for audience ease.

Passion Projects

Predicting Diabetes

December 2020

Analyzing medical data on adult females with Pima heritage, I determined top drivers of diabetes diagnosis rates. I also used KMeans Clustering to create new features and explore the interaction of these with the predictive variable. The final model of RandomForest correctly predicted a patient's diagnosis 3 out of 4 times on the unseen test data.

Analyzing Fitbit Data

November 2020

Data for this project simulated an individual's activity tracked through Fitbit over a period of eight months. I determined the characteristics of the individual by looking at trends with time series analysis. To predict the next two weeks of data, I tested models on each column to determine the best results. Final models were a weekly rolling average and Holt Optimized.

Office Space

October 2020

Taking inspiration from the movie Office Space, I created a python challenge to fix a bug in a function designed to skim fractional cents from bank transactions' calculated interest. Focus was on well-documented python coding and creating a visually appealing Readme file.