




MATTHEW SANTOS

DATA SCIENTIST

(805)674-5594  mattsantos541@gmail.com  in/mattsantos541  Mattsantos541

I am enamored by the problems we can solve and the stories that can be told with data. In data science, we can find different ways to approach traditional problems that plague every industry. I want to help solve these problems with data driven approaches that can maximize results for everyone.

CORE COMPETENCIES

Applied Statistics - SQL - Python - Pandas -
Matplotlib - Machine Learning - Distributed Data -
Data Storytelling - Git -- Anaconda- NLP

ACADEMIC BACKGROUND

CODEUP

Certificate of completion | 2020

Fully-immersive, project-based 18-week career accelerator that provides students with 600+ hours of expert instruction data science.

UNIVERSITY OF PHOENIX- 2014

Bachelor's in Finance | 2014

EMPLOYMENT HISTORY

COORDINATOR/ ACCOUNTANT

City of San Antonio. | Nov 2018- July 2019

Helped streamline three departments (aviation, city council, housing) to digitize their accounting processes. served as a liason between council members and districts in funding several projects including: Barkitecture 2019, youth soccer league, and several SA fiesta activities

ACCOUNTING/ INVENTORY

The Block SA | 2016 - 2019

Helped with a successful rebrand in 2018, conducted research for the new inventory based on ratings and sales. Will be consulted for the expansion coming in 2020

ACCOUNTING

University of Florida- Nephrology| 2014 - 2015

Helped improve and streamline the process for both travel reimbursements and project proposals. Assisted in picking the location for a new dialysis center in expansion proposal (5 year plan),

DATA SCIENCE PROJECTS

ZILLOW ZESTIMATOR

Using clustering to attempt to improve upon the zestimate from the real estate app, Zillow. We utilized clustering, regression and some feature engineering.

FITBIT TIME SERIES ANALYSIS

Using Fitbit data, used Facebook's prophet and sklearn to predict the future two weeks of activity. Used Fitbit data from a single person to predict their steps and weight loss for the two weeks post data acquisition.

NLP ANALYSIS OF README

Used NLP to scrape data from GitHub repository README files. The goal will be to build a model that can predict what programming language a repository is, given the text of the README file.

TMDB BOX OFFICE-PERSONAL SIDE PROJECT

Using metadata to predict a film's international box office return. using regression and cross validation to predict the return on 4300 films based on a training a given 7,000 films.

YELP/HEALTH SCORES- CAPSTONE PROJECT

Compared food establishment health inspection scores from the San Antonio Health Department website to investigate any links/anomalies as well as creating a model to predict an establishment's inspection score based on Yelp reviews.

MLB CONTRACT/ LINEUP- PERSONAL SIDE PROJECT

Using data per baseballreference.com, used clustering and regression to come up with a logical salary cap for the MLB (no salary cap). Used PS simulation to check performance of teams constructed within those salary restrictions