

JOSHUA CHAVEZ

DATA SCIENTIST



[/joshuamchavez2](#)



[/joshuamchavez2](#)



joshua.m.chavez2@gmail.com



(805) 621-4021

TECHNICAL SKILLS

SQL - Python - Pandas - Matplotlib - Seaborn - Plotly
- Machine Learning - Natural Language Processing -
Apache Spark - Data Storytelling - Git - Jupyter
Notebooks - Anaconda - Tableau - C - C++ - C# - Java
- HTML - CSS - Javascript - MongoDB - Docker

EDUCATION BACKGROUND

Codeup

Data Science Program 2021

Brandman University(Now UMass Global)

Bachelor's in Business Administration 2016-2018

Allan Hancock College

Associate's Degree Transfer Studies 2007-2010

WORK EXPERIENCE

Battleground

Business Entrepreneur 2012-2017

Operated and supervised all aspects of a brick-and-mortar gaming retail store. Researched and developed a business plan, biweekly budget, to meet business goals on a short and long term basis. Used an ecommerce software that provided business intelligence about past sales and current inventories. Built relationships with multiple distributors. Managed staff records and payroll. Actively worked to display a courteous and attentive attitude to customers.

United States Army

Human Resources Specialist 2004-2007

Prepared and reviewed personnel casualty documents. Monitored suspense actions. Initiated, monitored, and processed personnel evaluations. Prepared orders and request for orders. Prepared and maintained officer and enlisted personnel records. Processed recommendations for awards and decorations.

PERSONAL PROFILE

Proud Army Veteran with an inactive secret clearance. Recent training received in data science from a 22-week immersive bootcamp. Enjoy solving ambiguous problems and getting exposure to multiple projects. I would excel in a collaborative, cross-functional, and team-driven environment.

PROJECTS

CLASSIFICATION

2021

Esports is a booming global industry that will soon rival that of traditional professional sports. League of Legends is one of the largest esports in the world that generated \$1.75 billion dollars for Riot Games in 2020 alone. A professional match can last around 20 minutes, and we've created a model that predicts the winner of a match at the 10 minute mark. Our best Random Forest Classifier model performed 23% better than baseline and had an accuracy of 61% on unseen test data.

TIME SERIES ANALYSIS

2021

The goal of this project is to predict the temperature on the date of 09-01-2013 for the city of Los Angeles. Used data exploration and visualization tools to find key features used in my model. I was able to meet my goal with an accuracy of 99% using time series analysis.

LINEAR REGRESSION

2021

Using the zillow dataset, I created multiple regression models utilizing home features to predict home values. Used SQL to query the database. Through visualizations and statistical testing, I uncovered top drivers of home value. I created multiple machine learning models using regression algorithms that accurately predicted a California based home's value.

NATURAL LANGUAGE PROCESSING

2021

Using GitHub's API & Beautiful Soup, the team scraped over 100 repositories & pulled the text off their README's. We analyzed text data using Natural Language Processing techniques and created multiple models to predict the primary language of the repository. Our best performing model correctly predicted the programming language at a 62% accuracy rate.