

CINDY VILLANUEVA

Data Scientist



720-440-1925



in/cindyvillanuevads/



cindy.villanueva.ds@gmail.com



cindyvillanuevads

I am a detail-oriented Data Scientist who naturally craves diving deeper into data in search of insights that have real-world applications.

SKILLS

Python	Pandas
SQL	Numpy
Applied Statistics	Jupyter Notebook
Machine Learning	Visual Code
Matplotlib,	Natural Language Processing
Seaborn	Github

EDUCATION

Codeup Data Science Program, March 2021 - Sep 2021

Fully immersive, a project-based 22-week career accelerator that provides students with 670+ hours of expert instruction in applied data science.

Certified Professional Coder · American Academy of Professional Coders (AAPC). Aug 2015

Bachelor's Degree in Electronics and Communication Engineering · Autonomous University of Nuevo Leon, Mexico. Dec 2003

PROFESSIONAL EXPERIENCE

Hospital Medicine Coder

Alteon Health, February 2020 - March 2021

Assigned codes for diagnoses according to the ICD-10-CM and E&M levels according to the CPT Official Guidelines for Coding and Reporting.

Facility Coder II

Denver Health and Hospital Authority

January 2019 – May 2019

Assigned codes for diagnoses, treatments, and procedures according to the ICD-10-CM and CPT Official Guidelines for Coding and Reporting for Urgent Care, Ophthalmology, Cardiology, Orthopedic E&M

Facility Coder I

Denver Health and Hospital Authority

November 2015 – December 2018

Assigned codes for diagnoses, treatments, and procedures according to the ICD-10-CM and CPT Official Guidelines for Coding for Urgent Care and Ophthalmology.

SNAP-Ed Nutrition Educator,

Colorado State University · Extension, Denver CO

April 2012 - November 2013

DATA SCIENCE PROJECTS

Capstone Project: Attention Walmart Shoppers, Sep 2021

The retail stores are facing a challenge due to unforeseen demands and run out of stock. Our four-member team built a model that can forecast demand for the following week. This model can help managers maintain proper inventory levels and increase profits for the store. We used Walmart retail data acquired from Kaggle. After extensively cleaning and prepping the data, we were able to create a 2nd-degree polynomial regression model that outperformed our baseline RMSE by 29%.

NLP Project: Predicting Programming Languages from Github Repositories, August 2021

Build a model that can predict what programming language a repository is, given the text of the README file. Data was collected via web-scraping of the GitHub repository README files. I utilized a multiclassification classification predictive model. In order to create the model, the Readme content was normalized, tokenized, stemmed, lemmatized, and stopwords were removed to produce "clean" content and explored my data. My best model (decision tree) using a simple bag of words performed better than my baseline.

Individual Project: Predicting Loan Application Status, July 2021

Used the classification methodology that predicts the loan status based on customer information provided while completing the online application. I utilized a binary classification predictive model. In order to create the model, I cleaned my data and decided how to handle missing values, created new columns, scaled my data, and explored my data. My best model (KNN) performed better than my baseline.

Clustering: Zillow Log Error Prediction June 2021

SQL was used to join multiple tables for 2017 properties (Zillow database). I prepped and explored the data using pandas, Seaborn, and Matplotlib and used Scikit-learn to cluster and encode data for modeling to predict the error in Zillow's Zestimates.

Regression: Zillow Home Values Prediction June 2021

Examined Zillow's single-unit properties with transactions during May-Aug 2017, acquired from MySQL database. This dataset required cleaning data. I constructed a final regression model that predicts a home's tax assessed value.

Classification: Predicting Churn May 2021

SQL was used to join multiple tables that would be imported into a Jupyter Notebook where a classification-based machine learning model was constructed to predict customer churn.