

Lori

Segovia


Data Scientist

+ CONTACT

 (210) 844 - 1482

 segovialori@gmail.com

 github.com/segovialori

 in/lorisegovia

ABOUT ME

A naturally curious nature paired with a passion for helping others led me to the realm of Data Science. My experience in challenging environments has helped me cultivate an adaptive mindset that thrives off of finding unique ways to problem-solve. As a solution focused Data Scientist, I am determined to produce results that will help provide valuable insights and recommendations.

TECHNICAL SKILLS

- Scikit learn
- Pandas
- Seaborn
- Matplotlib
- Numpy
- Python
- SQL
- Git
- Jupyter
- Project planning
- Storytelling
- Applied statistics

EDUCATION

CODEUP JUNE 2021
CERTIFICATE OF COMPLETION

Fully immersive, project-based 22-week career accelerator providing 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.

UTSA AUGUST 2014
BACHELOR OF SCIENCE IN BIOLOGY

CONCENTRATION: Neurobiology
MINOR: Psychology

EXPERIENCE

CENTER FOR AUTISM AND RELATED DISORDERS

ASSISTANT SUPERVISOR APRIL 2017 - DECEMBER 2020
Created, analyzed, and updated behavior intervention plans and skill acquisition plans for each client.

SKILLS EXPERT NOVEMBER 2016 - APRIL 2017
Maintained and organized medical information, assessment scores, and SKILLS accounts for each client.

BEHAVIOR THERAPIST JULY 2016 - NOVEMBER 2016
Provided 1:1 Applied Behavioral Analysis (ABA) therapy to children with autism in the home, clinic, and community.

DATA SCIENCE PROJECTS

WHAT'S THE 311? PREDICTING SPEED OF RESPONSE TO 311 CALLS

San Antonio 311 call data was acquired from sanantonio.gov open datasets. Data was cleaned and prepared with Pandas library. Scikit learn library was used to create decision tree model that improved accuracy from baseline by ~10%. We also evaluated if inequality exists in 311 call response time across different issues and across different neighborhoods in San Antonio.

PREDICTING NASA PROGRAMMING LANGUAGES

Requests and BeautifulSoup libraries were used to scrape data from NASA Github repository README's to acquire and create a dataset. Data was cleaned with the Natural Language Toolkit and fed into several different regression models to predict programming language from repo README's.

PREDICTING PROPERTY VALUE

Zillow data was queried from SQL and made into a pandas Dataframe. Utilized SciKit Learn library to create and evaluate linear regression models that predict property value. In the first and second iteration, the polynomial regression model performed the best, with the best model beating baseline by 42%.

PREDICTING CUSTOMER CHURN

Query from SQL was used to obtain telco data. Data was cleaned and prepped using Pandas library. Utilized SciKit Learn library to create and evaluate classification models that predict customer churn. Decision Tree model was the top model used to identify whether a customer was going to churn, improving from baseline by ~8%.

PASSION PROJECTS

2021 CIVTECH DATATHON: 1ST PLACE WINNER

We analyzed pilot data from SMARTSA's streetlight program to determine the relationship between air quality and water consumption in SA's medical center. We analyzed all pilot data and determined several insights pertaining to the accuracy of the data collected by the street light sensors.

PREDICTING AUTISTIC TRAITS IN TODDLERS

In this project two different iterations were created using classification algorithms to predict autistic traits in toddlers. In both iterations, the best performing models were K-nearest Neighbors models. Both iterations outperformed the baseline accuracy, with the best model increasing accuracy by ~22%.