COREY SOLITAIRE

DATA SCIENCE / ANALYTICS

PERSONAL PROFILE

I'm an experienced and detail-oriented data professional with a deep understanding of real-world applications for statistical analysis. Through my work in educational leadership, I have developed effective communication skills. Meaningful collaboration with partners is central to my professional values. My versatile skill set allows me to use data science to provide actionable insights and results for optimized business applications.

AREAS OF EXPERTISE

- SQL and Sequel Pro
- Direct experience designing, analyzing and implementing machine learning models
- Python, and Python data science libraries
- Working knowledge of Tableau, Git, Apache Spark, Scrum, and Excel

SKILLS

- Ability to analyze complex technical information
- Scientific research and data collection
- Budgeting and resource allocation
- Scrappy problem solving

EDUCATION

Certificate of Completion, 2021
Codeup
San Antonio, TX

M.S. in Geology, 2008
University of Vermont
Burlington, VT

B.S. in Education, 2005
University of Vermont
Burlington,VT

CONTACT

210-243-9573

corey.solitaire@gmail.com

www.linkedin.com/in/corey-solitaire

github.com/CSolitaire

PROFESSIONAL PROJECTS

Capstone: Predicting COVID Utilizing Social Vulnerability Index — **Nov | 2020**

The CDC's social vulnerability index (SVI) is a scale that predicts the vulnerability of a population in the event of an emergency or natural disaster. As COVID is the first global pandemic since the development of this measure, we aim to evaluate the association between SVI score and COVID infection count in San Antonio, Texas. Features from this measure will be incorporated into a predictive model that can be used to guide recovery resource prioritization.

Anomaly Detection : Who's Fitbit is This? — **Nov | 2020**

Exploration of two years of website log data to uncover trends and identify trends in curriculum demographics, and anomalous events in regards to curriculum access and security. As a result of this analysis six specific user ids and several IP addresses were flagged as nefarious.

Zillow Error Prediction: Home Price Estimator — **Nov | 2020**

Using Linear Regression I developed a model that predicted the error of the Zillow home price estimate. Using Kmeans, I identified clusters in the data based on error, location, and size of the home. I also used pandas, matplotlib, and seaborn to analyzed what drives theerror in Zillow's home price estimates.

PASSION PROJECTS

Curriculum Access Time-series Analysis : — Nov | 2020

Exploration of two years of website log data to uncover trends and identify trends in curriculum demographics, and anomalous events in regards to curriculum access and security. As a result of this analysis six specific user ids and several IP addresses were flagged as nefarious.

Topic Modeling: Geology Current Events — Dec |2020

This project utilized web scraping and topic modeling to identify current trends in geologic research. A function was developed that takes in a list of GitHub URL addresses and collects README text data and the repository's primary programming language. Topic modeling was performed using the LatentDirichletAllocation (LDA) class from the sklearn.decomposition library. Five trends were identified in the corpus over the course of two rounds of modeling, with trend association becoming clearer when common words were removed.

PROFESSIONAL EXPERIENCE

Science Teacher, Northside Independent School District — 2011-2020

Solution-driven, collaborative educational professional with 15 years of experience in formal and informal education, community outreach, budgeting, and project and grant management. Deeply committed to optimizing project outcomes with an innate ability to connect, engage and build relationships. Skilled in team development and training.

- Assessment based data collection and analysis
- Database management
- Campus and district leadership team