

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 and 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 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

Natural Language Processing (NLP)

Nov | 2020

Used BeautifulSoup for web scrapping and natural language processing techniques to analyze README's from GitHub. Used statistical analysis and Python to gain insights from the data, and accurately predict the programming language being used. Leveraged these insights to create a function that accepts a URL and returns a prediction on that URL's primary programming language.

Anomaly Detection

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 (Clustering)

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 database on error, location, and size of the home. I also used pandas, matplotlib, and seaborn to analyze what drives the error 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.

Mahjong Predictive Model :

Ongoing | 2020

I am currently working a project that will take a series of 19 Mahjong tiles and return the statistical probability associated with achieving each of the 50 possible winning hands. The model will eventually be able to take in new information (as tiles are picked or dropped) to provide a playing with a running probability on their current hand of choice.

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