Kay Royo

+1 (510) 619 - 4645
kroyo@ucdavis.edu | https://www.linkedin.com/in/kayr/ | https://github.com/kayannr | https://kayannr.github.io/

SUMMARY

A self-motivated, assertive, and strategic Data Analyst/Scientist with comprehensive analytical, logical, and communication skills. An avid learner who is keen on utilizing technical tools and methods to solve and simplify complex problems. Adept at working independently and collaborating with teams across multiple functions. Highly detail-oriented, solution-focused, and versatile with 3 years of experience in research, data acquisition, data wrangling, and data analysis. Exceptional and effective time management, decision-making, and organization skills continuously enhanced by academic and professional experiences. Capable of building long-lasting relationships with clients and colleagues at all organizational levels. Driven to explore various strategies and methods to provide high-quality work and exceed expectations.

TECHNICAL SKILLS

Programming Languages: R, Python, SQL, VBA, C++
Web Technologies: JavaScript, HTML5, CSS3, AJAX, jQuery, D3, NodeJS, JSON, Plotly, Leaflet, Bootstrap, API
Database Management Systems: PostgreSQL, MongoDB
IDEs: Visual Studio, RStudio, Jupyter Notebook, JupyterLab, XCode, Adobe Photoshop
Software: Tableau, Git, ArcGIS, gINT
Office Tools: MS Word, MS Excel, MS Outlook, MS PowerPoint
Other: Hypothesis Testing**,** Regression Analysis**,** Web-Scraping, Regular Expression (RegEx), Spark, Big-Data, Machine Learning, Hadoop/Hive, Pandas, ETL, ERD, Relational databases, Flask, Scikit-learn, Statistical Modeling and Analysis, Time Series Analysis, Data Modeling, Data Manipulation and Cleaning, Data Wrangling, Data Visualization, Data Analysis, Natural Language Processing, Natural Language Generation, Github, Jupyter Notebooks, Project Management, Technical Writing

EDUCATION AND CERTIFICATIONS

University of California – Davis September 2021 – June 2023
Master of Science (Statistics –Data Science Concentration)

ESRI – MOOC November 2020 **-** December 2020
Spatial Data Science Certificate

- A 6-week online course focused on data engineering and visualization, suitability modeling, pattern detection, space-time pattern mining, and object detection with deep learning.

University of California - Berkeley (Extension) February 2020 - August 2020
Data Analytics Certificate

- A 24-week intensive program focused on gaining technical programming skills in Excel, VBA, Python, R, JavaScript, SQL Databases, Tableau, Big Data, and Machine Learning

California State University – East Bay September 2014 **-** August 2019
*Bachelor of Science (Geoscience**)***

- Honor's and Dean's Listed
- Research Assistant (Department of Earth and Environmental Sciences)
- Event Coordinator (Earth and Environmental Sciences Club)
- Scholarship Award Recipient (East Bay Mineral Society)

PROJECTS

Sports Stats September 2022
https://github.com/kayannr/sportstats
The project focuses on analyzing historical Olympics games dataset using SQL and PandaSQL.

- Independently performed data wrangling, data visualization, descriptive, exploratory, and inferential analysis using graphical and statistical methods.
- Methods and Tools used: SQL, Databricks, Python, PandaSQL, Pandas, SQL Window functions.

Word Wright December 2020

https://github.com/howec/wordwright | https://devpost.com/software/wordwright

The project aims to promote connecting with loved ones through an interactive story writing web application that uses a natural language generator to provide a story prompt based on a user's choice of writing style.

- Developed and implemented Natural Language Generator model using PyTorch and performed data wrangling using Python and Regular Expressions.
- Methods and Tools used: Natural Language Generator (NLG), SQL, Python, JavaScript, HTML, CSS, Bootstrap, Regular Expressions, PyTorch, ETL for Data Cleaning, Web API, React

Music Genre Prediction August 2020

https://github.com/etarakci/music-genre-prediction | https://music-genre-prediction.herokuapp.com/

The project delves into the relationship between song lyrics, titles, artists, and genre using machine learning models that determine the prevalent words in the song lyrics to categorize a song into a genre.

- Successfully completed front-end tasks including interactive web development, web design, and front-end to back-end connection using Python, HTML, CSS, Flask, and Javascript.
- Methods and Tools used: Machine Learning (NLP & KNN), SQL, Python, JavaScript, HTML, CSS, Bootstrap, ETL, Web API, Heroku, Flask

US Homeless Population July 2020

https://github.com/kayannr/US-Homelessness |
https://mengye22.github.io/US-homelessness/master/templates/index.html

The project explores the homeless population and homeless shelter locations for each state in the United States. The final datasets generated after performing Web-scraping and ETL are loaded to its destination, a SQL database.

- Gathered data using web scraping method and Python. Effectively performed ETL on the scraped dataset using Python and Regular Expressions. Designed interactive choropleth map to display homeless population data for each state using D3, Leaflet, Python, HTML, Javascript, and CSS.
- Methods and Tools used: Web-scraping, Data Wrangling, RegEx (Regular Expressions), ETL, JavaScript, Python, HTML, CSS, D3, Leaflet

---

PROFESSIONAL EXPERIENCE

Graduate Student Teaching Assistant September 2021 – December 2022

UC Davis

- Assisted faulty members by performing teaching-related duties in the Department of Biomedical Engineering and Graduate School of Management for business analytics, entrepreneurship, and accounting courses

Project Geologist September 2019 – June 2021

ATC Group Service LLC

- Modernized and automated detailed technical reports associated with Phase II Environmental Site Assessments to present data to clients which resulted in 30% increase in productivity
- Maintained client relations by clarifying needs and communicating progress and relevant information
- Provided project planning and management assistance which led to increase in efficiency and saving up to $1000 in project funds
- Analyzed historical subsurface environmental data to determine the locations for new soil vapor and groundwater water monitoring wells using Microsoft Excel

Research Assistant September 2018 – November 2020

CSU – East Bay

- Analyzed scientific data gathered to determine the correlation between sediment deposition and the changes in sea level and climate in the San Francisco Bay Area

Geophysical Survey Intern June 2018 - July 2018

Canary Islands Volcanological Institute (INVOLCAN) and Geotenerife

- Efficiently acquired subsurface electromagnetic data by operating a Multi-Channel Geophysical Survey System and installing Magnetotelluric (MT) and seismic survey stations

---