

Kay Royo

+1 (510) 619 – 4645 | kroyo@ucdavis.edu | San Francisco Bay Area
<https://www.linkedin.com/in/kayr/> | <https://github.com/kayannr> | <https://kayannr.github.io/>

SUMMARY

A self-motivated and strategic Data Scientist with comprehensive analytical, logical, and communication skills. An avid learner driven by curiosity who is keen on utilizing technical tools and methods to solve complex problems and develop data-driven solutions. 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, and data analysis. 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.

SKILLS

Programming Languages:	R, Python, SQL (Advanced)
Web Technologies:	HTML5, CSS, Javascript, jQuery, D3.js, Bootstrap
Database Management:	PostgreSQL, MongoDB, ERD, Relational Databases
Data Visualization:	Tableau, ArcGIS, Plotly, Leaflet, Matplotlib, Seaborn
Data Collection & Cleaning:	ETL, Web-Scraping, API, JSON, RegEx
Data Analysis:	Statistical Modeling and Analysis, Hypothesis Testing, Regression Analysis, Time Series Analysis, Descriptive Statistics, Exploratory Data Analysis, Topic Analysis
Machine Learning:	NLP, NLG, KNN, Scikit-learn, PyTorch
Libraries:	Pandas, NumPy, NLTK, SciPy, Statsmodels, tidyverse, dplyr, caret, forecast
IDEs:	Visual Studio, RStudio, Jupyter Notebook, JupyterLab
Soft Skills:	Teamwork, Communication, Adaptability, Problem-solving
Other:	Git, GitLab, Github, MS office Suite, Spark, Hadoop/Hive, Flask, Technical Writing, Project Management

EDUCATION AND CERTIFICATIONS

University of California – Davis	09/2021 – 06/2023
<i>Master of Science (Statistics – Data Science Concentration)</i>	
<ul style="list-style-type: none">Coursework: Time Series Analysis, Statistical Machine Learning, Artificial Intelligence, Statistical Data Science, Big Data and High-Performance Statistical Computing, Statistical Research Methods, Multivariate Data Analysis, Probability Theory, Mathematical Statistics, Data & Web Technologies for Data Analysis, Web Science Research Methods, Computational Statistics, Topics in Applied & Computational Statistics	
ESRI – MOOC	09/ 2020 – 12/2020
<i>Spatial Data Science Certificate</i>	
<ul style="list-style-type: none">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)	02/2020 – 08/2020
<i>Data Analytics Certificate</i>	
<ul style="list-style-type: none">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	09/2014 – 07/2019
<i>Bachelor of Science (Geoscience)</i>	
<ul style="list-style-type: none">Research Assistant (Department of Earth and Environmental Sciences)Event Coordinator (Earth and Environmental Sciences Club)Scholarship Award Recipient (East Bay Mineral Society)	

PROJECTS

Popular Music Analysis

03/2023

<https://kayannr.github.io/top-hits/> | <https://github.com/kayannr/top-hits>

Analysis of popular songs across various countries that involves data collection using API, data wrangling, data visualization, descriptive and exploratory analysis, topic modeling, feature selection, predictive modeling, and presenting findings in an HTML-based report.

- Methods and Tools used: **Python, RegEx, HTML5, NumPy, Pandas, API, Requests, NLTK**

Sports Statistics

09/2022

<https://github.com/kayannr/sportstats>

Historical Olympics games data analysis that involves data wrangling, data quality assessment, data deduplication, data visualization, hypothesis testing, descriptive and exploratory analysis, and inferential analysis using graphical and statistical methods.

- Methods and Tools used: **SQL, Databricks, Python, PandaSQL, Pandas, SQL Window functions**

Word Wright

12/2020

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

An interactive web application developed for Devpost Hackathon that includes a story prompt generator based on choice of writing style using a Natural Language Generator model.

- Methods and Tools used: **SQL, Python, PyTorch, JavaScript, HTML, CSS, Bootstrap, RegEx, ETL, Web API, React.js, Data Wrangling**

Music Genre Prediction

08/2020

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

A web-based interactive application that utilizes Natural Language Processing to predict song genres based on sound features and lyrics, which involves front-end tasks, front-end to back-end connectivity, as well as data collection and cleaning

- Methods and Tools used: **NLP, KNN, SQL, Python, JavaScript, HTML, CSS, Bootstrap, ETL, Web API, Heroku, Flask**

US Homeless Population

07/2020

<https://mengye22.github.io/US-homelessness/master/templates/index.html>

An analysis of homeless population in the US showcased on an interactive website that involves data collection, data wrangling, ETL, and interactive visualizations

- Methods and Tools used: **Web-scraping, RegEx, ETL, JavaScript, Python, HTML5, CSS, D3.js, Leaflet**

PROFESSIONAL EXPERIENCE

Graduate Student Assistant

07/2021 – 12/2022

UC Davis

- Provided effective teaching support for business analytics, accounting, and entrepreneurship courses in the Department of Biomedical Engineering and Graduate School of Management

Project Geologist

09/2019 – 06/2021

ATC Group Service LLC

- Maintained positive relationships with clients by effectively communicating progress and relevant information, ensuring a clear understanding of their requirements
- Assisted with project management and technical report writing, resulting in 30% increase in efficiency and cost savings of up to \$1000
- Conducted analysis of historical environmental data to provide optimal location recommendations

Research Assistant

09/2018 – 11/2020

CSU – East Bay

- Conducted a thorough analysis of scientific data collected to establish a connection between sediment deposition and the fluctuations in sea level and climate in the San Francisco Bay Area

Geophysical Survey Intern

06/2018 – 07/2018

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
-