DAVID LAPAGLIA

Denver, CO │ david.lapaglia@colorado.edu │ 240-808-9496

Detail-oriented, data and information scientist, who strives to generate actionable insights from data.

EDUCATION

Bachelor's of Science in Information Science │ GPA 3.9 │ Dean's List

- Minors in Mandarin Chinese and Business
- University of Colorado Boulder, Boulder CO

RELATED COURSEWORK

Programming for Information Science 1 & 2, Applied Machine Learning, Data Visualization, Web Data Science, Mathematical Concepts in Data Science, Information Exploration, Information Ecosystems, Statistics, Foundations of Data Science, Music as Information, GIS: Mapping (ArcGIS)

WORK AND RESEARCH EXPERIENCE   Summer 2024 - Present

Junior Data Manager │ National Oceanographic Atmospheric Administration (NOAA)/National Center Environmental Information (NCEI) – Cooperative Institute Research Environmental Science (CIRES)

- Simultaneously managed multiple ingestion scripts for passive acoustic data (PAD), archiving 50+ TBs
- Programmed 10+ Python scripts to aid the tasks of the data manager
- Led over four major UI/UX updates and bug testing for NCEI Passive Acoustic Collection Engine that led to increased productivity of data collection.
- Engineered revolutionary ways to manage hard drives in the PAD archive.

Center for Asian Studies Assistant │ UNIVERSITY OF COLORADO BOULDER   February 2022 - Present

- Implemented video editing requirements.
- Spearheaded design and integration of surveys and Google Sheets, improved productivity by 10%.
- Developed web scraping scripts that reduced the manual NRC reporting process from 5 hours to near-instant completion, achieving a 1800% increase in efficiency.

Code Sensei │ CODE NINJAS   February 2020 - January 2021

- Taught computational reasoning skills with JavaScript to children ages 8-12.

PROJECTS

Traffic Light Recognition Convolutional Neural Network (CNN)

- Engineered a CNN using Tensorflow, reaching 99% F1 score.
  - Practicing data cleaning and standardization techniques.
- Created a comparative model (support vector machine) with scikit-learn machine learning in Python.

Visualizing Emission Trends Across Countries

- Utilized the Altair package, created dashboards on carbon emissions by country.
- Analyzed type of emissions by region, identifying the most problematic places and emission type.

Beats and Bytes: Decoding the Data Behind Song Popularity

- Drawing from Python visualization packages, conducted an in-depth analysis on music data.
- Identified key song features that drive overall song popularity, creating interactive information artifacts.

RELEVANT SKILLS

Technical: Python, SQL, APIs, Data Mining (Web Scraping), Supervised and Unsupervised Machine Learning Models, TensorFlow, Data Visualization, Data Cleaning, Statistical Analysis
Software: PowerBI, Tableau, Figma, Microsoft Office (Word, Excel, PowerPoint), Google Suite
Languages: High Proficiency in Mandarin Chinese, Conversational French