Andy Pickering, Ph.D

Data Scientist

Golden, CO | 978-764-3009 | andypicke@gmail.com | [LinkedIn /andypicke](#) | [GitHub /andypicke](#)
[Data Science Blog](#) | [Google Scholar](#)

SUMMARY

Results-oriented energy analyst with 2+ years of data science experience with a background in climate and oceanography research. Expertise in quantitative analytics, data visualization, and effective stakeholder communication. Proficient in SQL, R, and Python for data analysis and research. Seeking a hands-on role at a Denver-based energy company to help drive the clean energy transition.

TECHNICAL SKILLS

Coding Languages: R (RStudio, tidyverse, Plotly, Leaflet, Shiny, R-Markdown), SQL, Python (Pandas, NumPy, SciPy Stats, MatPlotLib, Seaborn, Scikit-Learn), Matlab, Unix

Data Science: Statistics (Regression Models), Machine Learning (Decision trees, random forest)

Other: Github, VScode, Quarto, SVN, Docker, JIRA, AWS, Agile Methodology

PROFESSIONAL EXPERIENCE

Sabbatical   Nov 2019 - Nov 2023
- Took time off to upskill and pursue personal and professional development opportunities. Built multiple [data science projects](#) related to clean-energy transition, electric vehicles (EV), weather and climate sectors.
- Completed 6+ Data Science and professional development courses including Learning Git and GitHub, Unix Essential Training, Building Data Apps with R and Shiny: Essential Training, Intermediate SQL for Data Scientists, Creating Maps with R, andCommunicating with Emotional Intelligence

Data Scientist & Data Analytics Supervisor - ICF - Golden, CO   Sept 2017 - Oct 2019
- Analyzed energy utility customers' energy usage with R and SQL to create home energy reports and drive messaging for energy efficiency and rebate programs.
- Delivered relevant offers and recommendations to customers for reducing energy usage and upgrading equipment, based on modeling of their heating and cooling energy usage.
- Created automated reports in R-Markdown to summarize data analyses and communicate results to both technical and non-technical stakeholders.
- Analyzed smart thermostat and weather data to provide customized recommendations for saving energy.

Postdoctoral Research Associate - Oregon State University   Jan 2015 - June 2017
- Developed data-processing pipeline and analysis for turbulence measurements made by novel instruments during standard shipboard sampling, with the goal of greatly increasing our understanding of turbulent mixing across the global oceans.
- Collected and analyzed data from an autonomous research vessel operating in coastal and open-ocean waters, including location and speed from an arduino gps logger and water velocity from an Acoustic Doppler Current Profiler (ADCP).
- Created and maintained [Github repository](#) for standard, version-controlled codes related to ocean mixing, for use by the oceanography community.

Graduate Research Assistant - University of Washington   Sept 2008 - Dec 2014

- Analyzed time series and spatial data from a variety of instrument platforms and models.
- Applied time-series analysis techniques including spectral analysis and harmonic fitting to time series of oceanographic data (temperature, salinity, velocity) in order to isolate and study waves generated by different tidal frequencies.
- Used harmonic-fit techniques on velocity measured by multiple shipboard transects to extract the spatial structure of internal tide beams generated by tidal flow over the Hawaiian Ridge.
- Communicated results through peer-reviewed journals and presentations at scientific conferences.

DATA ANALYTICS PROJECTS

Calculating State Electricity Generation By Fuel Type | R (httr, dplyr, ggplot2, Plotly)   2023

- Used httr package to read data from the U.S. Energy Information Administration (EIA) API
- Calculated the percent of total electricity generation by fuel type for Colorado using dplyr, and visualized results with ggplot2 and Plotly.
- Utilized ggplot2 to visualize how electricity generation has transitioned to cleaner fuels over time.

Parking Prediction Analysis | Python (Pandas, Matplotlib, SciKit-Learn)   2020

- Analyzed LotSpot parking data and weather data to identify usage trends and predict the number of spaces available at popular trailheads.
- Built a random-forest regression model with R^2 of 0.64. The most important predictors were temperature and UV index.

EDUCATION

Certificate of Data Science | Galvanize | Remote   June 2020
12-week intensive Python-based curriculum covering best practices in machine learning, statistical analysis, natural language processing, and data visualization.

Ph.D. & M.S. in Physical Oceanography | University of Washington, Seattle WA   Dec 2014
Ph.D. Thesis: Investigation of the Spatial and Temporal Structure of Internal Waves

M.S. Thesis: Near-Inertial Waves Observed During the Internal Waves Across the Pacific Experiment

BS Physics & Geology | Northeastern University, Boston MA   June 2008