Emiko Sano

530-329-4631 | emiko.b.sano@gmail.com | ebsano.github.io | linkedin.com/in/ebsano

SKILLS

Programming Languages: Python | R | PostgreSQL

Tools: Git | Jupyter | Excel | Tableau

Python Libraries: Pandas | NumPy | Matplotlib | Seaborn | Statsmodels | Pingouin | Scikit-Learn | NLTK | BeautifulSoup

Machine Learning: Linear Regression | Logistic Regression | Decision Tree | Random Forest | Gradient Boosting | PCA

Statistical Methods: Statistical Inference | t-Test | ANOVA | Tukey's HSD Test | A/B Testing

EXPERIENCE

Correlation One   Remote
Data Science Teaching Assistant   2021 – Present
- o Guide 40 students through topics in **data analytics** (data wrangling, visualization, hypothesis testing, statistical inference, A/B testing, regression analysis) and debugging **python** codes
- o Support 7 team projects in sourcing data, project scoping, and analyzing data to solve business problems and deliver dashboards/datafolios and reports in 13 weeks

Freelance   Missoula, MT/Remote
Science Expert Reviewer @ Study.com | Private Biology Tutor   2018 – Present
- o Reviewer: Editing and fact-checking of educational science content (biology, chemistry, math) to bring the material to Study.com's standards while providing positive and constructive feedback to content writers
- o Tutor: Prepare material and break down complex concepts for biology high school students

The University of Montana   Missoula, MT
Postdoctoral Researcher   2013 - 2018
- o Led a microbial evolution research project from **experimental design** to **statistical analysis**
    - ⇒ Used **R** and **Excel** for analysis; t-test, ANOVA, Tukey's HSD test
    - ⇒ Resulted in a publication in *Nature Ecology and Evolution*
- o Produced preliminary findings that contributed to the award of a multi-lab, $8.2 million grant
- o Supervised all lab personnel (13 total) of all levels to execute experiments and meet project milestones

University of California, Davis   Davis, CA
Postdoctoral Researcher   2012 - 2012
Graduate Student Researcher   2005 - 2011
- o Conducted microbial genetics research (experimental design, data collection, and statistical analysis)
    - ⇒ Resulted in a publication in a peer-reviewed journal (*Genetics*) and issue-highlighted article
    - ⇒ Conducted **hypothesis testing** and applied **regression analysis** in Excel
- o Presented research results at local to international scientific conferences in oral or poster presentations

DATA SCIENCE PROJECTS

Identifying factors that affect students' performance using NYC schools survey data
- o Used **machine learning** models (PCA, linear regression, gradient boosting) in **Scikit-Learn** & found that the perception of safety in schools is important in how well students perform
- o Recommended schools to implement programs to address safety/bullying at schools

Predicting the origin of subreddits based on the text in posts
- o Cleaned and analyzed text from 15,000 posts using **NLTK**, **Countervectorizer**, or **TfidfVectorizer**
- o Optimized **classification models** (Logistic Regression, Decision Tree, Random Forest) reaching ~90% accuracy

EDUCATION

University of California, Davis   *Ph.D. Microbiology* | Davis, CA
San Diego State University   *B.S. Biology with an emphasis in Cellular and Molecular Biology* | San Diego, CA
General Assembly   *Certificate of Completion in Data Science* | Remote

VOLUNTEER EXPERIENCE

Dat0s | A Minority Data Community   Remote
Engagement Manager | Events Coordinator   2021-Present
- o Helping other data scientists with data projects, coordinating and hosting monthly social hour and career event