

Sarah Jenner has completed the 2020 Stanford Pre-Collegiate Summer Institutes course in

Introduction to Data Science: Visualization and Modeling, Session 3

from Monday, July 20 - Friday, July 31, 2020, online with Stanford University. Sarah's admission into the Summer Institutes, participation in and successful completion of the program, demonstrate a high level of academic ability and intellectual curiosity.

Course Description:

Data science has revolutionized the way our world works and how we understand it. Technology enables us to ask more questions of more data, but how do we go about using these tools effectively and ethically?

This course will introduce students to computer algorithms and the diversity of models they can generate, each with pros and cons. Students will use datasets from the natural and social sciences to answer real-world questions, pursuing questions and data relevant to their own lives. They will apply different facets of machine learning through R programming exercises deeply integrated into the course. By the end of the course, students will have developed a technical skillset that allows them to investigate any given dataset with strong coding abilities and a scientific approach.

Student Evaluation (from Instructor):

Sarah Jenner's did not have much coding or statistics experience before the start of the class, but her love of learning and boldly pursuing anything new was evident. She may apply the techniques of analyzing data and the process of executing scientific research to her Marine Biology interests in the future, but in the mean time, Sarah was true to the "learn something new" spirit: she approached data science by choosing to analyze obstacles on the show "American Ninja Warrior" for the final, open ended topic research project for the class.

In the very short two weeks of the class, she learned so much about how to read data into R, how to prepare it for machine learning analysis, the importance of splitting the data into training and test datasets, and how to choose a model for best predictive or classification analysis. Sarah also published the results of her analysis for the short project where students examined COVID-19 cases in California and visualized the data in her online dashboard.

Sincerely,



Victoria D'Urso