

Dolciani Mathematics Enrichment Grants

Proposal title

Mathematical Data Science for High School Students: A University Experience

Introduction

In the past the topics of Data Science and Statistics have been underrepresented in the K-12 Mathematical curriculum. Because data are increasingly important and widely available there is an increased need for students to acquire these new-century skills for analyzing and interpreting data. This trend is apparent in the new adaptation of the Common Core Standard where statistics and probability are given a greater emphasis. Students will be exposed to more statistical thinking and methods starting at an earlier age than more in depth as they enter high school. However, because Statistics was not given enough attention previously, teachers are not confident enough to give students the proper experience in the methods to analyze data with the new curriculum. This proposal aims to enrich students' mathematical experience with a more in-depth look at data science at a university environment.

In this project, Dr. Nicolas Christou, a faculty member of the UCLA Department of Statistics, will partner with Mr. Robert Montgomery, a Mathematics high school teacher from Edward Roybal Learning Center (Los Angeles Unified School District) to provide 30 high school students with an experience in Data Science at UCLA. This project will also involve graduate students in Statistics as well as undergraduate students in Mathematics and Statistics as classroom facilitators.

Settings

The Edward Roybal Learning Center is an urban high school located in the center of downtown Los Angeles. The school has a population of approximately 1200 students where 99 % are minority and 81 % are from economically disadvantaged families. The high school Mathematics teacher and students from Edward Roybal Learning Center will be invited to UCLA on three Saturdays in the Fall 2018 quarter with transportation provided through the grant.

Students will be picked up from the high school at 8:00 AM and travel to UCLA. Session will run from 9:00 AM to 2:00 PM each day and with the students returning to the school to be picked up by parents by 3:00 PM. The day's schedule will be as follows:

| | |
|---------------------|---|
| 9:00 AM - 10:30 AM | Whole class instruction planned by the university professor and the high school teacher |
| 10:30 AM - 10:45 AM | Break |
| 10:45 AM - 12:00 PM | Students will be divided into two groups for a small group exploration lead by one graduate and two undergraduates under the direction of the teacher and the professor |
| 12:00 PM - 01:00 PM | Lunch. |
| 01:00 PM - 02:00 PM | College experience which will include different activities on each Saturday (campus tour, student panel, small group mentoring.) |

To monitor the effectiveness of our efforts and to better plan each session to address the needs of our students, we will administer a pre- and post-surveys. There will be an overall

experience survey as well as evaluation survey at the end of each session. This data will provide us feedback to better plan for future events.

Topics covered

Day 1

The topics covered in day 1 will set the stage for the applications covered in day 2 and day 3. In day 1 students will learn the mathematical formulation of simple and multiple regression and properties of covariance and correlation. The topics include the method of least squares, the Gauss-Markov theorem, method of maximum likelihood, distribution theory in regression, and theory- and simulations-based inference. The assumptions of regression will be discussed as well as remedies to apply when they do not hold. Examples using the open source statistical software R and the Statistics Online Computational Resource (SOCR) will be used to strengthen students' understanding.

Day 2

Day 2 will focus on applications of the material learned in day 1, specifically, portfolio risk management. Most, if not all, students have some familiarity with the stock market (or at least they have heard about it) and therefore can better relate to the problem. During this session we explore the construction of efficient portfolios through optimization using examples of two and three stocks, and how covariance and correlation can help an investor to diversify their risk. We will discuss why diversification works, but also the problems that can arise in portfolio management.

Small group activity: Students will be using R and SOCR to access online stock market data and perform portfolio optimization.

Day 3

Similar to day 2 students will be exposed to another application of the material learned in day 1. Spatial statistics is a discipline that employs various techniques to analyze data collected at geographic locations defined by their coordinates (longitude and latitude). This type of data is found in the study of earth sciences, defense, cartography, biology, ecology, hydrology, epidemiology, sociology, and more. Spatial statistics is relevant to our efforts to address several problems facing society today. Important examples include, environmental hazards such as pollution and the effects of climate change which are becoming increasingly pressing. The field of spatial statistics provides the appropriate analytical tools to needed to draw informed conclusions, and the real world applications of spatial statistics have important ramifications to the future of many public policies and social issues.

Small group activity: Use environmental data to make predictions and construct a surface of the response variable of interest with SOCR and R.

Discussion

The collaboration between the university professor and the high school in planning the three sessions will help build content knowledge for the teacher to more confidently teach statistical topics in the classroom. The project also incorporates graduate and undergraduate students to facilitate discussions in small group settings and enhance their understanding of the materials learned. These university students will also act as mentors to the high school students during this time. The afternoon activities are designed to be an active experience

where students will be exposed to UCLA and truly learn about what the university has to offer and to encourage their interest in university life. Students from under-represented communities often think a university such as UCLA is out of their reach. We hope that these three days will help them feel UCLA is a part of their future.