

Introduction to Data Science

Grades: 11-12

Credit: 1

Level: CCP 1

Department: Mathematics

Prerequisite: Successful completion of Algebra II or Co-Enrolled in Algebra II.

COURSE DESCRIPTION

Introduction to Data Science (IDS) will introduce students to the main ideas in data science through free tools such as but not limited to Google Sheets. Students will learn to be data explorers in project-based units, through which they will develop their understanding of data analysis, sampling, correlation/causation, bias and uncertainty, probability, modeling with data, making and evaluating data-based arguments, the power of data in society, and more! This instructional model involves students in problem-solving investigations and other meaningful tasks, allows students to construct their own knowledge through inquiry, and culminates in a project to showcase their newly developed abilities. The CT Core Standards for High School Statistics and Probability relevant to data science are taught along with the data demands of good citizenship in the 21st century, including alignment with CSTA K-12 Computer Science Standards and Applied Computational Thinking Standards.

COURSE GOALS

Students will

- Engage in data analysis, computing, and mathematics through hands-on activities.
- Engage in practical applications of data analysis to give students concrete and applicable skills.
- Find and communicate meaning in data, and to think critically about arguments based on data.

COURSE STANDARDS

CT Core Content Standards for Mathematics:

- **HS Probability & Statistics**
- **Standards for Mathematical Practices**

The Computer Science Teacher's Association (CSTA): K-12 Computer Science Standards

Applied Computational Thinking Standards (ACT): Computational Thinking Competencies

ESSENTIAL QUESTIONS FOR THE COURSE

- What are variability, data, and models?
- What is the story I can tell from a set of data?
- How can univariate and bivariate data be described and summarized?
- How do statisticians make predictions based on data?

- How do you determine if a linear model is appropriate to represent the data?
- How can probability be applied to daily life?
- How can you best represent categorical data?
- How can a model be created to best represent the situation?
- What does it mean to normalize data in the real world?
- What are sources of bias and how can they be addressed?
- How can conditional probability be applied to algorithms to make predictions?
- How can you use data to make decisions and communicate findings effectively?

ENDURING UNDERSTANDINGS FOR THE COURSE

Students will keep considering.....

- The many ways to organize, display, interpret, and compare data.
- A relationship between two sets of data provides important information in many real world situations.
- Theoretical probabilities, experimental probabilities, and probability simulations can provide insight into real life situations.
- Mathematical models can be used to represent scenarios and make and justify decisions.
- The media has an impact on how data is interpreted.
- Providing different types of data show different aspects of the situation.
- How technology can impact and aid data analysis.
- How the data science process provides insight into the world.

COURSE CONTENT OUTLINE

- I. Unit 1: Data Tells a Story
- II. Unit 2: The Data of our Community
- III. Unit 3: Water in your life
- IV. Unit 4: Shuffle Songs
- V. Unit 5: Color Preferences and Representations
- VI. Unit 6: What's the best place for me?
- VII. Unit 7: Predicting my preferences
- VIII. Unit 8: Being a Data Scientist