

11 Data Science - Being a data Scientist

Content Area: **Math**
Course(s):
Time Period: **Full Year**
Length: **3 weeks**
Status: **Published**

General Overview, Course Description or Course Philosophy

This course combines the study of Statistics and Probability with Data Science. The goal is to have students think critically about data in today's data-driven world and understand its role in the 21st Century economy. Furthermore, students will become familiar with the concepts, topics, and techniques used by data scientists and statisticians in their day-to-day work.

Throughout this course, students will engage in project-based observational studies and experiments to develop their understanding of data analysis, sampling, correlation/causation, bias and uncertainty, probability, modeling with data, as well as making and evaluating data-based arguments. Students will also learn about the roles of data scientists, the power of data in society, machine learning, and how data scientists extract knowledge and insights from real-world data.

In this unit students will have an opportunity to work through the full cycle of data science: making their own decisions about the questions they are interested in exploring, finding data to answer that question, cleaning the data, creating and analyzing a model, communicating with the data visually and reflecting on their process. This will be an iterative process mirroring how data scientists work on a project. Students will gather their own data. They will make decisions about how to work with it and describe the choices they have made including what technology tools to use, cleaning moves, visualization selection, univariate or bivariate data choices, combining data, and other content relevant to their project of choice.

OBJECTIVES, ESSENTIAL QUESTIONS, ENDURING UNDERSTANDINGS

Essential Questions

- **How can I use the full cycle of data science to pose and solve a guiding question that interests me?**

Enduring Understandings

- **By asking questions, gathering and organizing data, modeling analyzing and synthesizing, it is possible to gain immense insight into an area of interest.**
- **Communicating the results of the project, along with method in which the data was**

obtained, and the process in which the conclusion was deduced is a vital component to the data science process.

CONTENT AREA STANDARDS

S.ID

- A. Summarize, represent, and interpret data on a single count or measurement variable**
- B. Summarize, represent, and interpret data on two categorical and quantitative variables**
- C. Interpret linear models**

S.IC

- A. Understand and evaluate random processes underlying statistical experiments**
- B. Make inferences and justify conclusions from sample surveys, experiments, and observational studies**

S.CP

- A. Understand independence and conditional probability and use them to interpret data**
- B. Use the rules of probability to compute probabilities of compound events in a uniform probability model**

S.MD

- A. Calculate expected values and use them to solve problems**
- B. Use probability to evaluate outcomes of decisions**

CS.K-12.2.a	Cultivate working relationships with individuals possessing diverse perspectives, skills, and personalities.
CS.K-12.2.b	Create team norms, expectations, and equitable workloads to increase efficiency and effectiveness.

CS.K-12.2.c	Solicit and incorporate feedback from, and provide constructive feedback to, team members and other stakeholders.
CS.K-12.2.d	Evaluate and select technological tools that can be used to collaborate on a project.
CS.K-12.3.b	Decompose complex real-world problems into manageable sub-problems that could integrate existing solutions or procedures.
CS.K-12.3.c	Evaluate whether it is appropriate and feasible to solve a problem computationally.

RELATED STANDARDS (Technology, 21st Century Life & Careers, ELA Companion Standards are Required)

9.1.12.PB.6: Describe and calculate interest and fees that are applied to various forms of spending, debt and saving.

LA.RI.11-12.7	Integrate and evaluate multiple sources of information presented in different media or formats (e.g., visually, quantitatively) as well as in words in order to address a question or solve a problem.
LA.11-12.SL.11-12.2	Integrate multiple sources of information presented in diverse formats and media (e.g., visually, quantitatively, orally) in order to make informed decisions and solve problems, evaluating the credibility and accuracy of each source and noting any discrepancies among the data.
CRP.K-12.CRP2	Apply appropriate academic and technical skills.
CRP.K-12.CRP4	Communicate clearly and effectively and with reason.
CRP.K-12.CRP7	Employ valid and reliable research strategies.
CRP.K-12.CRP8	Utilize critical thinking to make sense of problems and persevere in solving them.
CRP.K-12.CRP11	Use technology to enhance productivity.
TECH.8.1.12.A.4	Construct a spreadsheet workbook with multiple worksheets, rename tabs to reflect the data on the worksheet, and use mathematical or logical functions, charts and data from all worksheets to convey the results.
TECH.8.1.12.A.5	Create a report from a relational database consisting of at least two tables and describe the process, and explain the report results.
TECH.8.1.12.C.CS2	Communicate information and ideas to multiple audiences using a variety of media and formats.
TECH.8.1.12.C.CS4	Contribute to project teams to produce original works or solve problems.
TECH.8.1.12.E.CS2	Locate, organize, analyze, evaluate, synthesize, and ethically use information from a variety of sources and media.
TECH.8.1.12.E.CS4	Process data and report results.
TECH.8.1.12.F.CS3	Collect and analyze data to identify solutions and/or make informed decisions.

STUDENT LEARNING TARGETS

Declarative Knowledge

Students will understand that:

- Data can be used to solve and answer questions about a specific problem
- Peer feedback is an essential element of the design process
- The project design can be broken up into manageable parts, with each part being comprised of goals.
- The final element of the project will involve presenting the projects to the class, answering questions and receiving feedback from peers.

Procedural Knowledge

Students will be able to:

- Examine their portfolios and reflect on their learning over the course in support of designing their final project.
- Build from their reflection of their learning and an exploration of Data Resources available, and brainstorm topics for the final project.
- Work with a partner on their project proposal.
- Work through a feedback protocol to present their ideas and questions and receive feedback and support from their peers.
- Present their project to a peer and get feedback on their project based on the criteria.
- Present their final projects to the class.

EVIDENCE OF LEARNING

Alternate Assessments

- Portfolios
- Verbal Assessment (instead of written)
- Multiple choice
- Modified Rubrics
- Performance Based Assessments

Formative Assessments

Observations

Task completion

Student journals and notebooks

Cooperative team work

Summative Assessments

PBL Assessment

Unit assessments

RESOURCES (Instructional, Supplemental, Intervention Materials)

<https://hsdatascience.youcubed.org/curriculum/>

Unit 8

INTERDISCIPLINARY CONNECTIONS

Educational tech applications

Current Events

Experimentation

ACCOMMODATIONS & MODIFICATIONS FOR SUBGROUPS

See link to Accommodations & Modifications document in course folder.