

# Program Learning Outcomes

Short Title	
Math	Explain the mathematical foundations of data science, including probabilistic reasoning as well as Bayesian and Frequentist statistics, and utilize these techniques to solve data science problems.
Programming	Independently create computer programs which analyze complicated real-world datasets, as well as explain and modify data science programs written by others.
AI/ML	Identify and apply machine learning and artificial intelligence techniques to effectively solve real-world problems; describe how these techniques work on a technical level.
Tools	Identify the optimal data science tools required for various analytical procedures and data visualization tasks.
Limits	Describe the appropriate usage and limits of data science, e.g. explain what kinds of questions can be asked and answered versus those that cannot be addressed.
Types	Manipulate a wide variety of common data types to effectively accomplish an objective: e.g., traditional experiments, image collections, natural language text, and real-time time-series processing.
Security	Articulate privacy and security and ethical issues surrounding data of various types, and describe approaches to mitigate these issues
Formats	Retrieve and store data in a variety of different formats, including databases and popular file formats.
Domain	Apply and integrate knowledge from numerous disciplines, including those outside of mathematics and computer science such as business, natural sciences, health, and social sciences.
Software	Utilize existing data science and informatics software effectively.
Communication	Communicate insights from large datasets to others, in written, oral, and visual forms.

## B.S., Data Science, Curricular Map

I = Introduced, D = Developed & Practiced with Feedback, M = Demonstrated at the Mastery Level appropriate for graduation in this program

	Math	Programming	AI/ML	Tools	Limits	Types	Security	Formats	Domain	Software	Communication
DATA/CS 171 Data Science Fundamentals in R	I	I		I		I	I	I	I	I	I
DATA/CS 172 Python for Data Analysis		I		I		I	I	I	I		
MATH 211 Matrix Algebra With Applications	D				I	I					
MATH 241: Calculus 1	D										
DATA/MATH 271 Applied Statistics with R	D	I		D	D	I		I	I	I	I

DATA/CS 272 Machine Learning for Data Sci	D	D	D	I	D	D	I	I	D		D
DATA 315 Mathematical Methods for Data Sciences	M	D	I	I	M	D					
DATA/CS 370 Data Management		D						M		D	
DATA/MATH 371 Multivariate Modeling with R	M	D		D	D	M			D		I
DATA/CS 373 Data Security & Privacy	D	D	D	D			M				
DATA/QBA 465 Text Mining for Behavior and Social Sciences		D	D	M		D		D	D	D	
DATA 470 3D Mapping of Ecosystems		I	I	M		D		D	D	M	
DATA 474 Applied Informatics	I	I	I	M		D		D	D	M	D
CS 475(L) Data Visualization (Lab)		I	I	D		D		D	D	D	D
DATA/CS 483 Computer Vision	D	M	M	I		D		D	D		
DATA 490 Data Science Capstone	D	D	D	M	D	M	D	M	M	M	M
DATA 495 Data Science Seminar	D		D								D

**Rubric for Types, Formats, Domain, and Communication PLOs, to be assessed through final presentations/papers (in the capstone course, DATA 490)**

Score	Types	Formats	Domain	Written Communication*	Oral/Visual Communication*
Mastery 3	Presents an <i>effective and appropriate</i> analysis examining the relationships between multiple different types of data relevant to the project objective.	Presents systems that effectively read and produce in a wide variety of formats. Choice of data formats is well-justified and appropriate to the task.	Presentation evidences deep understanding of the domain as it relates to the project. The needs of the clients/domain experts are considered at every step in the design and evaluation of the system.	Report is well-written. Motivation is strong and assumptions are clearly stated. Relevant work is cited and discussed. Reasoning is clear without having to read large amounts of background material. Technical details are presented effectively without being glossed over. Results and conclusions are adequately discussed, including a discussion of limitations.	Presentation is clear and compelling. Introduction is strong and captures the audience's interest. Visuals are effective in communicating both the processes/systems used as well as their results. Presenters explain technical concepts well, and take into account the intended audience. Presenters seem to possess a deep knowledge of the topic, and can confidently answer questions about it.
Competent 2	Presents an effective and appropriate analysis of multiple types of data to achieve the project objective, but how the different types of data are related is not well-explored.	Presents systems that effectively read and produce data, but almost exclusively in one format. Choice of data format is well-justified and appropriate to the task.	Presentation evidences adequate surface-level understanding of the domain as it relates to the project. The needs of the clients/domain experts are considered at the most important places in the design and evaluation of the system, although there may be some places that do not fully align with domain needs.	Report is generally well-written, although some parts may be unclear. Motivation and assumptions are presented. Reasoning may require reading background material to grasp. Technical details are largely discussed, although a few may be glossed over. Important results, conclusions, and limitations are discussed, although some points may have been missed.	Presentation is clear, although it could be more compelling. Introduction makes clear what the presentation is about. Visuals are effective in communicating the results, but the visuals related to how the systems work may be less clear. Presenters explain technical concepts correctly, although sometimes not at the right level for the audience. Presenters seem to possess adequate knowledge of the topic, and can correctly answer questions about it.
Beginning 1	Only presents an analysis of a single type of data - or attempts to analyze multiple, but analysis is not executed competently or not relevant to the project objective.	Presented system may attempt to read and/or produce data in a given format, but may contain serious flaws. Data format may not be justified at all and/or may not be appropriate for the task.	Presentation indicates that team members have misunderstood key aspects of the domain, which have ultimately harmed the value of their project. The needs of the clients/domain experts have been misunderstood or largely disregarded during the design of the system.	Report is not well-written and many key aspects may be missing or unclear. This could include a lack of motivation, unstated assumptions. Reasoning may be difficult to following even with the aid of background material. Technical details are poorly explained or glossed over. Important results, conclusions, and limitations may be omitted or discussed only in a vague manner.	Key aspects of the presentation are unclear and/or incorrect. Introduction may give an unclear or incorrect sense of what the project is about. Visuals are inadequate to effectively communicate the results to the audience. Presenters give incorrect explanations for key technical concepts, or gloss over key aspects. Presenters cannot answer questions about the topic, or answer questions incorrectly.

\*For the purposes of rating, we feel it is most useful to assess the Written and Oral/Visual aspects of the communication PLO separately, since the students will submit a separate report and oral presentation.