Institution: Rutgers-New Brunswick (RUNB)

Course title: Data 101: Data Literacy

Course number: 01:198:142 or 01:960:142

Course type: Undergraduate

Course credits: 4

Prerequisite: placement in 01:640:026 (placement into Intermediate Algebra or higher, or completion of Math 025. No programming experience required).

Note: This course may not be used toward RUNB Computer Science (CS) major degree credit.

Expected Work: Weekly Data Puzzles, Presentations, Homework, Final Project.

Exams: Midterm exam and final exam.

Description: This class aims to provide you with a basic set of tools for data literacy as well as general view of the impact of data on society and elements of common-sense data analysis and reasoning. A significant piece of the class will be learning foundations of R. R is a statistical software environment and programming language that we'll use to analyze and visualize datasets. Learning simple R will take some work; however, if you're able to master the basics covered in this class, you'll gain a concrete, marketable skill that may very well be extremely useful in your academic and professional life.
On the statistical side, we'll cover basic topics from statistics and probability that are required to argue persuasively using data (a list of some of the topics to be covered can be found below). This is not a "typical" Statistics 101 class; instead of covering an exhaustive list of topics and asking you to memorize many formulas, our goal is to focus only on the most important topics for convincingly analyzing data now by solving "hands on" weekly data puzzles.
This class is taught in unique manner – students have to solve "data puzzles" (one or more weekly) and defend their solutions in class in the so called "Court of Data." Students compete in the semester long competition for the titles of Data Masters (aggregated score for all data puzzles and the project).
One of the objectives of the class is to show the danger of false, random conclusions from data and learning right methodology of "healthy skepticism."
We will also discuss how not to be fooled by data and show examples of rushed and ad hoc conclusions from so called "big data" in the news and on the web. In addition, we will examine both upside and downside of big data on the web. We will talk about privacy, anonymity vs personalization and data ownership when we increasingly rely on online services.
In Final project your data findings should have real consequences, preferably "actionable" and consequential in the real society.

Learning Outcomes:
- Analyze and critically assess information from traditional and emergent technologies - data analysis in meaningful way.
- Apply effective and efficient mathematical or other formal processes to reason and to solve problems - finding patterns in data.

- Formulate, evaluate, and communicate conclusions and inferences from quantitative information - present and defend your results.
- Analyze the relationship that science and technology have to a contemporary social issue.

Topics List/Calendar (weekly schedule subject to change).

Week 1:	Intro to Data Science, Setting Up R.
Week 2:	Basic R Instructions: working with data in R. Simple queries in R.
Week 3:	Data Visualization and Data Transformation-Exploration with Derived Attributes.
Week 4:	Data science vs Databases, Fooled by Randomness (law of small numbers).
Week 5:	Normal Distribution, Central Limit Theorem.
Week 6:	Hypothesis Testing: Permutation test, z-test.
Week 7:	Midterm Exams.
Week 8:	Multiple Hypothesis Testing.
Week 9:	Bayesian Reasoning, Common Sense Judgement and Probability (Kahneman availability, conjunctive fallacy etc.): Court of Data (student presentations on social issues such as Gun Control using data as argument).
Week 10:	Common Sense Judgement and Probability.
Week 11:	Prediction, rpart (recursive partitioning).
Week 12:	Linear Regression, other prediction models from R-library.
Week 13:	How can data fool us – paradoxes – Simpson paradox, Ecological fallacy, Prosecutorial Fallacy.
Week 14:	Power Law distribution.
Week 15:	Finals.