Course Name: MAT-114 Introduction to Data Science

Date Updated: 5/2025

Credit Hours/week: 3 hrs./wk. – 3 cr.

BEGINNING: FALL 2025

Catalog Description: Introduction to Data Science will provide students with data literacy skills in order to understand techniques in data manipulation, visualization and interpretation. This project based course will allow students to utilize a toolkit of statistical software to perform data science methods. Ethical issues related to data privacy, authenticity and security will be addressed alongside an introduction to artificial intelligence.

Prerequisite: MAT 016, MAT 120, or equivalent

Text: R for Data Science Authors: Garrett Grolemund and Hadley Wickham Free for students: **https://r4ds.hadley.nz/**

Supplementary Material: Big Data Demystified Author: David Stephenson PhD

Calculators are often (not always) permitted for use in mathematics classes. A scientific calculator is the most advanced calculator that you will need for any of your courses. An excellent (and affordable) model is the Texas Instruments TI-30XIIS 2 Line Scientific Calculator.

Graphing calculators are sometimes permitted. However, the TI Nspire is not an acceptable graphing calculator. Only TI 84 models or below are acceptable. Calculators that can solve equations or perform other advanced operations are not allowed. Therefore, no Casio model calculators are allowed. There are also some models of Texas Instrument calculators as well as other brands that will not be allowed due to their capabilities.

Use of a calculator that is not permitted on an exam can lead to issues with academic integrity. If you have a question as to whether or not your calculator is acceptable, you need to ask your professor before your first exam.

Syllabus:

Period	Text Chp.	Topics
1	1-2	Introduction- The Data Science Model
2	3.1-3.4	Data Visualization-Grammar of Graphics, problems with data
3	3.4-3.7	Data Visualization- Facets, objects, transformations
4	3.8-3.10	Data Visualization Exercises
5	4.1-4.3	Workflow Basics, coding, naming, functions
6	4.4	Workflow Basics Practice
7		**Test 1**
8-9	5.1-5.7	Data Transformation, filters, sub- setting, grouping, missing values
10-11	5.1-5.7	Data Transformation Exercise and Practice
12	6	Workflow Scripts
13-14	7.1-7.5	Exploratory Data Analysis-variation, missing values, covariation
15-16	7.6-7.8	Exploratory Data Analysis-patterns and models, ggplot2, examples
17		**Test 2 / Midterm Exam**
18-19	8-10	Wrangle Data - tibbles
20	11	Importing Data, parsing data, writing data
21-22	21.1-21.7	Tidy Data

23	13	Relational Data- Filtering and joins
24	14-16	Strings, Factors and Dates and Times
25		**Test 3**
26-27	Part 1	Big Data Demystified-define big data, artificial intelligence, neural networks, uses of big data, case studies
28-29	Part 2	Understanding the big data ecosystem-open source, strategies with big data, governance and legal compliance with big data
30		**Final Exam**
		Required technology (Excel, R, Tableau)

Students are expected to adhere to the policies of the County College of Morris. These can be accessed at: https://www.ccm.edu/compliance-information/academic-policies/

Statement of Expected Course LEARNING OUTCOMES

- Produce and interpret data visualizations, including dashboards, graphs, charts and maps to describe and explore large data sets.
- Produce and interpret numerical summaries to describe and explore large data.
- Investigate and explore relationships between more than two variables, multivariate analysis.
- Solve problems utilizing programming languages for data scientists.
- Apply exploratory data analysis and transformations to identify trends and problems within the data
- Clean and prepare data for analysis.
- Identify problems with messy and missing data.
- Recognize questions and problems that can be investigated using data.
- Communicate findings and outcomes based on data science techniques.
- Explain issues related to data privacy, security and authenticity.