

# Introduction to Exploratory Data Analysis

Lecture in which CLO is covered | Assignments (C-level) | Skill Checks (B-level) | Projects (A-level)

## Detailed Course Learning Objectives:

At the completion of this course, students should be able to:

1. Understand the basic structure and function of the *R* programming language.
  - 1.1. Understand the *R* Workspace.
    - 1.1.1. Understand how to use the command line. | L1.4 | SK1 | P1
    - 1.1.2. Understand how to use the help function of *R*. | L1.4 | SK1 | P1
    - 1.1.3. Understand the basic syntax of the *R* language. | L1.4 | SK1 | P1
    - 1.1.4. Execute inbuilt mathematical functions to perform calculations in *R*. | L1.4 | SK1 | P1
    - 1.1.5. Learn how to assign variables in *R*'s environment. | L1.4 | SK1 | P1
    - 1.1.6. Understand the basic syntax of functions in *R*. | L1.4, L1.6 | SK1 | P1
    - 1.1.7. Open, edit and save a script in RStudio's editor. | L1.4 | SK1, SK3 | P1
    - 1.1.8. Understand the concept of working directories. | L1.4 | SK1 | P1
    - 1.1.9. Create vectors, arrays, matrices, lists, and data frames. | L1.5, L1.6, L1.7, L1.8, L1.9 | SK1 | P1
    - 1.1.10. Understand vectors and vectorized calculations. | L1.5, L1.6, L1.7, L1.8, L1.9 | SK1 | P1
    - 1.1.11. Understand the data classes of *R*. | L1.5, L1.6, L1.7, L1.8, L1.9 | SK1 | P1
    - 1.1.12. Learn how to index vectors, arrays, matrices, lists, and data frames. | L1.5, L1.6, L1.7, L1.8, L1.9 | SK1 | P1
  - 1.2. Understand the basic principles of computer programming.
    - 1.2.1. Understand the way computers execute commands. | L1.3, L1.7, L3.4, L3.5 | SK1, SK3 | P1, P2
    - 1.2.2. Create functions in *R*. | L3.1, L3.7, L3.8 | SK3 | P2
    - 1.2.3. Use functions to reduce repetitive procedures in a script. | L3.7, L3.8 | SK3 | P2
    - 1.2.4. Use functions to automate and standardize the production of a product (e.g. a graph, an analysis). | L3.7, L3.8 | SK3 | P2
    - 1.2.5. Create a function that vectorizes a calculation. | L3.7, L3.8 | SK3 | P2
    - 1.2.6. Understand and successfully execute a **while** loop. | L3.5 | SK3 | P2
    - 1.2.7. Understand and successfully execute conditional **if/else** statements (vectorized and non-vectorized). | L3.4 | SK3 | P2
    - 1.2.8. Understand and successfully execute **repeat** and **for** loops. | L3.5 | SK3 | P2
    - 1.2.9. Use a loop to automate a calculation or procedure. | L3.6 | SK3 | P2
    - 1.2.10. Use a conditional statement to automate a calculation or procedure. | L3.6 | SK3 | P2

- 1.3. Understand the basic principles of software design.
  - 1.3.1. Understand why and how code should be documented. | L3.1 | SK3 | P2
  - 1.3.2. Create documentation within code and outside of code in the form of a README file. | L3.10 | SK3 | P2
  - 1.3.3. Use whitespace effectively to make scripts more readable. | L1.10, L3.6, L3.10 | SK3 | P2
  - 1.3.4. Create and maintain R Project files in RStudio. | L3.2, L3.3, L3.10 | SK4 | P3
  - 1.3.5. Understand the basic principles of refactoring. | L3.6, L3.11 | SK3 | P2
  - 1.3.6. Execute code refactoring toward a specific goal (e.g. improving speed, improving readability). | L3.11 | SK3 | P2
2. Understand and follow best practices in scientific computing.
  - 2.1. Properly organize your work.
    - 2.1.1. Understand how file systems are structured and organized. | L1.2, L3.1 | SK1, SK3 | P2
    - 2.1.2. Understand how to navigate file systems using a GUI interface. | L1.2 | SK1, SK3 | P2
    - 2.1.3. Use directories to organize course work. | L1.2 | SK1, SK3 | P2
    - 2.1.4. Use standard organization to organize an R project. | L3.1 | SK3 | P2
  - 2.2. Understand the importance of reproducibility in scientific data analysis.
    - 2.2.1. Create reproducible scripts in *R*. | L1.4, L2.2, L3.1, L3.10 | SK1, SK3, SK4 | P2
    - 2.2.2. Include effective documentation in scripts and projects. | L2.2, L3.1, L3.10 | SK3, SK4 | P2
    - 2.2.3. Understand the Open Science movement and the role of data repositories in research. | L3.1, L3.10 | SK4 | P2, P3
    - 2.2.4. Create and use Notebooks and documents using RMarkdown. | L2.2, L3.10, L4.4, L4.5 | P2 | SK4 | P2, P3
3. Independently perform basic data analysis and visualizations in a way that communicates ideas clearly.
  - 3.1. Load, clean, and organize data in *R*. | L2.3 | SK2.1 | P1
    - 3.1.1. Load text, CSV, Excel data files and built-in package data sets. | L2.3 | SK2.1 | P1
    - 3.1.2. Clean, arrange, and transform data sets. | L2.3 | SK2.1 | P1
    - 3.1.3. Learn to index specific values in data sets. | L2.3 | SK2.1 | P1
    - 3.1.4. Create and use data frames from other data types. | L2.3 | SK2.1 | P1
  - 3.2. Learn how to plot quickly using *R*'s base graphics. | L2.2 | SK2.1 | P1
    - 3.2.1. Create bar graphs, line, scatter, box, and histogram plots. | L2.2 | SK2.1 | P1
    - 3.2.2. Learn when to implement each type of plot above. | L2.2 | SK2.1 | P1
    - 3.2.3. Create a plot legend. | L2.2 | SK2.1 | P1

- 3.2.4. Alter the axes, title, and labels of a plot. | [L2.2](#) | [SK2.1](#) | [P1](#)
- 3.2.5. Create text labels on a plot. | [L2.2](#) | [SK2.1](#) | [P1](#)
- 3.3. Learn the basics of `ggplot2`. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.1. Understand the basic grammar of graphics used by `ggplot2`. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.2. Create line, scatter, box, and histogram plots with `ggplot2`. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.3. Make multi-panel plots. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.4. Create a plot legend. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.5. Alter the axes, title, and labels of a plot. | [L2.6](#) | [SK2.1](#) | [P1](#)
  - 3.3.6. Create text labels on a plot. | [L2.6](#) | [SK2.1](#) | [P1](#)
- 3.4. Execute qualitative and quantitative data analyses. | [L2.5](#) | [SK2.1](#) | [P1](#)
  - 3.4.1. Learn the the basic ways data are described. | [L2.5](#) | [SK2.1](#) | [P1](#)
  - 3.4.2. Learn the `head`, `tail`, and `summary` commands. | [L2.5](#) | [SK2.1](#) | [P1](#)
  - 3.4.3. Calculate the mean, median, quartiles, range, and standard deviation of a data set. | [L2.5](#) | [SK2.1](#) | [P1](#)
- 3.5. Think and work independently with code. | [L2.4](#), [L2.5](#), [L2.6](#), [L3.9](#), [L3.10](#), [L3.11](#) | [SK1](#), [SK2.1](#), [SK3](#), [SK4](#) | [P1](#), [P2](#), [P3](#)
  - 3.5.1. Learn basic skills in debugging and troubleshooting error messages. | [L1.3](#), [L1.10](#), [L2.4](#), [L3.8](#) | [SK1](#), [SK3](#), [SK4](#) | [P1](#), [P2](#)
  - 3.5.2. Search for effective solutions and tools using online resources. | [L2.4](#), [L3.8](#) | [SK1](#), [SK3](#), [SK4](#) | [P1](#), [P2](#)
  - 3.5.3. Learn to accept and provide constructive feedback to peers. | [L2.10](#), [L2.11](#), [L2.12](#) | [SK2.1](#), [SK2.2](#) | [P1](#), [P2](#)
- 3.6. Understand the basic principles of data visualization and communication.
  - 3.6.1. Understand the basic way our brain processes light stimuli into sight. | [L2.7](#) | [SK2.2](#) | [P1](#)
  - 3.6.2. Understand the role of contrast in low-level visual processing. | [L2.7](#), [2.9](#) | [SK2.2](#) | [P1](#)
  - 3.6.3. Understand the organization of the ‘what’ and ‘where’ visual processing systems. | [L2.7](#), [2.8](#) | [SK2.2](#) | [P1](#)
  - 3.6.4. Understand how acuity is focused centrally in vision and drops off as you move into the periphery. | [L2.7](#), [L2.8](#) | [SK2.2](#) | [P1](#)
  - 3.6.5. Understand the importance of attention and accessibility to communicating ideas. | [L2.8](#) | [SK2.2](#) | [P1](#)
  - 3.6.6. Understand the importance of ethical presentation of data through visualization. | [L2.8](#) | [SK2.2](#) | [P1](#)
  - 3.6.7. Understand the principle of the Curse of Knowledge and how it impedes communicating ideas. | [L2.9](#) | [SK2.2](#) | [P1](#)
- 3.7. Learn the most effective strategies for written and oral presentation of scientific ideas.
  - 3.7.1. Discuss the reasons for communicating scientific results to a variety of audiences. | [L4.1](#) | [SK4.1](#), [SK4.2](#) | [P3](#)

- 3.7.2. Practice reading and understanding scientific papers and grant proposals. | **L4.2** | **SK4.1** | **P3**
- 3.7.3. Practice watching and understanding oral presentations (slideshow and poster). | **L4.3** | **SK4.2** | **P3**