# Module 1 In-class Exercise



Miami University | College of Arts & Science | Department of Statistics | Spring 2021-22

Course Title: Nonparametric Statistics | Subject and Course Number: STA 333

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Module 1 In-class Exercise (Student Copy)

Due date: Thursday, 01/27/2022

Total: 100 points

**Objectives:** Module 1 In-class exercise will

* Provide you an opportunity to produce quality documents using R Markdown and export them as HTML document
* Help you to read in datasets and create a data vector
* Assist you to install R packages
* Prepare you for the future in-class exercises

**Submission guidelines:** All in-class exercises must be prepared using R Markdown and submitted electronically (HTML format). The following protocol should be followed:

* Your name must appear in print at the top of your submission.
* Do the work in problem order, with number/part indicated by headers (e.g. Question 1, Part 1, etc.).
* Start each problem’s solution by including the original question just before you begin your solution.
* Relevant R code chunks and output must be included in-line at the appropriate point in the solution.
* All mathematical notation must be rendered correctly with appropriate use of subscripts, superscripts and symbols.
* All assignments must be formatted neatly and executed professionally. This includes use of proper grammar and spelling, and clear organization. Serious deficiencies may result in grade lowering.

**Note:** This in-class exercise will be collected individually. I will walk through these tasks in the class.

**Task 1:** Changing the header size, bold, and italic **(5 points)**

Note: # symbol controls header size, \*\*your text here\*\* produce bold, \*your text here\* produce italic.

**Task 2:** Writing the most commonly used statistical symbols **(10 points)**

Note: Try the following statistical symbols

Population mean:

Population proportion:

Population standard deviation:

Level of significance equal 0.05:

*p*-value:

Chi-square distribution with three degrees of freedom:

**Task 3:** Writing the null and alternative hypotheses **(55 points)**

*H0*: = 4 miles

*Ha*: 4 miles

*Ha*: < 4 miles

*H0*: = 0.85

*Ha*: 0.85

*Ha*: 0.85

*H0*: =

*Ha*:

*Ha*: <

*H0*: =

*Ha*:

*Ha*: >

*H0*: =

*Ha*:

*Ha*: <

*H0*: =

*Ha*:

*Ha*:

**Task 4:** Formatting texts **(5 points)**

Note: Texts can be formatted for a nicer look. Backtick (`) is used for text formatting.

*t*-test (try both formatted and non-formatted)

Chi-square test (try both formatted and non-formatted)

Kruskal-Wallis test (try both formatted and non-formatted)

**Task 5:** Installing R packages **(5 points)**

Note: R packages such as, boot, coin, exactRankTests, randtests, agricolae, Tidyverse, lmPerm, rpart, rpart.plot, ecelR, ggpubr, nortest, etc. will be used in this course.

**Task 6:** Creating a data vector **(10 points)**

Note: Create a data vector from the following data and create a normal QQ plot.

142, 103, 86, 165, 134, 154, 119, 81, 98, 122, 161, 117,119, 93, 144, 128, 131, 137, 158, 103

**Task 7:** Reading in a csv file on R **(10 points)**

Note: Read the Engine dataset on R and create a normal QQ plot for the variable Emission.