



---

## Statistical Analysis and Applications Final Project

Statistical Analysis and Applications course has introduced you to the methods and procedures that allow us to: Explore Data and make Statistical Inference. This Final Project is designed to allow you to do:

### Project Outline:

#### Basic Idea

#### **(5 Marks Bonus + Up to 2 Marks Bonus)**

Implementation of the descriptive statistics part of the course using Python or R programming language.

- Analyze one dimensional data: Graphically and Numerically.
  - Analysis of central tendency and variability using different metrics.
  - Analysis of data using different graphs like bar charts, Pie charts, Histograms, etc...
- Perform correlation and regression analysis for two-dimensional data.

#### **Bonus**

- Perform Estimation and Inference.
  - Finding population mean if sample mean is known.

#### Advanced ideas

#### **(5 Marks Bonus + Up to 5 Marks Bonus)**

Using concepts introduced in the statistical analysis course and machine learning concepts to answer a research question:

- Identify a research question.
- Identify your dataset (Training and Testing) if found.
- Perform machine learning techniques.
- Present your findings.

#### **Note:**

The purpose of this project is to allow you to communicate your understanding of Statistical Analysis.



Reports should be prepared using Microsoft Word and follow the required structure.

Your final report should clearly indicate your understanding of data analysis and inference.

**Deadline: December 30, 2021**

**Submission Method: Will be announced later.**

## **Project Objectives and Skills:**

### **Research Question**

Identify a question that is interesting, appropriate, and worthy of investigation.

Your question must lend itself to data that can be analyzed using the methods learned in class.

### **Exploratory Data Analysis**

Analyze raw data using appropriate graphical and numerical procedures.

Describe Shape, Outliers, Center, and Spread of datasets in the context of your research question. Include appropriate graphical displays and numeric summaries/descriptions in your final report.

Show all applicable work: Sampling Distribution, Test Statistic Calculation, etc.

### **Inference (Bonus)**

- Starting by population parameters and find the of sample statistics.
- Starting with sample statistics and estimate the population parameters.
- Answer your research question based on your inferential calculations.

### **Final Report**

You are expected to write up your findings in a final report. This report should follow a standard academic format and should include a section for each task noted above.



## **Requirements of the Report:**

---

Your Report should consist of a summary of your research and experiment/survey as well as your personal conclusions. The goal is to enlighten the reader with words, numeric summaries, and appropriate graphs. Use the following format:

**Title Page (title of project, your name, class, my name, year)**

**Introduction (state topic, project goals and direction)**

Describe the project by discussing your question of interest, why you chose it, and how you planned on analyzing it.

**Summary of Research (data, statistics, charts, graphs, sources).**

**Grading Criteria for Basic Idea (7 points total (5 + Up to 2 Bonus))**

1. Calculating descriptive statistics (2 Marks)
2. Calculating correlations (2 Marks)
3. Quality of the report (1 Mark)
4. Conducting inferential tests (Up to 2 Marks)

**Grading Criteria Advanced Ideas (10 points total (5 + Up to 5 Bonus))**