

MATU-203 Final Project Instructions

Mastering the concepts in Introductory Statistics assists in building critical thinking skills, developing businesses and organizations, and solving all types of problems that require data. But an understanding of statistics extends beyond the ability to crunch numbers or use a software program. The ability to collect, organize, and analyze data is the beginning. The ability to clearly communicate your results to another person is the mark of true mastery.

The Assignment

In this assignment, you will choose a scenario with data – from one of 5 options provided at the end of these directions under **Project Topics** – and you will construct a paper that pulls together the statistics you have learned in order to answer a question. You will:

1. Introduce the main question, and explain the data that you will use to address it
2. Organize your data by providing appropriate charts, graphs and descriptive statistics
3. Analyze your data by conducting a hypothesis test
4. State your conclusions and recommendations

See the section **Outline Of Material To Present** below for a more detailed explanation of what you will submit for each of these four sections.

The topics that you may choose from are

- **Business** – Analysis of Home Sales
- **Health Sciences** – The Effect of Caffeine on Heart Rate
- **Psychology** – The Effect of Colored Paper on Reading Speed
- **Sociology** – Level of Education and its Effect on Homelessness in Veterans
- **Criminal Justice** – Analysis of Poverty and Crime

At the end of these directions, under **Project Topics**, are the full descriptions for each of these 5 topic options, including the data you must use in your analysis. Please scroll down and read through each of these 5 options in their entirety before making your selection. You must select one of these 5 topics. **You may not select your own topic.**

Outline of Material to Present

This assignment is broken down into four parts: *Collection of Data*, *Organization of Data*, *Analysis of Data*, and *Conclusion/Recommendations*. Each part has a subset of questions and issues you must address. Please ensure that your report includes a section for each of the 4 Parts listed below, and that each part addresses ALL of the sub-questions listed.

Again, you must select one of the 5 topics provided at the end of these instructions in the section **Project Topics**, and you must use the data and names provided within your chosen topic prompt.

Part 1: Collection of Data - Introduction and Primary Data Analysis (3 – 5 paragraphs):

1. Describe the objective: Before you can examine the data, you must understand the problem.
 - a. Discuss the importance of this issue or situation.
 - b. Introduce the company or organization you are preparing this report for, and explain why it is important to them.
 - c. What is the research question? In other words, what is the basic question you, as the researcher, want to address? Why should we care about it?
 - d. Was this an experimental or observational study? Explain.
2. Clearly and with sufficient detail, describe the population, sample, and collection methods in this study.
 - a. What is the population you are interested in?
 - b. What is the sample, specifically?
 - c. What is a plausible way the sample was chosen and why?
 - d. What problems or biases might have occurred from choosing that type of sampling method?
3. Discuss the type of data.
 - a. Was the data quantitative or qualitative? Explain.
 - b. What is the level of measurement (nominal, ordinal, interval and ratio)? Explain.
4. Describe the variables
 - a. What are the independent and dependent variables? Give the type, units, and more specific information.
 - b. Give examples of any confounding variables, lurking variables, and/or missing variables and explain how they may be affecting your study.

Part 2: Organization of Data - Examination of Descriptive Statistics (graphs and tables, and approximately 4 paragraphs)

Now that your data is collected, you need to organize it to identify characteristics and patterns.

1. Graph your data appropriately. Construct a scatterplot, bar graph, or other graph to show the nature of the data. For each graph, be sure you label the graph completely – that means give it a title, label the axes, and explain what that graph means in the context of this narrative.
2. Discuss whether the data is normally distributed. For this, use a visual inspection of a

Histogram and Normal Quantile Plot, as well as what you see in the data itself and what that means about the high and low ends of the data.

3. Calculate and present the three Measures of Central Tendency: mean, median, and mode. Provide both the value of the statistics as well as an analysis of what they mean in terms of understanding the sample.
4. Calculate and present the Measures of Variation: range and standard deviation. Provide both the value of the statistics as well as an analysis of what they mean in terms of understanding the sample.
5. Calculate and present the 5-Number Summary: minimum, Q1, median, Q3, maximum. Provide both the value of the statistics as well as an analysis of what they mean in terms of understanding the sample.
6. Identify any Outliers. You can do this using a visual inspection of the graph as well as the formulas from the textbook (HINT: $Q1 - 1.5 \cdot IQR$, and $Q3 + 1.5 \cdot IQR$).
7. Discuss any corrections: Based on your inspection of the outliers are there any errors that should be corrected? How would you correct them? Discuss the implications of this result.

Part 3: Analysis of Data - Examination of Inferential Statistics (tables of results, and appropriate hypothesis test steps)

Assuming that all assumptions have been met, it is now time to analyze the data. Present a complete hypothesis test.

1. Identify the claim.
2. State the null and alternative hypotheses, in words and in symbolic form.
3. Explain what type of test you will be performing (i.e. a test of two dependent means, a test for correlation, etc.) and why that test is appropriate to address the main question you are trying to answer.
4. Select the significance level and determine if it is a one or two-tailed test.
5. Identify the test statistic and compute the value of the test statistic and the p-value.
6. Make a decision of whether to Reject or Fail to Reject the null hypothesis.
7. Restate your decision in nontechnical terms. That means, state your conclusion in a way that anyone can understand; a final conclusion that just says "reject the null hypothesis" by itself without explanation is not helpful to those who hired you. Explain in ordinary terms what it means.

Part 4: Conclusion and Recommendations (approximately 3 paragraphs)

Summarize and explain your results. Provide recommendations.

1. What can you infer from the statistics?
2. What information might lead you to a different conclusion?
3. What variables are missing?
4. What additional information would be valuable to help draw a more certain conclusion?
5. What qualitative or quantitative data would you want to collect if you were hired to do a follow up study?

Formatting Requirements

Your paper should be a **Word document**, with **embedded charts, graphs, figures, and tables**. It should be **APA or MLA Format**, with name and page number on each page, and should include each of the following:

1. **Title page:** You should have a cover page. The cover page should have the specific title of your study (Note: "MATU 203 Paper" does not clearly define the topic you are presenting), your name, Brandman University, MATU 203, term, and year. You might find it helpful to include an image of something representative of the study to provide a visual context for your report.
2. **Write Up/Body:** The body should be 5 - 8 pages with graphs, images, screen shots of data output, and text included. Please use a "page break" to separate the cover from the body and the references from the body and the appendices from the reference page.
3. **References:** Please include all articles, books, websites, publications, or other information that helped you reach your conclusion. The references come after the main body and before the appendix. At least two references are required, not including the textbook.
4. **Appendices:** You must have an appendix. The appendix goes at the end of your paper and might be an additional 1-2 pages. In the appendix, include your given data as well as supplemental charts and graphs outside of the ones included in your paper. Label the appendices, Appendix 1, Appendix 2, and so on.

Academic Integrity

The University expects that students will conduct themselves in an honest and ethical manner and respect the intellectual work of others. Your paper will be submit through TurnItIn, which is a plagiarism detection service. Your paper will receive a Similarity Report that represents the percentage of your paper that is similar to or copied from other sources. Any type of plagiarism, including using work without referencing, copying another paper, not citing references in the body of the paper, using a paper from another course, etc., will be met with penalties determined by the instructor, including the possibility of failing the course, and having an Academic Integrity Violation Sanction Report submitted to University administration. Note: You may not use a paper you wrote from a prior course, even if it is the same class you are taking now, without prior permission from your individual instructor. If you have further questions about academic integrity, please contact your instructor and/or see the catalog.

Project Topics *(select one of the topics and data sets below)*

Study #1: Business

Mickey and Minnie Home Sales, an Orange County real estate company, wants to know if there is a relationship between the price a home sells for and the number of weeks the home was on the market prior to the sale.

Mickey and Minnie conjecture that the higher the price of the home, the longer the house will be on the market. They select a simple random sample of 31 recently sold homes in Anaheim for analysis and they want to know if the number of weeks to sale influences asking price. You are asked to organize and analyze this data, and present your results to the company with an appropriate hypothesis test.

Please present a report to the Mickey and Minnie executives that includes your data, results, and related evidence from national studies.

Asking Price (thousands), X	Weeks to Sale, Y
150	6
250	17
200	8
285	10
320	12
80	5
120	8
300	12
180	6
600	22
300	9
210	8
300	16
460	14
670	15
400	10
300	18
200	9
240	14
65	7
900	13
330	12
205	5
405	15
630	16
255	7
310	14
440	16
305	13
700	20
280	6

[Click here to Download the Data Set for Study #1](#)

Study #2: Nursing and Health Sciences

Premiere Hospital of Newport Beach (PHNB) wants to determine whether or not to serve caffeinated coffee for their patients in the waiting rooms. At this point, they are not concerned about the visitors, only people who will be treated at the hospital. As a state-of-the-art medical facility, they know that many similar facilities do serve coffee and they want to determine if there is sufficient evidence to show that caffeinated coffee increases heart rate.

The hospital tested 31 hospital patients who each recently had two cups of coffee while awaiting a doctor consultation. The patients' heart rates were recorded before they drank coffee, and again after they drank coffee. You are asked to organize and analyze this data, and present your results with an appropriate hypothesis test of whether caffeine significantly increases heart rate.

Please present a report to the Premiere Hospital administrators that includes your data, results, and related evidence from national studies.

Heart Rate Before Coffee	Heart Rate After Coffee
80	95
70	100
80	84
65	65
77	84
60	74
90	100
75	85
88	98
70	95
78	80
70	80
62	78
50	60
72	88
75	79
60	72
72	90
60	65
68	78
65	75
62	75
75	90
84	88
87	95
73	87
78	88
68	75
72	84
81	92
75	86

[Click here to Download the Data Set for Study #2](#)

Study #3: Psychology

Some people believe that reading text printed on colored paper is easier than reading text on white paper. For this reason, disability services locations often provide colored filters that are clear to put over top of books and imprinted pages.

Irvine University, a public research institution with 30,000 undergraduate students, is considering the possibility of providing colored filters for those people in disability services who would be helped by this tool. Irvine University randomly tested 31 students in their learning disability program during the first week of classes in Fall 2018. The students, of varying ages, abilities, majors, and socioeconomic status were timed as they read a passage printed in black ink printed on white paper and then timed reading a similar passage printed with black ink on light blue-colored paper. Students were tested on one day with a passage on white paper and then asked to willfully return the next day to read a similar, but different passage of the same length and difficulty level, on light blue-colored paper.

Is there sufficient evidence to show that reading text on light blue-color paper is more effective? You are asked to organize and analyze this data, and present your results to the university with a hypothesis test on whether or not there is significant improvement in reading speed.

Please present a report to Irvine University's Disability Services directors that includes your data, results, and related evidence from national studies.

Light Blue Paper (seconds taken to read passage)	White Paper (seconds taken to read passage)
70	65
113	100
50	40
138	120
77	74
120	100
100	90
79	68
67	60
73	71
67	74
70	68
110	90
95	85
49	44
75	69
60	58
82	75
88	78
80	74
100	85
102	100
96	90
64	64
73	73
93	94
85	75
69	65
91	92
112	100
124	105

[Click here to Download the Data Set for Study #3](#)

Study #4: Sociology

California Veteran Support, an agency delivering resources to support military veterans in California, wants to address the problem of veterans being homeless. To start, they want to determine whether or not there is a correlation between homelessness and other variables, such as education, crime, and mental illness, so that they can determine the most effective resources to offer. You are hired to specifically investigate whether or not there is a relationship between years of education, and length of time spent homeless.

You receive a survey of responses from 31 homeless veterans from various cities in California. Each veteran responded to indicate how many months they have lived on the streets, as well as the grade they finished in school (one year of college would equal 13 years of school, 4 years of college would equal 16 years of school, and so on. Some veterans received a GED and had less than 12 years of formal education).

Is there sufficient evidence to show that the years of education a veteran has received may affect the length of time that he or she is homeless? You are asked to organize and analyze this data, and present your results with a hypothesis test on whether or not there is a correlation between years of education and months spent homeless.

Please present a report to the California Veteran Support directors that includes your data, results, and related evidence from national studies.

Education in years of school	Homelessness in months on the street
16	5
13	10
12	4
12	12
18	7
12	10
16	20
16	18
12	20
13	17
14	4
9	8
10	9
15	15
12	14
13	9
10	18
12	35
16	28
15	14
12	18
14	4
14	10
15	10
12	15
13	24
16	5
16	30
18	6
12	14
13	18

[Click here to Download the Data Set for Study #4](#)

Study #5: Criminal Justice

Many public bodies seek to understand the issues surrounding criminal behavior so that they can find ways to reduce crime. You are asked to determine if there is evidence of a relationship between poverty and crime.

Since arrests do not equal convictions and number of convictions is not a good variable with the Three Strike Law in California, you decide to ask 31 random people in and around the Los Angeles courthouse over the month of July to voluntarily give you their family income and self-reported incidents of criminal activity for your study. Half of the respondents were awaiting trials of various types, and half were bystanders. You are asked to organize and analyze this data, and present your results with a hypothesis test on whether or not there is a correlation between family income and number of incidents of criminal activity.

Prepare this analysis for your local law enforcement agency. Please include your data, results, and related evidence from national studies.

Family Income (thousands)	Number of self-reported instances of criminal activity
160	1
5	8
12	0
100	1
18	7
200	0
16	2
25	10
40	5
32	13
14	20
90	1
30	9
40	6
15	14
28	10
30	8
120	0
30	10
40	4
35	4
47	2
51	3
75	0
130	1
55	3
25	9
15	3
35	0
65	1
52	3

[Click here to Download the Data Set for Study #5](#)