

Math 227 Sample Project

Project 1 Data Collection and Descriptive Statistics

In this project, you will demonstrate your understanding of data collection and descriptive statistics. You will collect the data, organize, summarize, and describe the data the way it is explained in Chapter 1-Chapter 3. You will, then, create a written report including all parts below and turn in by the stated due date according to the guidelines provided in this paper.

Required components:

1. Research Questions about a Population

Question 1: Come up with a research question on making an estimate about a population proportion.

(For example: What proportion of Los Angeles Mission College students are full time students?)

State your Question 1 (and your estimated prediction of the results):

Question 2: Come up with a research question on making an estimate about a population mean.

(For example: What is the average number of hours that Los Angeles Mission College students work each week?)

State your Question 2 (and your estimated prediction of the results):

2. Collecting Data

a) Create a survey and collect the data for responses to your proposed questions above.

For question 1, your survey should have a qualitative variable with only two possible responses, Write your variable and the question below (For example: full time student status. "Are you a full-time student, Yes or No?"):

For question 2, your survey should have a quantitate variable. Write your variable and the question below (For example: number of hours of work in a week. "How many hours do you work each week?"):

b) Describe the data collecting process including the population of interest. For each question, address if there was any of the following biases: Sampling bias, Non-response bias, and Response bias.

c) Enter your raw data in Column 1 (categorical data) and in Column 2 (numerical data) in StatCrunch. Name each column and save the data (Data>Save>Enter the File Name>Click Save). To later open the saved data in StatCrunch, (Data> Load>My Data>Choose the file name.)

3. Organizing, Summarizing, and Describing Qualitative Data using StatCrunch

a) Open your saved survey data. Organize your data for your categorical variable in column 1, into a frequency/relative frequency table. Save the distribution table. (To open the saved results later in StatCrunch, go to statcrunch.com> MyStatCrunch> My Results>Choose the name/notes.)

b) Construct a bar graph for a frequency distribution and for a relative frequency distribution for your categorical variable. Choose to display the value above bar. Save the graphs.

c) Construct a pie chart for the categorical variable. Save the result.

d) Create a dotplot and a histogram for your survey's numerical data in column 2. Make sure that your histogram has an appropriate width for the bins. Explain why your choice for the width makes sense for your specific data. Save your graphs.

4. Organizing, Summarizing, and Describing Quantitative Data using StatCrunch

a) Open your survey data and the dotplot and the histogram for your survey's numerical variable. Create a boxplot for the variable in column 2. Write a description of each graph (What does each graph show you about your data, that you won't be able to easily see in the other graphs?). Save the boxplot.

b) Find the following sample statistics for the data set: Minimum value, Maximum value, Mean, Median, Mode, Q1, Q3, Range, IQR, and Standard Deviation. For each of these values write a sentence explaining what it tells you about the results in your survey. Please do not write about how they are calculated. The purpose of this exercise is to make sure that you understand the meaning for each of these statistics (not how you find it). So please don't include steps for finding each of these. You will be graded on your explanation of what each value tells us specifically about your sample.

Here's an example of how I will grade the sentence about the median if it related to the price of cars assuming that it is worth 6 points:

-The median is 22,000. - I would give this 1 point

-The median price of a car is 22,000 - I would give this 2 point.

-The median price of a car is \$ 22,000. - I would give this 3 points.

-The median price of a car is \$ 22,000 which means half the cars are less than \$22,000 and half are more than \$22,000 - I would give this 5 points

-The median price of a car is \$ 22,000 which means half the cars sold cost less than \$22,000 and half cost more than \$22,000 -I would give this 6 points

c) Describe and analyze the numerical data. What is the overall shape of the data set? Based on your answer, determine which measure of center (mean, median or mode) you think is best and why. Determine which measure of spread (IQR / standard deviation or range) you think is best and why. Write a sentence explaining the real meaning of the center you found and a sentence explaining the real meaning of the spread you have found for your data.

5. Create a written report including all your findings from the questions above. What did you learn about the population of your interest from your data analysis? Please don't forget to address the answers to your initial research question 1 and question 2. Do you think your data represented the population well? Did the result match your initial predictions? Why or why not? You will be graded not only on your analysis and graphs, but also on the quality of your writing. You should have an introduction paragraph discussing why this topic was important in the world or important to you and a conclusion paragraph summarizing your findings. Your report should be at least one and a half pages double spaced (font 12) not counting graphs.

Grades will be based on:

1. Choosing and stating the appropriate research questions about a population in part 1. [8 pts],
2. Relevance and completeness of the survey and data collection in part 2. [12 pts],
3. Organizing, summarizing, and describing qualitative data using StatCrunch in part 3. [20 pts],
4. Organizing, summarizing, and describing quantitative data using StatCrunch in part 4. [20 pts]
5. Clear summary of the main findings and completeness of the report in part 5 [30 pts], and
6. Overall quality of the report and adherence of the project guidelines [10 pts].

Submission Guidelines and the due date:

The report including all graphs and symbols must be generated using computer and only the printed version will be accepted. The report should contain the title, your name, date, course name and instructor's name. The report should not contain any misspelled words. The report will not be returned to you.

The printed project score sheet along with the title of your project and two research questions (page 3-6 of this guidelines), and the electronic copy of your raw data must be submitted to the instructor by:

The project is due at the beginning of the lab on: _____

Note: Late submission will be accepted. However, there will be a 10% per day penalty from your final project grade.

Math 227 Project 1 Data Collection and Descriptive Statistics (Total 100 points)

Name: _____ Date: _____

Title of the project: _____

Research Question 1 on making an estimate about a population proportion:

Research Question 2 on making an estimate about a population mean:

Project Grade: _____

Required components	Grading criteria	Comments/Scores
1. Research Questions about a Population Question 1: Come up with a research question on making an estimate about a population proportion. (For example: What proportion of Los Angeles Mission College students are full time students?) Question 2: Come up with a research question on making an estimate about a population mean. (For example: What is the average number of hours that Los Angeles Mission College students work each week?) State your Question 2 (and your estimated prediction of the results):	2. Choosing and stating the appropriate research questions about a population in part 1. [8 pts]	
2. Collecting Data a) Create a survey and collect the data for responses to your proposed questions above. For question 1, your survey should have a qualitative variable with only two possible responses, Write your variable and the question below (For example: full time student status. "Are you a full-time student, Yes or No?"): For question 2, your survey should have a quantitate variable. Write your variable and the question below (For example: number of hours of work in a week. "How many hours do you work each week?"):	3. Relevance and completeness of the survey and data collection in part 2. [12 pts]	

b) Describe the data collecting process including the population of interest. For each question, address if there was any of the following biases: Sampling bias, Non-response bias, and Response bias. c) Enter your raw data in Column 1 (categorical data) and in Column 2 (numerical data) in StatCrunch. Name each column and save the data (Data>Save>Enter the File Name>Click Save). To later open the saved data in StatCrunch, (Data> Load>My Data>Choose the file name.)		
3. Organizing, Summarizing, and Describing Qualitative Data using StatCrunch a) Open your saved survey data. Organize your data for your categorical variable in column 1, into a frequency/relative frequency table. Save the distribution table. (To open the saved results later in StatCrunch, go to statcrunch.com> MyStatCrunch> My Results>Choose the name/notes.) b) Construct a bar graph for a frequency distribution and for a relative frequency distribution for your categorical variable. Choose to display the value above bar. Save the graphs. c) Construct a pie chart for the categorical variable. Save the result. d) Create a dotplot and a histogram for your survey's numerical data in column 2. Make sure that your histogram has an appropriate width for the bins. Explain why your choice for the width makes sense for your specific data. Save your graphs.	3. Organizing, summarizing, and describing qualitative data using StatCrunch in part 3. [20 pts]	
4. Organizing, Summarizing, and Describing Quantitative Data using StatCrunch a) Open your survey data and the dotplot and the histogram for your survey's numerical variable. Create a boxplot for the variable in column 2. Write a description of each graph (What does each graph show you about your data, that you won't be able to easily see in the other graphs?). Save the boxplot. b) Find the following sample statistics for the data set: Minimum value, Maximum value, Mean, Median, Mode, Q1, Q3, Range, IQR, and Standard Deviation. For each of these values write a sentence explaining what it tells you about the results in your survey. Please do not write about how they are calculated. The purpose of this exercise is to make sure that	4. Organizing, Summarizing, and describing quantitative data using StatCrunch in part 4. [20 pts]	

you understand the meaning for each of these statistics (not how you find it). So please don't include steps for finding each of these. You will be graded on your explanation of what each value tells us specifically about your sample. c) Describe and analyze the numerical data. What is the overall shape of the data set? Based on your answer, determine which measure of center (mean, median or mode) you think is best and why. Determine which measure of spread (IQR / standard deviation or range) you think is best and why. Write a sentence explaining the real meaning of the center you found and a sentence explaining the real meaning of the spread you have found for your data.		
5. Create a written report including all your findings from the questions above. Write a short summary of the main findings that you discovered. Create a written report including all your findings from the questions above. What did you learn about the population of your interest from your data analysis? Please don't forget to address the answers to your initial research question 1 and question 2. Do you think your data represented the population well? Did the result match your initial predictions? Why or why not? You will be graded not only on your analysis and graphs, but also on the quality of your writing. You should have an introduction paragraph discussing why this topic was important in the world or important to you and a conclusion paragraph summarizing your findings. Your report should be at least one and a half pages double spaced (font 12) not counting graphs.	5. Clear summary of the main findings and completeness of the report in part 5 [30 pts]	
6. Overall quality of the report	6. Overall quality of the report and adherence of the project guidelines [5 pts]	