MTH 245, Exam #1, Spring 2022   Name _____ KEY _____

Instructions: Show all work to receive full credit. You should note any formulas used or calculator functions used, their inputs and outputs, or attach a spreadsheet with your calculations. I cannot grade work if I don't know where an answer came from. Be sure complete all parts of each questions, including requests for interpretation and explanations. Be as thorough as possible.

This exam will be submitted in 2 parts. Part 1 are numerical or multiple-choice responses that will be submitted separately and graded by the computer. The second part will consist of explanatory responses, working with graphs and other questions that will be submitted as scanned documents and graded by hand.

Part 1: Answer these questions using your calculator or Excel. Show your work on this page or in Excel and submit along with part 2. Then submit your answers to these questions in the Exam #1 Part 1 submission tool in Canvas.

1. For each of the following variables, determine i) is the variable qualitative or quantitative? ii) the level of measurement: nominal, ordinal, interval, or ratio? iii) if the variable is quantitative, is it discrete or continuous? (6 points each)
   a. The time of day (i.e., 7:30 p.m.)

   quantitative, interval, continuous

   b. Car model

   qualitative, nominal, N/A

   c. Phone number

   qualitative, nominal, N/A

   d. Number of students in a class

   quantitative, ratio, discrete

2. For the Car Value data on Sheet 1 of the data file **245exam1data.xlsx**, find the following statistics:
   a. The mean, median and mode (6 points)

   Mean = $5908.48
   Median = $4175
   mode(s) = 2050, 1090

   b. The standard deviation and range (4 points)

   St. dev = $5533.46

   range = $33740

c. Calculate the five-number summary for this data. (5 points)

$$min = \$130 \qquad Q3 = \$7717.50$$
$$Q1 = \$2110 \qquad Max = \$33,870$$
$$median = \$4175$$

d. Based on this information, what percentile is a car valued at $10,500? (3 points)

$$84.8^{th} \text{ percentile}$$

3. The 100 tiles in Scrabble are distributed as follows:

Tile	Number	Tile	Number	Tile	Number	Tile	Number
Blank	2	A	9	B	2	C	2
D	4	E	12	F	2	G	3
H	2	I	9	J	1	K	1
L	4	M	2	N	6	O	8
P	2	Q	1	R	6	S	4
T	6	U	4	V	2	W	2
X	1	Y	2	Z	1		

a. What is the probability of selecting a T as the first tile in a game? (3 points)

$$\frac{6}{100} = \frac{3}{50}$$

b. What is the probability of selecting a vowel (not Y) as the first tile? (4 points)

$$9 + 12 + 9 + 8 + 4 = 42$$

$$\frac{42}{100} = \frac{21}{50}$$

c. What is the probability of not selecting a vowel? (3 points)

$$1 - \frac{21}{50} = \frac{29}{50}$$

d. What is the probability of selecting the word CAB in order from the first three tiles in a game? (5 points)

$$\frac{2}{100} \cdot \frac{9}{99} \cdot \frac{2}{98} = 3.71 \times 10^{-5}$$

4. Use the following table to calculate the probabilities requested. (4 points each)

| | | STUDENT'S CHOICE | | |
		Art degree	Science degree	Total
GROUP	Boys	25	50	75
	Girls	55	20	75
	Total	80	70	150

a. What is the probability of a randomly selected person from this study is a girl?

$$\frac{75}{150} = \frac{1}{2}$$

b. What is the probability of a randomly selected person from this study being a girl and wants a science degree?

$$\frac{20}{150} = \frac{2}{15}$$

c. What is the probability of a randomly selected person from this study being a girl or wanting a science degree?

$$70 + 75 - 20 = 125 \qquad \frac{125}{150} = \frac{5}{6}$$

d. What is the probability of being a girl given that the person wants a science degree?

$$\frac{20}{70} = \frac{2}{7}$$

e. Are the variables gender and degree choice independent? Why or why not? Show calculations to justify your answer.

not independent since $P(G) \neq P(G | SD)$

$$\frac{1}{2} \neq \frac{2}{7}$$

5. For the discrete probability distribution below, answer the following questions. (3 points each)

X	0	1	2	3	4
P(x)	0.06	0.25	0.33	0.28	0.08

a. What conditions are required to be satisfied for the table of values to represent a probability distribution?

probabilities must add to 1
no probabilities can be more than 1 or negative

b. Fill in the missing value in the table, i.e., what is P(X=0)?

0.06

c. What is the probability that x>2, i.e., P(X>2)?

0.36

d. What is the probability x is greater than 4, i.e., P(X>4)?

0

e. What is the probability that x is not 3?

0.72

Part 2: Answer these questions in this file, using Excel (copy and paste solutions into this document), show work, etc. Don't make me hunt through Excel looking for answers to these questions! Submit your work for Part 1, work and solutions for Part 2, and any Excel file(s) you used to get your answers in the Exam #1 Part 2 submission folder.

6. A researcher is interested in finding out about the buying habits of the American public. She decides to conduct a survey of 100 shoppers at a local mall and asks them how much money they are planning to spend on gifts for an upcoming holiday. (3 points each)
   a. What is the population of interest in this study?

   all American shoppers

b. What is the sample used in this study?

100 shoppers at a local mall

c. What is the parameter being studied?

amount of money spent on gifts

d. What statistic is the researcher likely to use to estimate the parameter?

mean of sampled values from surveyed customers

7. A researcher notes a relationship in a study between getting married and weight gain. What is an example of a possible confounding variable the researcher would want to control for? Why? (4 points)

answers may vary

pregnancy – if women get pregnant after marriage this could be the cause of the weight gain

time of year – many weddings occur in June, maybe weight gain is associated w/ that

previous weight loss – maybe people lost weight for wedding and now they gain back

8. Why does a voluntary response sample tend to introduce bias into the results of a study? (4 points)

the people who respond voluntarily tend to be those w/ strong opinions that may not be reflective of general population

9. What is a double-blind study and why are studies conducted this way? (4 points)

in double-blind studies neither patients nor those interacting w/ patients (doctors, nurses) know which group they are in so as not to give away unconscious information/signals that might taint the result w/ placebo or nocebo effects.

10. Using the data on Sheet 2 of the data file **245exam1data.xlsx**, construct a frequency table of the favorite day to eat out. Display the frequencies as both counts, and percentages. Use the data to construct a bar graph and a pie chart. (Paste your graphs and table(s) here.) (16 points)

See Excel for my graphs and table

11. The data on Sheet 3 of the data file 245exam1data.xlsx, construct a line graph of the data. Paste the graph here, and then describe the graph. Are there any trends? (8 points)

decreased through late 60s, then increased to peaks in 1983 and early 90s. decreased again to 2000, Then increased again to new peak in 2010.
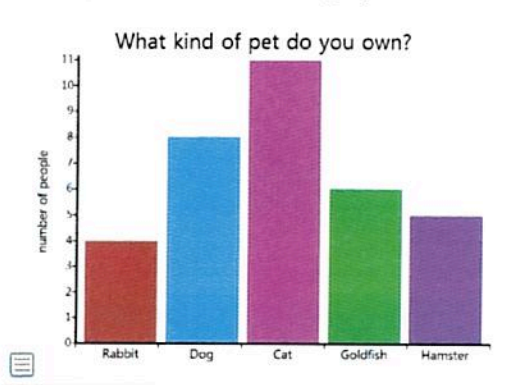
graph in Excel

12. For the Car Value data on Sheet 1 of the data file **245exam1data.xlsx**, construct a histogram and a boxplot of the data. Paste your graphs here. Describe the general shape of the distribution (symmetric, normal, right- or left-skewed, etc.) (10 points)

the graphs are right-skewed

my graphs are in Excel

13. Below is a bar chart of the type of pets a sample of people own. In the space next to the graph, or using Excel, convert this graph to a Pareto chart. (8 points)

What kind of pet do you own?

my graph in Excel
Sort The data Then graph The result.

14. Use the data in the table below answer the questions that follow.

Number of Texts	Adult Frequency	Teen Frequency
None	173	13
1-10	978	138
11-20	249	69
21-50	249	113
51-100	134	113
101+	153	181

a. Use this information to compare the texting habits of adults and teens. (3 points)

adults are more likely to text less frequently
teens are more likely to text very frequently
this is clearer in percentage form

b. Explain why breaking up the numerical categories in this fashion could be considered misleading, especially when making a graph. (3 points)

this can be considered misleading because
the categories are not the same width
which can make the categories seem
larger than they are