Name:
E number:
Math 1530 section number

MATH 1530 CAPSTONE TECHNOLOGY PROJECT SPRING 2018

Problem 1: Sampling. In the survey data, the variable "**BIRTHDAY**" is the days left until a student's next birthday.

a. "**BIRTHDAY.SUBSET**" is a variable that contains the first 50 observations from the variable "**BIRTHDAY.**" Calculate the mean number of days until a student's next birthday.

The mean number of days until a student's next birthday of the first 50 students is **157.6**. (Type the value into the space provided.)

b. What type of sampling method was used in part (a)? **convenience sampling**

c. Next, select a random sample of size n = 50 (Go to Calc > Random Data > Sample from Columns). Type the number 50 in the "Number of rows to Sample" slot. Enter the variable "**BIRTHDAY**" into the "From columns" slot. Enter C19 into the "Store samples in" slot. Calculate and report the mean age for your random sample of 50 students.

The sample mean number of days until a student's next birthday is **180.3**. (Type the value into the space provided.)

ANSWERS WILL VARY.

d. What type of sampling method was used in part (c)? **simple random sample**

e. Suppose we think of *all* the students who responded to the survey as a *population* for the purposes of this problem. In that case, the *population mean* number of days until a student's next birthday is 175.11. Discuss (two or more complete sentences) the **differences and similarities** between 175.11 and the answers you got in (a) and (c).

Instructors will need a bit of flexibility in how to interpret this one's answer.

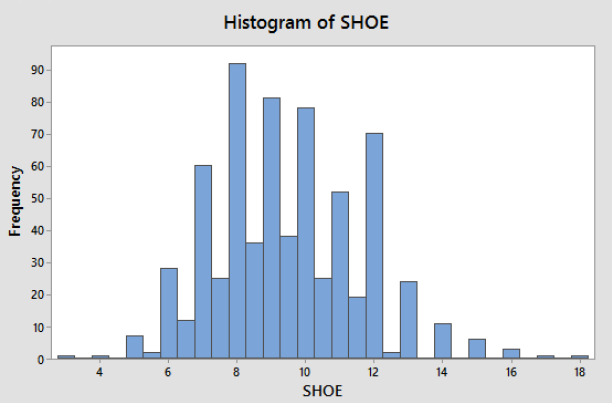
The 'convenience sample' mean found in (b.) *157.6* underestimates the 'population' mean 175.11, but that hardly matters. As it is not a random sample, there is no long-run guarantee that means from such samples would or would not come close to the population mean. The 'SRS' mean found in part (d.) of *180.3* is a little *above* the population mean of 175.11. However, in the 'long run,' the distribution of sample means centered around the population mean. *Students* may further remark that more samples would have a more even mix of $\bar{x}$ values above and below the population mean.

Name:
E number:
Math 1530 section number

Problem 2(e): If your E number ends in an even number (0, 2, 4, 6, or 8) then do this question. (Omit this page/problem if your E# ends with an odd number.)

1. Question 5 of the SPRING 2018 survey asked students, "What is your U.S. shoe size?"

a. Create an appropriate graph to display the *distribution* of the variable called **SHOE** and insert it here.



b. Which of the following best describes the shape of the distribution? Underline (or highlight) your answer.
Skewed left Uniform Skewed right Bimodal Symmetric

c. Using Minitab, calculate the basic statistics for the data collected on **SHOE**. Copy and paste all of the Minitab output here.

Statistics

Variable	N	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	IQR
SHOE	675	9.4452	0.0823	2.1377	3.0000	8.0000	9.0000	11.0000	18.0000	3.0000

Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **SHOE**.

d. Which statistic will you use to describe the center of the distribution? **Mean (or median)**

e. In one or two sentences, describe why this statistic was chosen. **Since the shape of the distribution is symmetric, the mean should be used to describe the center of the distribution. However, since the mean and the median are approximately the same, one could report either value.**

f. What is the value of that statistic? $\bar{x}$ **= 9.4452 (median = 9)**

g. Which statistic(s) will you use to describe the spread of the distribution? **Standard deviation**

h. What is (are) the value(s) of that (those) statistic(s)? **s = 2.1377**

Name:
E number:
Math 1530 section number

i. Are there any outliers in this distribution? Justify your answer using the IQR rule or an appropriate plot.

IQR rule says that any value *below* Q1 – 1.5*IQR or *above* Q3 + 1.5*IQR are outliers.

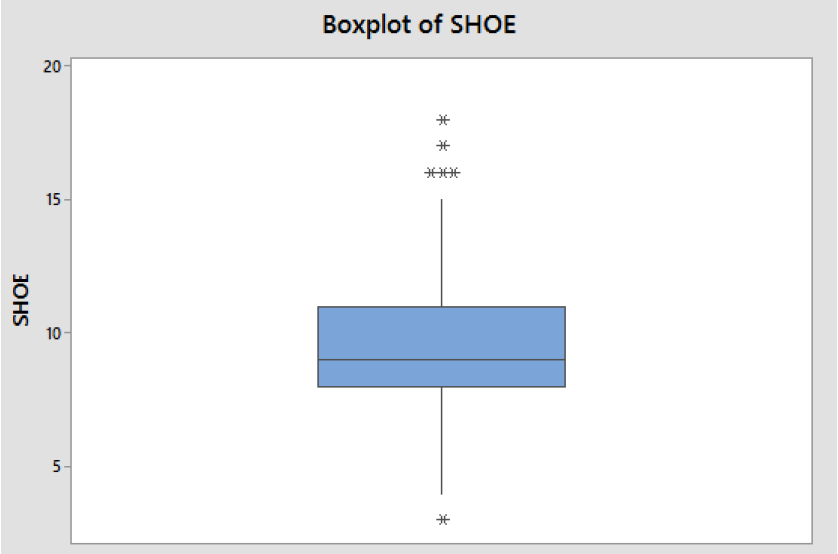
IQR = Q3 – Q1 = 11 – 8 = 3, so 1.5 * IQR = 1.5 * 3 = 4.5.

Q1 – 1.5 * IQR = 8 – 4.5 = 3.5 and Q3 + 1.5 * IQR = 11+ 4.5 = 15.5

Any value of 'SHOE' *below* 3.5 or *above* 15.5 would be considered outliers.

Yes, there are definitely outliers in the distribution of 'SHOE' on the upper end (5 outliers) and one on the lower end.

Minitab shows outliers with * on a boxplot:

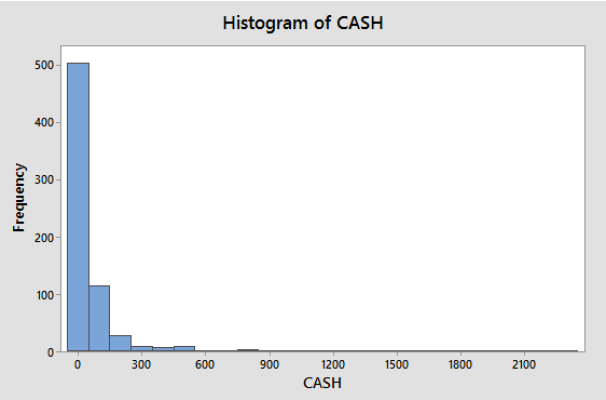


Name:
E number:
Math 1530 section number

Problem 2(o): If your E number ends in an odd number (1, 3, 5, 7, or 9) then do this question. (Omit this page/problem if your E# ends with an even number.)

Question 3 of the SPRING 2018 survey asked students, "How much cash do you have in your wallet or purse right now?"

a. Create an appropriate graph to display the *distribution* of the variable called **CASH** and insert it here.



b. Which of the following best describes the shape of the distribution? Underline (or highlight) your answer.

Skewed left Uniform Skewed right Bimodal Symmetric

c. Using Minitab, calculate the basic statistics for the data collected on **CASH**. Copy and paste all of the Minitab output here.

Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	IQR
CASH	675	0	56.82	6.72	174.57	0.00	0.00	10.00	50.00	2300.00	50.00

Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **CASH**.

d. Which statistic will you use to describe the center of the distribution? **Median**

e. In one or two sentences, describe why this statistic was chosen. **Since the shape of the distribution is skewed right, the median should be used to describe the center of the distribution instead of the mean because the median is robust to outliers while the mean is highly affected by outliers.**

f. What is the value of that statistic? **10**

g. Which statistic(s) will you use to describe the spread of the distribution? **Q1, Q3, and possible IQR**

h. What is (are) the value(s) of that (those) statistic(s)? **Q1 = 0, Q3 = 50, and possible IQR =50**

Name:
E number:
Math 1530 section number

i. Are there any outliers in this distribution? Justify your answer using the IQR rule or an appropriate plot.

IQR rule says that any value *below* Q1 – 1.5*IQR or *above* Q3 + 1.5*IQR are outliers.

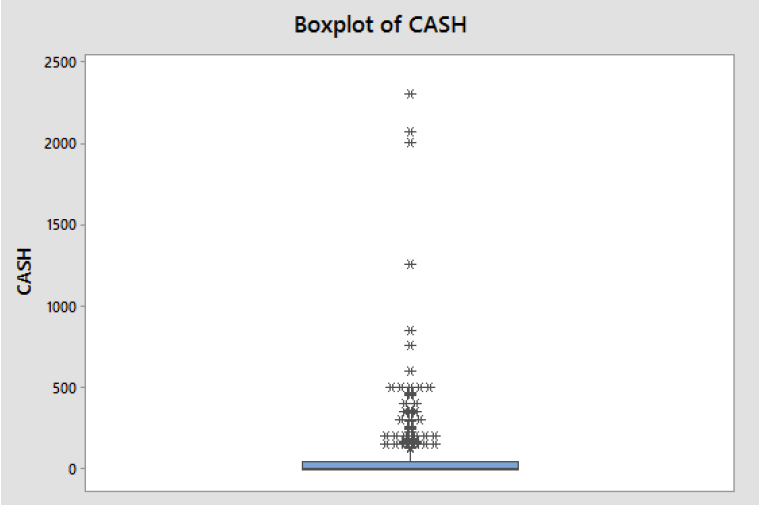
IQR = Q3 – Q1 = 50 – 0 = 50, so 1.5 * IQR = 1.5 * 50 = 75.

Q1 – 1.5 * IQR = 0 – 75 = –75 and Q3 + 1.5 * IQR = 50 + 75 = 125

Any value of 'CASH' *below* -75 or *above* 125 would be considered outliers.

Yes, there are definitely outliers in the distribution of 'CASH' on the upper end.

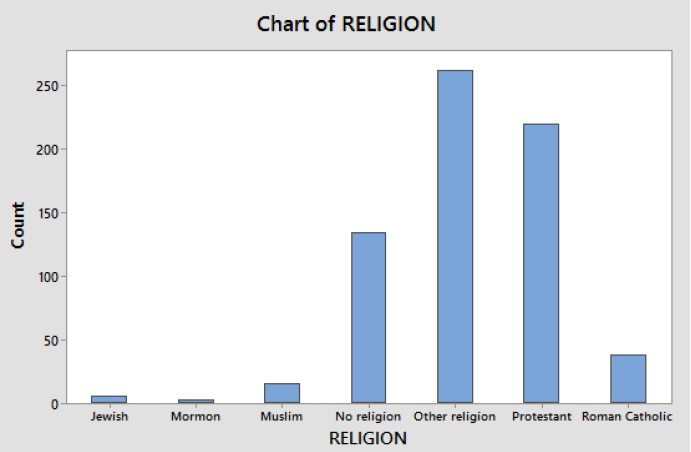
Minitab shows outliers with * on a boxplot:



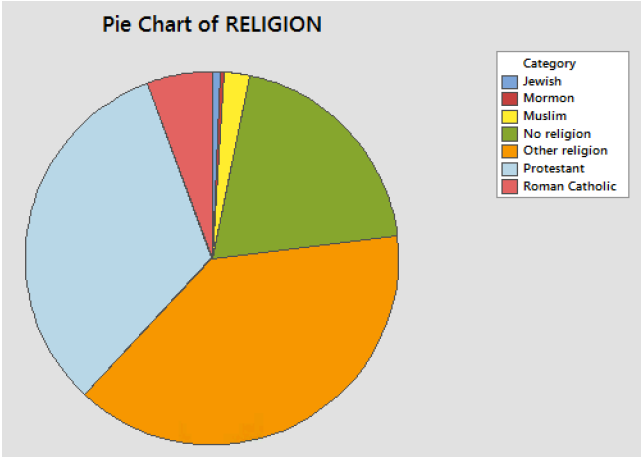
Name:
E number:
Math 1530 section number

Problem 3: RELIGION versus BIRTHDAY. Question 1 of the survey asked students, "How many days left until your next birthday?" Question 11 of the survey asked students, "What is your religious preference? (Jewish, Mormon, Muslim, No Religion, Other Religion, Protestant, Roman Catholic)?"

a. Create a suitable graph to display the *distribution* of **RELIGION** and insert it here.



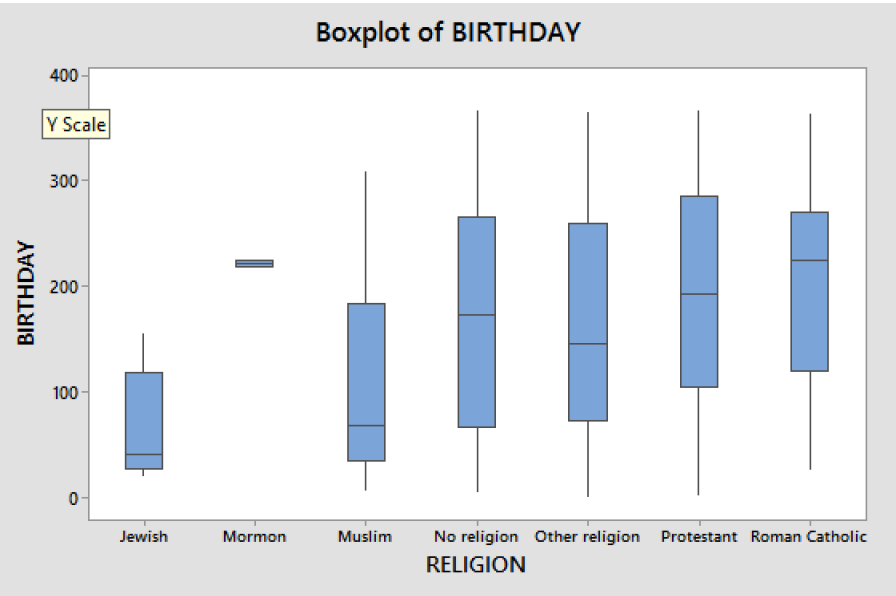
OR



b. What is the mode of this distribution? Underline (or highlight) one option.

Jewish Mormon Muslim No Religion Other Religion Protestant Roman Catholic

c. Create a side-by-side boxplot to display the age of students for the different levels of **RELIGION**. (Go to Graph > Boxplot > One Y with Groups > OK. Select **BIRTHDAY** for the "Graph variables" slot and **RELIGION** for the "Categorical variables for grouping" slot.) Insert your graph here.



Use the side-by-side boxplot found in part (c) to answer the following questions excluding Mormon as possible answer since there are only 2 observations for this religion group.

Name:
E number:
Math 1530 section number

d. Which religion group has the smallest variability in terms of the number of days until their next birthday? **Jewish**

e. Which religion group has the highest median number of days until their next birthday? **Roman Catholic**

f. Which religion group has the largest IQR in terms of the number of days until their next birthday? **No religion**

Name:
E number:
Math 1530 section number

Problem 4: REGRESSION We are interested in creating a regression model between two numeric variables asked from the survey.

a. Which two quantitative variables have the strongest correlation (excluding **ID**, **BIRTHDAY.SUBSET, GENDER.SUBSET, HAIR.SUBSET**)?

Correlations

	BIRTHDAY	CASH	HEIGHT	SHOE	SHOWER	HAIR
CASH	-0.048					
HEIGHT	0.075	0.085				
SHOE	-0.001	0.074	0.686			
SHOWER	-0.052	0.103	0.077	0.182		
HAIR	0.034	0.044	0.241	0.231	0.191	
DISTANCE	-0.025	0.006	-0.068	-0.056	0.128	-0.038

Cell Contents
Pearson correlation

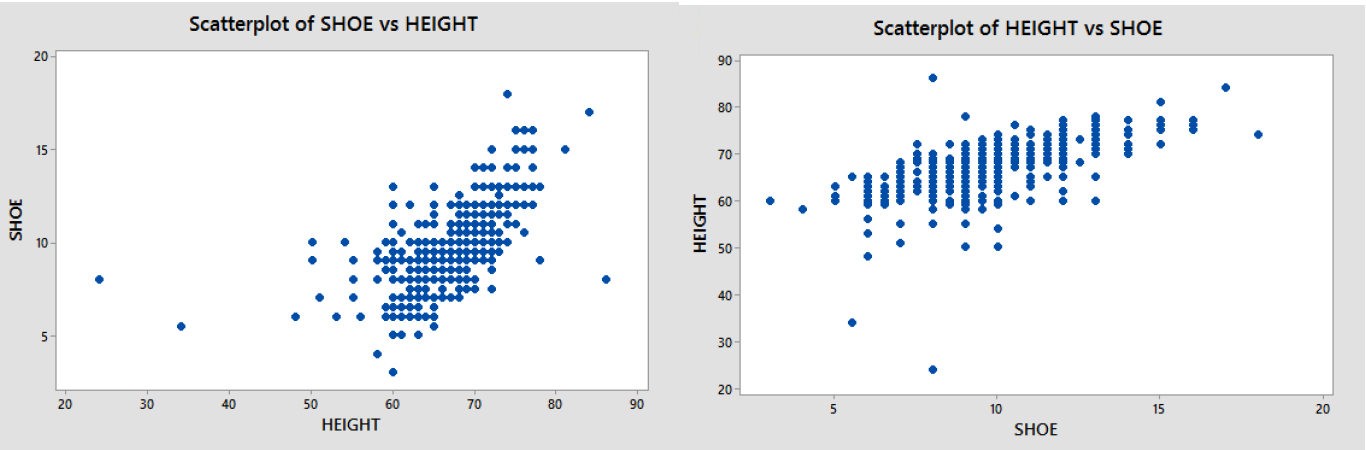
Based on the above correlation matrix, SHOE and HEIGHT have the strongest correlation

b. Report the value of the correlation between these two variables? ***r* = 0.686**

c. Create an appropriate graph to display the relationship between these two variables.

The first graph assumes HEIGHT is the explanatory variable and SHOE is the response variable. The second graph assumes SHOE is the explanatory variable and HEIGHT is the response variable.

Instructors: I feel a student could justify either situation in which variable they let be the response and explanatory.



d. Does the plot show a positive association, a negative association, or no association between these two variables? EXPLAIN what this means with respect to the variables being studied.

(Y = SHOE, X = HEIGHT): Positive association; As height increases, shoe size also increases.

(Y = HEIGHT, X = SHOE): Positive association; As shoe size increases, height also increases.

Name:
E number:
Math 1530 section number

e. Describe the *form* of the relationship between these two variables**.**

(Y = SHOE, X = HEIGHT): Linear (Possibly Curved)

(Y = HEIGHT, X = SHOE): Linear

f. Based on the information displayed in the graph and the correlation you just reported, how would you describe the *strength* of the association?

The strength is moderate to somewhat strong.

g. Using Minitab, obtain the equation for the least squares regression line. Use your best judgement to determine which is the response variable and which is the predictor variable.

SHOE = -8.802 + 0.2734 HEIGHT

OR

HEIGHT = 50.471 + 1.7227 SHOE

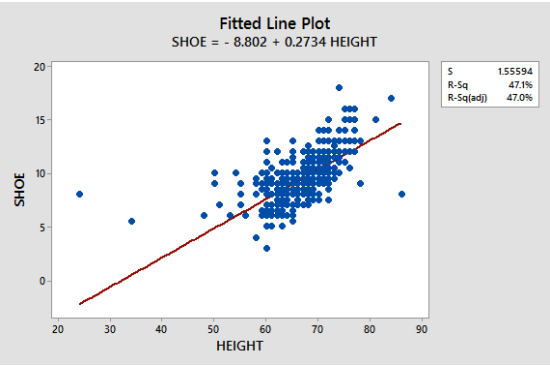
h. Interpret the value of the slope in the least squares regression equation you found in part (f).

(Y = SHOE, X = HEIGHT): As height increases by one inch, the estimated shoe size increases by 0.2734.

(Y = HEIGHT, X = SHOE): As shoe size increase by one size, the estimated height increases by 1.7227 inches.

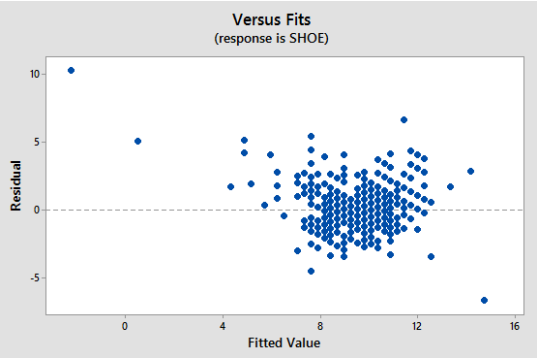
i. How well does the regression equation fit the data? Explain. Justify your answer with appropriate plot(s) and summary statistics.

(Y = SHOE, X = HEIGHT):



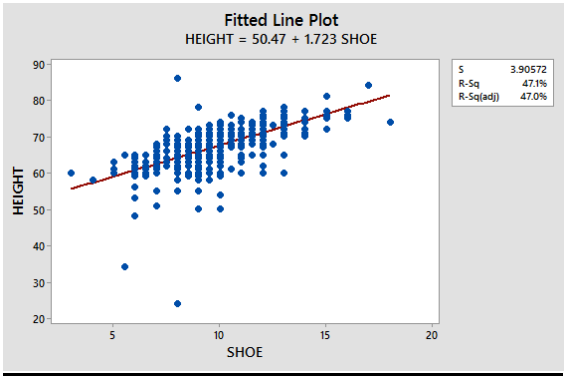
The association is somewhat moderate and can be seen clearly in the fitted line plot. There are a few points that are scattered far away from the regression line. The squared correlation (R^2) indicates that 47.1% of the variation we observed in shoe size is explained by the linear relationship with height.

Name:
E number:
Math 1530 section number

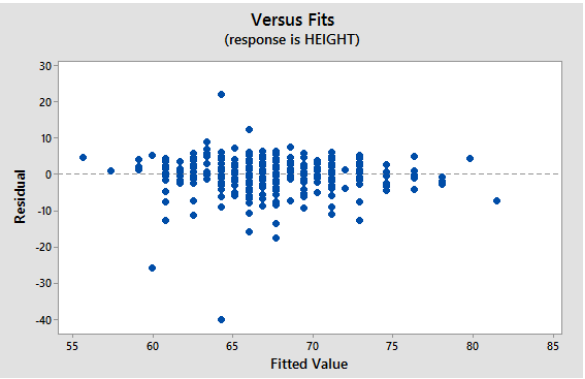


Note: Another scatterplot that is useful to see whether the model makes sense is the residual plot. This helps in determining the appropriateness of the regression model. Recall that the residuals are Residual = Observed Data – Predicted Data. The residual plot shouldn't have any interesting features, like direction or shape. It should stretch horizontally with about the same amount of scatter about the horizontal line at 0. There should be no bends and no outliers. We see that the plot above may possibly be cause to worry.

(Y = HEIGHT, X = SHOE):



The association is somewhat moderate and can be seen clearly in the fitted line plot. There are a few points that are scattered far away from the regression line. The squared correlation (R^2) indicates that 47.1% of the variation we observed in height is explained by the linear relationship with shoe size.



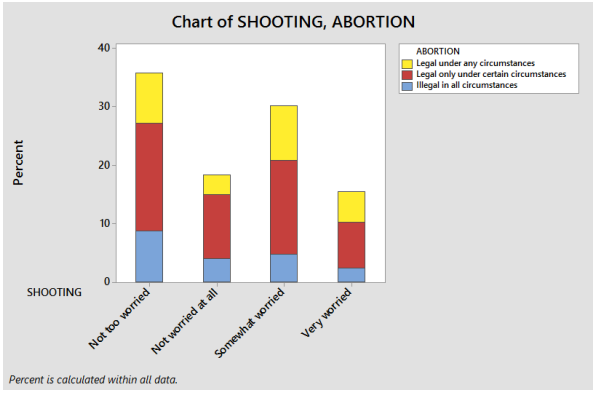
Note: Another scatterplot that is useful to see whether the model makes sense is the residual plot. This helps in determining the appropriateness of the regression model. Recall that the residuals are Residual = Observed Data – Predicted Data. The residual plot shouldn't have any interesting features, like direction or shape. It should stretch horizontally with about the same amount of scatter about the horizontal line at 0. There should be no bends and no outliers. We see that the plot above may possibly be cause to worry.

Name:
E number:
Math 1530 section number

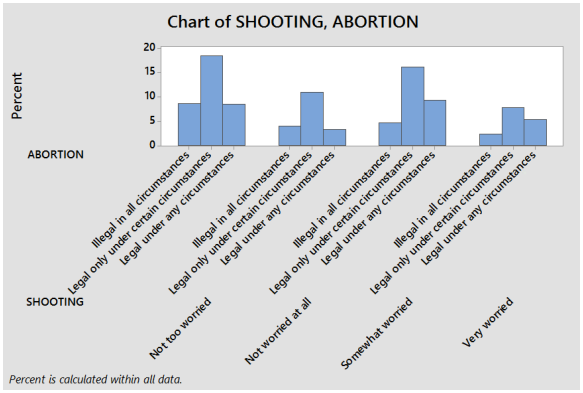
<u>Problem 5 (E):</u> If the sum of the digits in your E number is an even number then do this question. (Omit this page/problem if the sum of the digits in your E number is an odd number.)

SHOOTING and ABORTION: Question 9 from the SPRING 2018 Math 1530 survey asked students "How worried are you that you or someone in your family will become a victim of a mass shooting? (Very worried, Somewhat worried, Not too worried, Not worried at all)" and Question 12 of the survey asked students, "Do you think abortions should be legal under any circumstances, legal only under certain circumstances, or illegal in all circumstances? (Legal under any circumstances, Legal only under certain circumstances, Illegal in all circumstances)." We want to check if there is a relationship between **SHOOTING and ABORTION** among ETSU students. Assume the students who took the (SPRING 2018 Math 1530) class survey are from an SRS of ETSU students.

a. Create an appropriate **graph** to display the relationship between **SHOOTING** and **ABORTION**. Insert your graph here.



OR



b. Create an appropriate two-way table to summarize the data. Insert your table here. (**IN MINITAB: STAT → TABLES → CROSS TABULATION AND CHI-SQUARE. Make sure to select "Options" and click "No variables" under the "*Display missing values for*").**

Rows: SHOOTING Columns: ABORTION

	Illegal in all circumstances	Legal only under certain circum	Legal under any circumstances	All
Not too worried	59	125	58	242
Not worried at all	27	74	23	124
Somewhat worried	32	109	63	204
Very worried	16	53	36	105
All	134	361	180	675

Cell Contents
Count

SUPPOSE WE SELECT ONE STUDENT AT RANDOM: (Calculate the following probabilities and show your work.)

c. What is the probability that this student is very worried about being a victim of a mass shooting *and* feels abortion is legal under any circumstances?

P = 36/675 = 0.0533 = 5.33%

Name:
E number:
Math 1530 section number

d. What is the probability that this student is very worried about being a victim of a mass shooting *or* feels abortion is legal under any circumstances?

P = (105 + 180 – 36)/675 = 0.3688= 36.88%

e. What is the probability that this student is very worried about being a victim of a mass shooting *given* that the student feels abortion is legal under any circumstances?

P = 36/180 = 0.2 = 20%

f. BONUS: Carry out a test for the hypothesis that there is no relationship between **SHOOTING** and **ABORTION**. Use a significance level of $\alpha = 0.05$.

i. State the null and alternative hypotheses.

H_0: There is no relationship between SHOOTING and ABORTION
H_a: There is a relationship between SHOOTING and ABORTION

ii. Perform the test and include any output from Minitab here.

Rows: SHOOTING Columns: ABORTION

	Illegal in all circumstances	Legal only under certain circum	Legal under any circumstances	All
Not too worried	59	125	58	242
	48.04	129.43	64.53	
Not worried at all	27	74	23	124
	24.62	66.32	33.07	
Somewhat worried	32	109	63	204
	40.50	109.10	54.40	
Very worried	16	53	36	105
	20.84	56.16	28.00	
All	134	361	180	675

Cell Contents
Count
Expected count

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	14.230	6	0.027
Likelihood Ratio	14.398	6	0.025

Name:
E number:
Math 1530 section number

iii. Which test statistic are you using and what is its value? **A chi-square test statistic and its value is 14.230**

iv. State your decision and conclusion for the test. **Based on the chi-square test results, the p-value is 0.027. Therefore, at a 5% level of significance, we reject the null hypothesis and conclude there is a significant relationship between shooting and abortion.**

v. Examine the data. Are the conditions for inference in part (ii) violated? Explain.

Conditions for inference about a chi-square test:

*** No more than 20% of the expected counts are less than 5 and all individual expected counts are greater than 0.** **All the expected counts are greater than 5.**

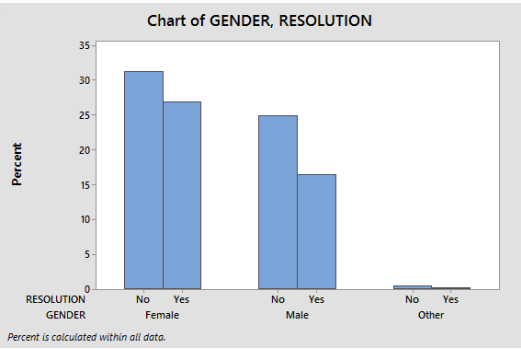
*** The data is a random sample from the population.** **Here the problem states to assume the students who took the class survey are from an SRS of ETSU Students.**

Name:
E number:
Math 1530 section number

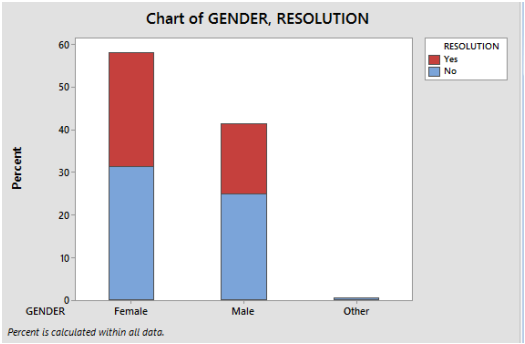
<u>Problem 5 (O):</u> If the sum of the digits in your E number is an odd number then do this question. (Omit this page/problem if the sum of the digits in your E number is an even number.)

GENDER and RESOLUTION: Question 2 from the SPRING 2018 Math 1530 survey asked students "What gender do you identify with? (Female, Male, Other)" and Question 13 from the SPRING 2018 Math 1530 survey asked students "Did you make a New Year's Resolution for 2018? (Yes, No)." We want to check if there is a relationship between **GENDER** and **RESOLUTION** among ETSU students. Assume the students who took the SPRING 2018 Math 1530 class survey are from an SRS of ETSU students.

a. Create an appropriate **graph** to display the relationship between **GENDER** and **RESOLUTION**. Insert your graph here.



OR



b. Create an appropriate two-way table to summarize the data. Insert your table here. (**IN MINITAB: STAT → TABLES → CROSS TABULATION AND CHI-SQUARE. Make sure to select "Options" and click "No variables" under the "*Display missing values for*").**

Rows: GENDER Columns: RESOLUTION

	No	Yes	All
Female	211	181	392
Male	168	111	279
Other	3	1	4
All	382	293	675

Cell Contents
Count

SUPPOSE WE SELECT ONE STUDENT AT RANDOM: (Calculate the following probabilities and show your work.)

c. What is the probability that this student is female *and* made a New Year's resolution for 2018?

P = 181/675 = 0.2681 = 26.81%

d. What is the probability that this student is female *or* made a New Year's resolution for 2018?

P = (392 + 293 – 181)/675 = 0.7467 = 74.67%

Name:
E number:
Math 1530 section number

e. What is the probability that this student is female *given* the student made a New Year's resolution for 2018?

P = 181/293 = 0.6177 = 61.77%

f. BONUS: Carry out a test for the hypothesis that there is no relationship between **GENDER** and **RESOLUTION**. Use a significance level of $\alpha = 0.05$.

i. State the null and alternative hypotheses.

H_0: There is no relationship between GENDER and RESOLUTION
H_a: There is a relationship between GENDER and RESOLUTION

ii. Perform the test and include any output from Minitab here.

Rows: GENDER Columns: RESOLUTION

	No	Yes	All
Female	211	181	392
	221.84	170.16	
Male	168	111	279
	157.89	121.11	
Other	3	1	4
	2.26	1.74	
All	382	293	675

Cell Contents
Count
Expected count

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	3.263	2	0.196
Likelihood Ratio	3.303	2	0.192

2 cell(s) with expected counts less than 5.

iii. Which test statistic are you using and what is its value? **A chi-square test statistic and its value is 3.263**

iv. State your decision and conclusion for the test. **Based on the chi-square test results, the p-value is 0.196. Therefore, at a 5% level of significance, we fail to reject the null hypothesis and conclude there is a not significant relationship between gender and resolution.**

Name:
E number:
Math 1530 section number

v. Examine the data. Are the conditions for inference in part (ii) violated? Explain.

Conditions for inference about a chi-square test:

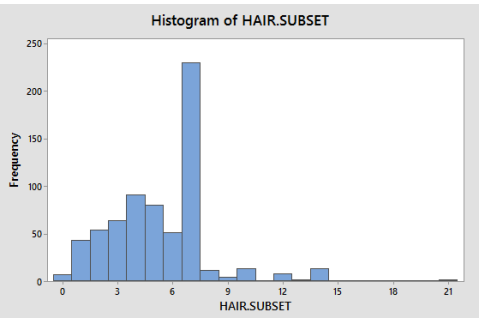
*** No more than 20% of the expected counts are less than 5 and all individual expected counts are greater than 0.** **Less than 20% of the expected counts are less than 5. In this case, there are two less than 5. Thus, violated the condition for this inference and thus, the results may not be valid.**

*** The data is a random sample from the population.** **Here the problem states to assume the students who took the class survey are from an SRS of ETSU Students.**

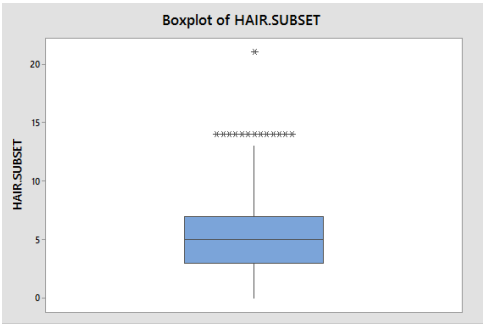
Name:
E number:
Math 1530 section number

Problem 6: Question 7 of the survey asked students, "How many times a week do you wash your hair?" and Question 2 of the survey asked students, "What gender do you identify with?" The variables **GENDER.SUBSET** and **HAIR.SUBSET** were created from a subset of corresponding values. The subset data is to be used to answer this question: On average, is there a difference in the number of times a female student washes their hair in a week versus a male student?

a. Create a suitable graph to display the distribution of **HAIR.SUBSET** reported by our sample of college students and insert it here.



OR



Perform a test of significance to see if, on average, there is a difference in the number of times a female student washes their hair in a week versus a male student using α = 0.05.

b. Write the correct null and alternative hypotheses for the test: **H_o: $\mu_F = \mu_M$ versus H_a: $\mu_F \neq \mu_M$**

c. Use Minitab to perform the appropriate test. Copy and paste the output for the test here.

Assuming Unequal variances:

Method

μ_1: mean of HAIR.SUBSET when GENDER.SUBSET = Female
μ_2: mean of HAIR.SUBSET when GENDER.SUBSET = Male
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics: HAIR.SUBSET

GENDER.SUBSET	N	Mean	StDev	SE Mean
Female	392	4.66	2.09	0.11
Male	279	6.31	3.11	0.19

Estimation for Difference

Difference	95% CI for Difference
-1.649	(-2.069, -1.228)

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-7.70	451	0.000

Name:
E number:
Math 1530 section number

Assuming Equal Variance:

Method

μ_1: mean of HAIR.SUBSET when GENDER.SUBSET = Female
μ_2: mean of HAIR.SUBSET when GENDER.SUBSET = Male
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics: HAIR.SUBSET

GENDER.SUBSET	N	Mean	StDev	SE Mean
Female	392	4.66	2.09	0.11
Male	279	6.31	3.11	0.19

Estimation for Difference

Difference	Pooled StDev	95% CI for Difference
-1.649	2.563	(-2.043, -1.254)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-8.21	669	0.000

d. What is the name of your test statistic and what is its value?

***t* test statistic, t = -7.70 assuming unequal variance and t = -8.21 assuming equal variance**

e. What is the P-value for the test? **P = 0**

f. State your decision regarding the hypotheses being tested.

Because the P-value = 0 is small, we *reject* the null hypothesis.

g. State your conclusion. USE COMPLETE SENTENCES.

Based on the sample data provided, we reject the null hypothesis. We believe, on average, there is a difference in the number of times a female student washes their hair in a week versus a male student.

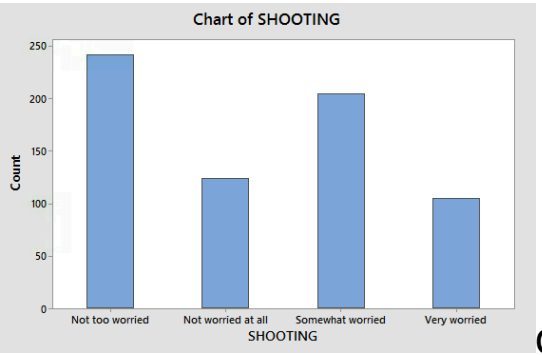
h. What assumptions are you making in order to carry out this test? Is the P-value valid in this case?

ASSUMING the sample of ETSU college students from the Math 1530 survey can be treated as a random/representative sample of college students, the sample size, n = 675, is large enough for the t-statistic to be valid.

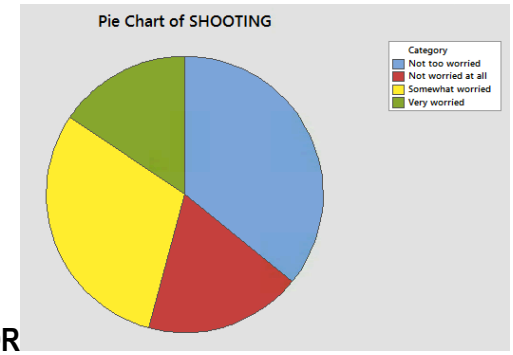
Name:
E number:
Math 1530 section number

Problem 7 Question 9 from the SPRING 2018 Math 1530 survey asked students "How worried are you that you or someone in your family will become a victim of a mass shooting? (Very worried, Somewhat worried, Not too worried, Not worried at all)." Gallup conducted a study and found that 29% of Americans are somewhat worried that they or their family will be a victim of a mass shooting (http://news.gallup.com/poll/220634/four-americans-fear-victim-mass-shooting.aspx?g_source=link_newsv9&g_campaign=item_226202&g_medium=copy). Is the same true for the population of all U.S. college/university students?

a. Create an appropriate graph to display the distribution of **SHOOTING** and insert it here.



OR



b. How many of the students surveyed said they were "Somewhat worried"? **204**

Tally

SHOOTING	Count
Not too worried	242
Not worried at all	124
Somewhat worried	204
Very worried	105
N=	675

c. What proportion of our sample said they were "Somewhat worried"? **204/675 = 0.3022 = 30.22%**

Tally

SHOOTING	Count	Percent
Not too worried	242	35.85
Not worried at all	124	18.37
Somewhat worried	204	30.22
Very worried	105	15.56
N=	675	

d. Assume (for the purpose of this problem) that we may treat the SPRING 2018 sample of Math 1530 students as a simple random sample drawn from the population of all U.S. college/university students. Use Minitab to calculate a 95% confidence interval for the proportion of students in the population who chose "Somewhat worried" to the survey question (based on our sample data). Copy and paste the Minitab output here.

Name:
E number:
Math 1530 section number

Exact:

Descriptive Statistics

N	Event	Sample p	95% CI for p
675	204	0.302222	(0.267767, 0.338417)

Normal Approximation:

Descriptive Statistics

N	Event	Sample p	95% CI for p
675	204	0.302222	(0.267579, 0.336865)

e. Interpret the confidence interval you reported in part (d).

With 95% confidence, the true proportion of students who would chose "Somewhat Worried" to the survey question is between 26.78% and 33.84%.

f. What do you think? Do our results contradict the results obtained from survey by Gallup or do they appear to agree with it? EXPLAIN.

Because the value 29% is within the calculated confidence interval, our sample suggests that the proportion of US college/university students that chose "Somewhat Worried" to the survey question is 29%. Therefore, 29% is within the 95% CI so our results did appear to be in agreement with the survey conducted by Gallup.

Instructors: A student may choose to answer this using a hypothesis test.

Test and CI for One Proportion

Exact:

Test

Null hypothesis	H_0: p = 0.29
Alternative hypothesis	H_1: p ≠ 0.29

P-Value
0.497

Normal Approximation:

Test

Null hypothesis	H_0: p = 0.29
Alternative hypothesis	H_1: p ≠ 0.29

Z-Value	P-Value
0.70	0.484

Because the P-value is greater than 0.05, we fail to reject the null hypothesis and thus, our results did appear to be in agreement with the survey conducted by Gallup.