Lab 13: Contingency Tables and Chi-Square Tests
STAT 301

Group Members: Type all names below:
Student 1 (will type answers to questions 1a, 1e, 2a, 3a, 3ciii): Syrena Hilgendorf
Student 2 (will type answers to questions 1b, 1f, 2b, 3a, 3ci): Marie Klapacz
Student 3 (will type answers to questions 1c, 1g, 2c, 3a): Lorpu Kokoi
Student 4 (will type answers to questions 1d, 1h, 3a, 3b, 3cii): Neha Maddali
If you have less than four students in your group, split up the remaining questions among your group.

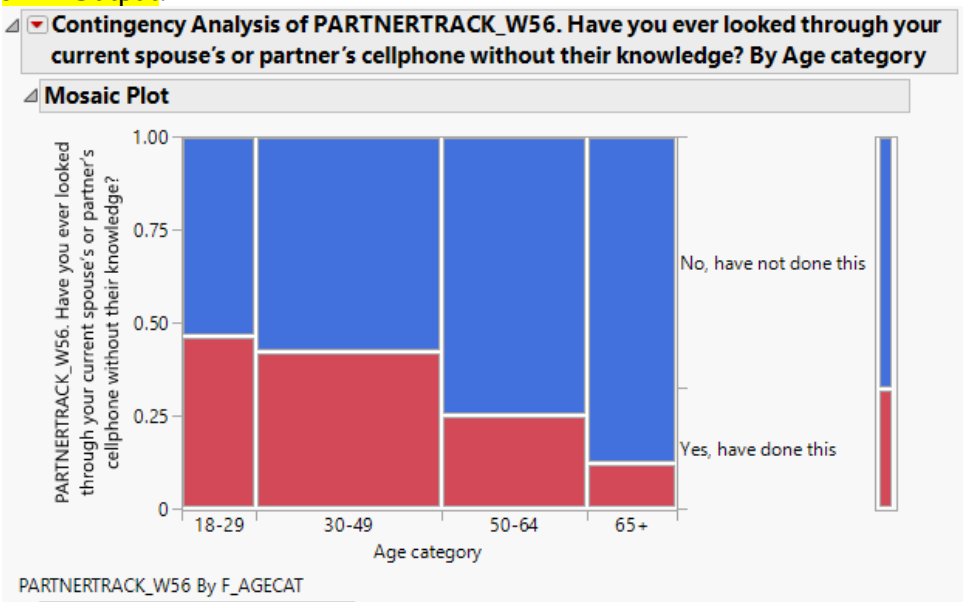
In this lab, you will work with the American Trends Panel-Wave 56 dataset.[1] According to the methodology documentation:

"The American Trends Panel (ATP) is a national, probability-based online panel of adults living in households in the United States. On behalf of the Pew Research Center, Ipsos Public Affairs ("Ipsos") conducted the 56th wave of the panel from October 16, 2019 to October 28, 2019. In total, 4,860 ATP and Knowledge Panel (KP) members (both English- and Spanish-language survey-takers) completed the Wave 56 survey." Note the "probability-based" sample is a random sample.

1. In this question you will examine the probabilities and odds for the PARTNERTRACK to Age category variables. PARTNERTRACK refers to whether or not they had ever looked through their current spouse's or partner's cellphone without their knowledge. Note that each probability and odds are estimates from October 16, 2019 to October 28, 2019 and not necessarily representative of current thoughts.

   (a) Follow the JMP instructions "Contingency Tables and Chi-Square Test for Independence using Fit Y by X" in the JMP Guide. Use PARTNERTRACK as the response variable and Age category as the explanatory variable.

JMP Output:



[1] Data from http://www.pewsocialtrends.org/datasets/

Contingency Table

PARTNERTRACK_W56. Have you ever looked through your ...

Age category	Count / Total % / Col % / Row % / Expected	Yes, have done this	No, have not done this	Total
18-29		199	230	429
		7.01	8.11	15.12
		21.61	12.00	
		46.39	53.61	
		139.27	289.73	
30-49		450	612	1062
		15.86	21.57	37.43
		48.86	31.94	
		42.37	57.63	
		344.766	717.234	
50-64		210	624	834
		7.40	22.00	29.40
		22.80	32.57	
		25.18	74.82	
		270.749	563.251	
65+		62	450	512
		2.19	15.86	18.05
		6.73	23.49	
		12.11	87.89	
		166.215	345.785	
Total		921	1916	2837
		32.46	67.54	

Tests

N	DF	-LogLike	RSquare (U)
2837	3	108.64454	0.0608

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	217.289	<.0001*
Pearson	202.425	<.0001*

(b) For a randomly selected person, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge?

Type answer here: The probability that a randomly selected adult has looked at their spouse's or partner's cellphone without their knowledge is 0.3246

(c) For a randomly selected person, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge and they are 18-29 years old?

Type answer here: The probability that a randomly selected adult has looked at their spouse's or partner's cellphone without their knowledge and they are 18-29 years old is 0.0701

(d) For a randomly selected person who has looked at their spouse's or partner's cellphone without their knowledge, what is the probability that they are 18-29 years old?

Type answer here: The probability that a randomly selected person who has looked at their spouse's or partner's cellphone without their knowledge is between the ages of 18 and 29 years old is 0.2161

(e) For a randomly selected person who is 18-29 years old, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge?

Type answer here: For a randomly selected person who is 18-29 years old, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge is 0.4639

(f) For a randomly selected person who is 50-64 years old, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge?

Type answer here: For a randomly selected person who is 50-64 years old, what is the probability that they have looked at their spouse's or partner's cellphone without their knowledge is 0.2518

(g) In this sample, what is the odds of a person having had looked at their spouse's or partner's cellphone without their knowledge? Show your calculations.

Type answer here: 0.3246 / (1-0.3246) = 0.4806

(h) For people who are 18-29 years old, what is the odds of a person having had looked at their spouse's or partner's cellphone without their knowledge? Show your calculations.

Type answer here: 0.4639 / (1-0.4639) = 0.8653

2. Are age and having had looked at their spouse's or partner's cellphone without their knowledge independent? Answer this question in three ways:

(a) Answer by referring to specific probabilities.

Type answer here: Because the conditional probabilities are not the same as the marginal probabilities, PARTNERTRACK and Age are dependent

P(A if B) = P(A) where A is 18-29 and B is Yes, have done this
$0.4639 \neq 0.3246$

P(A if B) = P(A) where A is 30-49 and B is Yes, have done this
$0.4237 \neq 0.3246$

P(A if B) = P(A) where A is 50-64 and B is Yes, have done this
$0.2518 \neq 0.3246$

P(A if B) = P(A) where A is 65+ and B is Yes, have done this
$0.1211 \neq 0.3246$

(b) Answer by referring to the mosaic plot.

Type answer here: The mosaic plot shows the condition properties of each age group and visibly differing from each other. So PARTNERTRACK and Age are dependent.

(c) Answer by conducting a chi-square test of independence. Make sure to check the conditions, write out your hypothesis statements, provide the test statistic and *p*-value, and write out a conclusion. *Note that you can use JMP to do all computations, including computing the expected counts.*

Type answer here:
H0: Age and having had looked at their spouse's or partner's phone without their knowledge are independent
Ha: Age and having had looked at their spouse's or partner's phone without their knowledge are dependent
Test Statistic: 202.425
Test Statistic p-value: <0.0001
Conclusion: There is overwhelming evidence to suggest that age and having had looked at their spouse's or partner's phone without their knowledge are dependent on one another.

3. Consider the other variables in the dataset. Choose 2 other categorical variables that you are interested in exploring the relationship between.

(a) Which 2 variables will you look at?

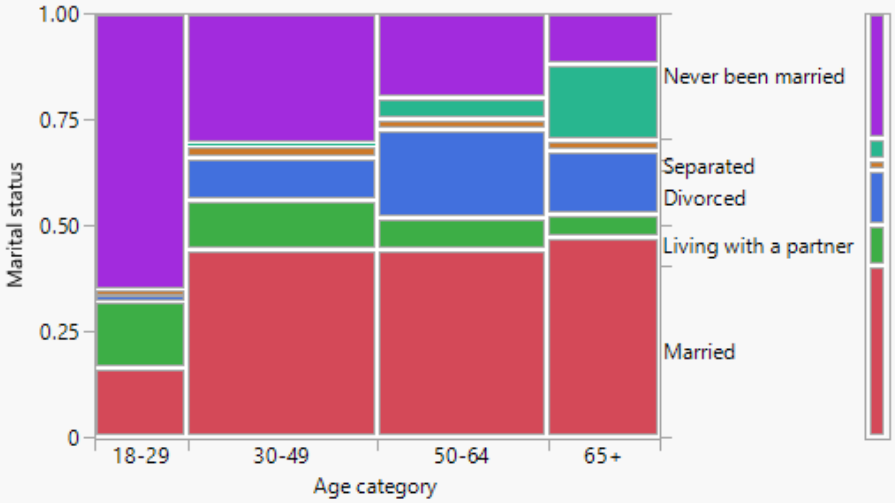
Type answer here: Marital Status (response), Age (explanatory)

(b) Create a contingency table to display your 2 variables.

JMP Output:

Contingency Analysis of Marital status By Age category

Mosaic Plot



F_MARITAL By F_AGECAT

Contingency Table

		Marital status						
	Count Total % Col % Row %	Married	Living with a partner	Divorced	Separated	Widowed	Never been married	Total
Age category	18-29	131 2.70 6.69 16.35	126 2.60 26.92 15.73	11 0.23 1.75 1.37	11 0.23 9.65 1.37	1 0.02 0.40 0.12	521 10.76 36.59 65.04	801 16.54
	30-49	724 14.95 36.96 44.55	187 3.86 39.96 11.51	165 3.41 26.27 10.15	42 0.87 36.84 2.58	8 0.17 3.20 0.49	499 10.30 35.04 30.71	1625 33.55
	50-64	648 13.38 33.08 44.51	106 2.19 22.65 7.28	305 6.30 48.57 20.95	39 0.81 34.21 2.68	68 1.40 27.20 4.67	290 5.99 20.37 19.92	1456 30.06
	65+	456 9.42 23.28 47.45	49 1.01 10.47 5.10	147 3.04 23.41 15.30	22 0.45 19.30 2.29	173 3.57 69.20 18.00	114 2.35 8.01 11.86	961 19.84
	Total	1959 40.45	468 9.66	628 12.97	114 2.35	250 5.16	1424 29.40	4843

Tests

N	DF	-LogLike	RSquare (U)
4843	15	632.32815	0.0896

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	1264.656	<.0001*
Pearson	1285.315	<.0001*

(c) Are your 2 chosen variables independent? Answer this question in three ways:

i. Answer by referring to specific probabilities.

Type answer here:
P(A if B) = P(A) where A is 18-29 and B is Married
$0.1635 \neq 0.4045$

P(A if B) = P(A) where A is 30-49 and B is Married
$0.4455 \neq 0.4045$

P(A if B) = P(A) where A is 50-64 and B is Married
$0.4451 \neq 0.4045$

P(A if B) = P(A) where A is 65+ and B is Married
$0.4745 \neq 0.4045$

ii. Answer by referring to the mosaic plot.

Type answer here: The variables are not independent. The Mosiac Plot displays multiple different layers in between variables, because these layers are not even our categorical variables are dependent on each other.

iii. Answer by conducting a chi-square test of independence. Make sure to show all steps of your hypothesis test.

Type answer here:
H0: Age and marital status are independent
Ha: Age and marital status are dependent
Test Statistic: 1285.315
Test Statistic p-value: <0.0001
Conclusion: There is overwhelming evidence to suggest that age and marital status are dependent.