1. John, the director of a dog shelter, noticed that out of all the dogs present, certain sizes (large, medium, small) are usually preferred more than others. Similarly, he also noticed that the status of the prospective owner (single, married, divorced, widowed) seems to play a part. Determine whether the various dog breeds are equally distributed between the owner's status when the level of significance is 5%.

		Single	Married	Divorced	Widowed	Row Totals
Large Dogs	Observed	8	6	1 A?	9	32
	Expected	6 B?	8.37	8.07	8.07	
Medium Dogs	Observed	11	7	12	10	2 C?
	Expected	9.35	7 D?	10.09	10.09	
Small Dogs	Observed	3 E?	15	6	8	35
	Expected	8.18	8 F?	8.83	8.83	
Column Totals		25	28	4 G?	27	**Overall Total** 5 H?

A = $32-(8+6+9)$
$= 9$

B = $(32\times25)/107$
$= 7.48$

C = $11+7+12+10$
$= 40$

D = $(40\times28)/107$
$= 10.47$

E = $25-(11+8)$
$= 6$

F = $(35\times28)/107$
$= 9.16$

G = $9+12+6$
$= 27$

H = $32+40+35$
$= 107$

a. What type of test is this?

Homogeneity

b. State the Hypotheses.

H_0: The dist. are Homogeneous. H_1: The dist. are not Homogeneous.

c. What is the Test Statistic? L_1: observed by row (L→R) ; L_2: expected by row

$L_3 = (L_1 - L_2)^2 / L_2$ → 1-Var Stat L_3 → $\Sigma x = TS = 8.01$

d. What is the degrees of freedom?

$df = (3-1)(4-1) = 6$

e. Calculate the p-value. What is the decision? $\alpha = 0.05$

$p = \chi^2cdf(8.01, E99, 6) = 0.2374 > 0.05$

Fail to Reject H_0

f. Choose the correct interpretation.

- ❑ At 5% level of significance, there is sufficient evidence to support that the variables are not homogeneous. ← H_1
- ~~☒~~ At 5% level of significance, there is insufficient evidence to support that the variables are not homogeneous. ← H_1

2. Jackson, at random, wondered whether there is a relationship between gender and pet species preference (cat, dog, other, or none). Complete the questions using the observed count table below. [Note: the level of significance is 1%.]

	Cat	Dog	Other	None	Row Totals
Male	13	22	17	20	72
Female	19	16	12	14	61
Column Totals	32	38	29	34	**Overall Total** 133

Enter into Matrix [A]

a. What type of test is this?

Independence

b. State the Hypotheses.

H_0: The variables are Independent. H_1: The variables are dependent.

c. Calculate the Expected Counts in the table below. χ^2-Test

	Cat	Dog	Other	None
Male	17.32	20.57	15.70	18.41
Female	14.68	17.43	13.30	15.59

d. What is the Test Statistic? χ^2-Test

TS = 3.10

e. What is the degrees of freedom? χ^2-Test

df = 3

from χ^2-Test

f. Calculate the p-value. What is the decision? $\alpha = 0.01$

$p = 0.3758 > 0.01 \rightarrow$ Fail to Reject H_0

g. Choose the correct interpretation.

- ❑ At 1% level of significance, there is sufficient evidence to support that there is an association between gender and preferred pet. ← H_1
- ~~❑~~ At 1% level of significance, there is insufficient evidence to support that there is an association between gender and preferred pet.

3. Ella, a beauty CEO, is performing an experiment comparing the time it takes (in minutes) for different face mask formulas to dry to determine which face mask is best for quick use. Use the table below to answer the questions, keeping in mind that we are testing against a 10% significance level.

Formula A → L_1	Formula B → L_2	Formula C → L_3
3.72	2.15	1.35
4.12	3	2
3.98	2.59	2.31
4	2.68	1.98
3.54	3.02	2.12
3.88	2.55	1.87
4.21		1.64
3.79		

a. What type of test is this?

ANOVA

b. Identify the Factor, Levels, and response variables.

Face Masks   Formulas A, B, & C   dry time

c. State the Hypotheses.

$H_0: \mu_A = \mu_B = \mu_C$ H_1: At least one differs.

d. Fill in the ANOVA table.

	SS	Df	MS	TS	p-value
Factor	A?	2	7.73	96.63	D?

(Variable)					
Residual (Error)	1.46	B?	C?		
Total	16.92	20			

A = $16.92 - 1.46 = 15.46$

B = $20 - 2 = 18$

C = $\frac{1.46}{18} = 0.08$

D = $p = Fcdf(96.63, E99, 2, 18) = 0.0000$

e. What is the decision? $\alpha = 0.10$

$0.0000 < 0.10 \rightarrow$ Reject H_0

f. Choose the appropriate interpretation.

- ☒ At 10% level of significance, there is sufficient evidence to support that at least one of the treatment means differ. $\leftarrow H_1$
- ❑ At 10% level of significance, there is insufficient evidence to support that at least one of the treatment means differ. $\leftarrow H_1$

f. Would we move onto Post-Hoc? If yes, continue.

Yes

g. Write the Post-Hoc Hypotheses.

$H_0: \mu_A = \mu_B$ $\quad H_1: \mu_A \neq \mu_B$

$\mu_A = \mu_C$ $\quad \mu_A \neq \mu_C$

$\mu_B = \mu_C$ $\quad \mu_B \neq \mu_C$

h. Analyze the R output below. $\alpha = 0.1$

	diff	lwr.ci	upr.ci	pval		
Formula B - Formula A	1.08	-1.455499	3.6154988	0.8149	> 0.1	Fail to Reject ($\mu_A = \mu_B$)
Formula C - Formula A	-1.42	-3.955499	1.1154988	0.4616	> 0.1	Fail to Reject ($\mu_A = \mu_C$)
Formula C - Formula B	-2.50	-4.814583	-0.1854169	0.0322	< 0.1	Reject ($\mu_B \neq \mu_C$)

i. Interpret the scenario.

At 10% l.o.s, we can say that while μ_A is equal to μ_B & μ_C, μ_B is not equal to μ_C.

4. In the previous years' track and field competition, out of the 55 students that participated from Coach Anderson's school, only 6 won gold, 9 won silver, 11 won bronze, and the remaining students got participation awards. Comparing the observed counts below, did Coach Anderson's school maintain it's previous years' achievements?

$\rightarrow 55-(6+9+11)=29$

	Gold	Silver	Bronze	Participation
Observed	8	7	15	25
Expected	6	9	11	29

a. What type of test is this?

Goodness of Fit

b. State the Hypotheses.

H_0: The claim is correct. H_1: The claim is incorrect.

c. Calculate the Expected Counts.

$\underbrace{E_G=(55)\left(\frac{6}{55}\right)}_{\text{This trend continues!}}=6$ $E_S=9$ $E_B=11$ $E_P=29$

d. What is the degrees of freedom?

$df=4-1=3$

e. What is the Test Statistic?

$$TS=\frac{(8-6)^2}{6}+\frac{(7-9)^2}{9}+\frac{(15-11)^2}{11}+\frac{(25-29)^2}{29}=3.12$$

f. Calculate the p-value. What is the decision? $\alpha=0.05$

$p=\chi^2cdf(3.12,E99,3)=0.3735>0.05$

g. Choose the correct interpretation. Fail to Reject Ho

- [ ] At 5% level of significance, there is sufficient evidence to support that the distribution proportions are not correct. ← H_1
- [x] At 5% level of significance, there is insufficient evidence to support that the distribution proportions are not correct. ← H_1

5. Doctor Smith noticed that a certain gender tends to require a Tonsillectomy before the age of 10, while others require it later in life. Thus, using the data pulled from 73 previous Tonsillectomies, determine if there is a relationship between gender and when a Tonsillectomy, on average, is required. [*Note: the level of significance is 10%]

		10 & Under	Over 10	Row Totals
Male	Observed	19	1 A?	34
	Expected	5 B?	14.9	
Female	Observed	2 C?	17	3 D?
	Expected	21.9	6 E?	
Column Totals		41	4 F?	Overall Total
				73

A = 34-19 = 15 B = (34x41)/73 = 19.10 C = 41-19 = 22

D = 73-34 = 39 E = (39x32)73 = 17.10 F = 73-41 = 32

a. What type of test is this?

Independence

b. State the Hypotheses.

H_0: The variables are Independent. H_1: The variable are dependent.

c. What is the Test Statistic?

$$TS = \frac{(19-19.10)^2}{19.10} + \frac{(15-14.9)^2}{14.9} + \frac{(22-21.9)^2}{21.9} + \frac{(17-17.10)^2}{17.10} = 0.002$$

d. What is the degrees of freedom?

$df = (2-1)(2-1) = 1$

e. Calculate the p-value. What is the decision? $\alpha = 0.1$

$p = \chi^2cdf(0.002, E99, 1) = 0.9643 > 0.1$, Fail to Reject H_0

f. Choose the correct interpretation.

- ❑ At 10% level of significance, there is sufficient evidence to support that there is an association between gender and when a Tonsillectomy is required. $\llcorner H_1$
- ❑ At 10% level of significance, there is insufficient evidence to support that there is an association between gender and when a Tonsillectomy is required. $\llcorner H_1$

6. Steven, a pharmacist, is trying to determine which ADHD medication is most effective out of the 4 most common types. To do so, the pharmacist asks 14 people for each medication how effective they would rate it. Use the statistics below to answer the questions, keeping in mind that we are testing against a 1% significance level.

Medication	Mean Score	Standard Deviation	Sample Size
#1	5.86	1.76	14
#2	3.29	3.21	14
#3	4.33	2.23	14
#4	3.72	1.43	14

a. What type of test is this?

ANOVA

$\bar{X}_{overall} = \frac{5.86+3.29+4.33+3.72}{4} = 4.3$

b. Identify the Factor, Levels, and response variables.

ADHD Meds — ADHD Med #1,2,3&4 — effectiveness score

c. State the ANOVA Hypotheses.

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: At least 1 differs.

d. Fill in the ANOVA table.

	SS	Df	MS	TS	p-value
between Factor (Variable)	53.07	3	17.69	3.47	0.0225

within Residual (Error)	265.45	52	5.10	[redacted]	

$SSB = (14(5.86-4.3)^2) + (14(3.29-4.3)^2) + (14(4.33-4.3)^2) + (14(3.72-4.3)^2) = 53.07$

$SSW = (\overset{13}{\cancel{14-1}})1.76^2 + (13 \times 3.21^2) + (13 \times 2.23^2) + (13 \times 1.43^2) = 265.45$

$DF_B = 4-1 = 3 \qquad Df_W = (14 \times 4) - 4 = 52$

$MSB = \frac{53.07}{3} = 17.69 \qquad MSW = \frac{265.45}{52} = 5.10$

$TS = \frac{17.69}{5.10} = 3.47 \qquad p = Fcdf(3.47, E99, 3, 52) = 0.0225$

e. What is the decision? $\alpha = 0.01$

$0.0225 > 0.01 \rightarrow$ Fail to Reject H_0

f. Choose the appropriate interpretation.

- ❑ At 1% level of significance, there is sufficient evidence to support that at least one of the treatment means differ. $\leftarrow H_1$
- ☒ At 1% level of significance, there is insufficient evidence to support that at least one of the treatment means differ. $\leftarrow H_1$

g. Would we move onto Post-Hoc? If yes, continue.

No

~~h. Write the Post-Hoc Hypotheses.~~

~~i. Analyze the R output below.~~

	diff	lwr	upr	p adj
2-1	0.36250000	0.12528287	0.59971713	0.0010358
3-1	0.07833333	-0.15888380	0.31555047	0.8143113
4-1	0.22000000	-0.01721713	0.45721713	0.0778376
3-2	-0.28416667	-0.52138380	-0.04694953	0.0131752
4-2	-0.14250000	-0.37971713	0.09471713	0.3869986
4-3	0.14166667	-0.09555047	0.37888380	0.3921830

~~j.~~ Interpret the scenario.

7. Sally wants to figure out which of the three highly recommended electric companies is the most recommended. To do so, she compares her survey ratings (shown below out of 90 people for 3 months each) to the company's claimed satisfactory percentages: company A is 70%, company B is 50%, and company C is 90%. Determine if her rates line up with the company's claims with a 5% level of significance.

	Company A	Company B	Company C
Observed	62	43	76
Expected	63	45	81

a. What type of test is this?

Goodness of Fit

b. State the Hypotheses.

$H_0: p_A = 0.70, p_B = 0.50, p_C = 0.90$ H_1: At least 1 differs.

c. Calculate the Expected Counts.

$E_A = (90)(0.70) = 63$ $E_B = (90)(0.50) = 45$ $E_C = (90)(0.90) = 81$

d. What is the degrees of freedom?

$df = 3 - 1 = 2$

e. What is the Test Statistic?

$$TS = \frac{(62-63)^2}{63} + \frac{(43-45)^2}{45} + \frac{(76-81)^2}{81} = 0.41$$

f. Calculate the p-value. What is the decision? $\alpha = 0.05$

$p = \chi^2cdf(0.41, E99, 2) = 0.8146 > 0.05$, Fail to Reject H_0

g. Choose the correct interpretation.

- ☐ At 5% level of significance, there is sufficient evidence to support that the distribution proportions are not correct. ← H_1
- ☐ ~~At 5% level of significance, there is insufficient evidence to support that the distribution proportions are not correct.~~ ← H_1

8. A student noticed that her English teacher, Mrs. Jones, has one of the highest rates of dress coding people throughout the year. The student decided to look into whether the distribution of dress coding (leveled by severity A or B) is equal between genders. [Notes: Severity A dress code is receiving clothes to cover up, Severity B dress code is being sent home to change.] Determine whether there is an equal distribution between genders in terms of dress codes, when the level of significance is 10%.

	Severity A Dress Code	**Severity B Dress Code**	
Male	32	23	55
Female	47	25	72
	79	48	127

a. What type of test is this?

Homogeneity

b. State the Hypotheses.

H_0: The dist. is homogeneous. H_1: The dist. isn't Homogeneous.

c. Calculate the expected counts.

$$E_{MA} = \frac{55 \times 79}{127} = 34.21 \qquad E_{MB} = \frac{55 \times 48}{127} = 20.79$$

$$E_{FA} = \frac{72 \times 79}{127} = 44.79 \qquad E_{FB} = \frac{72 \times 48}{127} = 27.21$$

d. What is the degrees of freedom?

$$df = (2-1)(2-1) = 1$$

e. What is the test statistic?

$$TS = \frac{(32-34.21)^2}{34.21} + \dots + \frac{(25-27.21)^2}{27.21} = 0.67$$

f. Calculate the p-value. What is the decision? $\alpha = 0.1$

$p = \chi^2cdf(0.67, E99, 1) = 0.4131 > 0.1 \rightarrow$ Fail to Reject H_0

g. Choose the correct interpretation.

- ❑ At 10% level of significance, there is sufficient evidence to support that the variables are not homogeneous. ← H_1
- ❑ At 10% level of significance, there is insufficient evidence to support that the variables are not homogeneous. ← H_1