MTH 324, Exam #2, At Home, Fall 2022 Name KEY

Instructions: This exam is in two parts: Part I is to be completed partly at home using the materials posted in the course for the at-home portion and you will answer questions about that work during the in-class portion of the exam; Part II is to be completed entirely in class. You may not use cell phones, and you may only access internet resources you are specifically directed to use.

At home, prepare for questions in Part I using R. Open the data file entitled **324exam2data.xlsx** posted in Blackboard. (Note: this file has multiple sheets of data. You may want to separate the data into separate files to upload to R, or look up how to access multiple sheets in R from a single upload.) Complete the calculations noted below. You will be asked for additional analysis and interpretation of this data in the in-class portion of the test. Print out the results of your analysis and code, and bring the pages with you to the exam. You will submit all this work along with the in-class exam.

From Sheet 1:

1. Husbands and wives go into a car dealership separately and look at purchasing the same vehicle. The data here records the offer of selling price received by each member of the couple. Conduct an appropriate hypothesis test of the data to determine if the selling price offered to men and women is different. Test your assumptions with normal probability plots.

From Sheet 2:

2. Conduct a one-way ANOVA test on whether the five different filling machines output different amounts of the product. Apply Tukey's method to determine how the machines group together. Which machine(s) are most in need of recalibration? Be sure to check your data for normality. Create a comparative boxplot to confirm your analysis.

From Sheet 3:

3. Use the data to determine if men and women in the dataset (of graduate business school students) is married at the same rate. You'll need to count the number of men and women in the data, and within each group, count the number of men that are married, and the number of women that are married. Conduct a two-sample proportion test to determine if the difference statistically significant. Check the assumptions of your test.

4. Using Gender, Marital Status (Married) and Number of Children, conduct a three-way ANOVA of school debt. Test main and interaction effects where possible. Test for the normality of school debt. Apply Tukey's method.

5. Conduct a one-sample hypothesis test of school debt to see if there is significant reason to believe school debt is less than $30,000 per student.

6. Build a sampling distribution of Previous Salary. Collect 1000 samples of 50 students each. Calculate the mean of the sample. Build a histogram of your sample mean data. Find the mean (of the means) and the standard deviation of your sample means (the standard error). Find the mean and standard deviation of the original data. Compare the results.

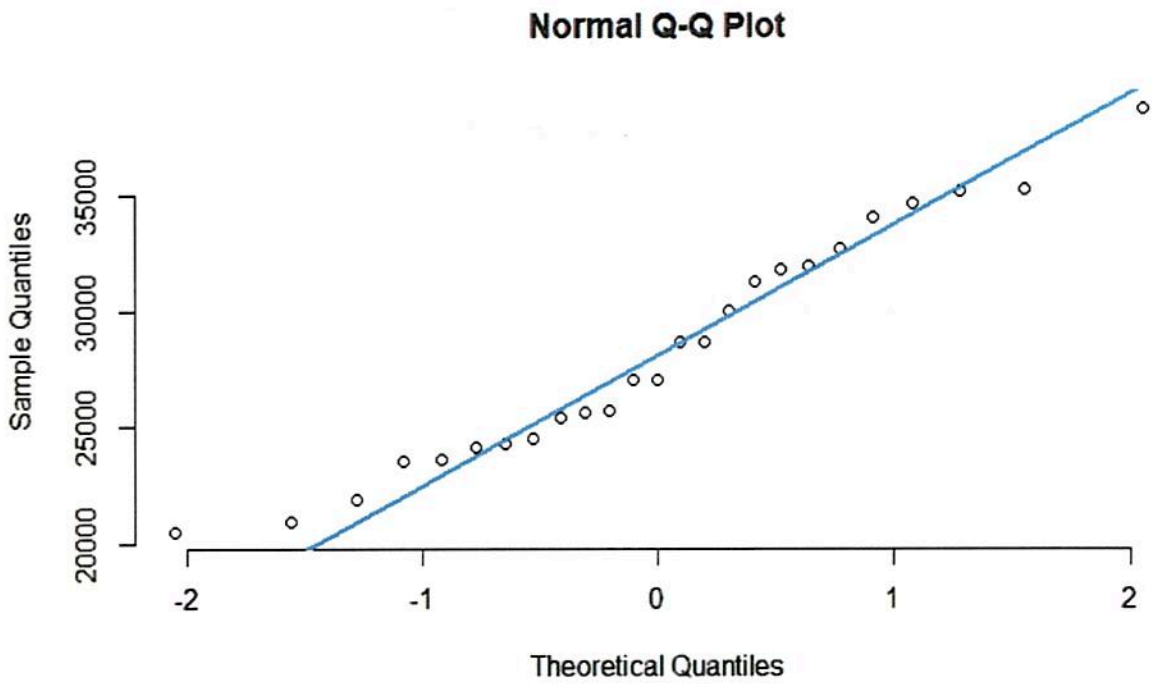
Exam #2 Part 1 Work Solutions

1.

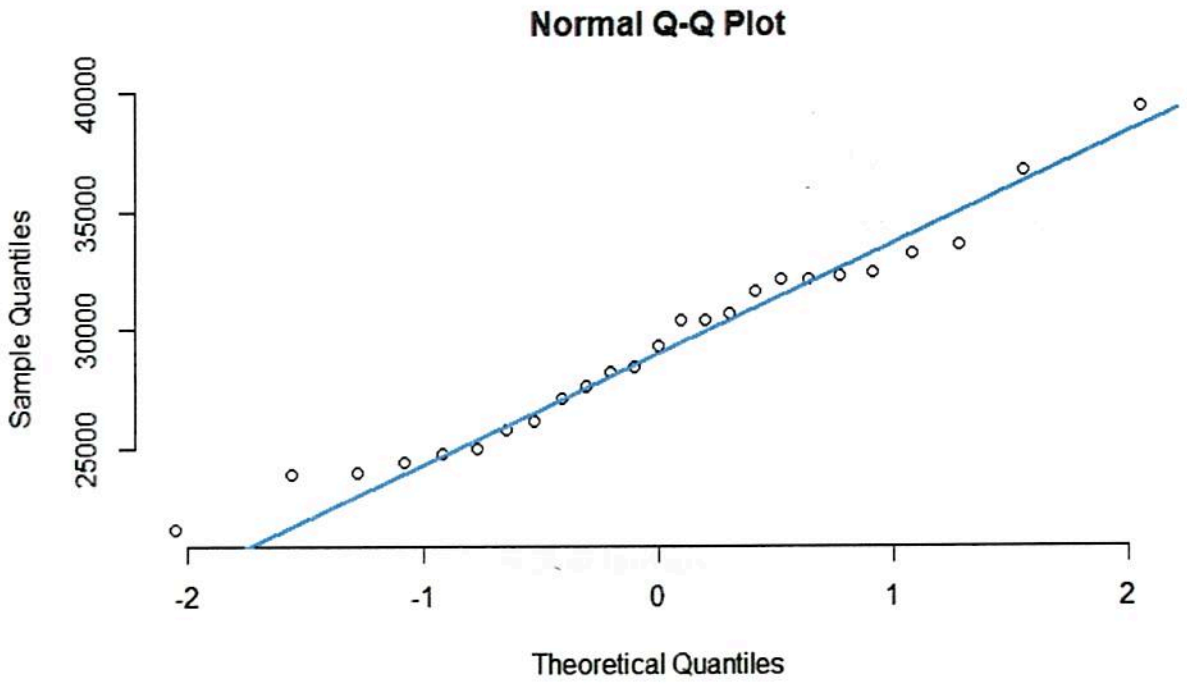
Paired t-test

data: data1$Wife and data1$Husband
t = -1.2978, df = 24, p-value = 0.2067
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2538.5281 578.5281
sample estimates:
mean of the differences
-980

Wife:



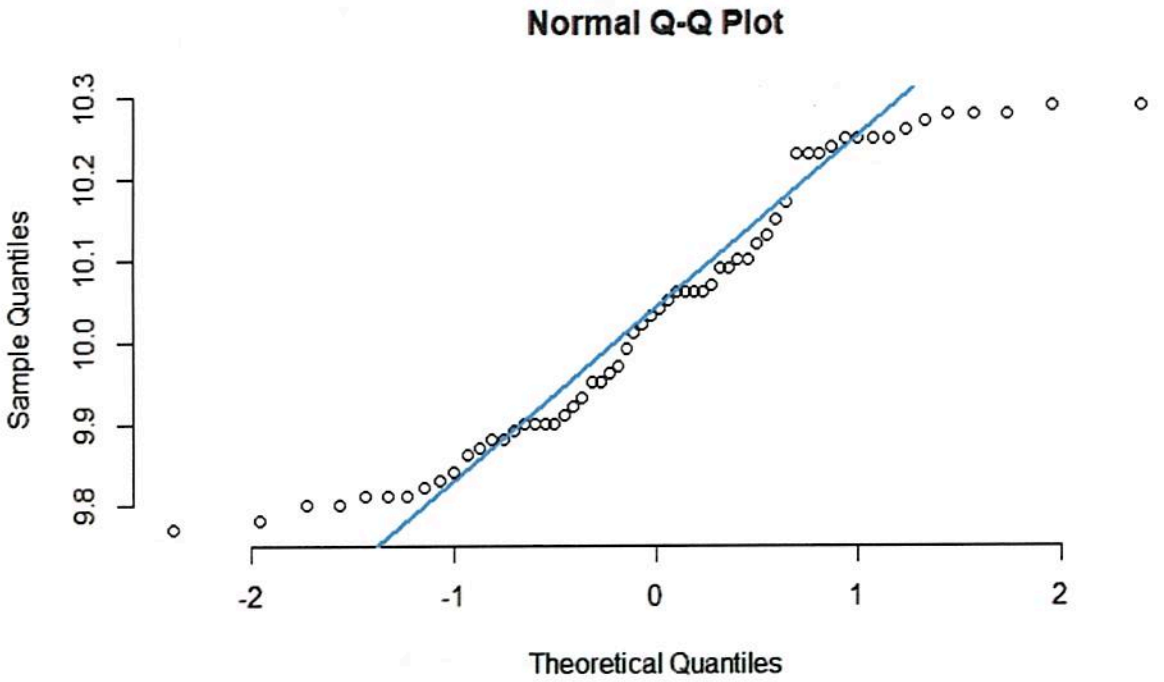
Husband:

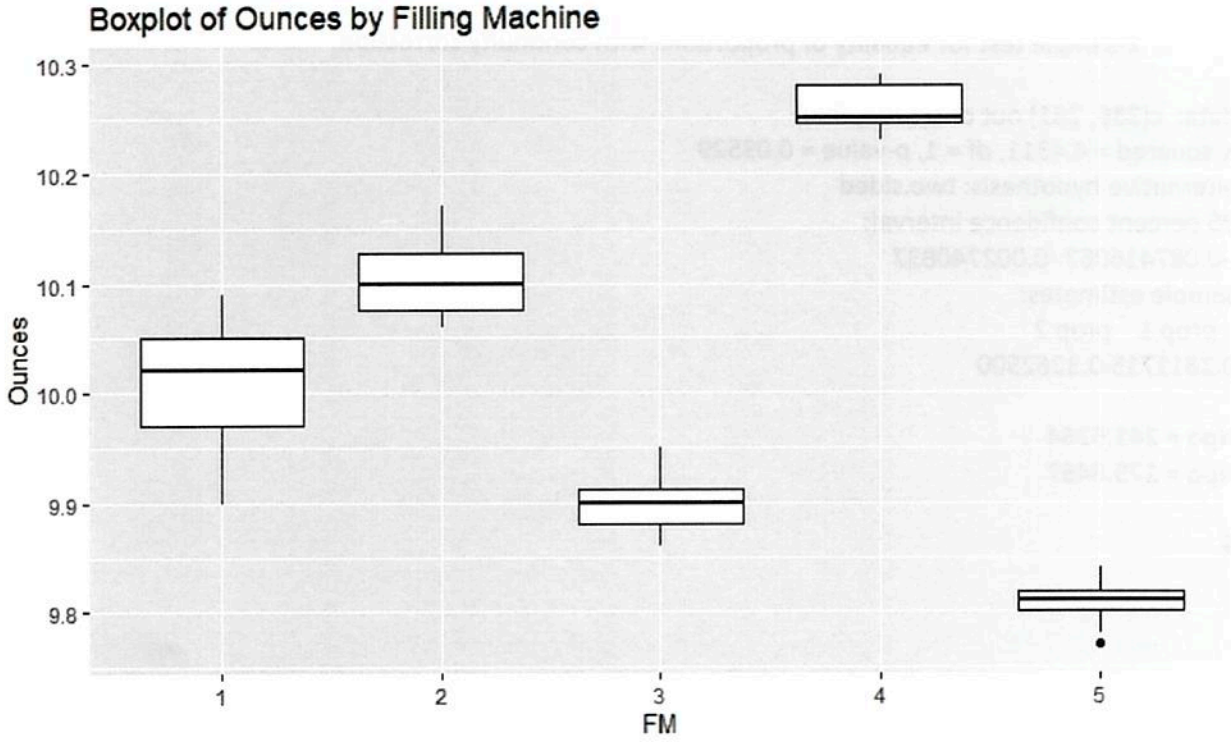


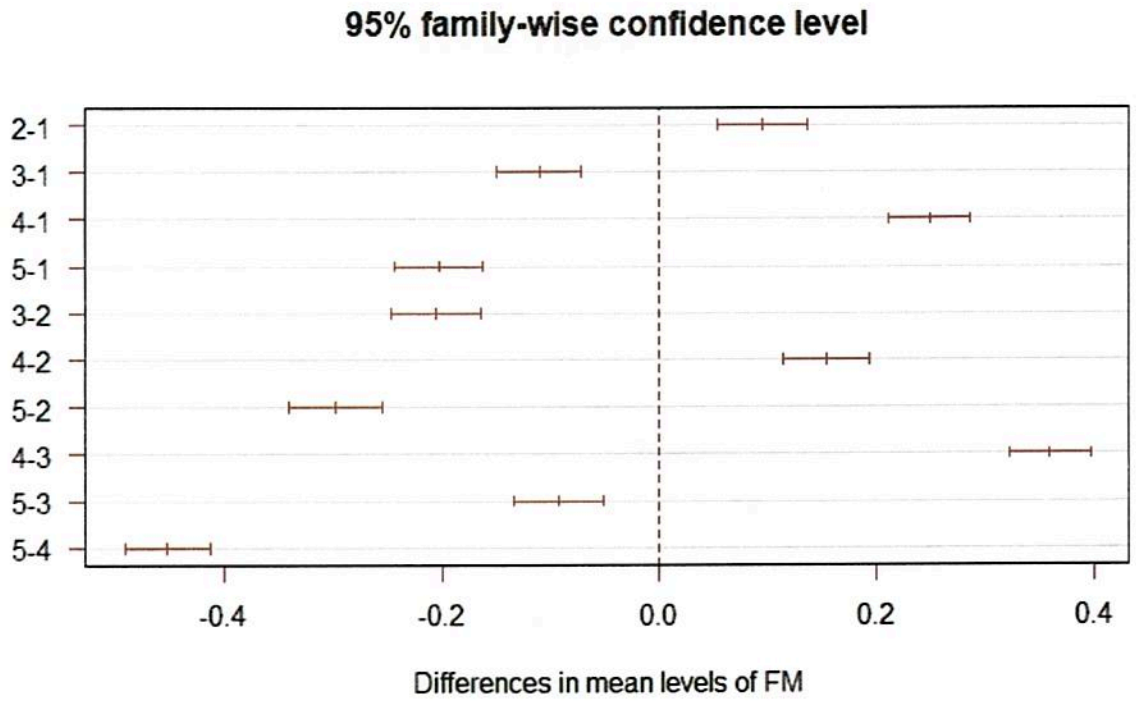
2.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
FM	4	1.5482	0.3871	326.9	<2e-16	***
Residuals	55	0.0651	0.0012			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Normal Q-Q Plot
Sample Quantiles
9.8
9.9
10.0
10.1
10.2
10.3
-2
-1
0
1
2
Theoretical Quantiles


Boxplot of Ounces by Filling Machine
Ounces
10.3
10.2
10.1
10.0
9.9
9.8
1
2
3
4
5
FM



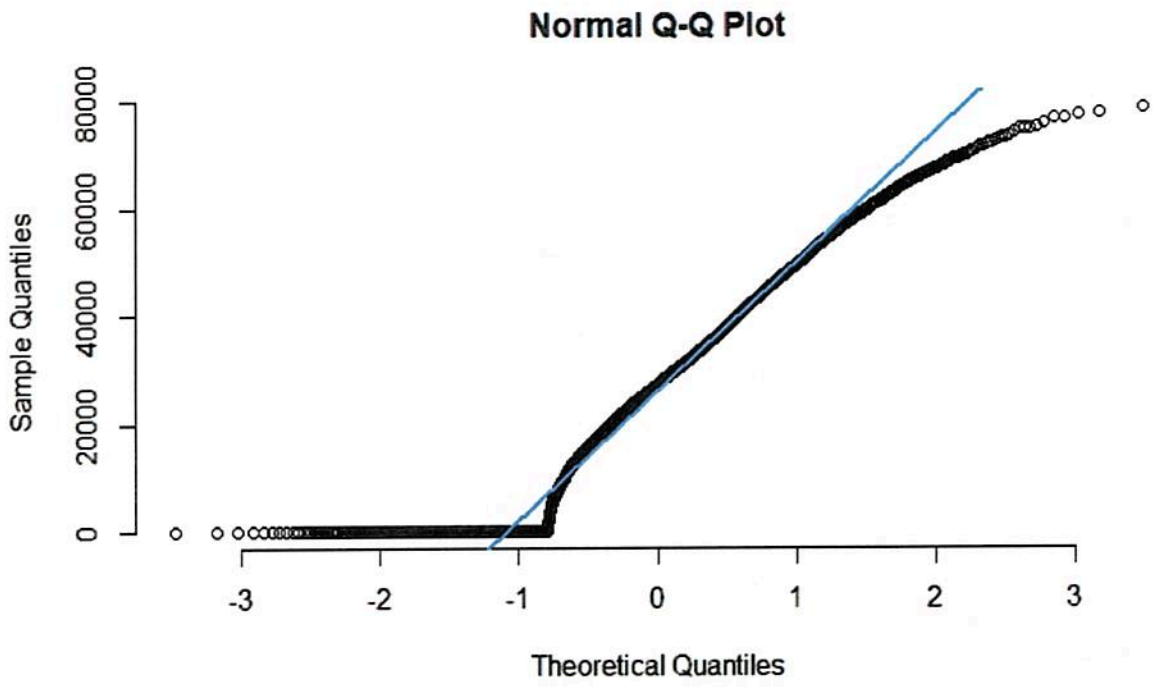
3.

2-sample test for equality of proportions with continuity correction

data: c(336, 261) out of c(1195, 800)
X-squared = 4.4311, df = 1, p-value = 0.03529
alternative hypothesis: two.sided
95 percent confidence interval:
-0.087416067 -0.002740837
sample estimates:
prop 1 prop 2
0.2811715 0.3262500

Npq = 241.5264
Npq = 175.8487

4.



	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Children	2	4.366e+10	2.183e+10	57.035	< 2e-16	***
Gender	1	7.172e+07	7.172e+07	0.187	0.66514	
Married	1	1.497e+10	1.497e+10	39.119	4.87e-10	***
Children:Gender	2	1.178e+08	5.888e+07	0.154	0.85740	
Children:Married	2	3.682e+09	1.841e+09	4.810	0.00825	**
Gender:Married	1	5.662e+08	5.662e+08	1.480	0.22399	
Children:Gender:Married	2	1.986e+09	9.930e+08	2.595	0.07493	.
Residuals	1983	7.589e+11	3.827e+08			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = `School Debt` ~ Children * Gender * Married, data = data3)

$Children

	diff	lwr	upr	p adj
1-0	9990.701	7158.7954	12822.61	0.0000000
2-0	14969.819	10324.3852	19615.25	0.0000000
2-1	4979.118	-210.0949	10168.33	0.0632535

$Gender

	diff	lwr	upr	p adj
Male-Female	386.6568	-1365.985	2139.299	0.6653096

$Married

	diff	lwr	upr	p adj
Yes-No	5047.821	3172.043	6923.599	1e-07

$`Children:Gender`

	diff	lwr	upr	p adj
1:Female-0:Female	10664.9357	5414.020	15915.852	0.0000001
2:Female-0:Female	14375.4666	5272.838	23478.095	0.0001026
0:Male-0:Female	525.6543	-2350.232	3401.541	0.9953477
1:Male-0:Female	10031.3046	5277.355	14785.254	0.0000000
2:Male-0:Female	15858.3548	8532.811	23183.899	0.0000000
2:Female-1:Female	3710.5309	-6310.294	13731.356	0.8984911
0:Male-1:Female	-10139.2814	-15221.585	-5056.978	0.0000002
1:Male-1:Female	-633.6312	-6970.754	5703.491	0.9997469
2:Male-1:Female	5193.4191	-3245.934	13632.772	0.4951550
0:Male-2:Female	-13849.8123	-22856.229	-4843.396	0.0001759
1:Male-2:Female	-4344.1620	-14113.750	5425.426	0.8022900
2:Male-2:Female	1482.8882	-9764.707	12730.484	0.9990242
1:Male-0:Male	9505.6503	4938.624	14072.676	0.0000001
2:Male-0:Male	15332.7005	8127.058	22538.343	0.0000000
2:Male-1:Male	5827.0502	-2312.397	13966.497	0.3187818

$`Children:Married`

	diff	lwr	upr	p adj
1:No-0:No	9696.3436	3676.0588	15716.628	0.0000676
2:No-0:No	25645.4115	2811.1328	48479.690	0.0173030
0:Yes-0:No	8776.4800	5078.0713	12474.889	0.0000000
1:Yes-0:No	12283.2355	8238.5119	16327.959	0.0000000
2:Yes-0:No	15947.3109	10101.6734	21792.948	0.0000000
2:No-1:No	15949.0678	-7563.8490	39461.985	0.3810136
0:Yes-1:No	-919.8637	-7637.8399	5798.113	0.9988281
1:Yes-1:No	2586.8919	-4327.7858	9501.570	0.8943917
2:Yes-1:No	6250.9672	-1849.9251	14351.860	0.2375716
0:Yes-2:No	-16868.9315	-39896.9939	6159.131	0.2930756
1:Yes-2:No	-13362.1759	-36448.3887	9724.037	0.5646056
2:Yes-2:No	-9698.1006	-33166.9077	13770.707	0.8471104
1:Yes-0:Yes	3506.7556	-1517.9001	8531.411	0.3478897
2:Yes-0:Yes	7170.8309	608.9064	13732.755	0.0227783
2:Yes-1:Yes	3664.0753	-3099.0908	10427.241	0.6347634

$`Gender:Married`

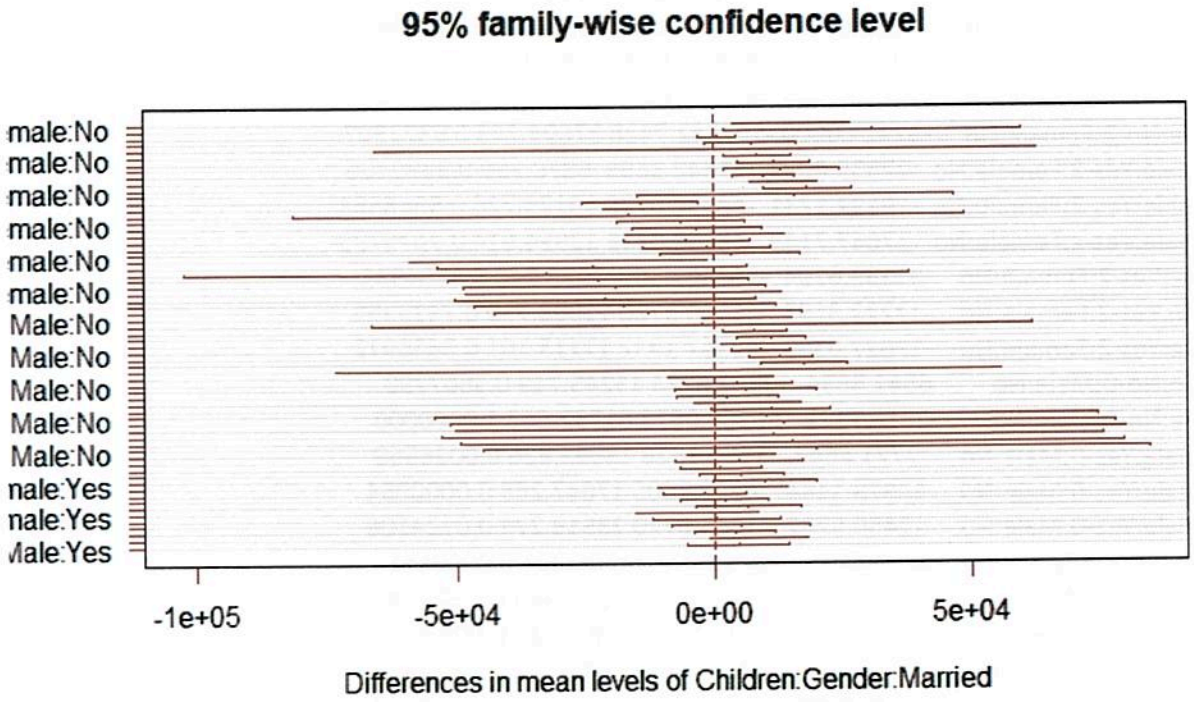
	diff	lwr	upr	p adj
Male:No-Female:No	52.0085	-2711.9899	2816.007	0.9999593
Female:Yes-Female:No	4003.9287	210.7306	7797.127	0.0338681
Male:Yes-Female:No	5888.0541	2391.7020	9384.406	0.0000921
Female:Yes-Male:No	3951.9202	396.6932	7507.147	0.0223499
Male:Yes-Male:No	5836.0456	2599.4168	9072.674	0.0000224

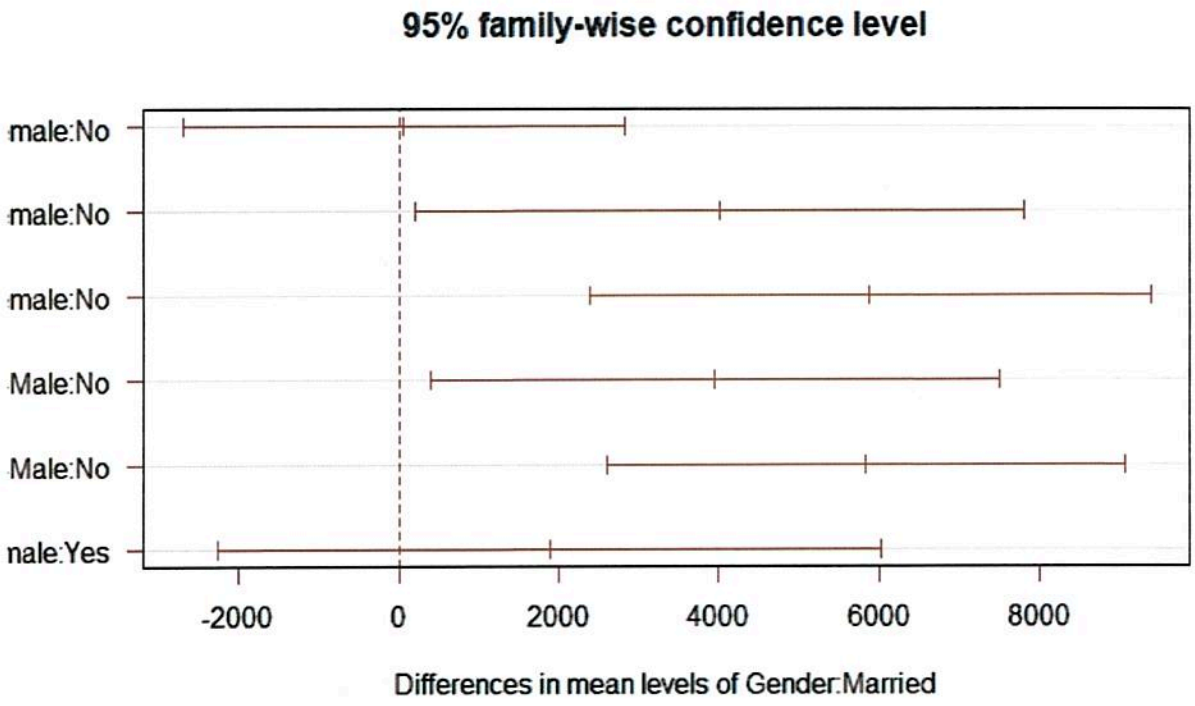
Male:Yes-Female:Yes 1884.1255 -2266.1071 6034.358 0.6476217

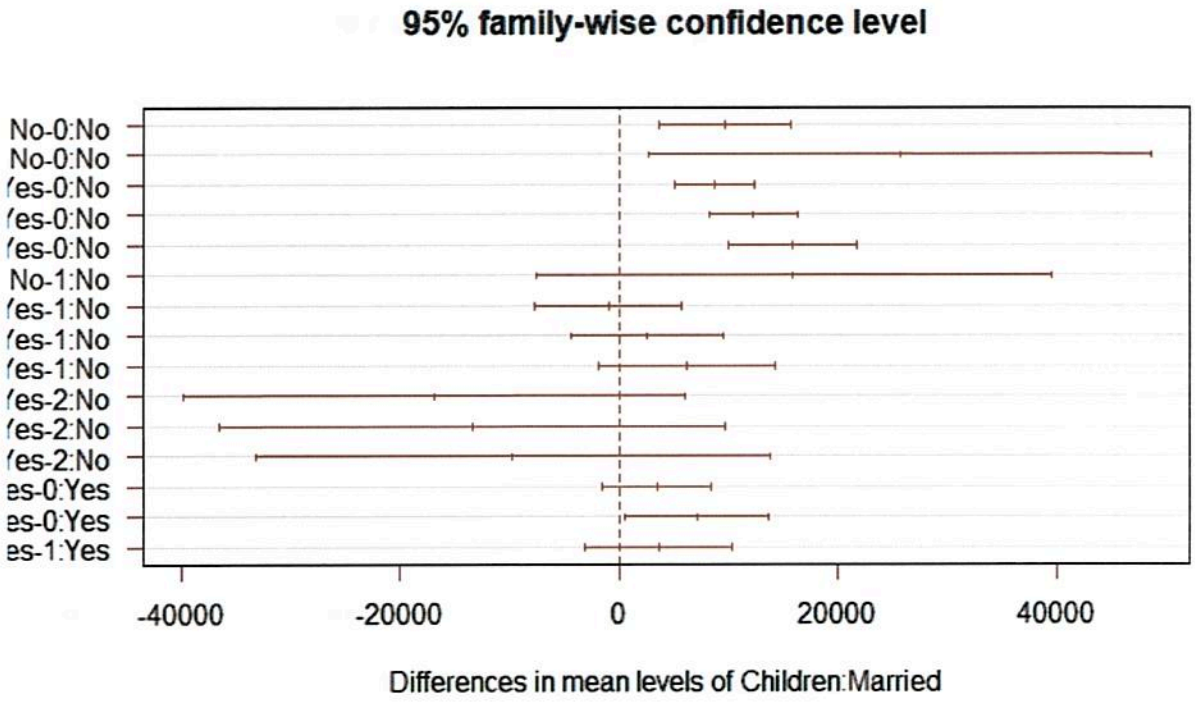
$`Children:Gender:Married`

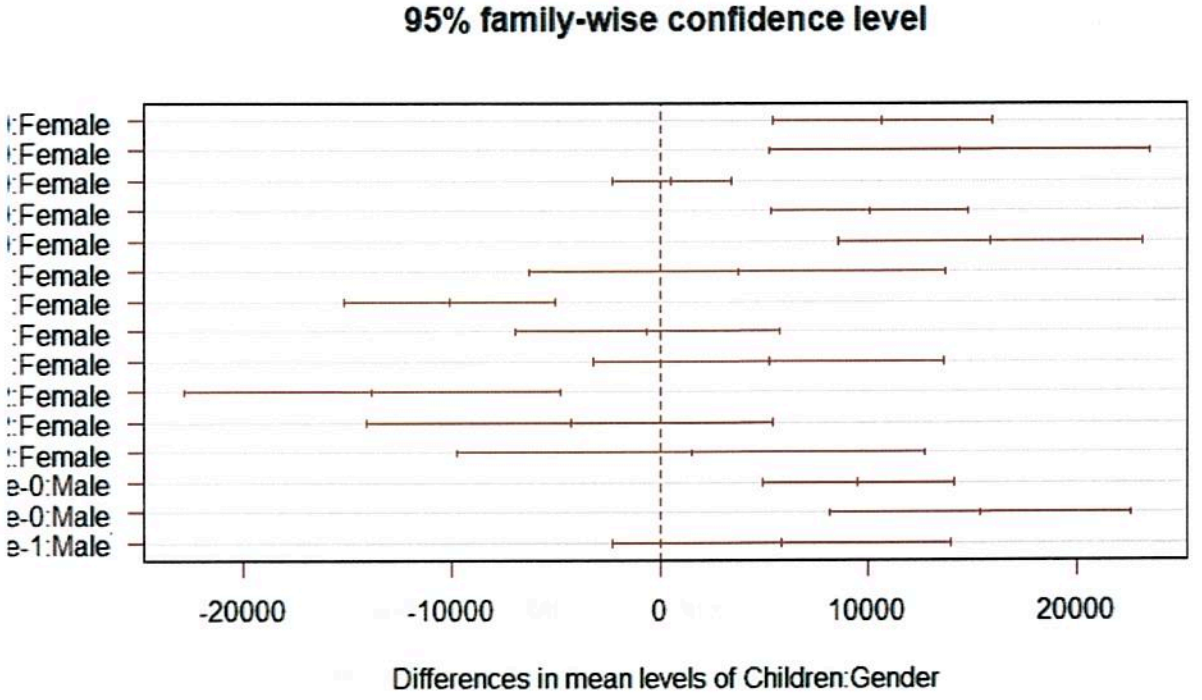
	diff	lwr	upr	p adj
1:Female:No-0:Female:No	14951.9341	3607.3216	26296.547	0.0010261
2:Female:No-0:Female:No	30860.6400	2092.0099	59629.270	0.0231808
0:Male:No-0:Female:No	753.3525	-2895.7426	4402.448	0.9999467
1:Male:No-0:Female:No	7294.9159	-1584.0339	16173.866	0.2327329
2:Male:No-0:Female:No	-1563.3600	-65636.6993	62509.979	1.0000000
0:Female:Yes-0:Female:No	8524.8367	2061.2471	14988.426	0.0010126
1:Female:Yes-0:Female:No	11728.4669	4829.8775	18627.056	0.0000020
2:Female:Yes-0:Female:No	13212.9257	2021.0890	24404.762	0.0064709
0:Male:Yes-0:Female:No	9707.0296	3807.9126	15606.147	0.0000053
1:Male:Yes-0:Female:No	13638.3207	7109.5662	20167.075	0.0000000
2:Male:Yes-0:Female:No	18233.6241	9676.2126	26791.036	0.0000000
2:Female:No-1:Female:No	15908.7059	-14749.8250	46567.237	0.8694776
0:Male:No-1:Female:No	-14198.5816	-25406.9416	-2990.222	0.0020938
1:Male:No-1:Female:No	-7657.0183	-21482.6173	6168.581	0.8115485
2:Male:No-1:Female:No	-16515.2941	-81459.1487	48428.560	0.9995867
0:Female:Yes-1:Female:No	-6427.0974	-18840.3737	5986.179	0.8711648
1:Female:Yes-1:Female:No	-3223.4672	-15868.7024	9421.768	0.9995770
2:Female:Yes-1:Female:No	-1739.0084	-17152.2705	13674.254	0.9999999
0:Male:Yes-1:Female:No	-5244.9045	-17373.8377	6884.029	0.9607909
1:Male:Yes-1:Female:No	-1313.6134	-13760.9455	11133.719	1.0000000
2:Male:Yes-1:Female:No	3281.6900	-10339.6437	16903.024	0.9997546
0:Male:No-2:Female:No	-30107.2875	-58822.4609	-1392.114	0.0302062
1:Male:No-2:Female:No	-23565.7241	-53399.9500	6268.502	0.2891049
2:Male:No-2:Female:No	-32424.0000	-102542.7428	37694.743	0.9372507
0:Female:Yes-2:Female:No	-22335.8033	-51542.3655	6870.759	0.3389406
1:Female:Yes-2:Female:No	-19132.1731	-48438.0739	10173.728	0.5966564
2:Female:Yes-2:Female:No	-17647.7143	-48250.0425	12954.614	0.7675795
0:Male:Yes-2:Female:No	-21153.6104	-50240.4608	7933.240	0.4204931
1:Male:Yes-2:Female:No	-17222.3193	-46443.3720	11998.733	0.7413285
2:Male:Yes-2:Female:No	-12627.0159	-42367.1332	17113.101	0.9656972
1:Male:No-0:Male:No	6541.5634	-2162.6224	15245.749	0.3667799
2:Male:No-0:Male:No	-2316.7125	-66366.0678	61732.643	1.0000000
0:Female:Yes-0:Male:No	7771.4842	1550.1426	13992.826	0.0026568
1:Female:Yes-0:Male:No	10975.1144	4302.9606	17647.268	0.0000054
2:Female:Yes-0:Male:No	12459.5732	1405.8720	23513.274	0.0124108
0:Male:Yes-0:Male:No	8953.6771	3321.0327	14586.321	0.0000141
1:Male:Yes-0:Male:No	12884.9682	6595.9510	19173.985	0.0000000
2:Male:Yes-0:Male:No	17480.2716	9104.3303	25856.213	0.0000000
2:Male:No-1:Male:No	-8858.2759	-73417.0844	55700.533	0.9999992
0:Female:Yes-1:Male:No	1229.9209	-8979.1411	11438.983	0.9999998
1:Female:Yes-1:Male:No	4433.5511	-6056.3246	14923.427	0.9668287
2:Female:Yes-1:Male:No	5918.0099	-7782.5071	19618.527	0.9611010
0:Male:Yes-1:Male:No	2412.1137	-7449.2521	12273.480	0.9997158
1:Male:Yes-1:Male:No	6343.4048	-3907.0386	16593.848	0.6763920

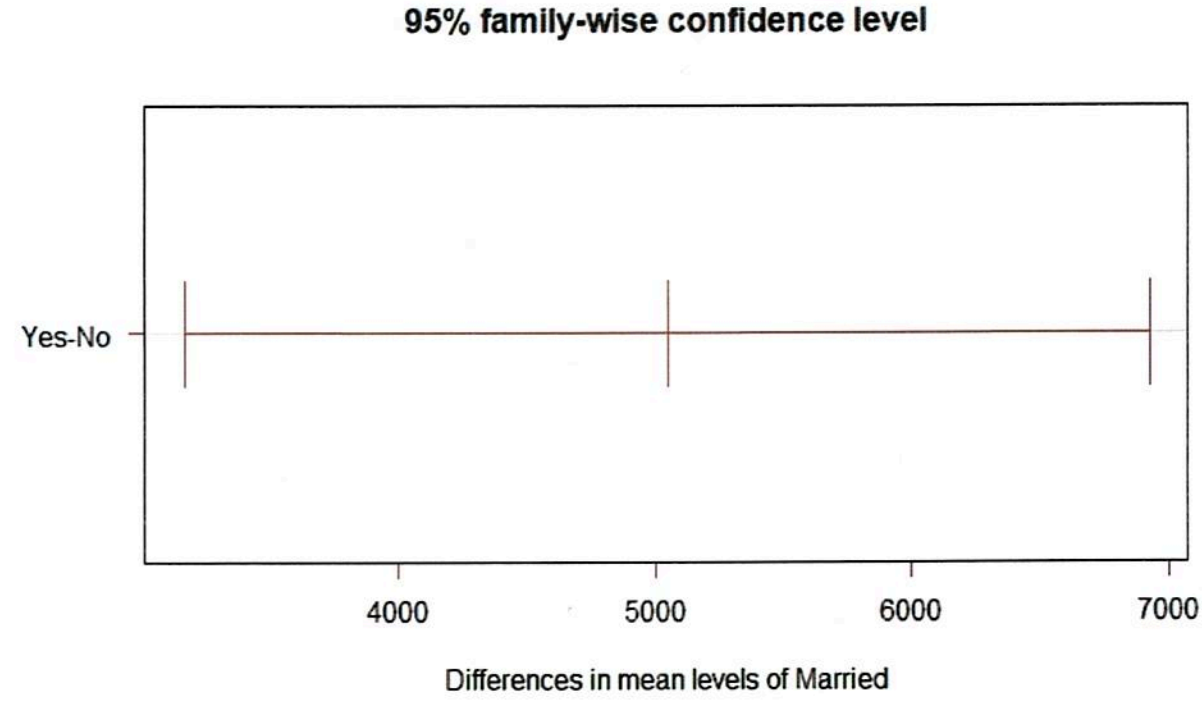
2:Male:Yes-1:Male:No	10938.7083	-709.3065	22586.723	0.0895561
0:Female:Yes-2:Male:No	10088.1967	-54182.9633	74359.357	0.9999968
1:Female:Yes-2:Male:No	13291.8269	-51024.5361	77608.190	0.9999462
2:Female:Yes-2:Male:No	14776.2857	-50141.0558	79693.627	0.9998593
0:Male:Yes-2:Male:No	11270.3896	-52946.4586	75487.238	0.9999898
1:Male:Yes-2:Male:No	15201.6807	-49076.0655	79479.427	0.9997951
2:Male:Yes-2:Male:No	19796.9841	-44718.3885	84312.357	0.9976320
1:Female:Yes-0:Female:Yes	3203.6302	-5339.1874	11746.448	0.9868523
2:Female:Yes-0:Female:Yes	4688.0890	-7585.7207	16961.899	0.9847652
0:Male:Yes-0:Female:Yes	1182.1929	-6575.9498	8940.336	0.9999977
1:Male:Yes-0:Female:Yes	5113.4840	-3133.5660	13360.534	0.6736573
2:Male:Yes-0:Female:Yes	9708.7874	-221.8968	19639.472	0.0623350
2:Female:Yes-1:Female:Yes	1484.4588	-11023.8964	13992.814	0.9999998
0:Male:Yes-1:Female:Yes	-2021.4373	-10145.5571	6102.683	0.9996647
1:Male:Yes-1:Female:Yes	1909.8537	-6682.3740	10502.081	0.9998887
2:Male:Yes-1:Female:Yes	6505.1572	-3713.9932	16724.308	0.6352733
0:Male:Yes-2:Female:Yes	-3505.8961	-15492.0546	8480.262	0.9984662
1:Male:Yes-2:Female:Yes	425.3950	-11882.8564	12733.646	1.0000000
2:Male:Yes-2:Female:Yes	5020.6984	-8473.6599	18515.057	0.9876662
1:Male:Yes-0:Male:Yes	3931.2911	-3881.2260	11743.808	0.8915471
2:Male:Yes-0:Male:Yes	8526.5945	-1046.2880	18099.477	0.1364890
2:Male:Yes-1:Male:Yes	4595.3035	-5377.9174	14568.524	0.9387579

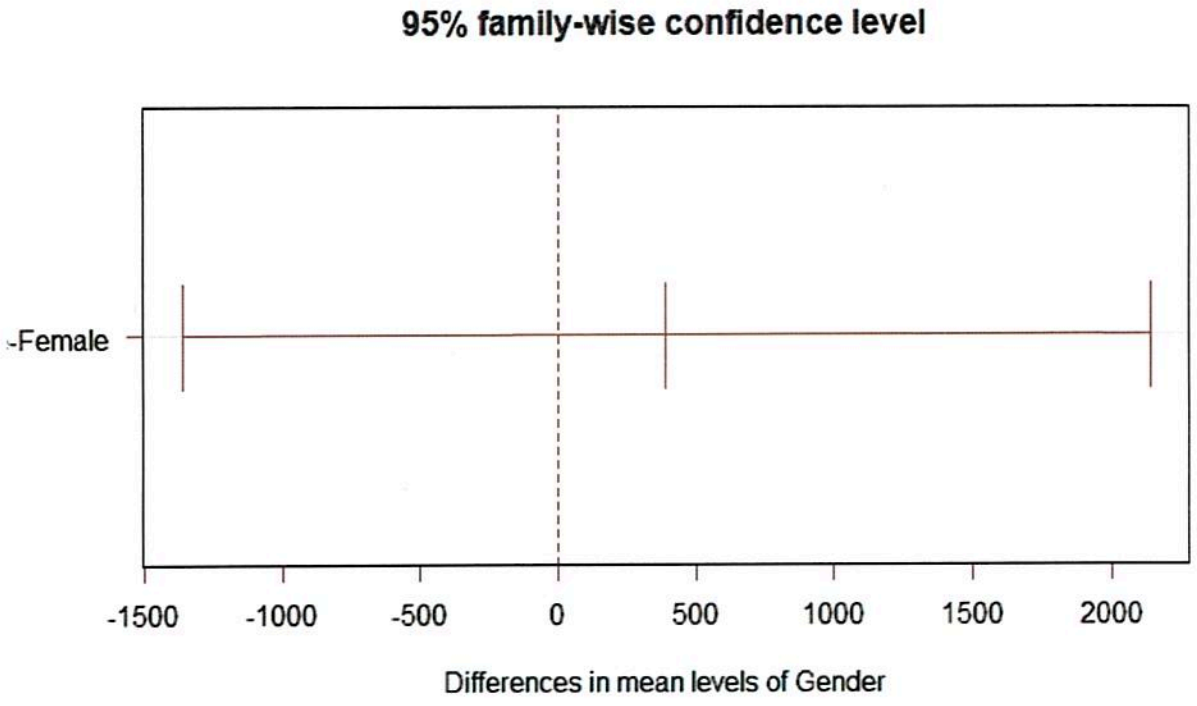


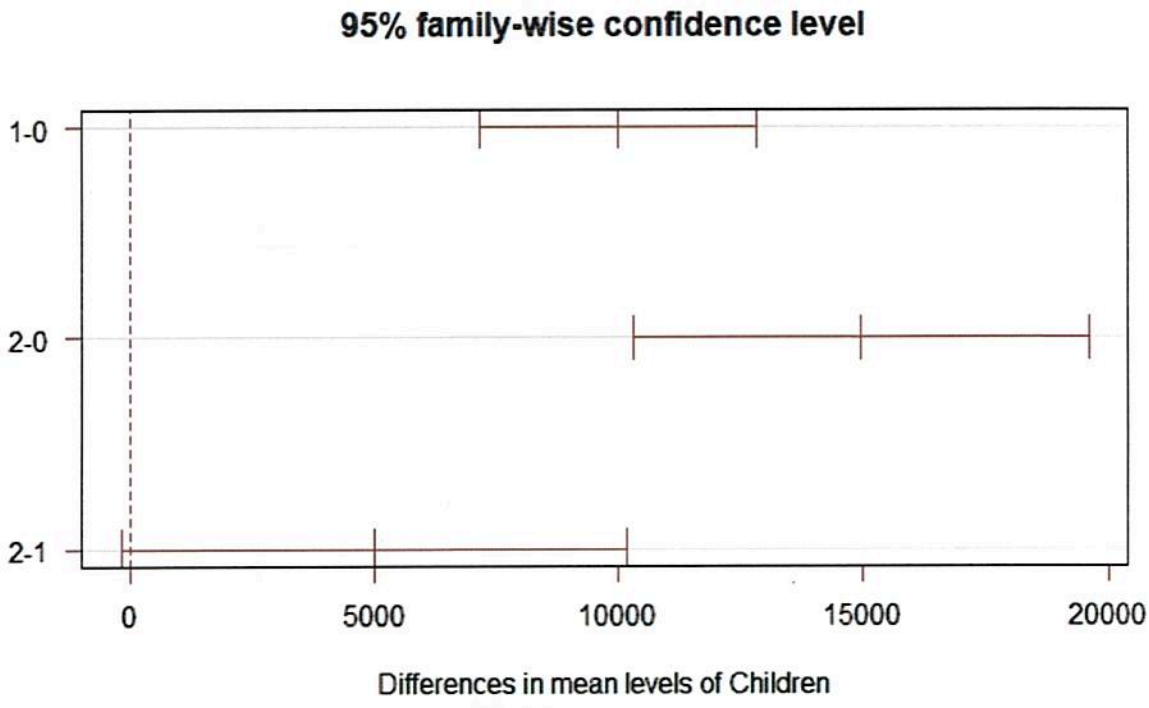

95% family-wise confidence level
male:No
male:No
male:No
Male:No
Male:No
nale:Yes
-2000
0
2000
4000
6000
8000
Differences in mean levels of Gender:Married

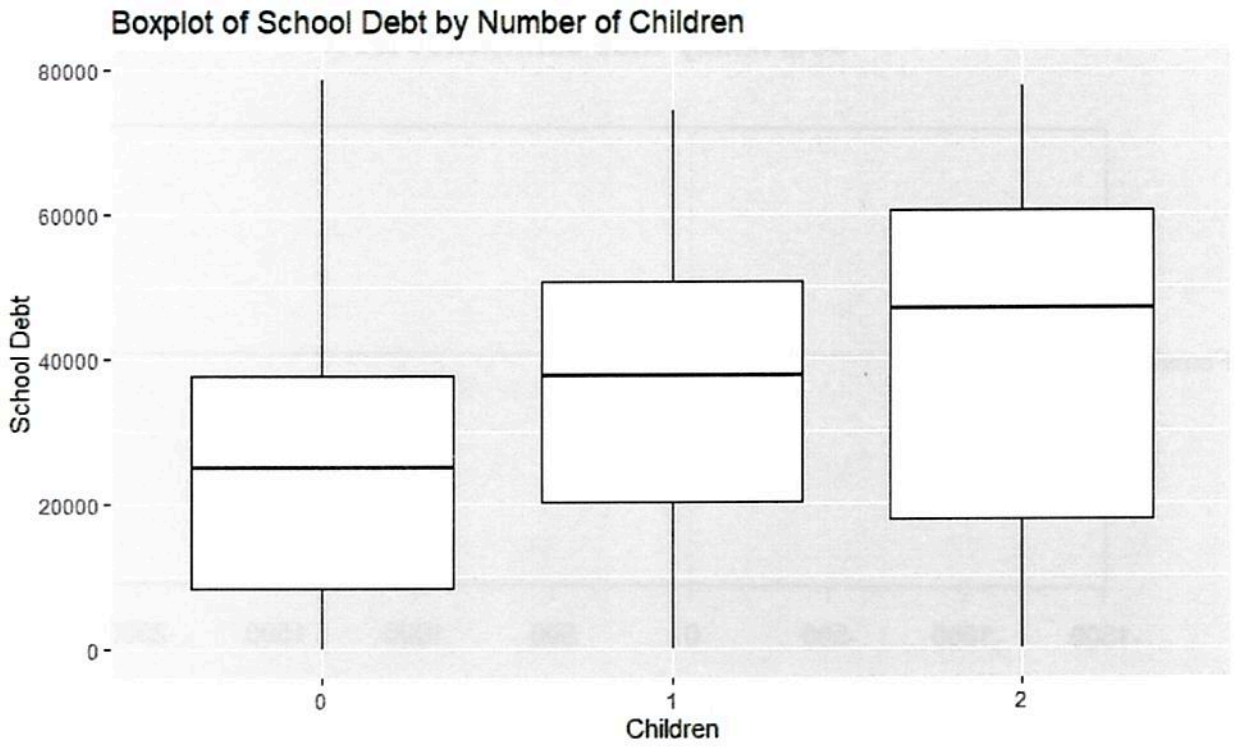

95% family-wise confidence level
No-0:No
No-0:No
/es-0:No
/es-0:No
/es-0:No
No-1:No
/es-1:No
/es-1:No
/es-1:No
/es-2:No
/es-2:No
/es-2:No
es-0:Yes
es-0:Yes
es-1:Yes
-40000
-20000
0
20000
40000
Differences in mean levels of Children:Married

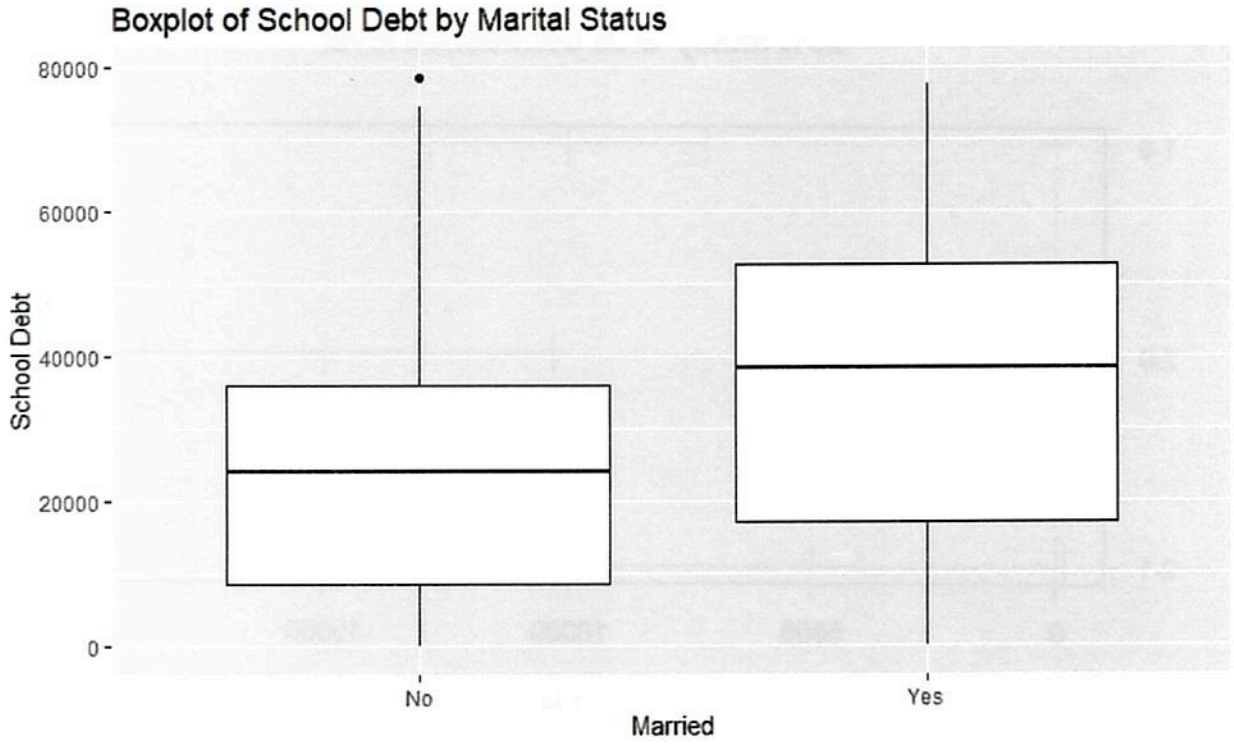

95% family-wise confidence level
:Female
:Female
:Female
:Female
:Female
:Female
:Female
:Female
:Female
:Female
:Female
:Female
e-0:Male
e-0:Male
e-1:Male
-20000
-10000
0
10000
20000
Differences in mean levels of Children:Gender

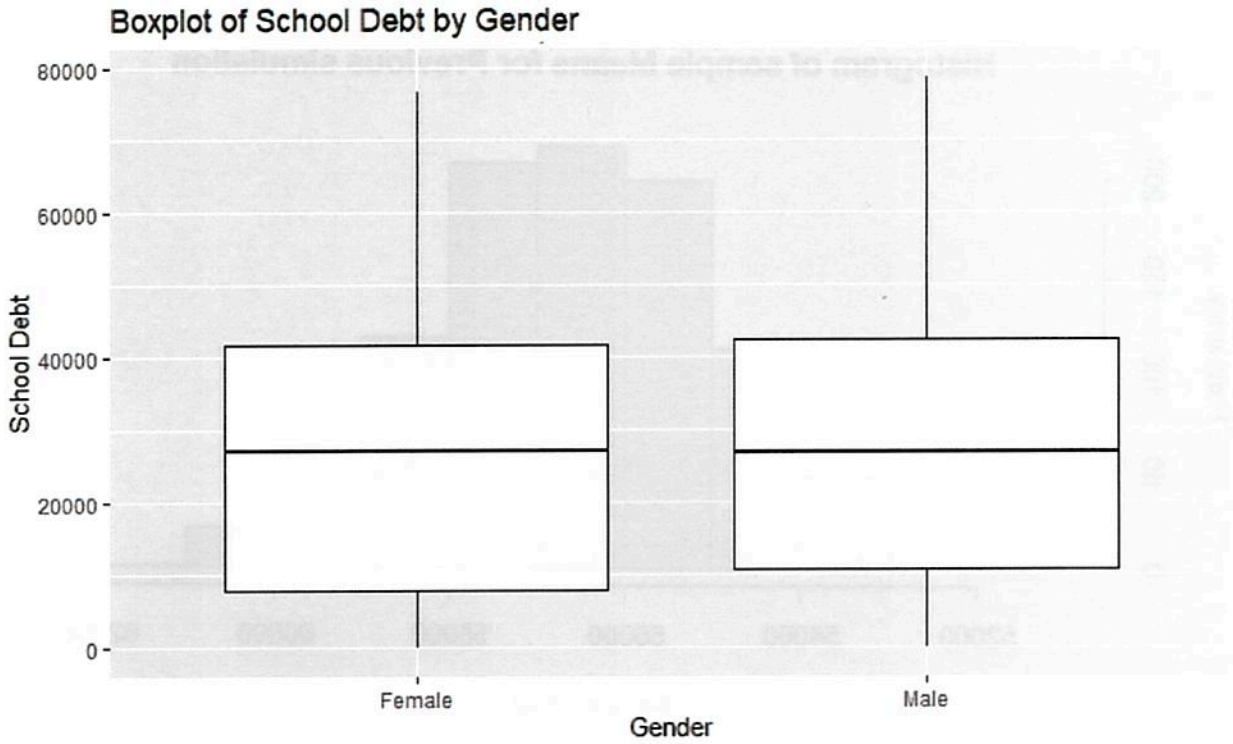

95% family-wise confidence level
Yes-No
4000
5000
6000
7000
Differences in mean levels of Married


95% family-wise confidence level
-Female
-1500
-1000
-500
0
500
1000
1500
2000
Differences in mean levels of Gender


95% family-wise confidence level
1-0
2-0
2-1
0
5000
10000
15000
20000
Differences in mean levels of Children


Boxplot of School Debt by Number of Children
80000
60000
40000
20000
0
School Debt
0
1
2
Children


Boxplot of School Debt by Marital Status
80000
60000
40000
20000
0
School Debt
No
Yes
Married

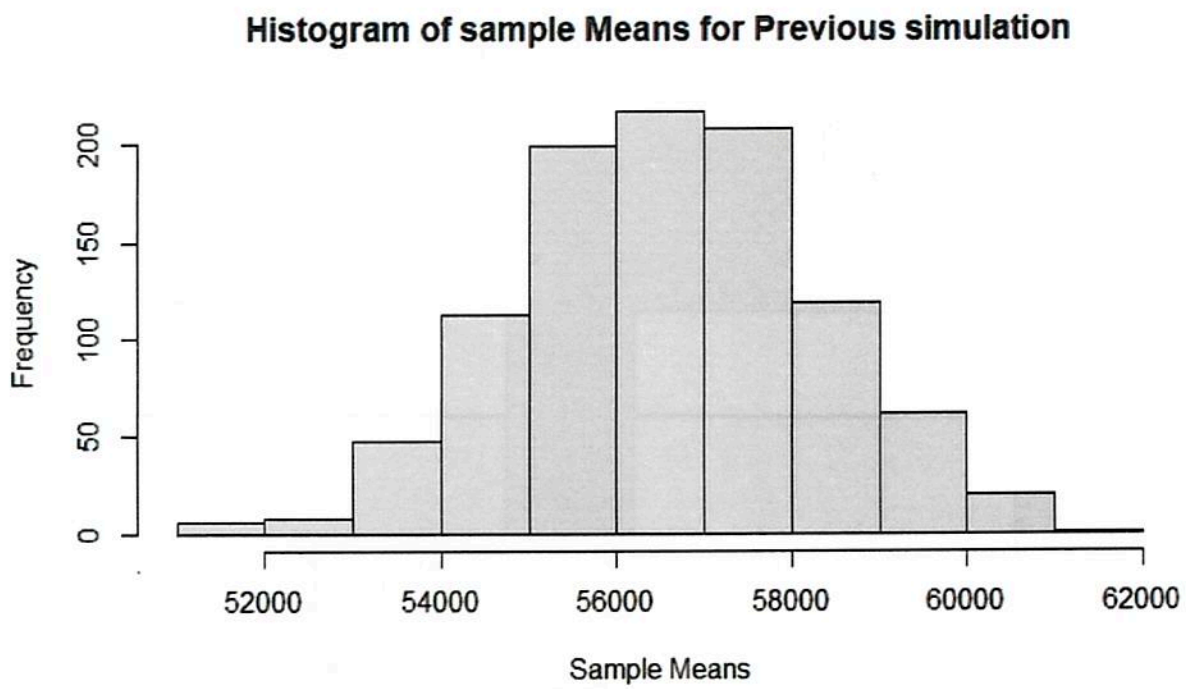


5.

One Sample t-test

data: data3$`School Debt`
t = -5.906, df = 1994, p-value = 2.055e-09
alternative hypothesis: true mean is less than 30000
95 percent confidence interval:
-Inf 28061.03
sample estimates:
mean of x
27312.08

6.



56603.24 (mean of means)
56648.47 (dataset mean)

1709.152 (sd of means)
11953.35 (sd of data)

MTH 324, Exam #2, Fall 2022 Name KEY

Instructions: Answer each question thoroughly. For questions in Part 1, use the work you did at home to answer the questions. Be sure to answer each part of each question. In Part 2, report exact answers unless directed to round.

Part I:
Use the work you did at home to answer these questions about tax paid and the neighborhoods in our dataset.

1. State your null and alternative hypotheses for the husbands and wives question.

$H_0: \delta = 0$
$H_a: \delta \neq 0$

2. What kind of test did you conduct? What is the P-value for your tests?

paired t-test P-value = 0.2067

3. What do you conclude from your test? State your conclusion in plain language in context.

fail to reject H_0.
based on this data, we cannot conclude husbands and wives pay different amounts

4. What is the null and alternative hypotheses for your one-way ANOVA?

$H_0: \mu_i = \mu_j \ \forall i \neq j$
$H_a: \mu_i \neq \mu_j$ for some $i \neq j$

5. What were the results of your test? Using Tukey's method and a box plot, which machines need to be recalibrated the most. Explain your reasoning.

p-value: $2 \times 10^{-16} < 0.05$ reject H_0
based on the box plot, machines 4 & 5 are furthest from the grand mean, and the most different from each other.

6. For the graduate business school data, state the null and alternative hypotheses for your test of proportions. What did you conclude about the differences in marital status between the genders?

$H_0: p_1 = p_2$ p-value: 0.03529 reject H_0

$H_a: p_1 \neq p_2$ there is a difference in marital status between men and women

7. What are the null and alternative hypotheses for your three-way ANOVA? List them all and label them clearly.

$H_0: \mu_i = \mu_j \quad \forall i \neq j$

$H_a: \mu_i \neq \mu_j$ for some $i \neq j$

they are the same for all 7 possible tests:
Children, Gender, Married, Children x Gender, Children x Married, Gender x Married, Children x Gender x Married.

8. Did any of the null hypotheses get rejected for your test? Which ones? State the form of your final model.

Children p-value: 2×10^{-16} reject H_0 α

Gender p-value: 0.66514 fail to reject β

Married p-value: 4.87×10^{-10} reject H_0 γ

Children x Gender p-value: 0.857 fail to reject δ

Children x Married p-value 0.00825 reject H_0 η

Gender x Married p-value 0.22399 fail to reject H_0 θ

Children x Gender x Married p-value 0.07493 fail to reject H_0 λ

$X_{ijk} = \mu + \alpha_i + \gamma_j + \eta_k + \varepsilon_{ijk}$

9. Describe the normality of the data? Is it normal? Are there significant deviations from normal?

the 0's are potentially problematic → yes.

10. Describe the Tukey intervals for the three main effects. (There are 5 intervals altogether. Explain what each one means.)

Married - one interval. min diff ~ 3000 no 0 in interval so diff is meaningful

Gender - one interval - spans 0 - no diff.

Children - 3 intervals diff in 2-1 spans 0

but diff in 2-0 and 1-0 have mins > 5000 There is a diff.

11. State the null and alternative hypotheses for your one-sample test of school debt levels. What is your P-value? What did you conclude?

$H_0: \mu_0 = 30{,}000$

$H_a: \mu_0 < 30{,}000$

p-value $2.055 \times 10^{-9} < 0.05$

reject H_0

there is good reason to think school debt is less than 30K.

12. For your sampling distribution, describe the shape of the distribution.

Symmetric, looks pretty normal

13. What is the mean of your means? What is the mean of the data? Are they similar?

$56,603.24 mean of means

$56,648.47 mean of data set They are similar

14. What is the standard deviation of your means? What is the standard deviation of the data? Given that you took samples of size 50, what is the predicted standard error from the central limit theorem? Is it similar to what you simulated?

$1709.15 sd. of means

$\frac{11953.35}{\sqrt{50}}$ - sd of data = $1690.46

it is close to what was predicted

Part II:

15. Describe the main results of the Central Limit Theorem.

Sample statistics (means) have a normal distribution (roughly) that gets more normal and more narrow as the sample size increases.

16. Consider a binomial experiment that conducted 15 trials and resulted in 11 successes and 4 failures. Find the maximum likelihood function for this outcome and then use it to find the maximum likelihood estimate of p (the proportion of successes).

$$L = p^{11}(1-p)^4$$

$$\frac{dL}{dp} = 11p^{10}(1-p)^4 + p^{11}(4)(1-p)^3(-1) = 0$$

$$p^{10}(1-p)^3[11(1-p) - 4p] = 0$$

$$11 - 11p - 4p = 0$$

$$11 - 15p = 0$$

$$11 = 15p$$

$$p = 11/15$$

17. At the beginning of the semester a representative sample of 501 students were surveyed and asked if they owned a dog. The sample proportion was 0.559. Use this information to construct a 95% confidence interval for the proportion of all STAT 200 students who own a dog.

$501 \times (.559) = 280$

$ME = 0.04336$

$$= 1.96 \times \sqrt{\frac{0.559(1-0.559)}{501}}$$

$(0.5154, 0.60236)$

18. A sample of 80 students is selected and it is determined that their mean math SAT score is 510. If the true mean math score is actually 500 (with a standard deviation of 100), what is the power of the one-sample test to detect this 10-point difference?

$$\frac{510 - 500}{100/\sqrt{80}} = t$$

Power is $\approx$ 14%

19. If you want to determine the appropriate sample size needed to conduct a poll with just at 2% margin of error for a proportion, with a 95% level of confidence, use the formula $n = p(1-p)\left(\frac{z^*}{E}\right)^2$. Use this formula with $p = 0.5$ to estimate the sample size needed.

$$n \approx 0.5^2\left(\frac{1.96}{0.02}\right)^2 = 2401$$

20. Describe what a Latin Square design is. Give an example of a Latin Square design for three ~~levels~~ variables of data, each with 4 levels each.

Example

A, B, C, D Var1

Var3 ↓ \ Var2				
	A	B	C	D
	B	C	D	A
	C	D	A	B
	D	A	B	C

Latin square designs help experimenters measure the effects (main) of many variables w/o having to collect one observation for every possible combination of variable levels.