

ANOVA

A comparison of 2022 Wisconsin home sales using select data - *Lake Country, Waukesha County*

BUS674—Data Analytics in Practice



CARROLL
UNIVERSITY

Melanie Fitzgerald

Executive Summary

This report examines the 2022 home selling prices in the Lake Country region of Waukesha County and the effect of the number of garage spaces and the Lake Country cities – Pewaukee, Delafield, Merton, Hartland, Oconomowoc, Nashotah, and Summit – have on the prices. A one-way ANOVA was done to determine the relationship garage space numbers have on home selling prices. A factorial ANOVA was done to determine if garage space numbers and the Lake Country cities independently have an effect on home selling prices and if their interaction also has an effect. The analysis can offer insight to the home market in Lake Country for a buyer, seller, and real estate professional.

Overview/Description of the Dataset

The dataset used in this report is from the Milwaukee MLS and contains all home sales in the Milwaukee 7 region since 1996. A new dataset was extracted to include only the cities of Lake Country for the year 2022.

Methods of Analysis Used

The data was cleaned for analysis which included recoding the outliers in variable *GarageSpaces*, recoding numerical values from scale to ordinal in *GarageSpaces*, and adding a new variable *LCCity* which assigns a number to each Lake Country city. SPSS is used for the analysis which includes descriptive statistics and ANOVA. The variables used are *SoldPrice* as the dependent variable, *GarageSpaces* and *LCCity* as independent variables.

Null hypothesis: There is no significant difference in the mean selling price between groups.

Alternative hypothesis: There is a difference in the mean price between groups.

Vernacular hypothesis: The null hypothesis states that the mean selling price based on a group has no significant difference. The alternative hypothesis states that the mean selling price for at least one of the groups is different.

Results from (type of data analysis used) Data Analysis

One-way ANOVA - Number of garage spaces (GarageSpace)

ANOVA is used to do a test comparison of mean home selling prices (*SoldPrice*) and the number of garage spaces (*GarageSpaces*) of at least two to a maximum of five spaces for the year 2022. A

significance level of 0.05 is assumed. Descriptive statistics was done on *SoldPrice* relative to each number of garage spaces (*GarageSpaces*) with total entries N=822. It is observed that the mean selling price rises as the number of garage spaces increases.

Descriptive Statistics			
Dependent Variable: SoldPrice			
GarageSpaces	Mean	Std. Deviation	N
2.0	472128.22	297042.089	244
2.5	516509.13	327748.018	231
2.8	410950.00	108419.400	12
3.0	765265.75	543535.797	187
3.5	776421.72	355326.126	100
4.0	904021.00	763515.190	30
4.5	1332909.09	1152044.657	11
5.0	1233557.14	636502.715	7
Total	621178.46	457321.124	822

ANOVA is used to do a test comparison of mean home selling prices of *GarageSpaces* for all Lake Country cities. A significance level of 0.05 is assumed. The ANOVA shows an $R^2=0.148$ which means that 14.8% of the variance in the dependent variable *SoldPrice* can be explained by the independent variable *GarageSpaces*. The observed power is 1.000 which means that there is a 100% chance of the ANOVA detecting a significant effect between the two variables. The ANOVA shows an F-statistic between *SoldPrice* and *GarageSpaces* (GS) of 20.162 and a significance level less than 0.05 ($p < 0.001$) indicating that the differences between the groups are statistically significant and the null hypothesis can be rejected for the alternative hypothesis.

Tests of Between-Subjects Effects								
Dependent Variable: SoldPrice								
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	2.537E+13 ^a	7	3.624E+12	20.162	<.001	.148	141.131	1.000
Intercept	1.099E+14	1	1.099E+14	611.110	<.001	.429	611.110	1.000
GarageSpaces	2.537E+13	7	3.624E+12	20.162	<.001	.148	141.131	1.000
Error	1.463E+14	814	1.798E+11					
Total	4.889E+14	822						
Corrected Total	1.717E+14	821						

a. R Squared = .148 (Adjusted R Squared = .140)

b. Computed using alpha = .05

A comparison of the groups using a Tukey Post-Hoc analysis shows the relationship of means between the groups. 2GS shows a statistically significant relationship ($p < 0.001$) between the other *GarageSpaces* except 2.5GS ($p = 0.948$) and 2.8GS ($p = 1.000$). 2.8GS does not have a statistically significant relationship with 3GS ($p = 0.095$), 3.5GS ($p = 0.091$) or 4GS ($p = 0.160$). 3GS is significant with 4.5GS ($p < 0.001$), but not with 3.5GS ($p = 1.000$), 4GS ($p = 0.7110$) or 5GS ($p = 0.080$). 3.5GS is not significant with 4GS ($p = 0.836$) or 5GS ($p = 0.107$) but is significant with 4.5GS ($p = 0.001$). 4GS does not have a statistically significant relationship with either 4.5GS ($p = 0.080$) and 5GS ($p = 0.585$). 4.5GS is not statistically significant to 5GS ($p = 1.000$).

Multiple Comparisons						
Dependent Variable: SoldPrice						
Tukey HSD						
(I) GarageSpaces	(J) GarageSpaces	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2.0	2.5	-44380.91	38923.091	.948	-162660.96	73899.14
	2.8	61178.22	125370.648	1.000	-319799.91	442156.36
	3.0	-293137.53*	41208.266	<.001	-418361.80	-167913.25
	3.5	-304293.50*	50343.770	<.001	-457278.87	-151308.12
	4.0	-431892.78*	82031.597	<.001	-681171.58	-182613.98
	4.5	-860780.87*	130689.368	<.001	-1257921.61	-463640.13
	5.0	-761428.92*	162537.790	<.001	-1255351.10	-267506.74
2.5	2.0	44380.91	38923.091	.948	-73899.14	162660.96
	2.8	105559.13	125535.902	.991	-275921.18	487039.44
	3.0	-248756.62*	41708.327	<.001	-375500.49	-122012.75
	3.5	-259912.59*	50753.901	<.001	-414144.28	-105680.90
	4.0	-387511.87*	82283.937	<.001	-637557.48	-137466.26
	4.5	-816399.96*	130847.905	<.001	-1214022.46	-418777.46
	5.0	-717048.01*	162665.290	<.001	-1211357.64	-222738.38
2.8	2.0	-61178.22	125370.648	1.000	-442156.36	319799.91
	2.5	-105559.13	125535.902	.991	-487039.44	275921.18
	3.0	-354315.75	126263.126	.095	-738005.96	29374.46
	3.5	-365471.72	129532.818	.091	-759097.91	28154.47
	4.0	-493071.00*	144822.093	.016	-933158.46	-52983.54
	4.5	-921959.09*	176985.781	<.001	-1459786.04	-384132.15
	5.0	-822607.14*	201650.165	.001	-1435384.57	-209829.71
3.0	2.0	293137.53*	41208.266	<.001	167913.25	418361.80
	2.5	248756.62*	41708.327	<.001	122012.75	375500.49
	2.8	354315.75	126263.126	.095	-29374.46	738005.96
	3.5	-11155.97	52526.876	1.000	-170775.40	148463.46
	4.0	-138755.25	83389.210	.711	-392159.59	114649.08
	4.5	-567643.34*	131545.766	<.001	-967386.51	-167900.17
	5.0	-468291.39	163227.176	.080	-964308.49	27725.70
3.5	2.0	304293.50*	50343.770	<.001	151308.12	457278.87
	2.5	259912.59*	50753.901	<.001	105680.90	414144.28
	2.8	365471.72	129532.818	.091	-28154.47	759097.91
	3.0	11155.97	52526.876	1.000	-148463.46	170775.40
	4.0	-127599.28	88261.738	.836	-395810.32	140611.76
	4.5	-556487.37*	134687.276	.001	-965777.01	-147197.73
	5.0	-457135.42	165769.372	.107	-960877.78	46606.94
4.0	2.0	431892.78*	82031.597	<.001	182613.98	681171.58
	2.5	387511.87*	82283.937	<.001	137466.26	637557.48
	2.8	493071.00*	144822.093	.016	52983.54	933158.46
	3.0	138755.25	83389.210	.711	-114649.08	392159.59
	3.5	127599.28	88261.738	.836	-140611.76	395810.32
	4.5	-428888.09	149450.159	.080	-883039.39	25263.21
	5.0	-329536.14	177972.392	.585	-870361.22	211288.93
4.5	2.0	860780.87*	130689.368	<.001	463640.13	1257921.61
	2.5	816399.96*	130847.905	<.001	418777.46	1214022.46
	2.8	921959.09*	176985.781	<.001	384132.15	1459786.04
	3.0	567643.34*	131545.766	<.001	167900.17	967386.51
	3.5	556487.37*	134687.276	.001	147197.73	965777.01
	4.0	428888.09	149450.159	.080	-25263.21	883039.39
	5.0	99351.95	204999.270	1.000	-523602.79	722306.69
5.0	2.0	761428.92*	162537.790	<.001	267506.74	1255351.10
	2.5	717048.01*	162665.290	<.001	222738.38	1211357.64
	2.8	822607.14*	201650.165	.001	209829.71	1435384.57
	3.0	468291.39	163227.176	.080	-27725.70	964308.49
	3.5	457135.42	165769.372	.107	-46606.94	960877.78
	4.0	329536.14	177972.392	.585	-211288.93	870361.22
	4.5	-99351.95	204999.270	1.000	-722306.69	523602.79
Based on observed means.						
The error term is Mean Square(Error) = 179772330338.017.						
*. The mean difference is significant at the .05 level.						

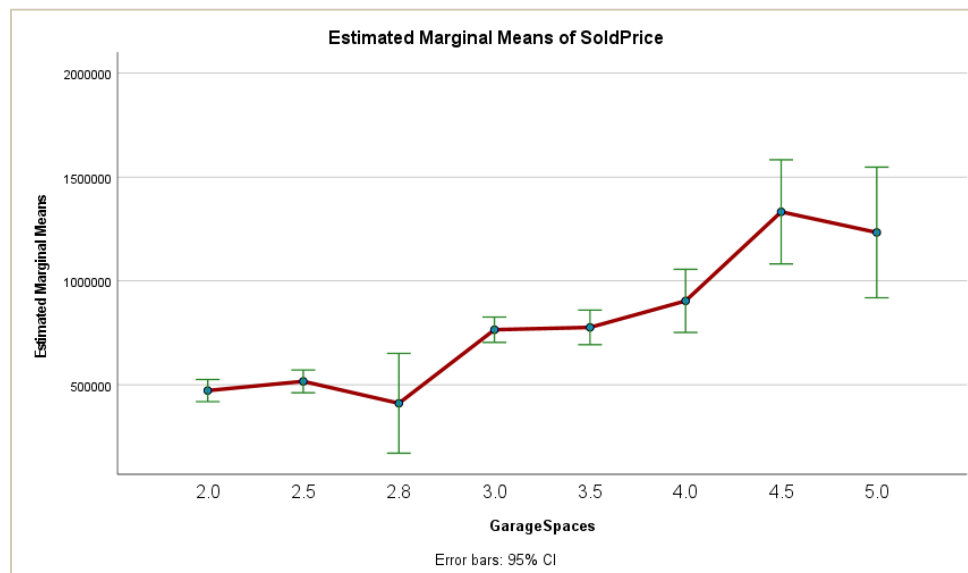
The table below shows *GarageSpaces* grouped into four subsets where *GarageSpaces* 2.9, 2, 2.5, 3, and 3.5 are in Subset 1; *GarageSpaces* 2.5, 3, 3.5, and 4 are in Subset 2; *GarageSpaces* 4 and 5 are in Subset 3; and *GarageSpaces* 5 and 4.5 are in Subset 4. Within each subset, there is no significant difference between the means. It can be concluded that 2.8GS has a lower mean than Subsets 2,3, and 4. However, it cannot be concluded that 2.8GS and the other number of garage spaces in Subset 1 have significantly different means.

SoldPrice					
Tukey HSD ^{a,b,c}					
GarageSpaces	N	Subset			
		1	2	3	4
2.8	12	410950.00			
2.0	244	472128.22			
2.5	231	516509.13	516509.13		
3.0	187	765265.75	765265.75		
3.5	100	776421.72	776421.72		
4.0	30		904021.00	904021.00	
5.0	7			1233557.14	1233557.14
4.5	11				1332909.09
Sig.		.092	.058	.180	.995

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = 179772330338.017.

a. Uses Harmonic Mean Sample Size = 21.378.
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.
c. Alpha = .05.

The graph below shows how the dependent variable *SoldPrice* changes across the different garage spaces with the error bars showing the range which the actual mean is likely to occur for the given confidence level of 95%.



Factorial ANOVA - Number of garage spaces (GarageSpaces) and Lake Country cities (City)

In the factorial ANOVA, *GarageSpaces* and *LCCity* are used as factors with *SoldPrice* as the dependent variable. A significance level of 0.05 is assumed. Descriptive statistics show the mean *SoldPrice* for the different interactions between *GarageSpaces* and *LCCity*. A home with

Descriptive Statistics				
Dependent Variable: SoldPrice				
GarageSpaces	Lake Country City	Mean	Std. Deviation	N
2.0	Pewaukee	406977.02	263123.194	86
	Delafield	545611.11	244785.894	18
	Hartland	433771.43	117408.792	35
	Merton	511785.71	323837.609	14
	Nashotah	292975.00	60273.287	4
	Oconomowoc	533278.80	360798.812	82
	Summit	626100.00	515001.262	5
	Total	472128.22	297042.089	244
2.5	Pewaukee	481202.77	303236.154	75
	Delafield	505240.76	345088.021	25
	Hartland	407804.23	107471.327	26
	Merton	530522.73	311684.841	22
	Nashotah	555500.00	205768.073	2
	Oconomowoc	567240.97	381677.571	75
	Summit	777316.67	419533.076	6
	Total	516509.13	327748.018	231
2.8	Pewaukee	495000.00	160234.204	3
	Delafield	360333.33	21571.586	3
	Merton	324900.00	.	1
	Nashotah	325500.00	.	1
	Oconomowoc	480000.00	18027.756	3
	Summit	275000.00	.	1
	Total	410950.00	108419.400	12
3.0	Pewaukee	716382.42	247542.565	31
	Delafield	1104171.43	938717.415	28
	Hartland	849363.15	275210.444	13
	Merton	839577.78	809135.230	27
	Nashotah	725333.33	101352.520	3
	Oconomowoc	620789.89	270294.533	73
	Summit	731338.08	306651.449	12
	Total	765265.75	543535.797	187
3.5	Pewaukee	743189.59	274299.570	22
	Delafield	816504.28	199119.599	18
	Hartland	808223.53	182402.520	17
	Merton	711637.82	368760.957	17
	Nashotah	774616.67	70692.580	3
	Oconomowoc	690396.37	236688.454	19
	Summit	1328975.00	1292960.145	4
	Total	776421.72	355326.126	100
4.0	Pewaukee	973600.00	1082751.967	10
	Delafield	1363000.00	859473.676	5
	Hartland	1379000.00	.	1
	Merton	362865.00	166686.282	2
	Oconomowoc	582625.00	141470.590	8
	Summit	950975.00	401527.171	4
	Total	904021.00	763515.190	30
4.5	Pewaukee	730500.00	480125.504	2
	Delafield	1050250.00	636199.951	4
	Hartland	1080000.00	.	1
	Merton	2900000.00	2227386.361	2
	Oconomowoc	1060000.00	834386.002	2
	Total	1332909.09	1152044.657	11
5.0	Pewaukee	1124666.67	690724.499	3
	Delafield	600900.00	.	1
	Merton	1277500.00	526794.552	2
	Summit	2105000.00	.	1
	Total	1233557.14	636502.715	7
Total	Pewaukee	541828.70	386639.426	232
	Delafield	791801.92	636304.863	102
	Hartland	570165.92	264289.318	93
	Merton	724236.47	683595.123	87
	Nashotah	546788.46	231289.141	13
	Oconomowoc	583695.14	335975.078	262
	Summit	850613.85	606981.571	33
	Total	621178.46	457321.124	822

The factorial ANOVA shows an $R^2=0.258$ which means that only 25.85% of the variance in the dependent variable *SoldPrice* can be explained by the independent variables *GarageSpaces* and *LCCity*. The observed power for *LCCity* is 0.822 which means that there is an 82.2% chance of the ANOVA detecting a significant difference. The observed power of the interaction between *GarageSpaces* and *LCCity* is 1.000. Both *GarageSpaces* ($F = 13.004$, $p < 0.001$) and *LCCity* ($F = 2.403$, $p = 0.026$) show to have a statistically significant effect on *SoldPrice* suggesting that the difference in means between the two groups are significant. The interaction between *GarageSpaces* and *LCCity* ($F = 2.711$, $p < 0.001$) also has a statistically significant with a p-value less than the significance level of 0.05 suggesting that their interactions do have an effect on *SoldPrice* and the null hypothesis can be rejected in favor of the alternative hypothesis.

Tests of Between-Subjects Effects								
Dependent Variable: SoldPrice								
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	4.424E+13 ^a	48	9.217E+11	5.590	<.001	.258	268.318	1.000
Intercept	7.652E+13	1	7.652E+13	464.054	<.001	.375	464.054	1.000
GarageSpaces	1.501E+13	7	2.144E+12	13.004	<.001	.105	91.028	1.000
LCCity	2.377E+12	6	3.962E+11	2.403	.026	.018	14.417	.822
GarageSpaces * LCCity	1.565E+13	35	4.471E+11	2.711	<.001	.109	94.897	1.000
Error	1.275E+14	773	1.649E+11					
Total	4.889E+14	822						
Corrected Total	1.717E+14	821						

a. R Squared = .258 (Adjusted R Squared = .212)
b. Computed using alpha = .05

A comparison of the groups using a Tukey Post-Hoc analysis shows the relationship of means between the group *LCCity*. Pewaukee shows a statistically significant relationship with Delafield ($p < 0.001$), Merton ($p = 0.007$) and Summit ($p < 0.001$) but not with Hartland ($p = 0.998$), Nashotah ($p = 1.000$), and Oconomowoc ($p = 0.914$). Delafield is significant with Hartland ($p = 0.003$) and Oconomowoc ($p < 0.001$) but not with Merton ($p = 0.916$), Nashotah ($p = 0.385$), and Summit ($p = 0.991$). Hartland has a statistically significant relationship with Summit ($p = 0.012$) but not with Merton ($p = 0.145$), Nashotah ($p = 1.000$), and Oconomowoc ($p = 1.000$). Merton is not significant with Nashotah ($p = 0.763$), Oconomowoc ($p = 0.077$), and Summit ($p = 0.731$). Nashotah is not significant with Oconomowoc ($p = 1.000$) and Summit ($p = 0.253$). Oconomowoc has a statistically significant relationship with Summit ($p = 0.007$).

Multiple Comparisons

Dependent Variable: SoldPrice

Tukey HSD

(I) Lake Country City	(J) Lake Country City	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Pewaukee	Delafield	-249973.23 [*]	48242.561	<.001	-392584.10	-107362.35
	Hartland	-28337.23	49837.633	.998	-175663.34	118988.88
	Merton	-182407.78 [*]	51049.668	.007	-333316.80	-31498.75
	Nashotah	-4959.77	115736.027	1.000	-347089.54	337170.01
	Oconomowoc	-41866.45	36607.497	.914	-150082.66	66349.77
	Summit	-308785.15 [*]	75548.034	<.001	-532114.33	-85455.97
Delafield	Pewaukee	249973.23 [*]	48242.561	<.001	107362.35	392584.10
	Hartland	221636.00 [*]	58220.635	.003	49528.72	393743.28
	Merton	67565.45	59261.465	.916	-107618.65	242749.55
	Nashotah	245013.46	119585.453	.385	-108495.69	598522.61
	Oconomowoc	208106.78 [*]	47391.576	<.001	68011.51	348202.04
	Summit	-58811.93	81322.557	.991	-299211.30	181587.44
Hartland	Pewaukee	28337.23	49837.633	.998	-118988.88	175663.34
	Delafield	-221636.00 [*]	58220.635	.003	-393743.28	-49528.72
	Merton	-154070.55	60567.040	.145	-333114.08	24972.99
	Nashotah	23377.46	120237.788	1.000	-332060.06	378814.99
	Oconomowoc	-13529.22	49014.350	1.000	-158421.60	131363.16
	Summit	-280447.92 [*]	82278.814	.012	-523674.11	-37221.74
Merton	Pewaukee	182407.78 [*]	51049.668	.007	31498.75	333316.80
	Delafield	-67565.45	59261.465	.916	-242749.55	107618.65
	Hartland	154070.55	60567.040	.145	-24972.99	333114.08
	Nashotah	177448.01	120745.205	.763	-179489.50	534385.52
	Oconomowoc	140541.33	50246.248	.077	-7992.69	289075.35
	Summit	-126377.38	83018.565	.731	-371790.35	119035.60
Nashotah	Pewaukee	4959.77	115736.027	1.000	-337170.01	347089.54
	Delafield	-245013.46	119585.453	.385	-598522.61	108495.69
	Hartland	-23377.46	120237.788	1.000	-378814.99	332060.06
	Merton	-177448.01	120745.205	.763	-534385.52	179489.50
	Oconomowoc	-36906.68	115383.902	1.000	-377995.53	304182.17
	Summit	-303825.39	132969.325	.253	-696898.89	89248.11
Oconomowoc	Pewaukee	41866.45	36607.497	.914	-66349.77	150082.66
	Delafield	-208106.78 [*]	47391.576	<.001	-348202.04	-68011.51
	Hartland	13529.22	49014.350	1.000	-131363.16	158421.60
	Merton	-140541.33	50246.248	.077	-289075.35	7992.69
	Nashotah	36906.68	115383.902	1.000	-304182.17	377995.53
	Summit	-266918.71 [*]	75007.482	.007	-488649.95	-45187.46
Summit	Pewaukee	308785.15 [*]	75548.034	<.001	85455.97	532114.33
	Delafield	58811.93	81322.557	.991	-181587.44	299211.30
	Hartland	280447.92 [*]	82278.814	.012	37221.74	523674.11
	Merton	126377.38	83018.565	.731	-119035.60	371790.35
	Nashotah	303825.39	132969.325	.253	-89248.11	696898.89
	Oconomowoc	266918.71 [*]	75007.482	.007	45187.46	488649.95

Based on observed means.

The error term is Mean Square(Error) = 164893064898.645.

*. The mean difference is significant at the .05 level.

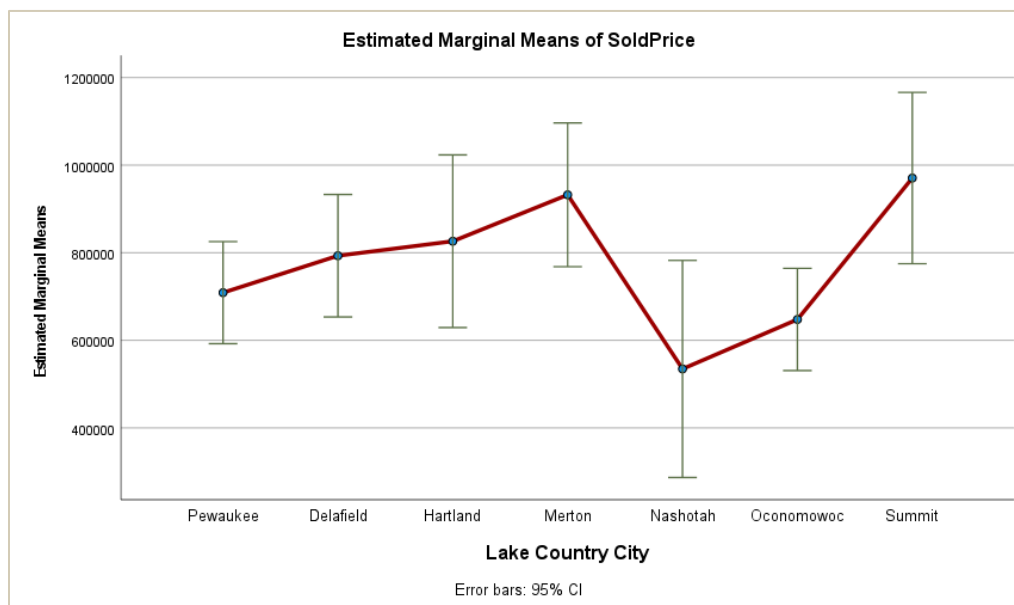
The table below shows *LCCity* grouped into three subsets with Pewaukee, Nashotah, Hartland, Oconomowoc, and Merton in Subset 1; Nashotah, Hartland, Oconomowoc, Merton, and Delafield in Subset 2; and Merton, Delafield, and Summit in Subset 3. Within each subset, there is no significant difference between the means. It can be concluded that Pewaukee has a lower mean than Subset 2 and Subset 3. However, it cannot be concluded that Pewaukee and the other cities in Subset 1 have significantly different means. (Table for *GarageSpaces* is located earlier in this report.)

SoldPrice				
Tukey HSD ^{a,b,c}				
Lake Country City	N	Subset		
		1	2	3
Pewaukee	232	541828.70		
Nashotah	13	546788.46	546788.46	
Hartland	93	570165.92	570165.92	
Oconomowoc	262	583695.14	583695.14	
Merton	87	724236.47	724236.47	724236.47
Delafield	102		791801.92	791801.92
Summit	33			850613.85
Sig.		.303	.052	.735

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = 164893064898.645.

a. Uses Harmonic Mean Sample Size = 47.488.
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.
c. Alpha = .05.

The graph below shows how the dependent variable *SoldPrice* changes across the Lake Country cities with the error bars showing the range which the actual mean is likely to occur for the given confidence level of 95%. (Graph for *GarageSpaces* is located earlier in this report.)



Discussion of Results and Recommendations

The ANOVA analysis of home sales in Lake Country for 2022 show that the number of garage spaces and the Lake Country cities significantly influence home selling prices. The one-way ANOVA showed that an increase in the number of spaces had a higher mean selling price. Factorial ANOVA showed that both the number of garage spaces and city significantly impact home selling price and their interaction also has an effect. Pewaukee has a lower mean compared to the other Lake Country cities while Delafield and Summit have higher means. The results from both analyses could provide insights for homebuyers who want to purchase a home in Lake Country and for real estate professionals to have a better understanding of the Lake Country home market.