4.0 Hypothesis

In statistical analysis, hypothesis testing is a crucial step that involves formulating a null hypothesis (H0) assuming no effect or difference, and an alternative hypothesis (H1) assuming a relationship or significant difference. Statistical tests, such as “chi-square tests” is carried out for categorical variables. For the continuous variable, “t-tests” is carried out to analyse data and investigate if there is sufficient evidence to reject the H0 and support H1 (Mishra et al., 2019).

ANOVA, or Analysis of Variance, is a statistical approach employed to compare the average values of several treatments (Mishra et al., 2019). It evaluates whether the observed differences in group means are likely due to genuine differences or random chance. ANOVA calculates an F-statistic by comparing the difference within groups and between group means. High F-statistic and corresponding p-value is below a predetermined cutoff (usually 0.05).

Levene's test for homogeneity of variance is employed for checking if assumption of equal variances holds true when conducting ANOVA. It assesses whether the variances of a continuous variable are similar across different groups or categories. A significant result suggests that there are significant differences in variances between the groups. On the other hand, Welch's ANOVA is a substitute for regular ANOVA when the rule of equal variances is broken. It adjusts the F-statistic and degrees of freedom to accommodate unequal variances. Both Levene's test and Welch's ANOVA play essential roles in statistical analysis, especially when comparing means between groups using ANOVA. They ensure that the underlying assumptions are met and yield more accurate results when there are variations in variances among the groups.

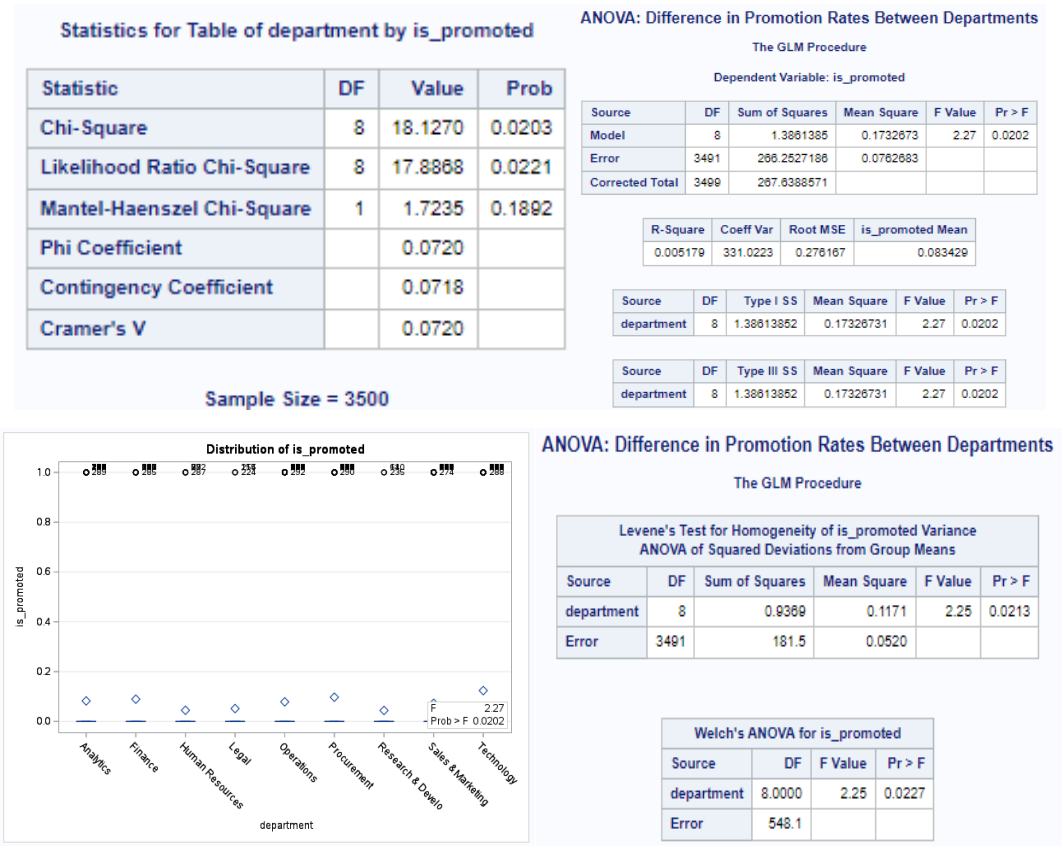
4.1 Hypothesis 1

H0: There is no significant difference in the “promotion status” between employees in different departments.

H1: There is a significant difference in the “promotion status” between employees in different departments.

Based on the chi-square test, the p-value is 0.0203 which is below the significance level of 0.05 (Figure 28). Hence, the null hypothesis is rejected while an alternative hypothesis is accepted. Based on ANOVA analysis, the p-value is 0.202. Similar to the chi-square test, the p-value is less than 0.05, the null hypothesis is rejected. Based on Levene’s test and Welch’s ANOVA, the p value 0.0213 and 0.0227, respectively, are below the typical significance level

of 0.05 which rejects the null hypothesis. For Tukey-Kramer adjusted pairwise comparison, between Sales & Marketing and Technology yields a p-value of 0.242 which is less than the significance level of 0.05 (Figure 29). It is suggested the occurrence of significant difference in the promotion status between employees in these two departments which support the alternative hypothesis. Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is rejected while the alternative hypothesis is accepted. This summarizes that there is a significant difference in the promotion rates between employees in different departments.



Statistics for Table of department by is_promoted

Statistic	DF	Value	Prob
Chi-Square	8	18.1270	0.0203
Likelihood Ratio Chi-Square	8	17.8868	0.0221
Mantel-Haenszel Chi-Square	1	1.7235	0.1892
Phi Coefficient		0.0720	
Contingency Coefficient		0.0718	
Cramer's V		0.0720	

Sample Size = 3500

ANOVA: Difference in Promotion Rates Between Departments

The GLM Procedure

Dependent Variable: is_promoted

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	1.3861385	0.1732673	2.27	0.0202
Error	3491	266.2527186	0.0762683		
Corrected Total	3499	267.6388571			

R-Square	Coeff Var	Root MSE	is_promoted Mean
0.005179	331.0223	0.276167	0.083429

Source	DF	Type I SS	Mean Square	F Value	Pr > F
department	8	1.38613852	0.17326731	2.27	0.0202

Source	DF	Type III SS	Mean Square	F Value	Pr > F
department	8	1.38613852	0.17326731	2.27	0.0202

ANOVA: Difference in Promotion Rates Between Departments

The GLM Procedure

Levene's Test for Homogeneity of is_promoted Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
department	8	0.9369	0.1171	2.25	0.0213
Error	3491	181.5	0.0520		

Welch's ANOVA for is_promoted

Source	DF	F Value	Pr > F
department	8.0000	2.25	0.0227
Error	548.1		

Figure 28. The chi-square and ANOVA test of the department by promotion status.

ANOVA: Difference in Promotion Rates Between Departments

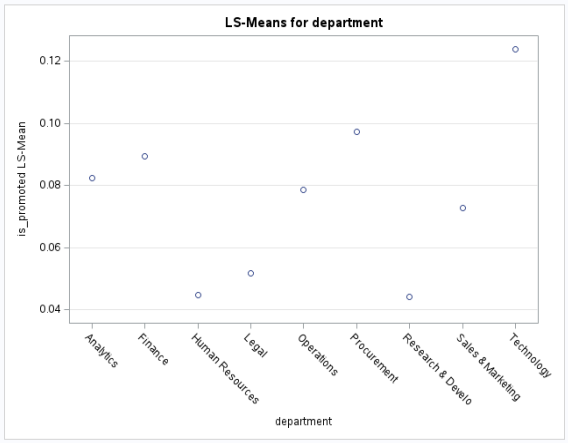
The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

department	is_promoted LSMEAN	LSMEAN Number
Analytics	0.08238636	1
Finance	0.08938547	2
Human Resources	0.04487179	3
Legal	0.05172414	4
Operations	0.07866667	5
Procurement	0.09732360	6
Research & Develo	0.04411765	7
Sales & Marketing	0.07277883	8
Technology	0.12393162	9

Least Squares Means for effect department
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: is_promoted

i/j	1	2	3	4	5	6	7	8	9
1		1.0000	0.8934	0.9973	1.0000	0.9981	0.9813	0.9998	0.4512
2	1.0000		0.8688	0.9929	0.9999	1.0000	0.9663	0.9981	0.8891
3	0.8934	0.8688		1.0000	0.9017	0.5295	1.0000	0.9611	0.0511
4	0.9973	0.9929	1.0000		0.9986	0.9613	1.0000	0.9998	0.6289
5	1.0000	0.9999	0.9017	0.9986		0.9742	0.9870	1.0000	0.1209
6	0.9981	1.0000	0.5295	0.9613	0.9742		0.8688	0.8420	0.8883
7	0.9813	0.9663	1.0000	1.0000	0.9870	0.8688		0.9960	0.3882
8	0.9998	0.9981	0.9611	0.9998	1.0000	0.8420	0.9960		0.0242
9	0.4512	0.8891	0.0511	0.6289	0.1209	0.8883	0.3882	0.0242	



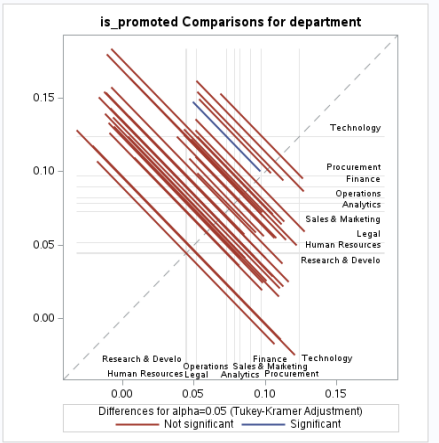


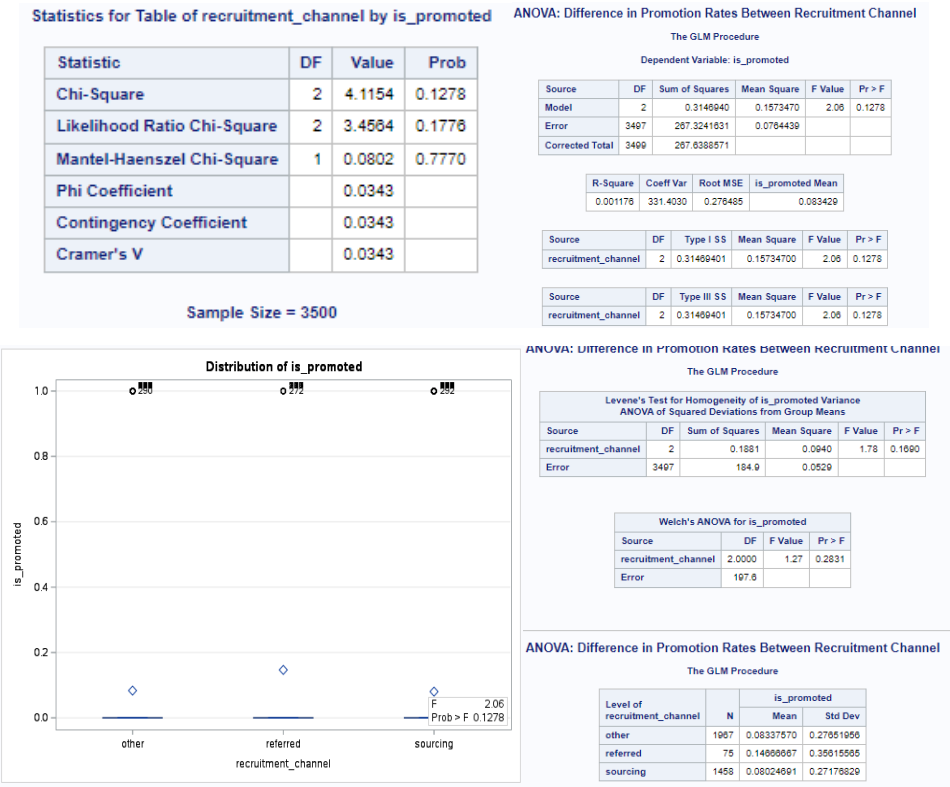
Figure 29. The Tukey-Kramer adjusted pairwise comparison of the department by promotion status.

4.2 Hypothesis 2

H0: There is no relationship between the recruitment channel and employee promotions.

H1: There is a relationship between the recruitment channel and employee promotions.

Based on the chi-square test, the p-value is 0.1278 which is more than the significance level of 0.05 (Figure 30). Hence, the null hypothesis failed to be rejected. Based on ANOVA analysis, the p-value is 0.1278. Based on Levene's test and Welch's ANOVA, the p value 0.1690 and 0.2831, respectively, are above the typical significance level of 0.05 which do not reject the null hypothesis. Similar to the chi-square test, the p-value is greater than 0.05, the null hypothesis is not rejected. For pairwise comparison, the p-values are greater than 0.05, indicating no significant difference in promotion rates between employees recruited from different channels (Figure 31). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis failed to be rejected and concludes that no statistical test to support significant relationship between the recruitment channel and employee promotions.



Statistics for Table of recruitment_channel by is_promoted

Statistic	DF	Value	Prob
Chi-Square	2	4.1154	0.1278
Likelihood Ratio Chi-Square	2	3.4564	0.1776
Mantel-Haenszel Chi-Square	1	0.0802	0.7770
Phi Coefficient		0.0343	
Contingency Coefficient		0.0343	
Cramer's V		0.0343	

Sample Size = 3500

ANOVA: Difference in Promotion Rates Between Recruitment Channel

The GLM Procedure

Dependent Variable: is_promoted

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.3146940	0.1573470	2.06	0.1278
Error	3497	267.3241631	0.0764439		
Corrected Total	3499	267.6388571			

R-Square	Coeff Var	Root MSE	is_promoted Mean
0.001176	331.4030	0.276485	0.083429

Source	DF	Type I SS	Mean Square	F Value	Pr > F
recruitment_channel	2	0.31469401	0.15734700	2.06	0.1278

Source	DF	Type III SS	Mean Square	F Value	Pr > F
recruitment_channel	2	0.31469401	0.15734700	2.06	0.1278

ANOVA: Difference in Promotion Rates Between Recruitment Channel

The GLM Procedure

Levene's Test for Homogeneity of is_promoted Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
recruitment_channel	2	0.1881	0.0940	1.78	0.1690
Error	3497	184.9	0.0529		

Welch's ANOVA for is_promoted

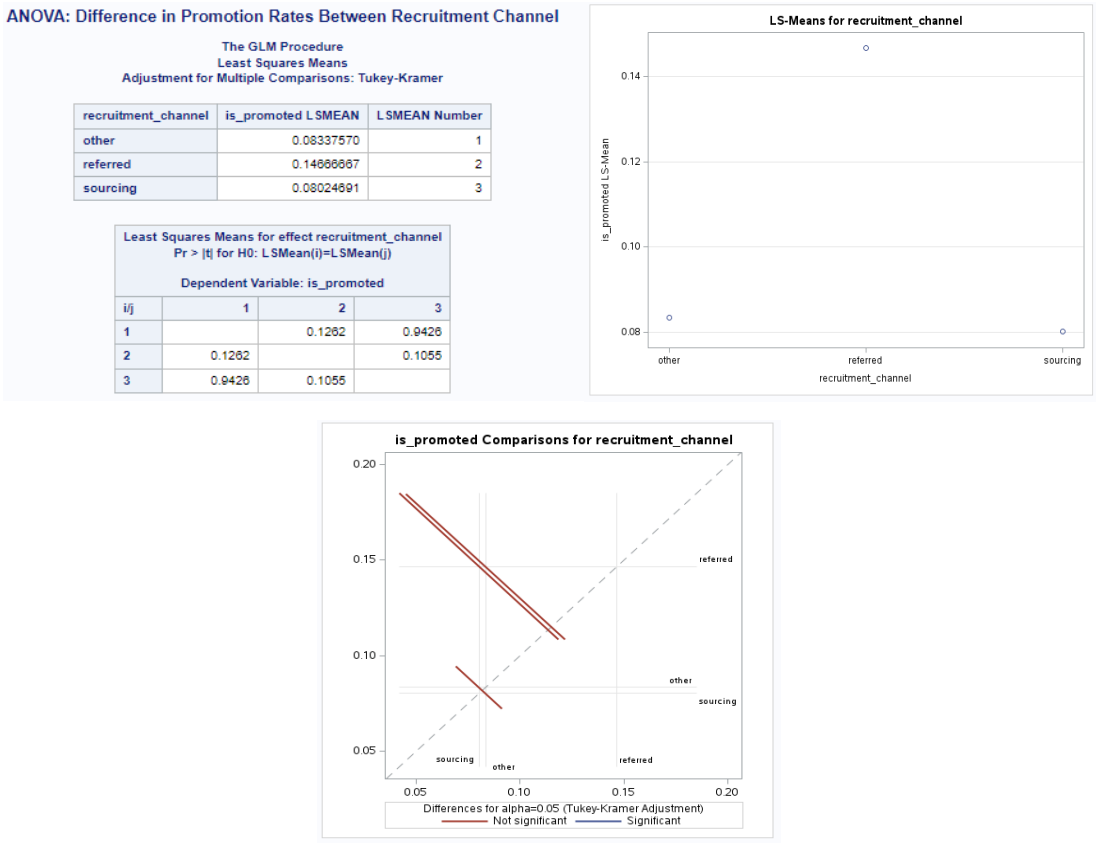
Source	DF	F Value	Pr > F
recruitment_channel	2.0000	1.27	0.2831
Error	197.6		

ANOVA: Difference in Promotion Rates Between Recruitment Channel

The GLM Procedure

Level of recruitment_channel	N	is_promoted Mean	is_promoted Std Dev
other	1967	0.08337570	0.27651956
referred	75	0.14666667	0.35515565
sourcing	1458	0.08024691	0.27176829

Figure 30. The chi-square and ANOVA test of the recruitment channel by promotion status.



ANOVA: Difference in Promotion Rates Between Recruitment Channel

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

recruitment_channel	is_promoted LSMEAN	LSMEAN Number
other	0.08337570	1
referred	0.14666667	2
sourcing	0.08024691	3

Least Squares Means for effect recruitment_channel
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: is_promoted

i/j	1	2	3
1		0.1262	0.9426
2	0.1262		0.1055
3	0.9426	0.1055	

Figure 31. The Tukey-Kramer adjusted pairwise comparison of the recruitment channel by promotion status.

4.3 Hypothesis 3

H0: There is no significant difference in the promotion rates between male and female employees.

H1: There is a significant difference in the promotion rates between male and female employees.

Based on the chi-square test, the p-value is 0.6371 which is more than the significance level of 0.05 (Figure 32). Hence, the null hypothesis failed to be rejected. Based on ANOVA analysis, the p-value is 0.6372. Similar to the chi-square test, the p-value is above 0.05, the null hypothesis is not rejected. According to Levene's test and Welch's ANOVA, the p value 0.6380 and 0.6410, respectively, are above the typical significance level of 0.05 which do not reject the null hypothesis. For pairwise comparison, the p-values are greater than 0.05, indicating no significant difference in promotion rates between employees' gender (Figure 33). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis failed to be rejected and concludes with no significant relationship between the gender and employee promotions.

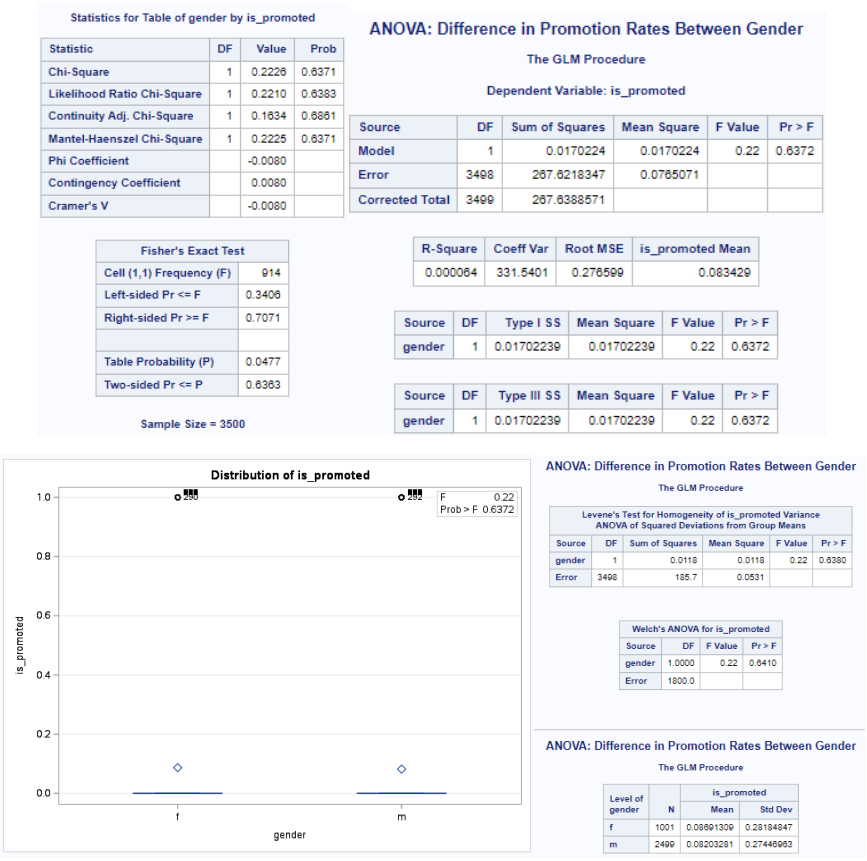


Figure 32. The chi-square and ANOVA test of gender by promotion status.

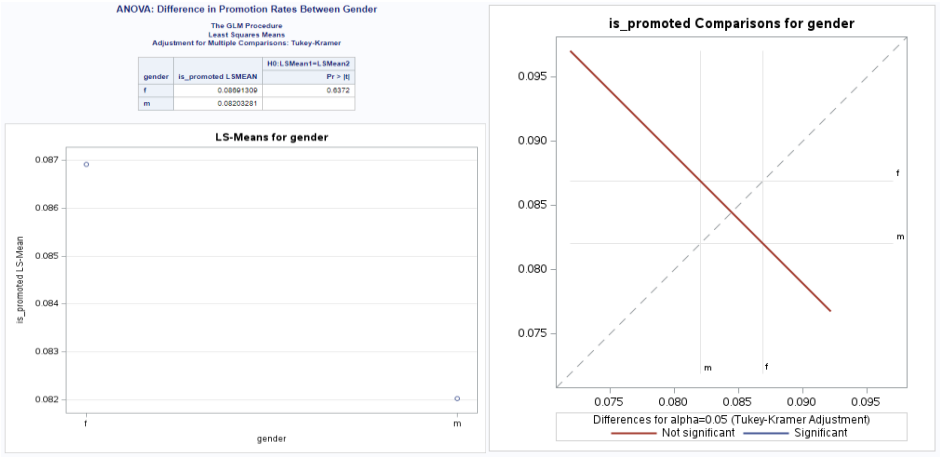


Figure 33. The Tukey-Kramer adjusted pairwise comparison of the gender by promotion status.

4.4 Hypothesis 4

H0: Age does not have a significant impact on the likelihood of promotion.

H1: Employees with higher age are more likely to be promoted.

Based on the t-test, logarithm of age was used as the original age data exhibit skewed distribution and nonlinear relationship. The logarithm of age help to stabilize variance which enhance the interpretability of results. The p-value of Pooled method is 0.1052 while for Satterthwaite method is 0.1033 which is above significance level of 0.05 (Figure 34). Hence, the null hypothesis failed to be rejected with the conclusion of no significant effect of age on employee promotions. Binning of age was used for ANOVA analysis as the categorical representation allows a straightforward comparison of mean difference across different age group. It enhances interpretability of the ANOVA result. Based on ANOVA analysis, the p-value is 0.1992. Similar to the t-test, the p-value is above 0.05, the null hypothesis is not rejected. Based on Levene's test and Welch's ANOVA, the p value 0.2051 and 0.2329, respectively, are above the typical significance level of 0.05 which do not reject the null hypothesis. For pairwise comparison, the p-values are greater than 0.05, indicating no significant difference in promotion status between employees' age (Figure 35). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis failed to be rejected and concludes no significant relationship between the age and employee promotions.

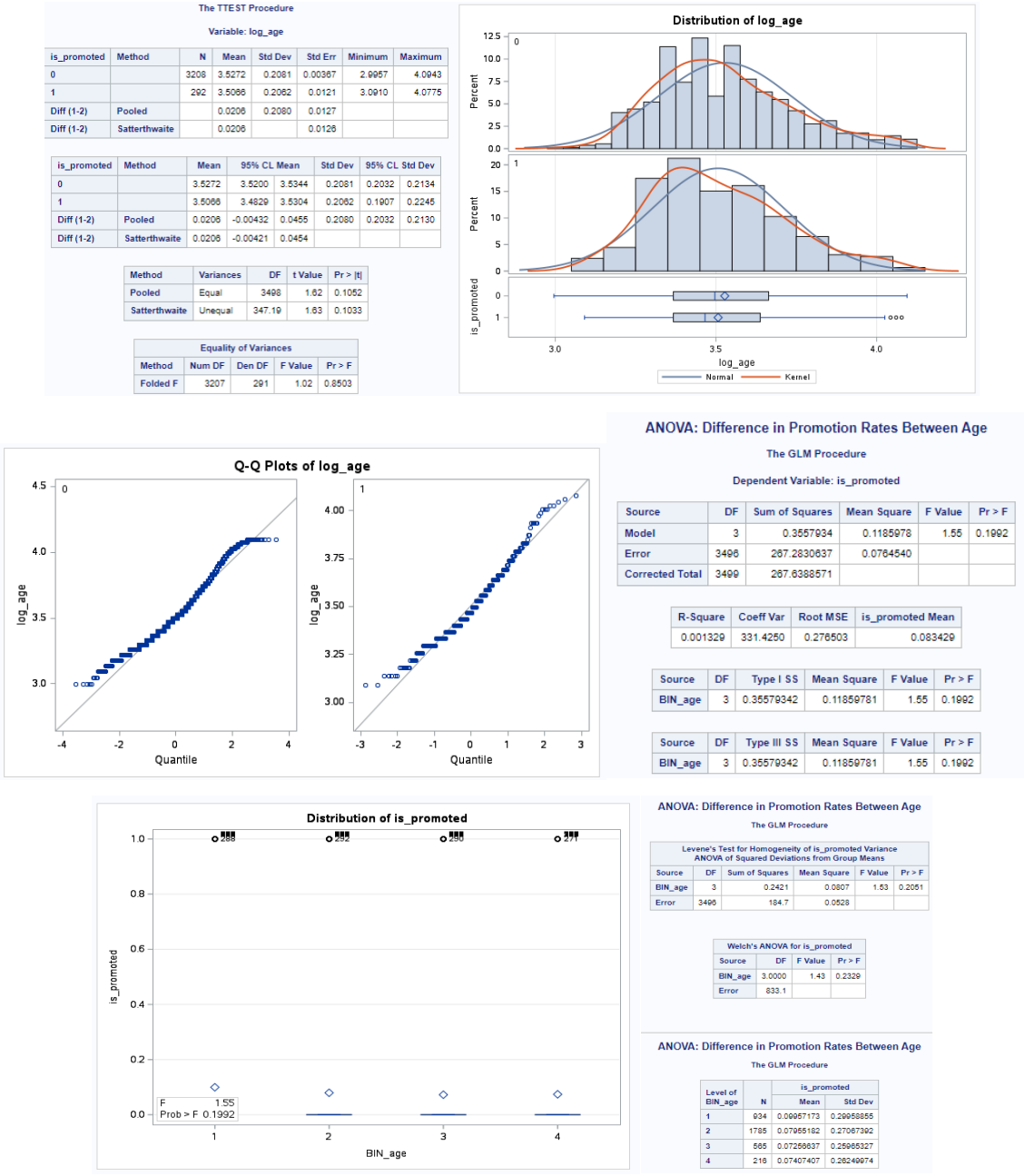


Figure 34. The t-test and ANOVA test of age by promotion status.

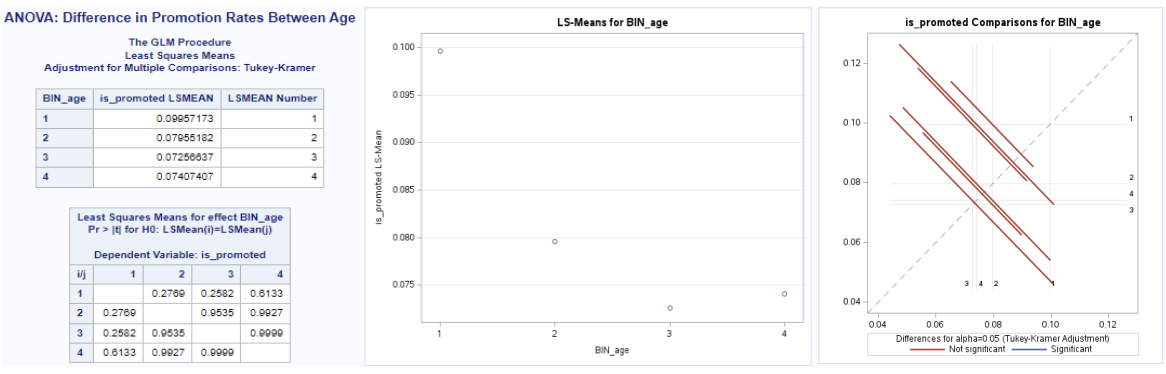


Figure 35. The Tukey-Kramer adjusted pairwise comparison of the age by promotion status.

4.5 Hypothesis 5

H0: The length of service does not have a significant impact on the likelihood of promotion.

H1: Employees with longer length of service are more likely to be promoted.

Based on the t-test, the p-value of Pooled method is 0.0006 while for Satterthwaite method is 0.0007 which is lesser than significance level of 0.05 (Figure 36). Hence, the null hypothesis is rejected with acceptance in alternative hypothesis. The p-value of equality of variances is 0.8512 which indicates no significant difference in variances. Based on ANOVA analysis, the p-value is 0.1746 which is greater than 0.05. The null hypothesis is not rejected. Based on Levene's test and Welch's ANOVA, the p value 0.1579 and 0.1164, respectively, are above the typical significance level of 0.05 which do not reject the null hypothesis of significant difference in variances and means. For pairwise comparison, the p-values are above 0.05, supporting no significant difference in promotion status between employees' length of service (Figure 37). Overall, the t-tests provide evidence that there is a significant impact of length of service on the likelihood of promotion. However, the ANOVA tests do not provide sufficient evidence to support this conclusion.

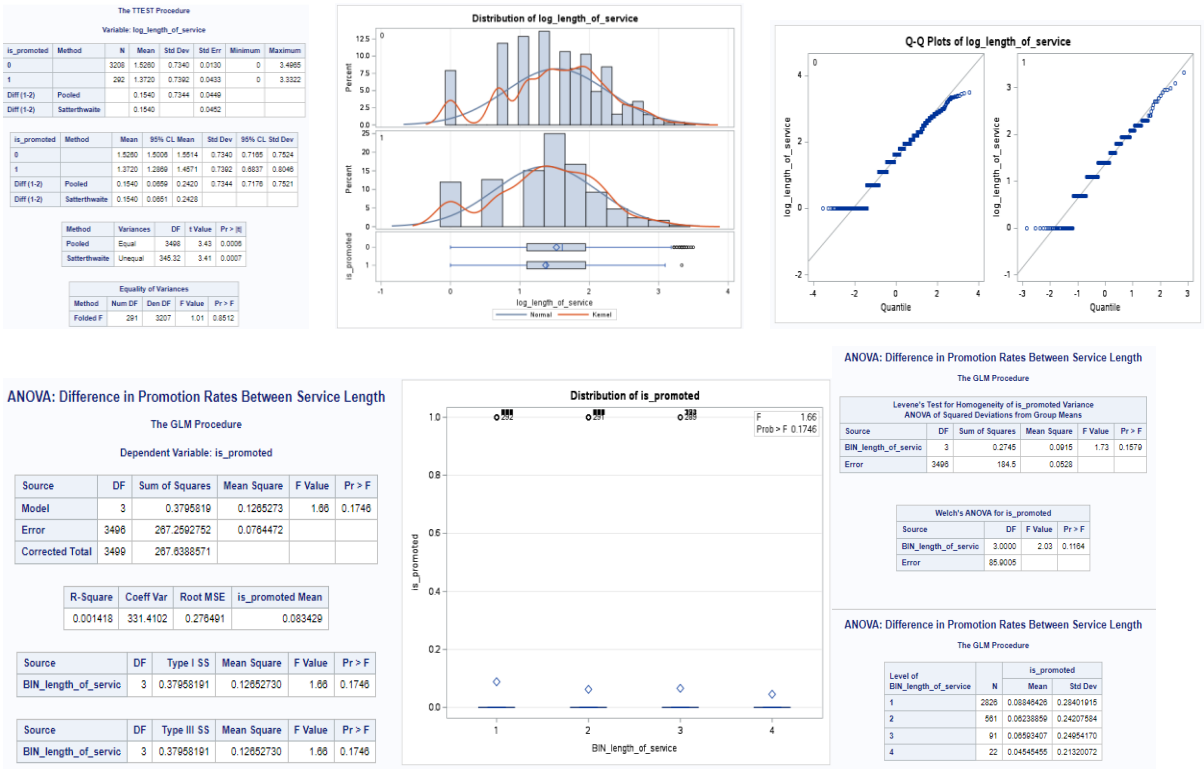


Figure 36. The t-test and ANOVA test of service length by promotion status.

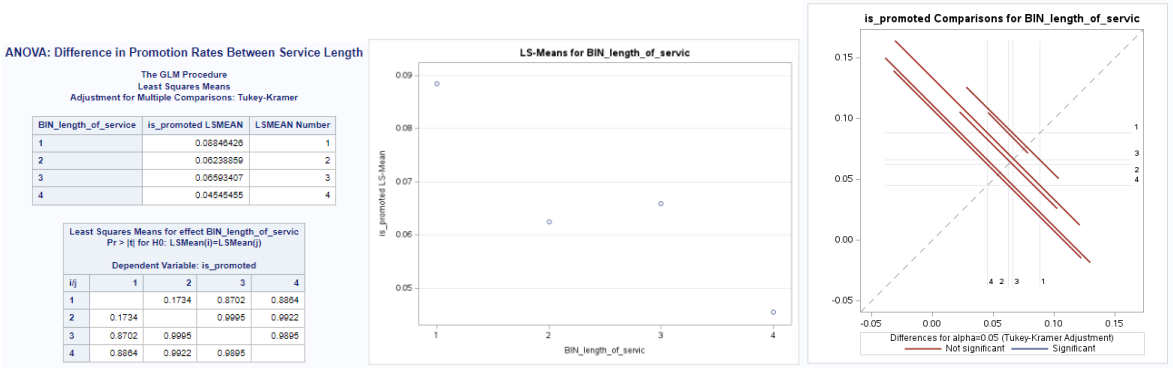


Figure 37. The Tukey-Kramer adjusted pairwise comparison of the service length by promotion status.

4.6 Hypothesis 6

H0: There is no significant difference in promotion rates between employees with different previous year ratings.

H1: Employees with higher previous year ratings are more likely to be promoted.

Based on the chi-square test, the p-value is less than 0.0001 which is below the significance level of 0.05 (Figure 38). Hence, the null hypothesis is rejected with acceptance in alternative hypothesis. Based on ANOVA analysis, the p-value is less than 0.0001. Similar to the chi-square test, the p-value is less than 0.05, the null hypothesis is rejected. Based on Levene's test and Welch's ANOVA, the p values are less than 0.0001 which are below the typical significance level of 0.05 which reject the null hypothesis. For pairwise comparison, the p-values are less than 0.0001, indicating employees with higher previous year ratings are more likely to be promoted (Figure 39). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is rejected. This summarizes a significant relationship between the previous year rating and employee promotions.

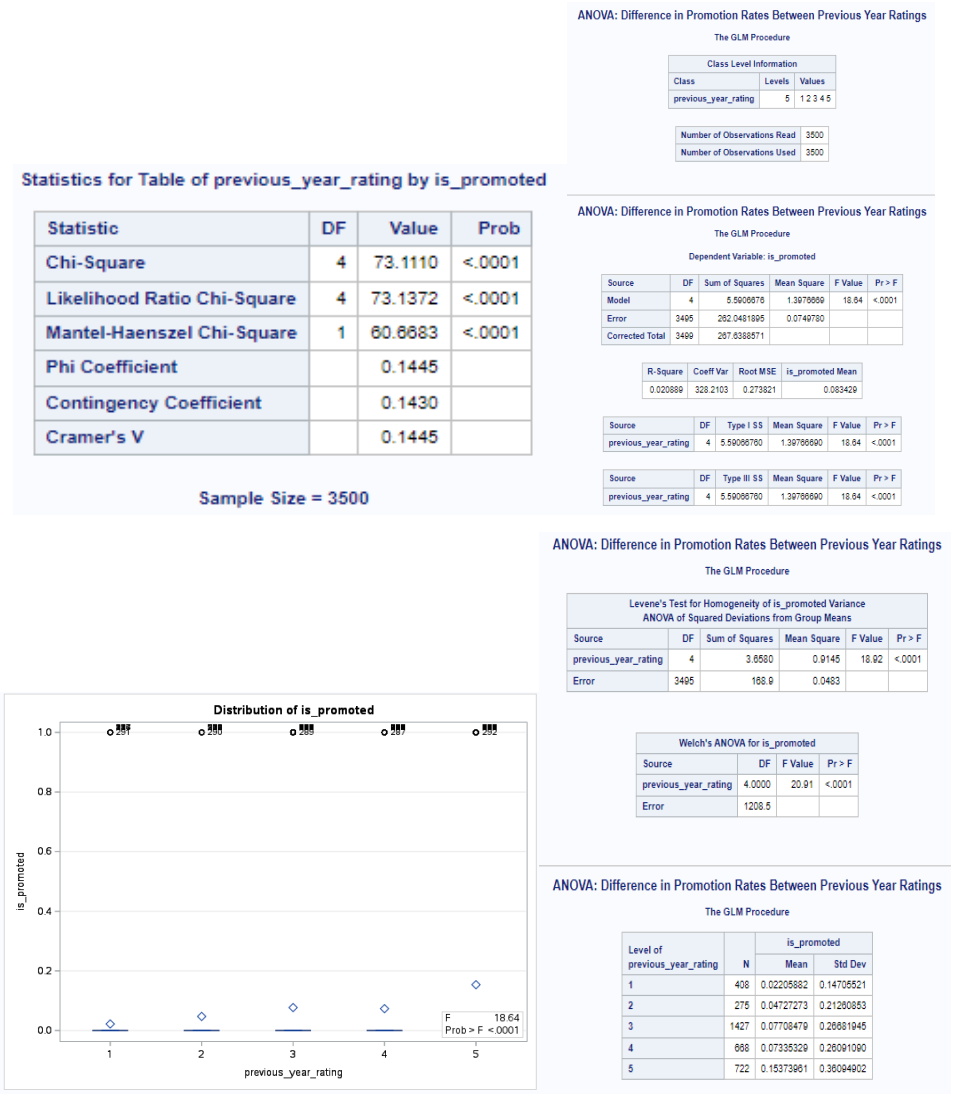


Figure 38. The chi-square test and ANOVA test of previous year rating by promotion status.

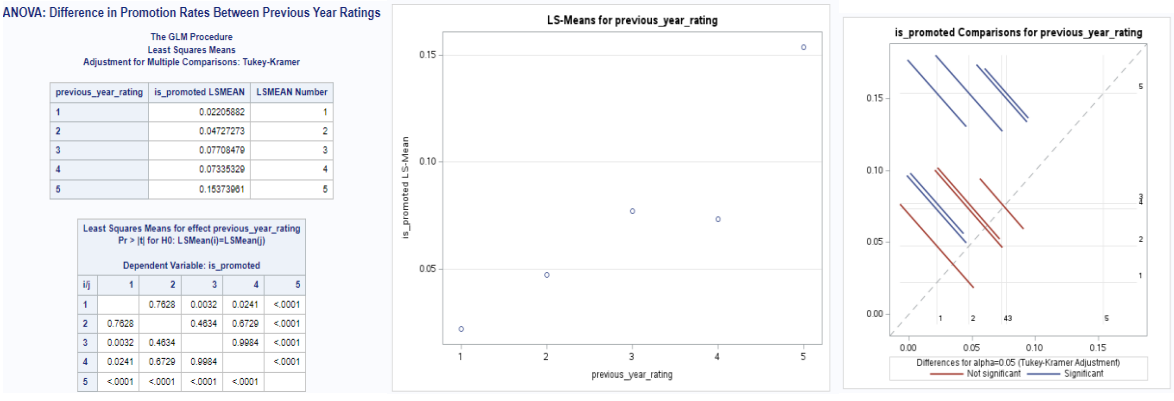


Figure 39. The Tukey-Kramer adjusted pairwise comparison of the previous year rating by promotion status.

4.7 Hypothesis 7

H0: The awards won do not have a significant impact on the likelihood of promotion.

H1: Employees with awards won are more likely to be promoted.

Based on the chi-square test, the p-value is less than 0.0001 which is below the significance level of 0.05 (Figure 40). Hence, the null hypothesis is rejected while the alternative hypothesis is accepted. Based on ANOVA analysis, the p-value is less than 0.0001. Similar to the chi-square test, the p-value is less than 0.05, the null hypothesis is rejected. Based on Levene's test and Welch's ANOVA, the p values are less than 0.0001 which are below the typical significance level of 0.05 which reject the null hypothesis. For pairwise comparison, the p-value is less than 0.0001, indicating employees with awards won are more likely to be promoted (Figure 41). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is rejected and an alternative hypothesis is accepted. This concludes that there is a significant relationship between the awards won and employee promotions.

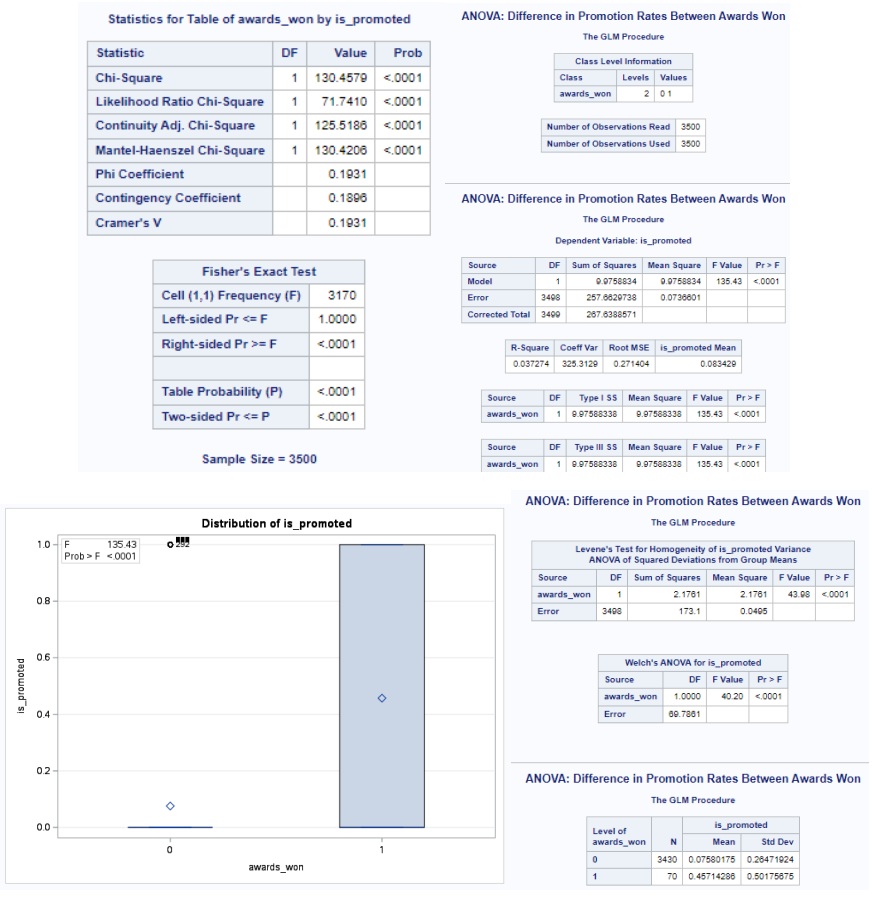


Figure 40. The chi-square test and ANOVA test of awards won by promotion status.

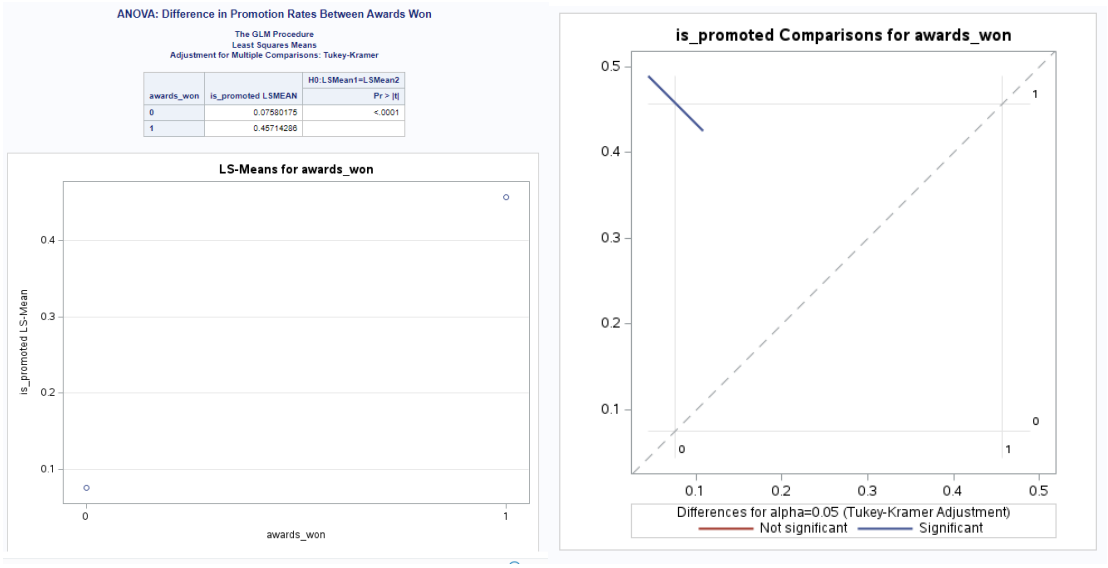


Figure 41. The Tukey-Kramer adjusted pairwise comparison of the awards won by promotion status.

4.8 Hypothesis 8

H0: There is no significant difference in promotion rates between employees with higher and lower average training scores.

H1: Employees with higher average training scores are more likely to be promoted.

Based on the t-test, the p-value of Pooled method is below 0.0001 while for Satterthwaite method is less than 0.0001 which is less than significance level of 0.05 (Figure 42). Hence, the null hypothesis is rejected and an alternative hypothesis is accepted. Based on ANOVA analysis, the p-value is less than 0.0001. Similar to the t-test, the p-value is less than 0.05, the null hypothesis is rejected. Based on Levene's test and Welch's ANOVA, the p values are less than 0.0001 which are below the typical significance level of 0.05 which reject the null hypothesis. For pairwise comparison, the p-value is below 0.05, indicating employees with high average training scores are more likely to be promoted (Figure 43). Overall, based on the t-test, ANOVA results and pairwise comparison, the null hypothesis is rejected. This summarizes a significant relationship between the average training score and employee promotions.

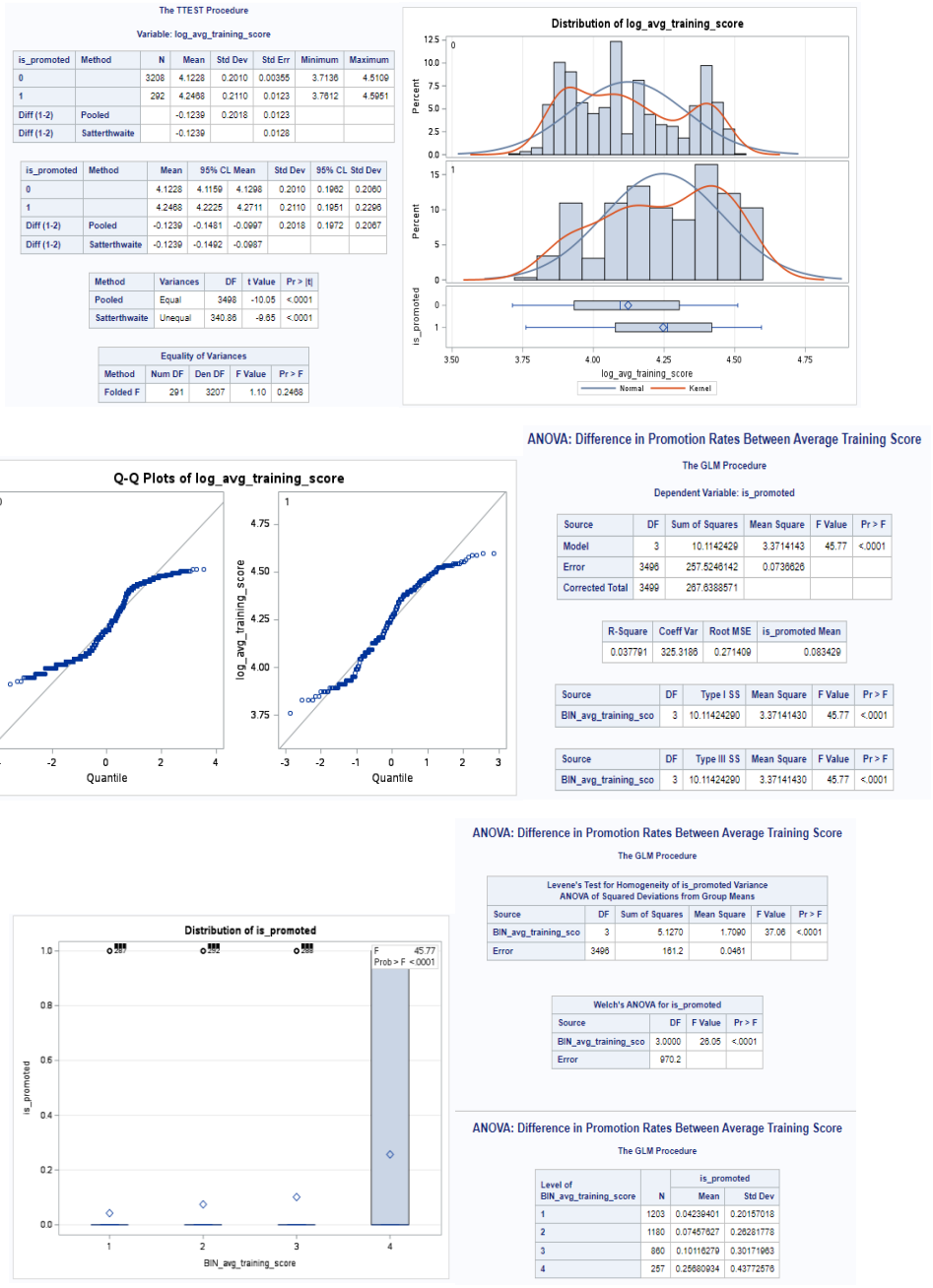


Figure 42. The t-test and ANOVA test of average training score by promotion status.

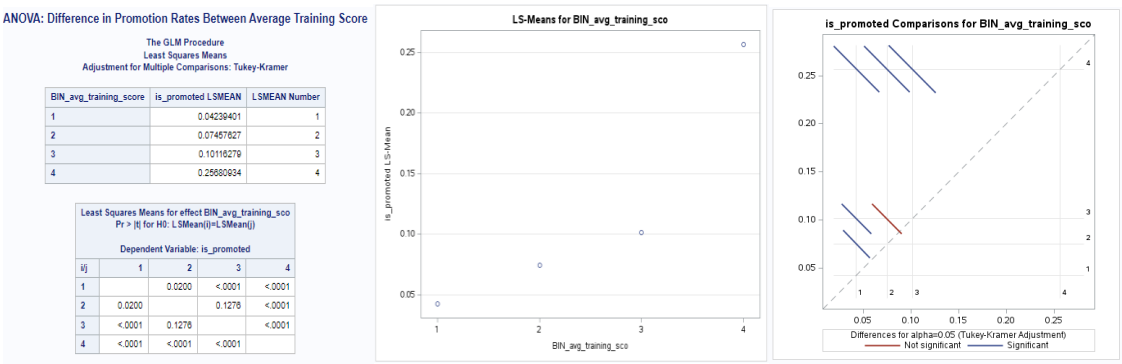


Figure 43. The Tukey-Kramer adjusted pairwise comparison of the average training score by promotion status.

4.9 Hypothesis 9

H0: There is no association between the education level of employees and their promotion status.

H1: There is an association between the education level of employees and their promotion status.

Based on the chi-square test, the p-value is 0.0276 which is less than the significance level of 0.05 (Figure 44). Hence, the null hypothesis is rejected and the alternative hypothesis is accepted. Based on ANOVA analysis, the p-value is 0.0276. Similar to the chi-square test, the p-value is less than 0.05, the null hypothesis is rejected. Based on Levene's test and Welch's ANOVA, the p values are 0.0317 which is below 0.05 indicating there is a significant difference in variances between the education levels. Welch's ANOVA is 0.0530, which is slightly above the significance level of 0.05, suggesting that there is weak evidence to reject the null hypothesis of no significant difference in means. For pairwise comparison, the p-value of Master's and above and Bachelor's is 0.0486, indicating there is a significant difference within these education levels (Figure 45). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is rejected and concludes with significant relationship between the education and employee promotions.

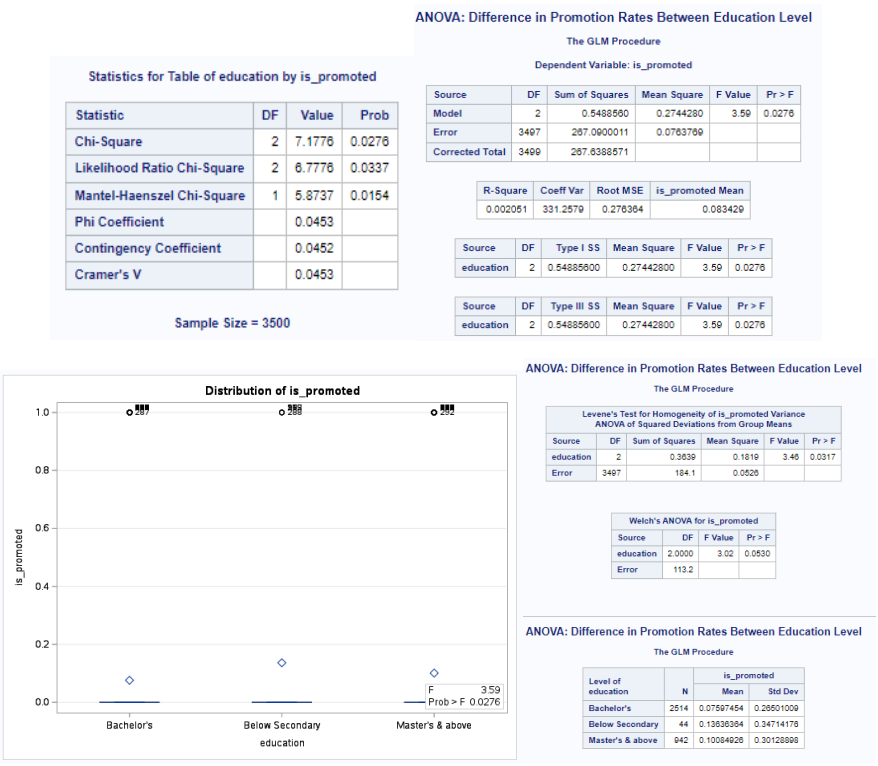


Figure 44. The chi-square test and ANOVA test of education level by promotion status.

ANOVA: Difference in Promotion Rates Between Education Level

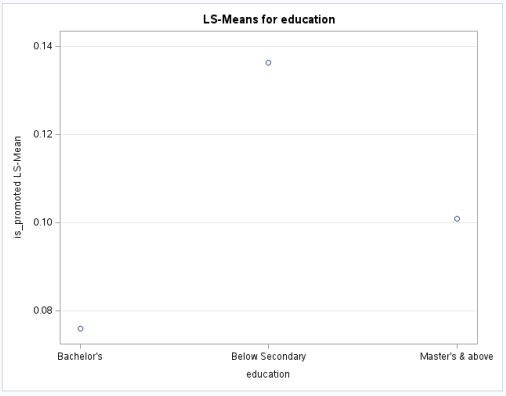
The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

education	is_promoted LSMEAN	LSMEAN Number
Bachelor's	0.07597454	1
Below Secondary	0.13636364	2
Master's & above	0.10084926	3

Least Squares Means for effect education
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: is_promoted

i/j	1	2	3
1		0.3220	0.0486
2	0.3220		0.6824
3	0.0486	0.6824	



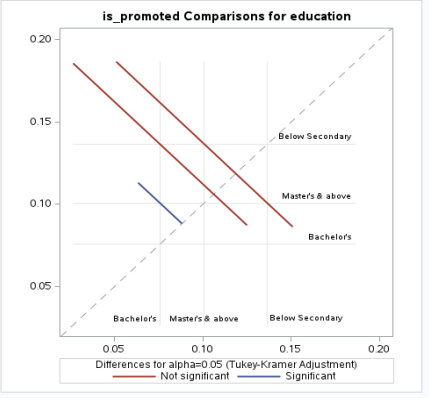


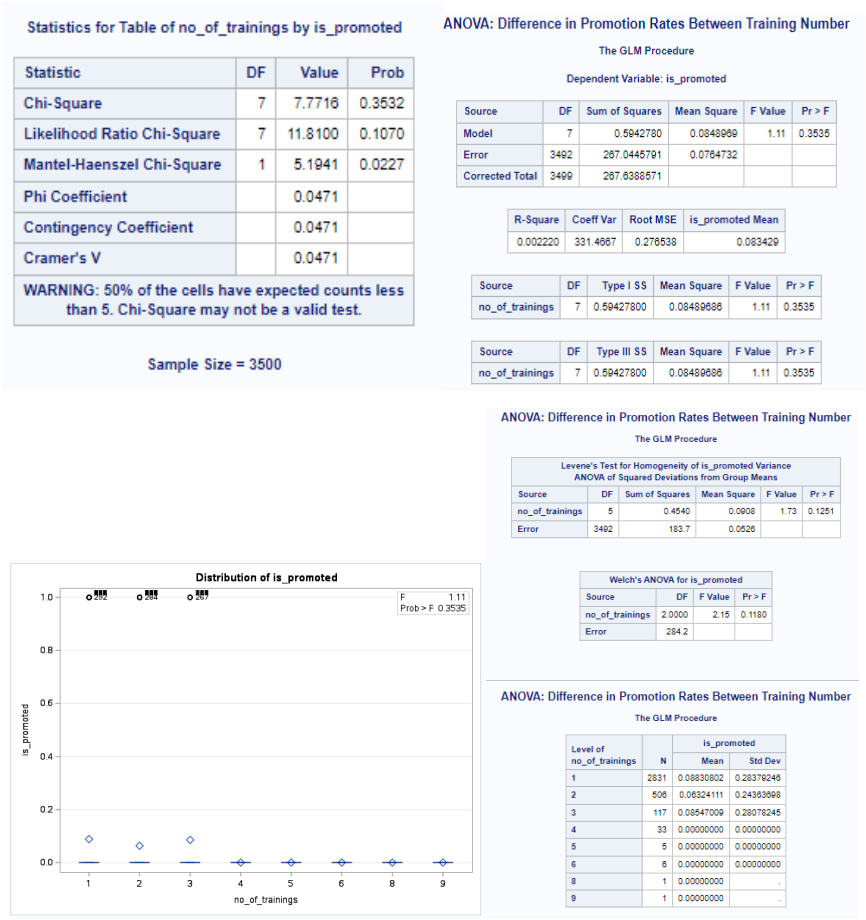
Figure 45. The Tukey-Kramer adjusted pairwise comparison of the education level by promotion status.

4.10 Hypothesis 10

H0: The number of training sessions does not have a significant influence on the likelihood of promotion.

H1: Employees with a higher number of training are more likely to be promoted.

Based on the chi-square test, the p-value is less than 0.3535 which is greater than the significance level of 0.05 (Figure 46). Hence, the null hypothesis is not rejected. Based on ANOVA analysis, the p-value is 0.3535. Similar to the chi-square test, the p-value is more than 0.05, the null hypothesis is not rejected. According to Levene's test and Welch's ANOVA, the p values are 0.1251 and 0.1180, respectively, which are above the typical significance level of 0.05 which do not reject the null hypothesis. For pairwise comparison, the p-values are more than 0.05, indicating training numbers have no significant impact on promotion status (Figure 47). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is not rejected and concludes with no significant relationship between the training numbers and employee promotions.



Statistics for Table of no_of_trainings by is_promoted

Statistic	DF	Value	Prob
Chi-Square	7	7.7716	0.3532
Likelihood Ratio Chi-Square	7	11.8100	0.1070
Mantel-Haenszel Chi-Square	1	5.1941	0.0227
Phi Coefficient		0.0471	
Contingency Coefficient		0.0471	
Cramer's V		0.0471	

WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 3500

ANOVA: Difference in Promotion Rates Between Training Number

The GLM Procedure

Dependent Variable: is_promoted

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	0.5942780	0.0848969	1.11	0.3535
Error	3492	267.0445791	0.0764732		
Corrected Total	3499	267.6388571			

R-Square	Coeff Var	Root MSE	is_promoted Mean
0.002220	331.4667	0.276538	0.083429

Source	DF	Type I SS	Mean Square	F Value	Pr > F
no_of_trainings	7	0.59427800	0.08489686	1.11	0.3535

Source	DF	Type III SS	Mean Square	F Value	Pr > F
no_of_trainings	7	0.59427800	0.08489686	1.11	0.3535

ANOVA: Difference in Promotion Rates Between Training Number

The GLM Procedure

Levene's Test for Homogeneity of is_promoted Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
no_of_trainings	5	0.4540	0.0908	1.73	0.1251
Error	3492	183.7	0.0526		

Welch's ANOVA for is_promoted

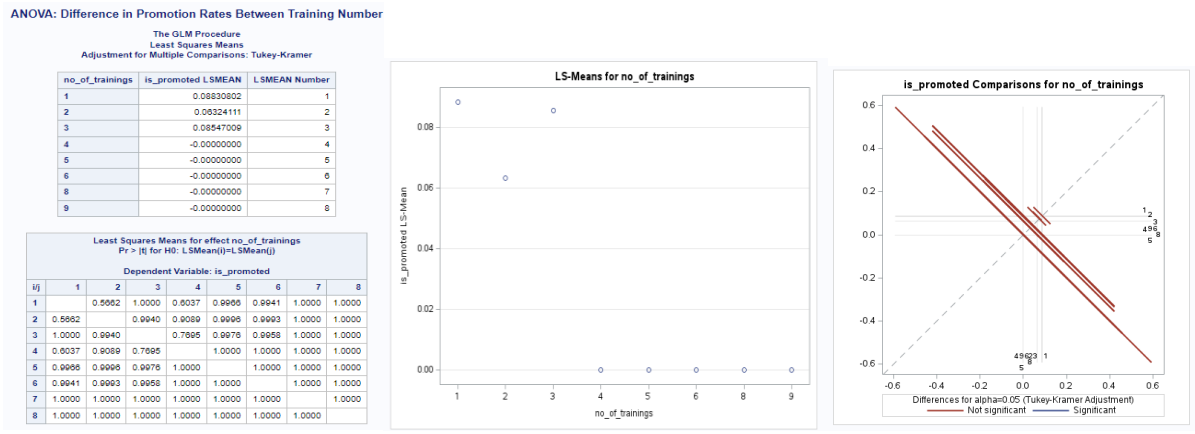
Source	DF	F Value	Pr > F
no_of_trainings	2.0000	2.15	0.1180
Error	284.2		

ANOVA: Difference in Promotion Rates Between Training Number

The GLM Procedure

Level of no_of_trainings	N	is_promoted Mean	is_promoted Std Dev
1	2831	0.08830802	0.28379246
2	506	0.06324111	0.24363698
3	117	0.08547009	0.28078245
4	33	0.00000000	0.00000000
5	5	0.00000000	0.00000000
6	6	0.00000000	0.00000000
8	1	0.00000000	.
9	1	0.00000000	.

Figure 46. The chi-square test and ANOVA test of training numbers by promotion status.



ANOVA: Difference in Promotion Rates Between Training Number

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

no_of_trainings	is_promoted LSMEAN	LSMEAN Number
1	0.08830802	1
2	0.06324111	2
3	0.08547009	3
4	-0.00000000	4
5	-0.00000000	5
6	-0.00000000	6
8	-0.00000000	7
9	-0.00000000	8

Least Squares Means for effect no_of_trainings
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: is_promoted

i/j	1	2	3	4	5	6	7	8
1		0.5662	1.0000	0.6037	0.9966	0.9941	1.0000	1.0000
2	0.5662		0.9940	0.9089	0.9996	0.9993	1.0000	1.0000
3	1.0000	0.9940		0.7695	0.9976	0.9958	1.0000	1.0000
4	0.6037	0.9089	0.7695		1.0000	1.0000	1.0000	1.0000
5	0.9966	0.9996	0.9976	1.0000		1.0000	1.0000	1.0000
6	0.9941	0.9993	0.9958	1.0000	1.0000		1.0000	1.0000
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		1.0000
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	

Figure 47. The Tukey-Kramer adjusted pairwise comparison of the training numbers by promotion status.

4.11 Hypothesis 11

H0: There is no association between the region of employees and their promotion status.

H1: There is an association between the region of employees and their promotion status.

Based on the chi-square test, the p-value is less than 0.2242 which is greater than the significance level of 0.05 (Figure 48). Therefore, the null hypothesis is not rejected with the conclusion of no significant effect of region on employee promotions. Based on ANOVA analysis, the p-value is 0.2238. Similar to the chi-square test, the p-value is more than 0.05, the null hypothesis is not rejected. Based on Levene's test and Welch's ANOVA, the p values are 0.1136 and 0.2431, respectively, which are above the typical significance level of 0.05 which do not reject the null hypothesis. For pairwise comparison, the p-values are more than 0.05, indicating the region has no significant impact on promotion status (Figure 49). Overall, based on the chi-square test, ANOVA results and pairwise comparison, the null hypothesis is not rejected and concludes with no significant relationship between the region and employee promotions.

Statistics for Table of region by is_promoted

Statistic	DF	Value	Prob
Chi-Square	33	38.8114	0.2242
Likelihood Ratio Chi-Square	33	46.5018	0.0597
Mantel-Haenszel Chi-Square	1	0.0990	0.7530
Phi Coefficient		0.1053	
Contingency Coefficient		0.1047	
Cramer's V		0.1053	

WARNING: 28% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 3500

ANOVA: Difference in Promotion Rates Between Region

The GLM Procedure

Dependent Variable: is_promoted

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	33	2.9678385	0.0899345	1.18	0.2238
Error	3466	264.6710187	0.0763621		
Corrected Total	3499	267.6388571			

R-Square	Coeff Var	Root MSE	is_promoted Mean
0.011089	331.2258	0.276337	0.083429

Source	DF	Type I SS	Mean Square	F Value	Pr > F
region	33	2.96783847	0.08993450	1.18	0.2238

Source	DF	Type III SS	Mean Square	F Value	Pr > F
region	33	2.96783847	0.08993450	1.18	0.2238

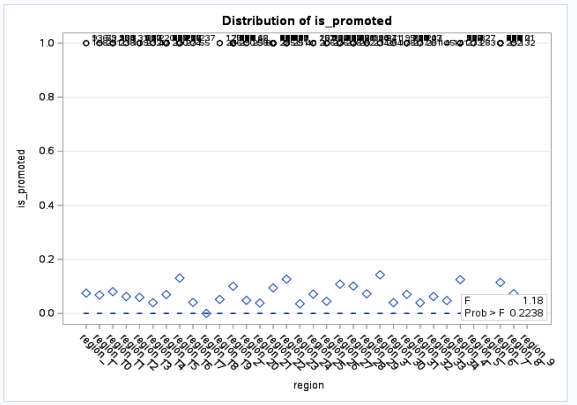


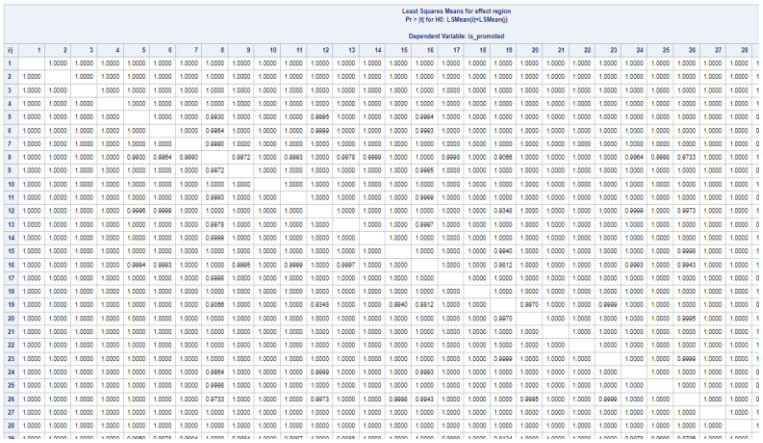
Figure 48. The chi-square test and ANOVA test of region by promotion status.

ANOVA: Difference in Promotion Rates Between Region

The GLM Procedure

Levene's Test for Homogeneity of is_promoted Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
region	32	2.1311	0.0666	1.31	0.1136
Error	3465	176.0	0.0508		

Welch's ANOVA for is_promoted			
Source	DF	F Value	Pr > F
region	30.0000	1.17	0.2431
Error	542.7		



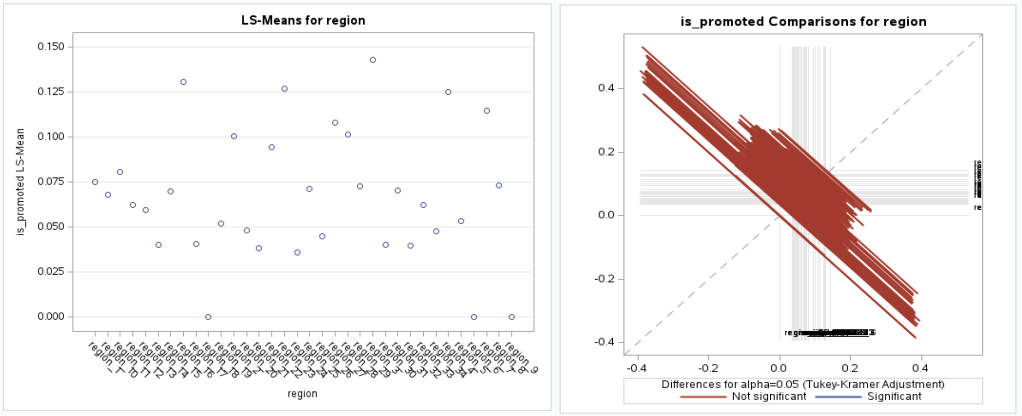


Figure 49. The Tukey-Kramer adjusted pairwise comparison of the region by promotion status.