Firefighter Pay & Overtime Analysis

For the following assignment you will be working with the raw data at the bottom of this file that describes differences in pay and overtime pay of different rank firefighters, in total there are 38 rows of data. In total there are three main parts to this assignment.

Part 1) Create a new categorical variable called OverTimePay in the data set that categorizes those who make over 25000 dollars in overtime pay as 'High' and those who make under 25000 dollars in overtime pay as 'Low'. Use a test to determine if the proportion of High vs Low categories are significantly different from 40% and 60%, respectively. Report the p-value and its interpretation, along with the actual observed proportions of the data.

A new categorical variable, OverTimePay, was created to classify firefighters with overtime pay above $25,000 as "High" and those with less as "Low". Among the 38 individuals in the dataset, the distribution was evenly split: 19 (50%) in each group. A one-sample binomial test was conducted to determine whether this observed proportion of 50% High significantly differs from the hypothesized 40%. The two-sided p-value from the test was 0.2083, indicating no statistically significant difference at the 0.05 level. This suggests that the observed proportions do not differ significantly from the expected 40% High and 60% Low distribution.

The FREQ Procedure

OverTimePay	Frequency	Percent	Cumulative Frequency	Cumulative Percent
High	19	50.00	19	50.00
Low	19	50.00	38	100.00

Binomial Proportion	
OverTimePay = High	
Proportion	0.5000
ASE	0.0811
95% Lower Conf Limit	0.3410
95% Upper Conf Limit	0.6590
Exact Conf Limits	
95% Lower Conf Limit	0.3338
95% Upper Conf Limit	0.6662

Test of H0: Proportion = 0.4	
ASE under H0	0.0795
Z	1.2583
One-sided Pr > Z	0.1041
Two-sided Pr > |Z|	0.2083

Sample Size = 38

Part 2) Using OverTimePay and Rank run a two-way chi square test for independence. State the null hypothesis of this test versus the alternative hypothesis of this test, in context. Report the p-value and state whether you should reject the null hypothesis.

A two-way chi-square test was conducted to evaluate whether the OverTimePay category (High vs. Low) is independent of firefighter Rank (Captain, Engineer, Firefighter). The test yielded a Chi-Square statistic of 0.1578 with a p-value of 0.9242. Since this p-value is much greater than the significance level of 0.05, we fail to reject the null hypothesis, indicating no statistically significant association between firefighter rank and overtime pay category. This means the distribution of High and Low overtime earners appears independent of their rank.

The FREQ Procedure

Frequency
Expected
Percent
Row Pct
Col Pct

Table of OverTimePay by Rank

OverTimePay	Rank			
	CAPTAIN	ENGINEER	FIREFIGHTER	Total
High	6 5.5 15.79 31.58 54.55	6 6 15.79 31.58 50.00	7 7.5 18.42 36.84 46.67	19 50.00
Low	5 5.5 13.16 26.32 45.45	6 6 15.79 31.58 50.00	8 7.5 21.05 42.11 53.33	19 50.00
Total	11 28.95	12 31.58	15 39.47	38 100.00

Statistics for Table of OverTimePay by Rank

Statistic	DF	Value	Prob
Chi-Square	2	0.1576	0.9242
Likelihood Ratio Chi-Square	2	0.1578	0.9242
Mantel-Haenszel Chi-Square	1	0.1523	0.6964
Phi Coefficient		0.0644	
Contingency Coefficient		0.0643	
Cramer's V		0.0644	

Sample Size = 38

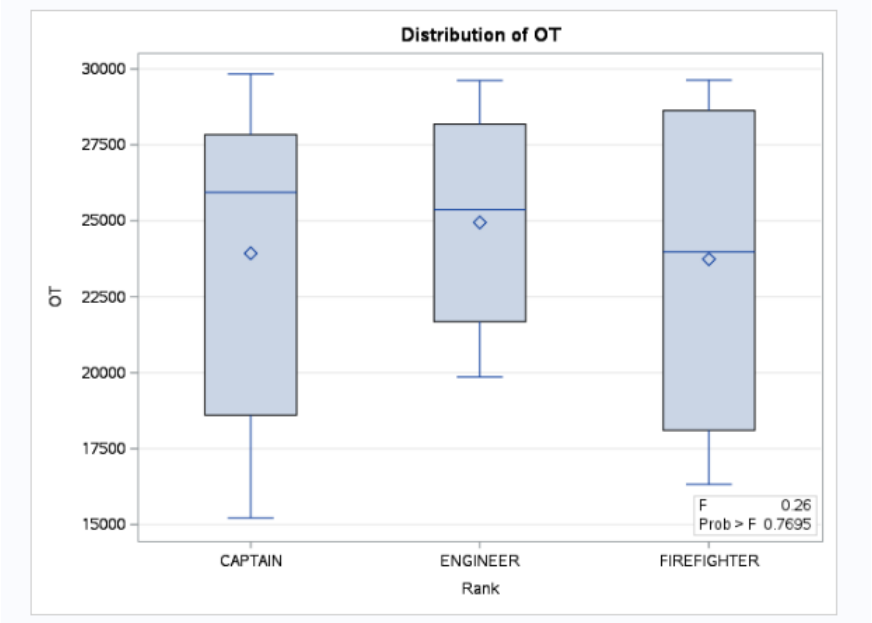
Part 3) Run a one-way anova test using the variable OT as the response and Rank as the independent variable to determine if a specific rank of Firefighter on average gets paid more in overtime. In addition, run any additional tests necessary to see if there are any significant differences amongst any of the rank categories. State the null hypothesis and whether you should reject it as well as if there were any significant differences between pairs. Report the p-value and any other statistic deemed necessary to justify your conclusions.

A one-way ANOVA was performed to test whether the mean overtime pay (OT) differs significantly by firefighter Rank (Captain, Engineer, Firefighter). The ANOVA output shows an F-value of 0.26 with a p-value of 0.7695, indicating no statistically significant difference in mean overtime pay across ranks. Thus, we fail to reject the null hypothesis, which states that all rank groups have equal mean OT.

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Rank	3	CAPTAIN ENGINEER FIREFIGHTER

Number of Observations Read	38
Number of Observations Used	38



The ANOVA Procedure

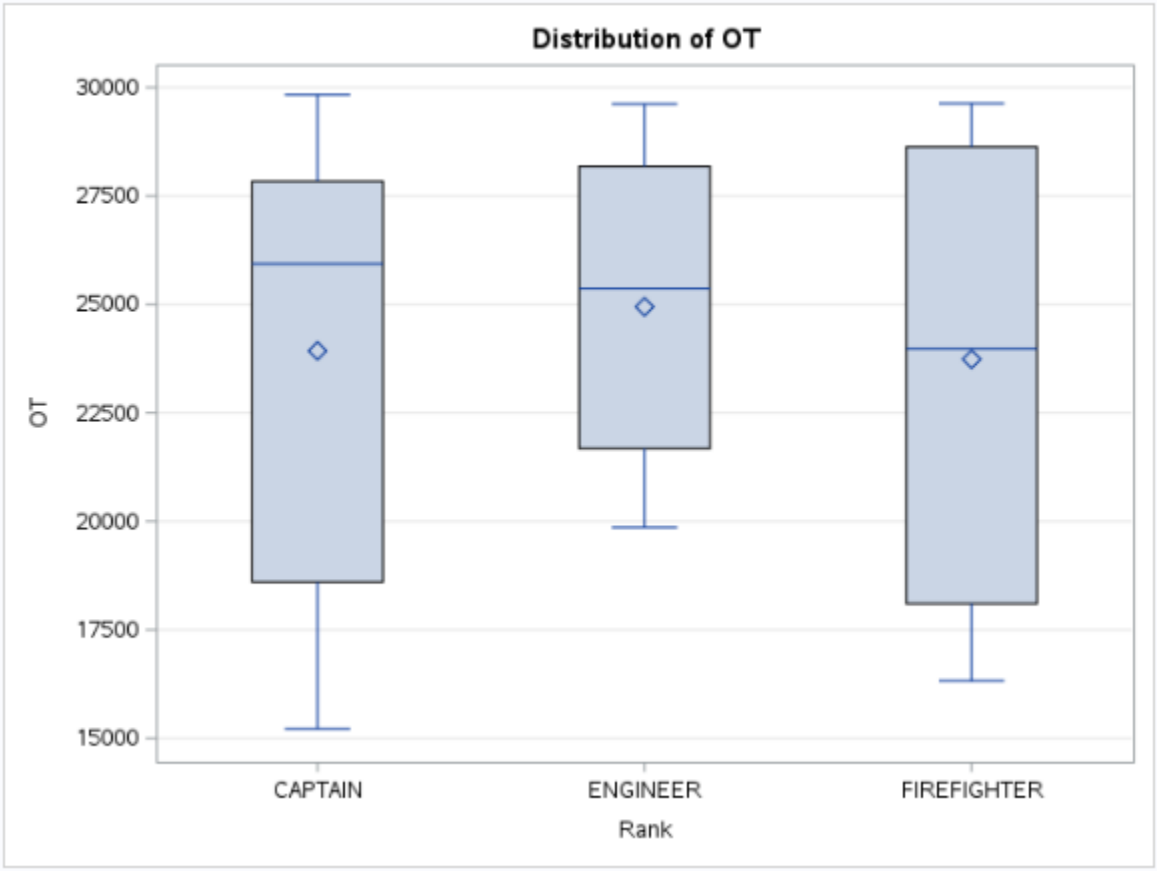
Dependent Variable: OT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	10661584.5	5330792.2	0.26	0.7695
Error	35	706832257.0	20195207.3		
Corrected Total	37	717493841.5			

R-Square	Coeff Var	Root MSE	OT Mean
0.014859	18.58713	4493.908	24177.53

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Rank	2	10661584.50	5330792.25	0.26	0.7695

The ANOVA Procedure



The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for OT

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	35
Error Mean Square	20195207
Critical Value of Studentized Range	3.46096

Comparisons significant at the 0.05 level are indicated by ***.

Rank Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
ENGINEER - CAPTAIN	1016	-3575	5607	
ENGINEER - FIREFIGHTER	1208	-3051	5468	
CAPTAIN - ENGINEER	-1016	-5607	3575	
CAPTAIN - FIREFIGHTER	192	-4173	4558	
FIREFIGHTER - ENGINEER	-1208	-5468	3051	
FIREFIGHTER - CAPTAIN	-192	-4558	4173	