University of the People

MATH 1281 - Statistical Inference

Unit 5 Written Assignment 5

Liang Xiao

# Assignment: ANOVA Analysis of WickedWitch Complaints

## a) Descriptive Statistics for Complaints Across Regions

The data from WickedWitch.csv was imported into JASP, and descriptive statistics were computed for the variable Complaints, grouped by Region. The summary is as follows:

|  |  |  |
| --- | --- | --- |
| Region | Mean | Standard Deviation |
| 1 | 1.0 | 1.22 |
| 2 | 1.4 | 1.14 |
| 3 | 3.6 | 1.14 |
| 4 | 4.2 | 0.84 |

1. Region with Highest Average Complaints:  
 - Region 4 has the highest average number of complaints, 4.2.  
  
2. Region with Largest Standard Deviation:  
 - Region 1 has the highest standard deviation of complaints, approximately 1.22, indicating more variability in complaint counts.  
  
3. Do We Expect Differences Across Regions?  
 - Yes. The means differ noticeably (ranging from 1.0 to 4.2), which suggests a likely difference in average complaints across the four regions. This observed spread justifies performing an ANOVA test to statistically evaluate the differences.

## b) Hypotheses for ANOVA

To determine whether there is a statistically significant difference in the average number of complaints among the regions, the following hypotheses are set:  
  
- Null Hypothesis (H₀):  
 The average number of complaints is equal across all four regions.  
 H₀: μ₁ = μ₂ = μ₃ = μ₄  
  
- Alternative Hypothesis (H₁):  
 At least one region has a different average number of complaints.  
 H₁: At least one μᵢ ≠ μⱼ

## c) ANOVA Output and Interpretation

Using JASP, a one-way ANOVA was conducted with Complaints as the dependent variable and Region as the fixed factor. The output was as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Sum of Squares | df | Mean Square | F | p-value |
| Region | 37.75 | 3 | 12.583 | 10.486 | < .001 |
| Residual | 19.2 | 16 | 1.2 |  |  |

1. F-statistic and p-value:  
 - F = 10.486, p < .001  
  
2. Conclusion:  
 - Since the p-value is less than the significance level of 0.05, we reject the null hypothesis.  
 - There is strong evidence to conclude that the average number of complaints differs significantly among the four regions.