11/12/2018

Andrew Marshall

Minitab Project-3.2

INFS 608 Applied Statistics

# Overview

We have been tasked with analyzing data for a 2 -part study done in the following 3 locations: New York, North Carolina and Florida. For the first part, there were 20 people in good health selected at each location and they were asked to take a standardized test to measure depression. For the second part, 20 people in relatively poor health were selected at each of the 3 locations. A higher score on said exam indicates a higher level of depression.

# Initial Findings

Based on the data below it appears that location does have an effect on a person’s depression level. You will note that the mean score gets lower the further south the person is located, though the group of healthy people have an overall lower score than the non-healthy people. Further analysis will be necessary to determine any further statistical significance.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Healthy |  |  |  |  |  |  |  |  |  |  |
| **Variable** | **N** | **Mean** | **SE Mean** | **StDev** | **Minimum** | **Q1** | **Median** | **Q3** | **Maximum** | **IQR** |
| Florida | 20 | 5.55 | 0.478 | 2.139 | 2 | 3.25 | 6 | 7 | 9 | 3.75 |
| New York | 20 | 8 | 0.492 | 2.2 | 4 | 7 | 8 | 8.75 | 13 | 1.75 |
| North Carolina | 20 | 7.05 | 0.634 | 2.837 | 3 | 4.25 | 7.5 | 8.75 | 12 | 4.5 |

Figure 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Non-Healthy | |  |  |  |  |  |  |  |  |  |
| **Variable** | **N** | **Mean** | **SE Mean** | **StDev** | **Minimum** | **Q1** | **Median** | **Q3** | **Maximum** | **IQR** |
| Florida | 20 | 14.5 | 0.709 | 3.171 | 9 | 12 | 14.5 | 17 | 21 | 5 |
| New York | 20 | 15.25 | 0.923 | 4.128 | 9 | 12.25 | 14.5 | 17.75 | 24 | 5.5 |
| North Carolina | 20 | 13.95 | 0.659 | 2.946 | 8 | 12 | 14 | 16.75 | 19 | 4.75 |

Figure 2

# Part I Variance Analysis (One-way ANOVA)

Our analysis will be based on the following hypothesis test:

H0 : All groups in the sample are equal.

H1: At least one group in sample is not equal to the others.

Significance level: α = 0.05

**Part I: Healthy**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Analysis of Variance** |  |  |  |  |  |
| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
| Factor | 2 | 61.03 | 30.517 | 5.24 | 0.008 |
| Error | 57 | 331.9 | 5.823 |  |  |
| Total | 59 | 392.93 |  |  |  |

Figure 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Means |  |  |  |  |
| Factor | N | Mean | StDev | 95% CI |
| Florida | 20 | 5.55 | 2.139 | (4.470, 6.630) |
| New York | 20 | 8 | 2.2 | (6.920, 9.080) |
| North Carolina | 20 | 7.05 | 2.837 | (5.970, 8.130) |
| *Pooled StDev = 2.41305* |  |  |  |  |

Figure 4



Figure 5

# Part I Conclusion

Based on the data provided above and the fact that the p-value of 0.008 is less than 0.05, we are required to reject the null hypothesis. This means that at least on the means in the group is not equal to the others.

# Part II Variance Analysis (One-way ANOVA)

Our analysis will be based on the following hypothesis test:

H0 : All groups in the sample are equal.

H1: At least one group in sample is not equal to the others.

Significance level: α = 0.05

**Part II: Non-Healthy**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Analysis of Variance** |  |  |  |  |  |
| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
| Factor | 2 | 17.03 | 8.517 | 0.71 | 0.494 |
| Error | 57 | 679.7 | 11.925 |  |  |
| Total | 59 | 696.73 |  |  |  |

Figure 6

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Means** |  |  |  |  |
| Factor | N | Mean | StDev | 95% CI |
| Florida | 20 | 14.5 | 3.171 | (12.954, 16.046) |
| New York | 20 | 15.25 | 4.128 | (13.704, 16.796) |
| North Carolina | 20 | 13.95 | 2.946 | (12.404, 15.496) |

Figure 7



Figure 8

# Part II Conclusion

Based on the data provided above and the fact that the p-value of 0.494 is greater than 0.05, we will fail to reject the null hypothesis at a 95% confidence level. This means that the group means are equal based on a 95% confidence interval.

# Part III: Secondary Analysis (Pairwise)

Given that the null hypothesis was rejected for our first sample group, we need to run an additional analysis on the 3 groups to see which 1 of the groups differs from the others.

# Analysis 1: New York/Florida

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Two-Sample T-Test and CI: New York, Florida** |  |  |  |  |
| Method |  |  |  |  |
| μ₁: mean of New York |  |  |  |  |
| µ₂: mean of Florida |  |  |  |  |
| Difference: μ₁ - µ₂ |  |  |  |  |
| *Equal variances are not assumed for this analysis.* |  |  |  |  |
| Descriptive Statistics |  |  |  |  |
| Sample | N | Mean | StDev | SE Mean |
| New York | 20 | 8 | 2.2 | 0.49 |
| Florida | 20 | 5.55 | 2.14 | 0.48 |
| Estimation for Difference |  |  |  |  |
| Difference | 95% CI for |  |  |  |
| Difference |  |  |  |
| 2.45 | (1.060, 3.840) |  |  |  |
| Test |  |  |  |  |
| Null hypothesis | H₀: μ₁ - µ₂ = 0 |  |  |  |
| Alternative hypothesis | H₁: μ₁ - µ₂ ≠ 0 |  |  |  |
| T-Value | DF | P-Value |  |  |
| 3.57 | 37 | 0.001 |  |  |

Figure 9

# Secondary Analysis Conclusion 1

Based on the data provided above and the fact that the p-value of 0.001 is less than 0.05, we are required to reject the null hypothesis. This verifies the possibility of the scores in Florida being lower than the scores in New York.

# Analysis 2: New York/North Carolina

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Two-Sample T-Test and CI: New York, North Carolina** |  |  |  |  |
| Method |  |  |  |  |
| μ₁: mean of New York |  |  |  |  |
| µ₂: mean of North Carolina |  |  |  |  |
| Difference: μ₁ - µ₂ |  |  |  |  |
| *Equal variances are not assumed for this analysis.* |  |  |  |  |
| Descriptive Statistics |  |  |  |  |
| Sample | N | Mean | StDev | SE Mean |
| New York | 20 | 8 | 2.2 | 0.49 |
| North Carolina | 20 | 7.05 | 2.84 | 0.63 |
| Estimation for Difference |  |  |  |  |
| Difference | 95% CI for |  |  |  |
|  | Difference |  |  |  |
| 0.95 | (-0.680, 2.580) |  |  |  |
| Test |  |  |  |  |
| Null hypothesis | H₀: μ₁ - µ₂ = 0 |  |  |  |
| Alternative hypothesis | H₁: μ₁ - µ₂ ≠ 0 |  |  |  |
| T-Value | DF | P-Value |  |  |
| 1.18 | 35 | 0.245 |  |  |

# Secondary Analysis Conclusion 2

Based on the data provided above and the fact that the p-value of 0.245 is greater than 0.05, we fail to reject the null hypothesis. This means there is no statistically significant difference between the means for these 2 groups.