**Data Analytics**

**Immersive Technologies**

**Pair Project**

**John Jordan D00252073**

**Farad Raissa D00252027**

# **Q1: Determine whether the data provided is appropriate for the test(s) available and that any analysis is achievable. GSR**

A group of 50 males and 50 females due for dental surgery were measured for Anxiety (GSR/Galvanic Skin Response) and Perceived pain (VNRS/verbal numerical Rating Scale), both on a scale of 0 to 10 respectively. We note that there is random sampling used. We have a control group (no meditation) and meditation (treatment) group. This Means that there is an experiment, and that we can move forward to set up a Null Hypothesis.

We will look at the data made available to determine whether the results are appropriate to create a test and consider if any analysis is possible. We will look at GSR first then do the same procedures for VNRS.

* We note that there is random sampling used.
* We have a control group (no meditation) and meditation (treatment) group.
* Both groups are independent from each other.
* The sample size is n > 10 so our Normality check is met.

**Graph outline**

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

**Conclusion from first 2 graphs:**

It could be argued that the 1st graph is normally distributed, but the 2nd graph is not.

|  |  |
| --- | --- |
|  |  |

**Box Plots of male GSR control and meditation.**

**Chart, box and whisker chart

Description automatically generated**

**Source code in R Studio to extract value from CSV file:**

We have split the participants by gender using a split method. We also applied a filter to the boxplot command to extract male, control and meditation.

Text

Description automatically generated

**Q2. Formulate a hypothesis test**

**Null hypothesis:** There is no difference in the average(Mean) of the two methods (Meditation & Control), which indicate the Mean difference between the two methods is equal to zero or there is no difference in Means.

**Alternative Hypothesis:** There is a difference in average for the two methods (Meditation & Control), therefore the Mean difference is not equal to zero or the two Means are not equal.

Let’s write our Hypothesis as follow:

H0: μMeditationGSR = μControlGSR

HA: μMeditationGSR ≠μControlGSR

**Q3. Analyse the data to provide the hypothesis testing conclusion.**

We separate our testing into Male and Female and focus on Male GSR difference in Mean. We consider what if any, differences there are between the Control and Meditation groups. We gather details on Mean (x̄), standard deviation (s) and the sample number(n).

|  |  |  |
| --- | --- | --- |
|  | **Male Meditation GSR** | **Male Control GSR** |
| **x̄ (Mean)** | 6.826 | 7.052 |
| **s(standard Deviation)** | 0.869 | 1.003 |
| **n (sample number)** | 23 | 23 |

We can now move forward to test our Hypothesis. We assume H0 is true until proven elsewise.

From the table above we can now calculate **Point Estimate (PE)** and **Standard Error (SE),** and

**Test Statistic (T).**

*Difference between the 2 Means of Meditation and Control for GSR*

**Point Estimate (PE) =** x̄1 - x̄2 = 6.826 -7.0521 = -0.226

*Standard Error is calculated from the standard deviation of the 2 groups.*

**Standard Error (SE) =** **≈** = 0.2767≈ 0.28

*This allows us to work out the Test Statistic. When we don’t know the SD of the population, we use a*

*T distribution;*

**Test Statistic (T) =** = = -0.8167

*We then get our degrees of freedom (df)*

**Degree of Freedom (*df*) =** min (*n1 -1, n2-1*) = min (*23 -1, 23-1*) = min (*22, 22*) = 22

*Using T and our (df) we calculate our P value; (the p value is a number describing how likely your data occurred by random chance, i.e. that H0 is true.)*

**P value:** 2 x [pt(q = -0.8167, df = 22)] = 2 x 0.211428 = 0.41833 ≈ 0.419

*We look up our critical value from T – Tables and RStudio.*

**Critical Value:** qt(p = 0.95, df = 22) = 1.71744 ≈ 1.72

**Margin of Error:** t\* x SE = 1.96 x 0.28 = 0.5488 ≈ 0.55

*We can then get our 95% confidence interval; a confidence interval is a range of values that we are reasonably certain our true value falls in.*

**95% Confidence Interval:** PE ± (t\* x SE)

= -0.226 ± (1.96 x 0.28)

= -0.226 ± 0.55

= (-0.78, 0.32)

**4. Provide descriptive statistics (graphs and tables) of the data (GSR).**

***Initial Conclusion:*** *Since t is in the critical region, we do not accept H0.*

***Final Conclusion:*** *We accept the claim that the average of people using meditation is not the same as people using control.*

Chart, histogram

Description automatically generated

# **Q1: Determine whether the data provided is appropriate for the test(s) available and that any analysis is achievable.**

Again we can use the logic as the previous sample;

* We note that there is random sampling used.
* We have a control group (no meditation) and meditation (treatment) group.
* Both groups are independent from each other.
* The sample size is n > 10 so our Normality check is met.

**Graph Outline**

Chart, box and whisker chart

Description automatically generated

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

**Conclusion from first 2 graphs:**

Again, it could be argued that the 1st graph is normally distributed, but the 2nd graph is not.

**Code in R**

Text

Description automatically generated

**Q2. Formulate a hypothesis test**

H0: μMeditationVNRS = μControlVNRS

HA: μMeditationVNRS ≠μControlVNRS

**Q3. Analyse the data to provide the hypothesis testing conclusion.**

|  |  |  |
| --- | --- | --- |
|  | **Male Meditation VNRS** | **Male Control VNRS** |
| **x̄ (Mean)** | 8.04 | 8.26 |
| **s(standard Deviation)** | 0.56 | 0.69 |
| **n (sample number)** | 23 | 23 |

**Point Estimate(PE) =** x̄1 - x̄2 = 8.04 – 8.26= -0.22

**Standard Error(SE) =** **≈** = 0.19

**Test Statistic(T) =** = = -1.16

**Degree of Freedom(*df*) =** min (*n1 -1, n2-1*) = min (*23 -1, 23-1*) = min (*22, 22*) = 22

**P value:** 2 x [pt(q = -1.16, df = 22)] = 2 x 0.13 = 0.26

**Critical Value:** qt(p = 0.95, df = 22) = 1.71744 ≈ 1.72

**Margin of Error:** t\* x SE = 1.96 x 0.19 = 0.3724 ≈ 0.37

**95% Confidence Interval:** PE ± (t\* x SE)

= -0.22 ± (1.96 x 0.19)

= -0.22 ± 0.37

= (-0.59, 0.15)

**4. Provide descriptive statistics (graphs and tables) of the data(VNRS).**

**Initial Conclusion:** Since t is in the critical region, we do not accept H0.

**Final Conclusion:** We accept the claim that the average of people using meditation is not the same as people using control.

Chart, histogram

Description automatically generated