6MMCS009W1 Usability Testing & Evaluation

Individual Coursework

Student: Pasindu Manodya (2018484)

Degree: BSc (Hons) in Computer Science

School of Computer Science & Engineering
College of Design, Creative and Digital Industries
University of Westminster

Date

Contents

1 Introduction

This is the extended document on the "Westminster Navigation Game" used by Westminster University. For this document the "Guides and Policies" section was chosen from the Westminster navigation game. Research questions and hypothesises were created in align with independent sample, paired sample, and Pearson correlation tests. The gathered data using the online questionnaire were used for this document when answering the questions.

2 Hypothesis testing

This section contains three subsections that present how data collected from experimental design aiming to test whether the formed hypotheses from the CW1 are valid or not in regard to the research questions is being organised and analysed statistically with the following methods:

- *independent-sample-test*
- *paired-sample t-test*
- *Pearson correlation r-test.*

2.1 Independent-sample-test

This section is in regard to the independent-sample-test of the document. In this section research questions, hypothesis testing, data organizing, and statistical results are discussed.

2.1.1 Hypothesis testing for Independent-sample-test

Research question	Is there a tendency which level 6 students are scoring more points than level 4 students in the Guides and Policies quiz room?
Hypothesis statement	H_o: There is no significant difference between the overall points scored by level 6 students compared to the overall points scored by level 4 students in the guides and policies quiz room.
	H_a: There is a significant difference between the overall points scored by level 6 students compared to the overall points scored by level 4 students in the guides and policies quiz room.
Mathematical expression of the Hypothesis	H_o: μLevelSixStudent – μLevelFourStudent = 0
	H_a: μLevelSixStudent ≠ μLevelFourStudent
Dependent variable	Overall points
Dependent variable data type	*Ratio*

Independent variable	level 6 or level 4 student
Independent variable type	*Nominal*
Experimental design	Between Group
Study method	independent-sample-test

2.1.2 Data organisation

RAW DATA

	A	B
	Level 4 Students	**Level 6 Students**
2	130	170
3	130	165
4	130	170
5	125	170
6	110	190
7	140	160
8	145	180
9	180	185
10	190	165
11	175	180
12	140	190
13	90	175
14	180	185
15	170	110
16	160	140
17	170	200
18	160	130
19	170	100
20	90	120

Above snapshot shows the data gathered for this experimental design to calculate the mean, median and mode.

RAW DATA FOR T TEST

1	Student_Level	Participants	Overall_Points	Level_Code
2	Level_	1	170	0
3	Level_6	2	165	0
4	Level_6	3	170	0
5	Level_6	4	170	0
6	Level_6	5	190	0
7	Level_6	6	160	0
8	Level_6	7	180	0
9	Level_6	8	185	0
10	Level_6	9	165	0
11	Level_6	10	180	0
12	Level_6	11	190	0
13	Level_6	12	175	0
14	Level_6	13	185	0
15	Level_6	14	110	0
16	Level_6	15	140	0
17	Level_6	16	200	0
18	Level_6	17	130	0
19	Level_6	18	100	0
20	Level_6	19	120	0
21	Level_4	20	130	1
22	Level_4	21	130	1
23	Level_4	22	130	1
24	Level_4	23	125	1
25	Level_4	24	110	1
26	Level_4	25	140	1
27	Level_4	26	145	1
28	Level_4	27	180	1
29	Level_4	28	190	1
30	Level_4	29	175	1
31	Level_4	30	140	1
32	Level_4	31	90	1
33	Level_4	32	180	1
34	Level_4	33	170	1
35	Level_4	34	160	1
36	Level_4	35	170	1
37	Level_4	36	160	1
38	Level_4	37	170	1

Above snapshot shows the data gathered for this experimental design for the independent sample t test.

SNAPSHOTS OF ORGANIZED DATA SHOWED IN PSPP

Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure	Role
1	Level_4	Numeric	8	0		None	None	8	Right	Scale	Input
2	Level_6	Numeric	8	0		None	None	8	Right	Scale	Input

Case	Level_4	Level_6
1	130	170
2	130	165
3	130	170
4	125	170
5	110	190
6	140	160
7	145	180
8	180	185
9	190	165
10	175	180
11	140	190
12	90	175
13	180	185
14	170	110
15	160	140
16	170	200
17	160	130
18	170	100
19	90	120

Above snapshots show the tables which are used to calculate the mean, median and mode values using pspp.

Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure	Role
1	Student_Lev	String	8	0		None	None	8	Right	Scale	Input
2	Participants	Numeric	8	0		None	None	8	Right	Scale	Input
3	Overall_Poir	Numeric	8	0		None	None	10	Right	Scale	Input
4	Level_Code	Numeric	8	0		None	None	8	Right	Scale	Input

Case	Student_Level	Participants	Overall_Points	Level_Code
1	Level_6	1	170	0
2	Level_6	2	165	0
3	Level_6	3	170	0
4	Level_6	4	170	0
5	Level_6	5	190	0
6	Level_6	6	160	0
7	Level_6	7	180	0
8	Level_6	8	185	0
9	Level_6	9	165	0
10	Level_6	10	180	0
11	Level_6	11	190	0
12	Level_6	12	175	0
13	Level_6	13	185	0
14	Level_6	14	110	0
15	Level_6	15	140	0
16	Level_6	16	200	0
17	Level_6	17	130	0
18	Level_6	18	100	0
19	Level_6	19	120	0
20	Level_4	20	130	1
21	Level_4	21	130	1
22	Level_4	22	130	1
23	Level_4	23	125	1
24	Level_4	24	110	1

Above snapshots show the variable and data view in the pspp which is used to calculate the t values.

2.1.3 Descriptive statistics and discussion of results

Statistics

|		Level_4	Level_6
N	Valid	19	19
	Missing	0	0
Mean		146.58	162.37
Median		145.00	170.00
Mode		.	170

Above snapshot shows the calculated values for the mean, median and mode variables which belongs to the between group independent sample t test. According to the above snapshot the mean of the overall points scored by the level 6 students is 162.37 and for the level 4 students is 146.58. When it comes to the median of the overall points, level 6 students' median is 170 and level 4 students' median is 145. Considering the above stated means, we can come to a conclusion that level 6 students have scored more points in overall than level 4 students who took part in this test.

Descriptive Statistics

|	N	Std Dev	Variance	Skewness	S.E. Skew	Range
Level_4	19	29.86	891.81	-.50	.52	100.00
Level_6	19	28.79	828.80	-.97	.52	100.00
Valid N (listwise)	19					
Missing N (listwise)	0					

Above snapshots shows the calculated values for standard deviation, variance, skewness, and range variables which belongs to the between group independent sample t test. Standard deviation is used to calculate the spread of the data from the mean. When it comes to level 4 standard deviation score it is shown as 29.86 and in level 6 standard deviation it is shown as 28.79. Since the level 4 students' standard deviation score is higher than level 6 students' standard deviation score, we can state that level 6 student overall points data are spread around the mean and level 4 students' overall points data are spread out away from the mean value. Using the variance variable, we can calculate the variability of the data. The variance of level 4 students' overall points is 891.81 and level 6 students' variance is 828.80. Considering the given variance scores, we can come to the conclusion that level 4 overall points have a higher data spread than level 6 overall points. The skewness scores of the level 4 overall points and level 6 overall points is stated as -0.50 and -0.97 respectively.

HISTOGRAM OF LEVEL 4 OVERALL POINTS
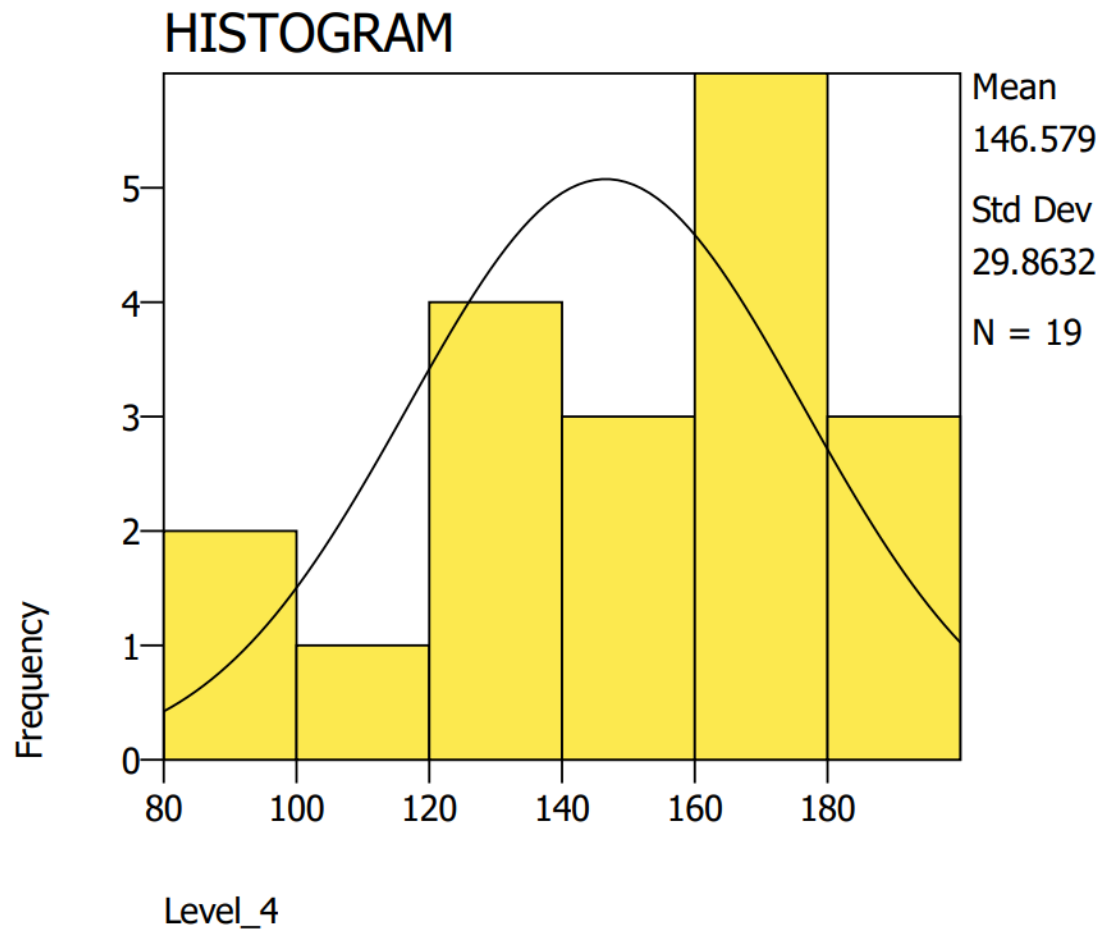
HISTOGRAM

This histogram shows the overall points scored by the level 4 students who participated in this study. According to the above table this diagram has a skewness of -0.50. Since the skewness is between -0.50 and +0.50 the distribution is approximately skewed. Since the distribution is approximately skewed this can be considered as a normal distribution.

HISTOGRAM OF LEVEL 6 OVERALL POINTS

HISTOGRAM

Mean 162.368
Std Dev 28.7889
N = 19
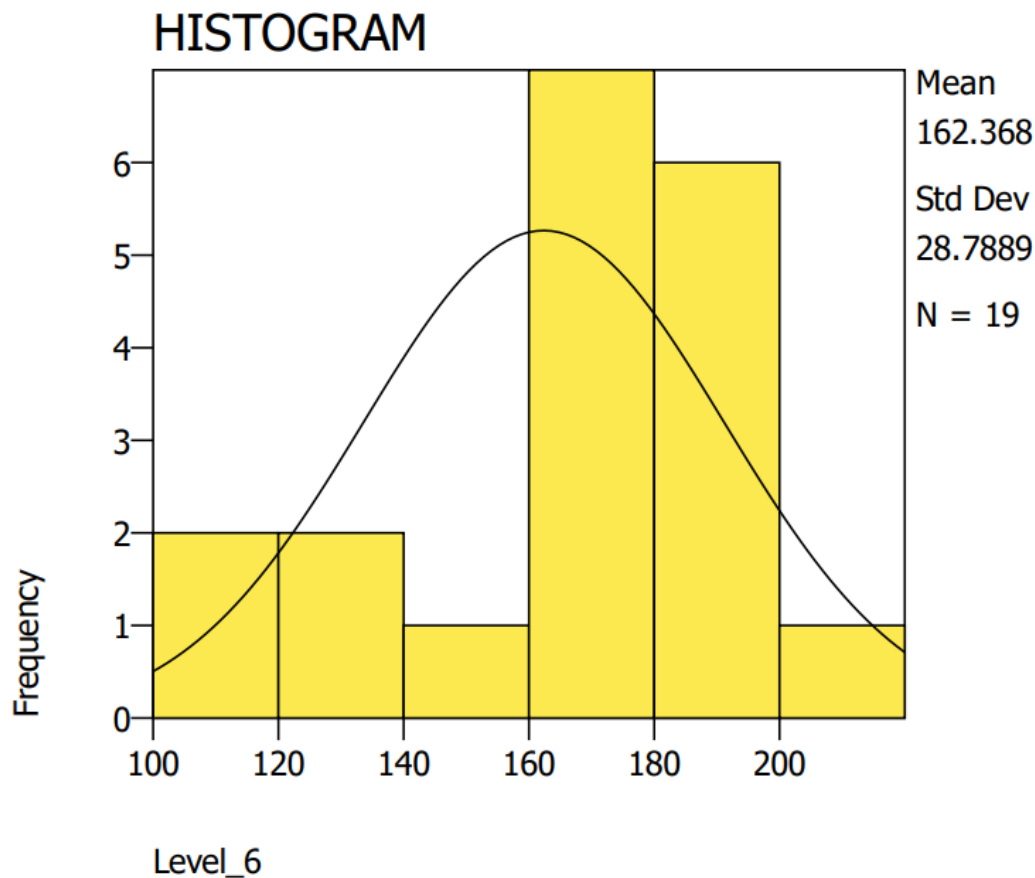
Level_6

This histogram shows the overall points scored by the level 6 students who participated in this study. According to the above table this diagram has a skewness of − 0.97. If the skewness is between +1 and 0.5 or -1 and -0.5 the skewness is moderately skewed. Since the skewness falls under this category this is a moderately skewed distribution. Since this is a moderately skewed this cannot be considered as a normal distribution.

What can you tell about the means you are studying?

Two variables were selected for this study. The means of the selected level 4 and level 6 are 146.5 and 162.3 respectively. Level 6 has a higher mean than level 4.

Can you proceed to parametric statistics to study the significance of the results?

If the data is normally distributed, we can proceed with a parametric test. Since the data is not normally distributed the answer is no. Cannot proceed with parametric tests.

If you cannot conduct parametric statistics, what would be the equivalent non-parametric test that you would have to conduct?

If not mann-whitney test has to be used to proceed.

2.1.4 Conducting independent sample t-test, interpret and report the results

Independent Samples Test

		Levene's Test for Equality of Variances		T-Test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Overall_Points	Equal variances assumed	.18	.671	1.66	36.00	.106	15.79	9.52	-3.51	35.09
	Equal variances not assumed			1.66	35.95	.106	15.79	9.52	-3.51	35.09

What is the p and t values you get?

P value = 0.10

T value = 1.66

Report the results:

- **Provide a short recap of method**
  Between group independent sample t test is done by comparing the two means which is level 4 overall points and level 6 overall points as show in the above snapshots. The aim of this test is to identify whether if there is significant difference between the means or not. In this test the level 4 overall points skewness is – 0.50 and it is between -0.50 and +0.50. Hence it is a normal distribution. Also, in the level 6 overall points the skewness is – 0.97 and it is between +1 and 0.5 or -1 and -0.5. Since it is between the designated values it is not a normal distribution. Since this cannot be a parametric test.

- **Descriptive recap with standard deviations**
  There are two standard deviations in this independent sample t test. Level 4 overall points and level 6 overall points and their standard deviations are 29.86 and 28.79 respectively. Since level 6 standard deviation is closer to 0 than the level 4 standard deviation it is safe to say that level 6 data points are spread around the mean and level 4 data points are spread out of the mean.

- **Significant test (with t-statistics and p-value)**
  We can calculate both t value and p values using this independent sample t-test. In this test calculated t-value is 1.66 and the calculated p-value is 0.10. When doing these tests, the confidence level was taken as 95%. The estimated p-value of 0.10 is larger than 0.05 for the degree of freedom 37. Since this result can be considered as an insignificant result.

- **Hypothesis test recap**
  After conducting this independent sample t-test the returned p-value is 0.10. Since 0.10 is larger than the selected 0.05 value, the significance value is lower than the p-value. Hence the author is able to come to the conclusion that the null hypotheses will be accepted.

2.1.5 Independent sample t-test files

[PSPP files and output files]

2.2   Paired-sample-test

This section is in regard to the paired-sample-test of the document. In this section research questions, hypothesis testing, data organizing, and statistical results are discussed which relates to the within group study.

2.2.1   Hypothesis testing for paired-sample-test

Research question	Is there an effect on the average time taken by students to complete the guides and policies quiz room?
Hypothesis statement	H_o: There is no significant difference between the time, students complete the guides quiz room in higher time and students complete the policies quiz in lesser time
	H_a: There is a significant difference between the time, students complete the guides quiz room in higher time and students complete the policies quiz in lesser time
Mathematical expression of the Hypothesis	H_o: μTimeTakenToCompleteTheGuidesQuiz – μTimeTakenToCompleteThePoliciesQuiz = 0
	H_a: μTimeTakenToCompleteTheGuidesQuiz ≠ μTimeTakenToCompleteThePoliciesQuiz
Dependent variable	average time taken by a student
Dependent variable data type	*Ratio*
Independent variable	guides and policies quiz room
Independent variable type	*Nominal*
Experimental design	Within Group
Study method	paired-sample-test

2.2.2 Data organisation

RAW DATA

	Guides_Time	Policies_Time
1	Guides_Time	Policies_Time
2	200	160
3	220	160
4	220	170
5	230	180
6	210	176
7	180	150
8	238	169
9	209	172
10	245	179
11	220	160
12	185	160
13	186	157
14	209	163
15	185	154
16	185	153
17	172	148
18	169	134
19	230	197
20	247	178
21	158	128
22	205	163
23	140	187
24	159	168
25	129	138
26	173	178
27	198	175
28	217	185
29	200	205
30	148	163
31	174	186
32	71	173
33	222	176
34	220	156
35	220	170
36	230	190

Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure	Role
1	Guides_Time	Numeric	8	0		None	None	8	Right	Scale	Input
2	Policies_Tim	Numeric	8	0		None	None	8	Right	Scale	Input

Case	Guides_Time	Policies_Time
1	200	160
2	220	160
3	220	170
4	230	180
5	210	176
6	180	150
7	238	169
8	209	172
9	245	179
10	220	160
11	185	160
12	186	157
13	209	163
14	185	154
15	185	153
16	172	148
17	169	134
18	230	197
19	247	178
20	158	128
21	205	163
22	140	187
23	159	168
24	129	138

Data View   Variable View

The above snapshots show the raw data which belongs to the within group test. And also, pspp variable view and data view snapshots. Two variables were taken as guides time and policies time because these are mentioned in the above hypotheses.

2.2.3 Descriptive statistics and discussion of results

Statistics

|		Guides_Time	Policies_Time
N	Valid	38	38
	Missing	1	1
Mean		195.11	167.66
Median		202.50	169.50
Mode		220	160
Std Dev		35.75	16.41
Variance		1277.93	269.31
Skewness		-1.26	-.22
Range		176.00	77.00

Above snapshot shows the calculated values for the mean, median and mode variables which belongs to the within group paired sample test. According to the above snapshot the mean of the guides time is 195.11 and for the policies time is 167.66. When it comes to the median of the completion time, guides time median is 202.50 and policies time median is 169.50. Considering the above stated means, we can come to a conclusion that students have spent more time complete guides quiz game than the time they spent on policies quiz game.

Descriptive Statistics

|	N	Mean	Std Dev	Variance	Skewness	S.E. Skew	Range
Guides_Time	38	195.11	35.75	1277.93	-1.26	.38	176.00
Policies_Time	38	167.66	16.41	269.31	-.22	.38	77.00
Valid N (listwise)	38						
Missing N (listwise)	1						

Above snapshots shows the calculated values for standard deviation, variance, skewness, and range variables which belongs to the within group paired sample test. Standard deviation is used to calculate the spread of the data from the mean. When it comes to guides time standard deviation score is shown as 35.75 and in policies time standard deviation is shown as 16.41. Since the guides time standard deviation score is higher than policies time standard deviation score, we can state that policies time data are spread around the mean and guides time data are spread out away from the mean value. Using the variance variable, we can calculate the variability of the data. The variance of guides time is 1277.93 and policies time variance is 269.31. Considering the given variance scores, we can come to the conclusion that guides time have a higher data spread than policies time. The skewness scores of the guides time and policies time are stated as – 1.26 and – 0.22 respectively.

GUIDES TIME HISTOGRAM

HISTOGRAM

Guides_Time

Mean 195.105
Std Dev 35.7482
N = 38
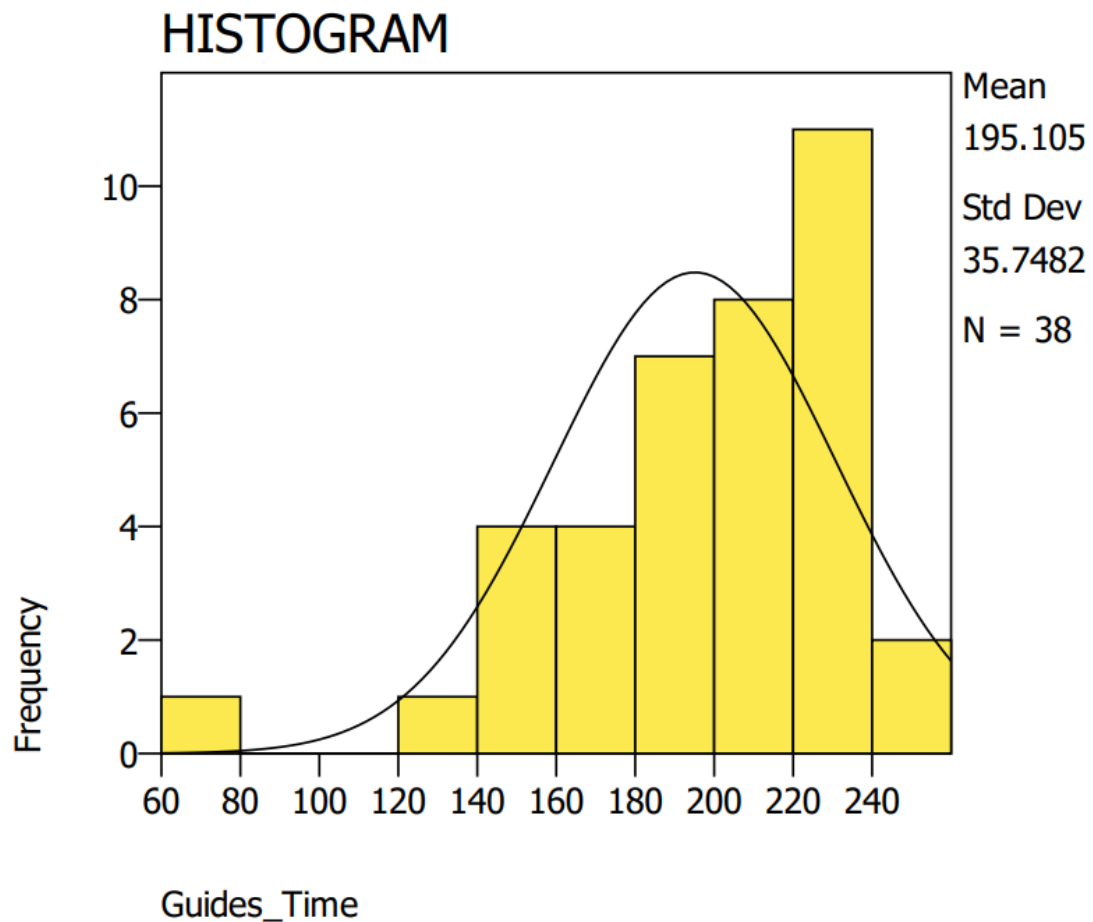
This histogram shows the overall time spent by students to complete the guides game. According to the above table this diagram has a skewness of − 1.26. If the skewness is higher tha +1 or lesser than -1 it is a highly skewed distribution. According to the above stated facts the author can state that this diagram does not presents a normal distribution because this is a highly skewed distribution and parametric test cannot be used.

POLICIES TIME HISTOGRAM

HISTOGRAM
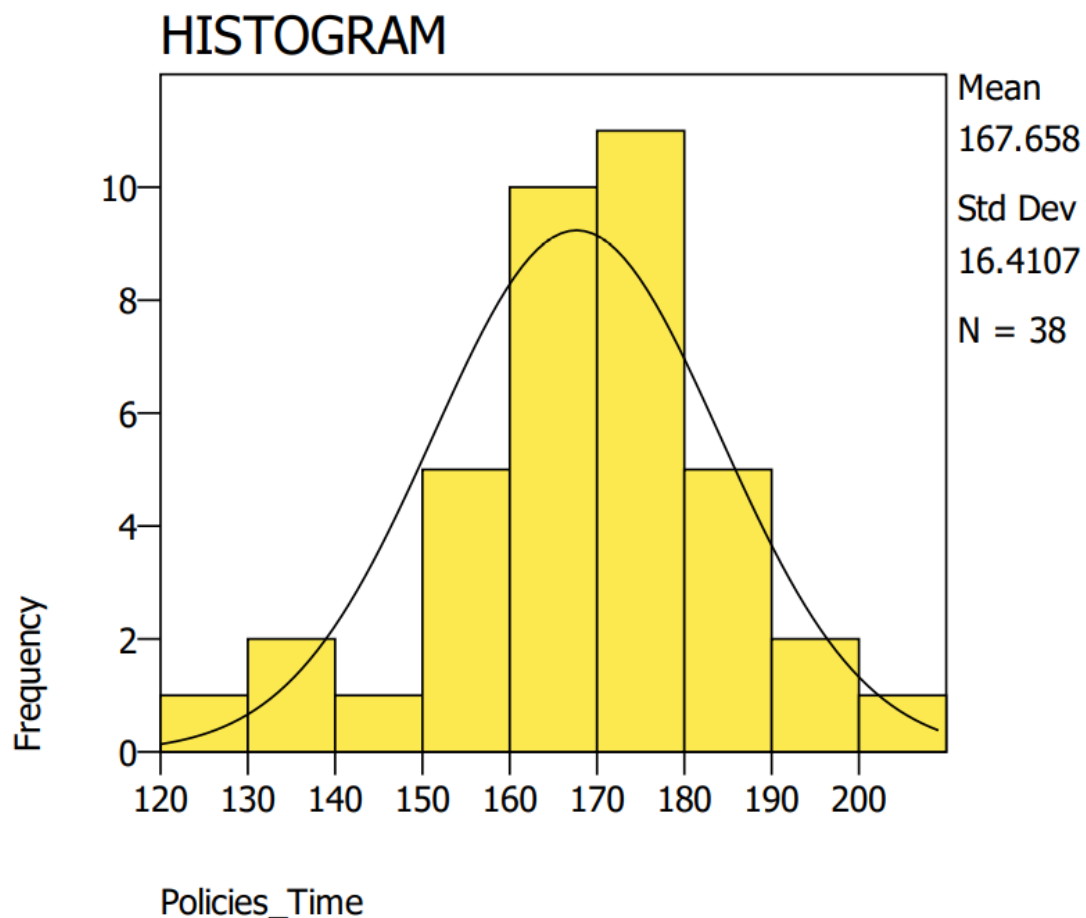
Policies_Time

This histogram shows the overall time spent by students to complete the policies game. According to the above table this diagram has a skewness of − 0.22. If the skewness is between +0.5 and -0.5 it is an approximately skewed distribution. Since this fall under this category we can consider a diagram presents a normal distribution.

What can you tell about the means you are studying?

Two variables were selected for this study. The means of the selected guides time and policies time are 195.1 and 167.6 respectively. Guides time has a higher mean than policies time.

Can you proceed to parametric statistics to study the significance of the results?

If the data is normally distributed, we can proceed with a parametric test. Since the data is not normally distributed the answer is no. Cannot proceed with parametric tests.

If you cannot conduct parametric statistics, what would be the equivalent non-parametric test that you would have to conduct?

Wilcoxon test is needed.

2.2.4 Conducting paired sample t-test, interpret and report the results

Paired Sample Statistics

		N	Mean	Std. Deviation	S.E. Mean
Pair 1	Guides_Time	38	195.11	35.75	5.80
	Policies_Time	38	167.66	16.41	2.66

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Guides_Time & Policies_Time	38	.341	.036

Paired Samples Test

| | | Paired Differences | | | | | | | |
| | | | | | 95% Confidence Interval of the Difference | | | | |
		Mean	Std. Deviation	S.E. Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Guides_Time - Policies_Time	27.45	33.87	5.50	16.31	38.58	4.99	37	.000

What is the p and t values you get?

P value = 0.00001

T value = 4.99

Report the results:

- **Provide a short recap of method**
  Within group paired sample test is done by comparing the two means which is guides game completion time and policies game completion time as show in the above snapshots. The aim of this test is to identify whether if there is significant difference between the means or not. In this test the guides time skewness is – 1.26 and it is not a normal distribution since it is is a highly skewed distribution. But in the policies time the skewness is – 0.22 and it falls under approximately skewed and it ia normal distribution. Since parametric test cannot be used.

- **Descriptive recap with standard deviations**
  There are two standard deviations in this paired sample test. Guides time and policies time and their standard deviations are 35.75 and 16.41 respectively. Since policies time standard deviation is closer to 0 than the guides time standard deviation it is safe to say that policies time data points are spread around the mean and guides time data points are spread out of the mean.

- **Significant test (with t-statistics and p-value)**
  We can calculate both t value and p values using this paired sample test. In this test calculated t-value is 4.99 and the calculated p-value is 0.00001. When doing these tests, the confidence level was taken as 95%. The estimated p-value of 0.00001 is lower than 0.05 for the degree of freedom 37. Since this result can be considered as a significant result.

- ***Hypothesis test recap***
  After conducting this paired sample test the returned p-value is 0.00001. Since 0.00001 is lower than the selected 0.05 value the significance value is higher than the p-value. Hence the author is able to come to the conclusion that the null hypotheses will be rejected.

In case the data is not normally distributed is this analysis reliable?

2.2.5 Paired sample t-test files

[PSPP files and output files](#)

2.3   Correlation Analysis

This section is in regard to the pearson-correlation-test of the document. In this section research questions, hypothesis testing, data organizing, and statistical results are discussed.

2.3.1   Hypothesis testing for Correlation Coefficient

Research question	Is there a leniency towards the age and, guides and policies quiz satisfaction of students who completed the guides and policies quiz room?
Hypothesis statement	H_o: There is no relationship between age and the students' satisfaction on policies and guides quiz room
	H_a: There is a relationship between age and the students' satisfaction on policies and guides quiz room
Variables studied	Independent variable - students' age Dependent variable – satisfaction on the quiz
Variables studied data type	Independent variable - Ratio Dependent variable - Ordinal
Study method	Pearson correlation r-test

2.3.2 Data organisation

RAW DATA

	A	B
2	20	2
3	23	5
4	22	3
5	24	5
6	23	5
7	23	4
8	19	2
9	22	5
10	20	1
11	23	5
12	25	5
13	23	4
14	24	4
15	19	1
16	22	4
17	25	5
18	23	4
19	20	2
20	19	3
21	24	5
22	20	5
23	19	4
24	23	3
25	20	5
26	20	4
27	23	5
28	20	3
29	19	3
30	24	2
31	20	5
32	21	5
33	20	4
34	22	4
35	21	4
36	24	5
37	19	4
	20	4

PSPP DATA

Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure	Role
1	Age	Numeric	8	0		None	None	8	Right	Scale	Input
2	Satisfaction	Numeric	8	0		None	None	8	Right	Scale	Input

| Case	Age	Satisfaction
1	20	2
2	23	5
3	22	3
4	24	5
5	23	5
6	23	4
7	19	2
8	22	5
9	20	1
10	23	5
11	25	5
12	23	4
13	24	4
14	19	1
15	22	4
16	25	5
17	23	4
18	20	2
19	19	3
20	24	5
21	20	5
22	19	4
23	23	3
24	20	5

The gathered data using the questionnaire were put into a excel for first and then they were migrated to the pspp. After that they were arranged according to the pearson correlation r test.

### 2.3.3 Descriptive statistics and discussion of results

Descriptive Statistics

	N	Mean	Std Dev
Age	38	21.61	1.92
Satisfaction	38	3.89	1.20
Valid N (listwise)	39		
Missing N (listwise)	1		

Above snapshot shows the mean values and the standard deviation values which relates to the pearson correlation r test. Mean of the age of the participants who took part in this study is calculated as 21.61 and mean of their satisfaction is calculated as 3.89. And the standard deviation is calculated as 1.92 and 1.20 respectively for age and satisfaction.
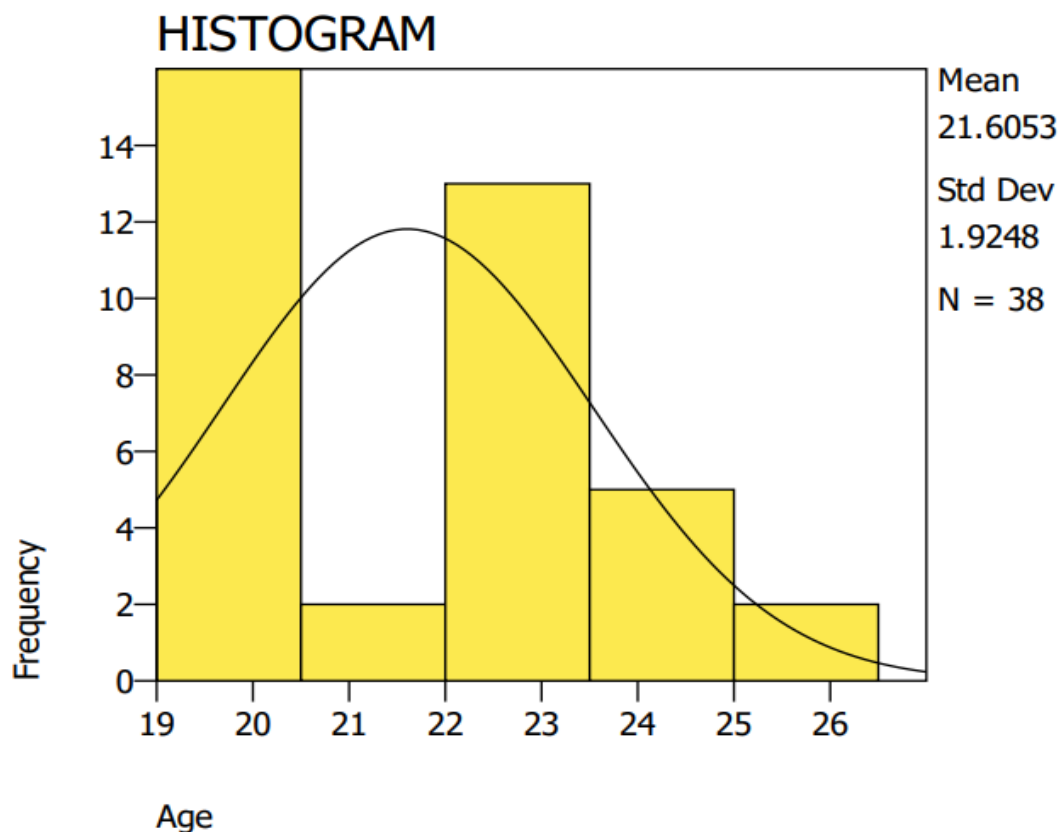
This histogram shows the age of the participants who took part in this study. This diagram has a skewness level of 0.09. If the skewness is between 0.5 and -0.5 it is a approximately skewed. Since we can consider a diagram presents a normal distribution.

HISTOGRAM

Mean
3.89474

Std Dev
1.20336

N = 38
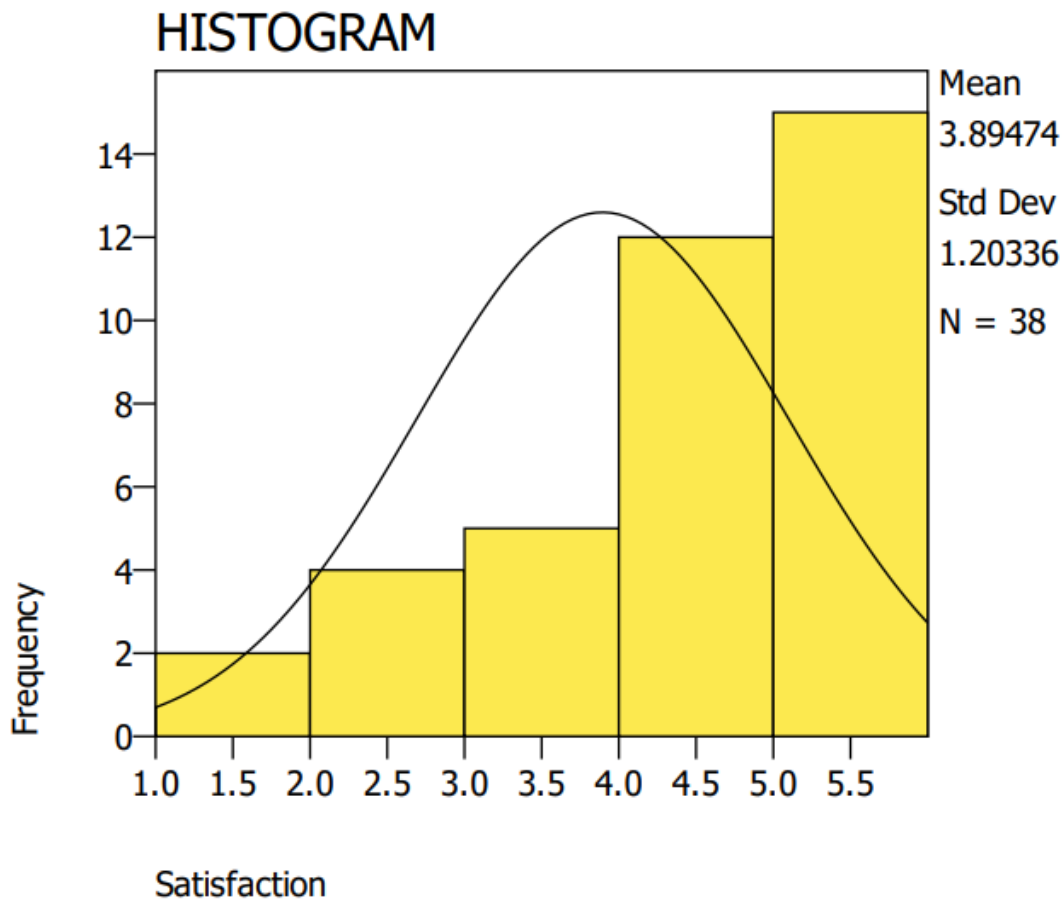
Satisfaction

This histogram shows the satisfaction level of the participants who took part in this study. This diagram has a skewness level of – 0.97. If the skewness is between +1 or 0.5 and -1 or -0.5 moderately skewed distribution, since we can consider a diagram does not present a normal distribution.

What can you tell about the means you are studying?

Two variables were selected for this study. The means of the selected guides time and policies time are 195.1 and 167.6 respectively. Guides time has a higher mean than policies time.

Can you proceed to parametric statistics to study the correlation coefficient?

If the data is normally distributed, we can proceed with a parametric test. Since the data is normally distributed the answer is yes.

If you cannot conduct parametric statistics, what would be the equivalent non-parametric test that you would have to conduct for correlation coefficient?

the Spearman correlation is needed.

2.3.4 Conducting Correlation Coefficient test, interpret and report the results

Correlations

		Age	Satisfaction
Age	Pearson Correlation	1.000	.472
	Sig. (2-tailed)		.003
	N	38	38
Satisfaction	Pearson Correlation	.472	1.000
	Sig. (2-tailed)	.003	
	N	38	38

What is the p and r values you get.

P = 0.003

R = 0.472

Report the results:

- **Provide a short recap of method**
  Pearson correlation r test is used to measure the statistical relationship between two selected variables. In this test those two variables are Age and Satisfaction. In this test the age skewness is 0.09 and it is between +0.5 and -0.5 it is a moderately skewed distribution.Also, in the satisfaction the skewness is – 0.97 and it is between +1 or 0.5 and -1 or -0.5 is it moderately skewed distribution.

- **Descriptive recap with standard deviations**
  There are two standard deviations in this pearson correlation r test. Age and satisfaction and their standard deviations are 1.92 and 1.20 respectively. Since satisfaction standard deviation is closer to 0 than the age standard deviation it is safe to say that satisfaction data points are spread around the mean and age data points are spread out of the mean.

- **Significant test (with t-statistics and p-value)**
  We can calculate both the r value and p values using this paired sample test. In this test calculated r-value is 0.472 and the calculated p-value is 0.003. When doing these tests, the confidence level was taken as 95%. The estimated p-value of 0.003 is lower than 0.05 for the degree of freedom 37. Since this result can be considered as a significant result.

- **Hypothesis test recap**
  After conducting this pearson correlation r test the returned p-value is 0.003. Since 0.003 is lower than the selected 0.05 value the significance value is higher than the p-value. Hence the author is able to come to the conclusion that the null hypothesis will be rejected.

2.3.5   Correlation Coefficient test files

[PSPP files and output files](#)

3 Conclusions

This is the extended document on the "Westminster Navigation Game" used by Westminster University. For this document the "Guides and Policies" section was chosen from the Westminster navigation game. Testing the hypotheses which were created previously was done in this document. Independent sample t test, paired sample t test and pearson correlation r test were used when investigating the hypothesis. Using those tests p values, t values, and r values were identified according to the appropriate tests, and they were explained.