Group 2

Project Number 6

Automated Classroom Performance Evaluation System

Group Members

Name	Roll Number
Kartikay Goel	180101033
Alay Chirag Shah	180101005
Rahul Choudhary	180101060
Falak Chhikara	180101023

Table of Contents:

Research Question:

The testable question that we posed is as follows:
In order to ensure that we are taking into consideration as many factors that can influence the empirical data as possible, we considered three factors in this research issue.

RQ: How does the **aesthetic score** (on a scale of 1-10) depend on the "**number of objects**", "**layout of interface objects**" and "**no of questions per page**"?

Hypothesis from RQ:
Ho : the aesthetic score (on a scale of 1-10) does **not** depend on the "**number of objects**", "**layout of interface objects**" and "**no of questions per page**"
H1 : The aesthetic score (on a scale of 1-10) depend on the "**number of objects**", "**layout of interface objects**" and "**no of questions per page**"

Determination of Variables:

Independent Variables:

Three independent variables were used, each of which was calculated on a nominal scale. The following table lists the factors and their related levels:

Factors	Levels		
No. of objects	5	10	15
Layout of objects	1(Layout1)	2(Layout2)	
No. of questions per page	1	3	

Dependent Variable: The aesthetic rating (on a scale of 1-10) is the dependent variable.

Design of Experiment:

Participants : We have used 12 participant for our study
Task: Each participant was shown 12 interfaces one by one and was asked to rate them on a scale of 1 to 10.
According to the above mentioned factors and levels, we created 12 (3x2x2) interfaces to use in testing. Since we created the test conditions ourselves, the influence of factors not accounted for is minimal. **The interfaces are present in this link:**
https://drive.google.com/drive/folders/1f_Fa5ELEdJoSyjHCVEQ0FDBcxNsbn54T?usp=sharing

The interfaces are as follows:

Interface #1: <N=5,L=1,Q=1>
Interface #2: <N=10,L=1,Q=1>
Interface #3: <N=15,L=1,Q=1>
Interface #4: <N=5,L=1,Q=3>
Interface #5: <N=10,L=1,Q=3>
Interface #6: <N=15,L=1,Q=3>
Interface #7: <N=5,L=2,Q=1>
Interface #8: <N=10,L=2,Q=1>
Interface #9: <N=15,L=2,Q=1>
Interface #10: <N=5,L=2,Q=3>
Interface #11: <N=10,L=2,Q=3>
Interface #12: <N=15,L=2,Q=1>

We have used “within-subject” study design
We have used Latin Square method for taking care for of practice effect

P#1	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12
P#2	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R1
P#3	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R1	R2
P#4	R4	R5	R6	R7	R8	R9	R10	R11	R12	R1	R2	R3
P#5	R5	R6	R7	R8	R9	R10	R11	R12	R1	R2	R3	R4

P#6	R6	R7	R8	R9	R10	R11	R12	R1	R2	R3	R4	R5
P#7	R7	R8	R9	R10	R11	R12	R1	R2	R3	R4	R5	R6
P#8	R8	R9	R10	R11	R12	R1	R2	R3	R4	R5	R6	R7
P#9	R9	R10	R11	R12	R1	R2	R3	R4	R5	R6	R7	R8
P#10	R10	R11	R12	R1	R2	R3	R4	R5	R6	R7	R8	R9
P#11	R11	R12	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
P#12	R12	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11

Data Obtained

The table below shows the results of each participant's ranking for each interface. The participant names are listed in the leftmost column, and the Interface numbers are listed in the topmost row.The ratings provided by the corresponding user to the corresponding interface are listed in the other cells. The second row contains additional information about the corresponding gui.

The first value represents the number of objects, the second is for layout (1 for horizontal, 2 for vertical) and the third represents the number of questions.

NAMING CONVENTION

T_A is interface 1, T_B is interface 2,.....,

.

T_L is interface 12

Aditya Kandekar	6	7	5	4	7	8	5	5	9	7	7	9
Anant Bansal	8	10	5	8	8	8	9	9	9	9	10	7
Himanshu	6	3	7	8	6	8	9	7	9	10	7	9
Nayan Tiwari	3	7	8	7	10	9	8	10	10	7	10	8
Rahul Mala	9	7	8	9	8	9	10	10	7	9	6	7
Rishikesh Songra	9	4	7	7	8	10	10	4	6	6	4	8
Aman	6	6	10	7	6	9	7	6	4	3	7	8

Kumar												
Pranav Gupta	7	10	6	7	8	6	9	8	9	7	8	4
Ritik Mandloi	9	5	9	10	7	8	6	5	8	8	6	8
Jatin Dhngra	7	9	10	6	7	6	4	8	9	6	8	9
Sparsh Batra	10	10	5	6	6	4	9	10	5	8	8	7
Eshwar Nukala	9	7	8	8	6	8	9	6	7	8	6	8

Data Analysis:

The data collected by us does not follow the normal distribution so we cannot apply the parametric test (t -test). So we can only apply some non-parametric test to our data and since we have used the within subject method for the experiment and the factor (number of elements) has 3 levels this makes the data eligible for the Friedman Test for repeated measures. While analysing the data using a Friedman test, part of the process involves checking to make sure that the data we want to analyse can actually be analysed using a Friedman test. We confirm the following 4 assumptions before doing the Friedman's test.
Assumption 1: One factor that is measured on three or more different occasions.
Assumption 2: Group is a random sample from the population.
Assumption 3: Your dependent variable should be measured at the ordinal or continuous level. Examples of continuous variables include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth.
Assumption 4: Data samples do NOT need to be normally distributed.

We used an online calculator to perform the Friedman test. The calculator takes the mean ratios (interface-wise) as its input, and computes some relevant values to judge the significance of the collected data.

RESULTS:

Friedman rank sum test for multiple correlated samples in a two-way balanced complete block design
Friedman chi-squared statistic: 20.681428
degrees of freedom df : 11
This is the number of correlated samples (or groups), minus 1

p-value: 0.02153
This p-value is for rejection of the omnibus null hypothesis, that all samples (groups) are from the same distribution, i.e. there is no effect of groups. The alternate hypothesis that one or more of the correlated samples (groups) is different, i.e. there is a group effect.

The omnibus p-value is at or below the respectable critical threshold of 0.05, so post-hoc pairwise multiple comparison tests are conducted to discern which of the pairs have significantly differences. Three of many possible post-hoc tests are conducted: the methods of (1) Conover and (2) Nemenyi. For the (1) Conover method, the pvalue is adjusted in two ways, first according to the family-wise error rate (FWER) procedure of Holm, and next by the false discovery rate (FDR) procedure of Benjaminyi-Hochberg.

Post-hoc p-values of all possible pairs (of samples/ groups) are compactly represented as a lower triangular matrix. Each numerical entry is the p-value of row/column pair, i.e. the null hypothesis that the group represented by a particular column name is different from the group represented by a particular row name.

Conclusion:
$F(11)=20.681428$ ans $p<0.05$ and we know that it is significantly small and se we reject NULL Hypothesis.We also came to the conclusion that the greater the number of objects, the higher the aesthetic score.
Layout 2 is the preferred choice.
Less questions result in a more aesthetically pleasing gui.