Name: Tejal Yadav
Branch: Computer
Year: Third year
UID: 2019130071
Subject: Data Analytics
Date: 19th May 2022

Experiment 3: One Way ANOVA on SAS platform

Objective: To perform One Way ANOVA with SAS online platform.

What is One Way ANOVA?

One way ANOVA tests the differences among means of a single categorical variable on a single dependent variable

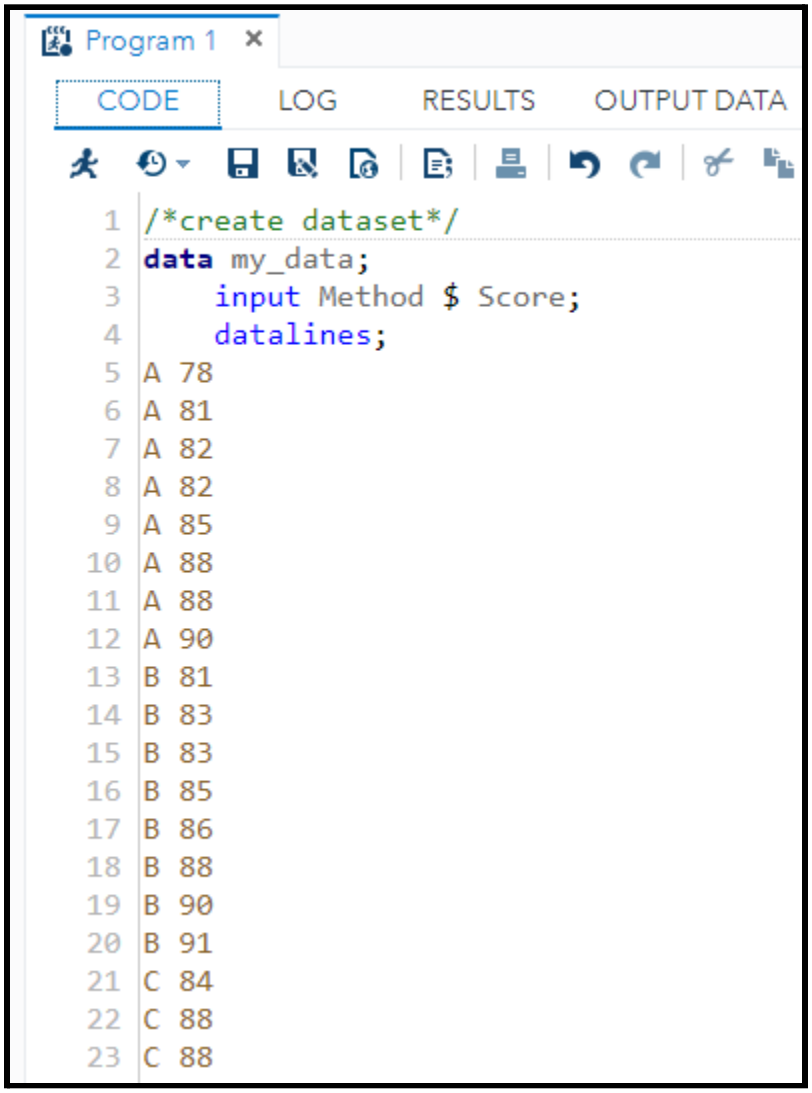
<u>Steps followed to perform One Way ANOVA in SAS</u>

Step 1: Create Data
30 students participated in a study. The students are randomly assigned to use one of three studying methods to prepare for an exam.

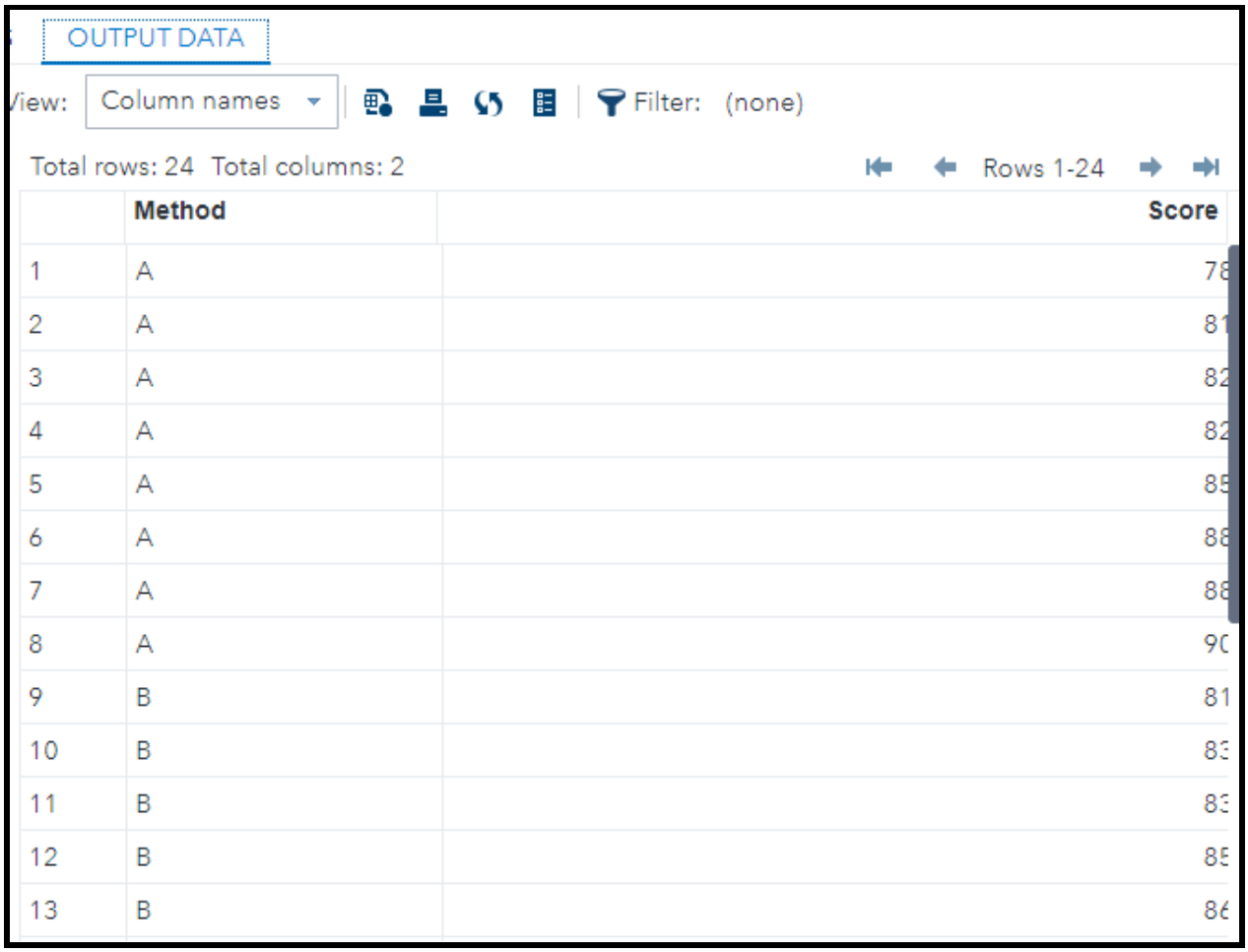
The exam results for each student are shown below:

Method A	Method B	Method C
78	81	84
81	83	88
82	83	88
82	85	89
85	86	90
88	88	93
88	90	95
90	91	98

Code to generate the data:



Dataset is generated in SAS as shown below:



OUTPUT DATA

View: Column names | Filter: (none)

Total rows: 24 Total columns: 2 | Rows 1-24

	Method	Score
1	A	78
2	A	81
3	A	82
4	A	82
5	A	85
6	A	88
7	A	88
8	A	90
9	B	81
10	B	83
11	B	83
12	B	85
13	B	86

Step 2: Perform the One-Way ANOVA

Code to perform One-Way ANOVA:

```
/*perform one-way ANOVA*/
proc ANOVA data=my_data;
class Method;
model Score = Method;
means Method / tukey cldiff;
run;
```

Step 3: Interpret the Results

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Method	3	A B C

Number of Observations Read	24
Number of Observations Used	24

The ANOVA Procedure

Dependent Variable: Score

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	175.5833333	87.7916667	5.26	0.0140
Error	21	350.2500000	16.6785714		
Corrected Total	23	525.8333333			

R-Square	Coeff Var	Root MSE	Score Mean
0.333914	4.698685	4.083941	86.91667

From this table we can see:

- The overall F Value: 5.26
- The corresponding p-value: 0.0140

A one-way ANOVA uses the following null and alternative hypotheses:

- H0: All group means are equal.
- HA: At least one group means different from the rest.

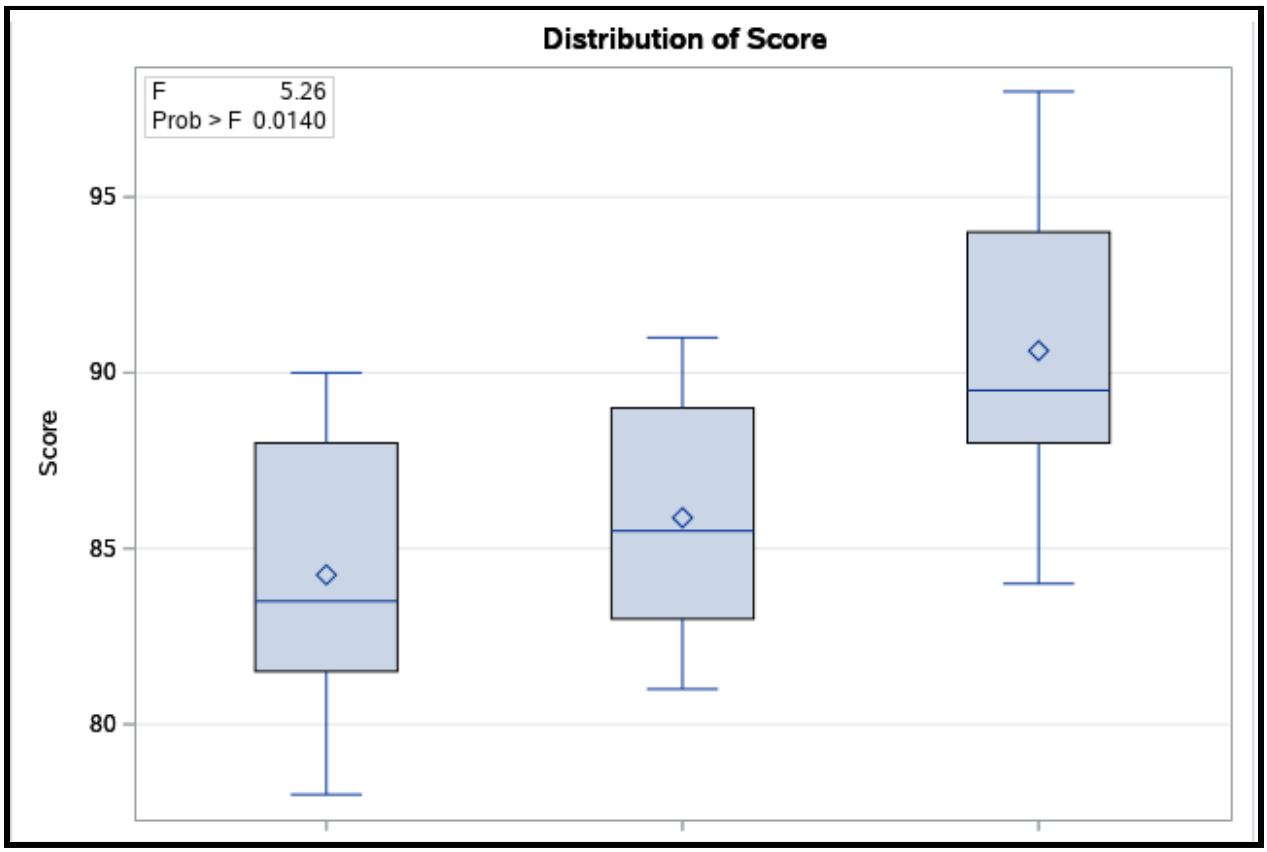
Since the p-value from the ANOVA table (0.0140) is less than $\alpha = .05$, we reject the null hypothesis.

This tells us that the mean exam score is not equal between the three studying methods.

Step 5: Visualizing Graphs and tables:

1. **Boxplot**

Boxplots are used to visualize the distribution of exam scores for each of the three studying methods:



From the boxplots we can see that the exam scores tend to be higher among students who used studying method C compared to methods B and C.

2. **Comparison Table**

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Score

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	21
Error Mean Square	16.67857
Critical Value of Studentized Range	3.56462
Minimum Significant Difference	5.1469

Comparisons significant at the 0.05 level are indicated by ***.

Method Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
C - B	4.750	-0.397	9.897	
C - A	6.375	1.228	11.522	***
B - C	-4.750	-9.897	0.397	
B - A	1.625	-3.522	6.772	
A - C	-6.375	-11.522	-1.228	***
A - B	-1.625	-6.772	3.522	

From the table we can see that the mean values for groups A and C are statistically significantly different.
We can also see the 95% confidence interval for the difference in mean exam scores between group A and C:

95% Confidence Interval for Difference in Means: [1.228, 11.522]

Conclusion:

1. A one-way ANOVA was performed to compare the effect of three different studying methods on exam scores.

2. A one-way ANOVA revealed that there was a statistically significant difference in mean exam score between at least two groups ($F(2, 21) = [5.26]$, $p = 0.014$).

3. The mean value of exam score was significantly different between method C and method A (95% C.I. = [1.228,11.522]).

4. There was no statistically significant difference in mean exam scores between method A and method B or between method B and method C.