|  |  |  |
| --- | --- | --- |
| Gender | Education Level | Monthly Income (in Thousands) |
| Female | Masters | 60 |
| Female | Bachelors | 50 |
| Female | SEE | 40 |
| Male | Plus2 | 45 |
| Female | Masters | 65 |
| Female | Bachelors | 45 |
| Female | Bachelors | 48 |
| Male | Bachelors | 51 |
| Male | Masters | 56 |
| Male | Plus2 | 35 |
| Male | Plus2 | 33 |
| Female | Plus2 | 33 |
| Female | Masters | 52 |
| Male | Masters | 55 |
| Male | SEE | 45 |
| Male | Bachelors | 46 |
| Female | SEE | 45 |
| Male | Bachelors | 35 |
| Female | Masters | 58 |
| Male | Masters | 57 |
| Female | Bachelors | 54 |
| Male | Masters | 47 |
| Female | Bachelors | 44 |
| Male | Masters | 49 |
| Female | Bachelors | 48 |
| Male | Bachelors | 46 |
| Male | Plus2 | 39 |
| Male | SEE | 38 |
| Male | Plus2 | 45 |
| Male | SEE | 46 |
| Female | Plus2 | 44 |
| Male | Plus2 | 32 |
| Male | Plus2 | 34 |
| Female | Masters | 60 |
| Male | Bachelors | 47 |
| Male | Bachelors | 38 |
| Female | Bachelors | 49 |
| Female | Masters | 47 |
| Female | Masters | 65 |

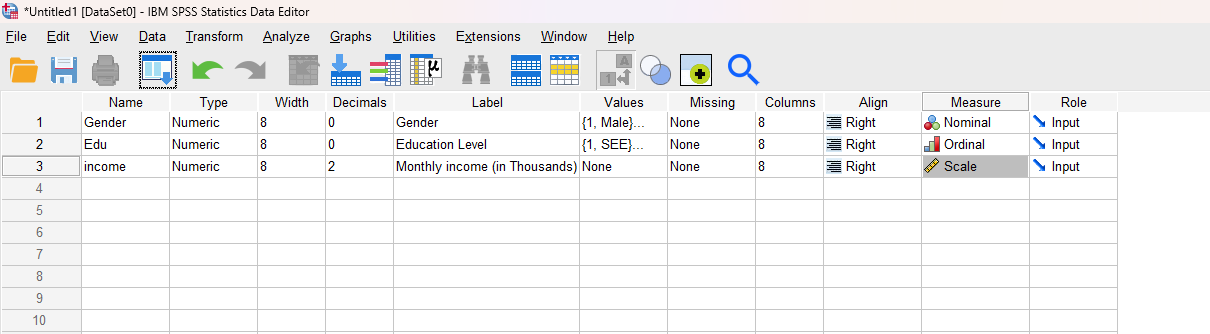
# Lab 1: Entering data in SPSS

## Enter the following data in SPSS

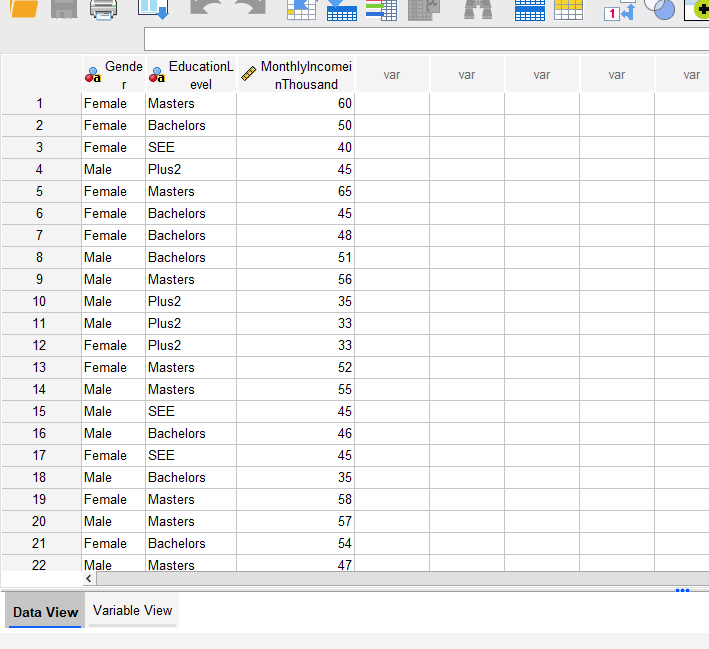
## SPSS Procedure

1. Open SPSS window
2. In the variable view create the following variables.

* Name:Gender, Type:Numeric, Width:8, Decimals:0, Label:Gender,values:{1:Male,2:Female}, measure:Nominal.
* Name:Edu, Type:Numeric, Width:8, Decimals:0, Label:Education Level,values:{1:SEE,2:plush2,3:Bachelors,5:Masters}, measure:Ordinal.
* Name:income, Type:Numeric, Width:8, Decimals:0, Label:Monthly income(in thousands), measure:Scale.



1. In the Data View after inserting the data as required:



# Lab 2: Data Visualization

# For the data of Lab 1

1. Represent the gender in a simple bar diagram.
2. Represent the educational level in a simple bar diagram.
3. Represent the educational level in a multiple bar diagram .
4. Represent the educational level in a pie chart.
5. Represent the income in a histogram with bin size 10 along with the normal curve.

## SPSS Procedure

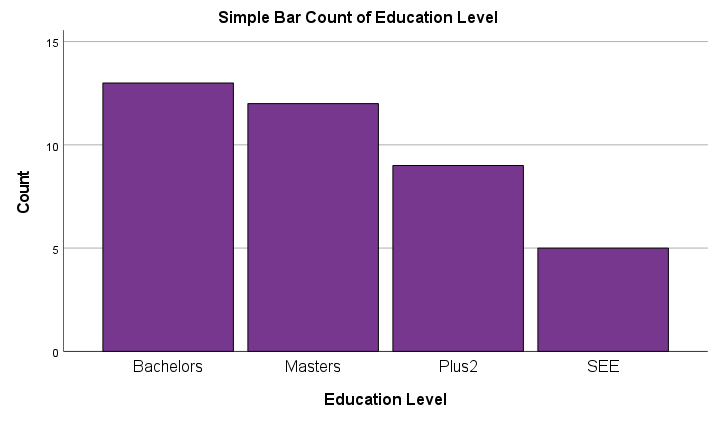
i.

1. In the Graphs menu choose Chart Builder
2. Choose Bar from Gallery
3. Double Click simple bar
4. Drag Gender to X-axis
5. Click Ok.



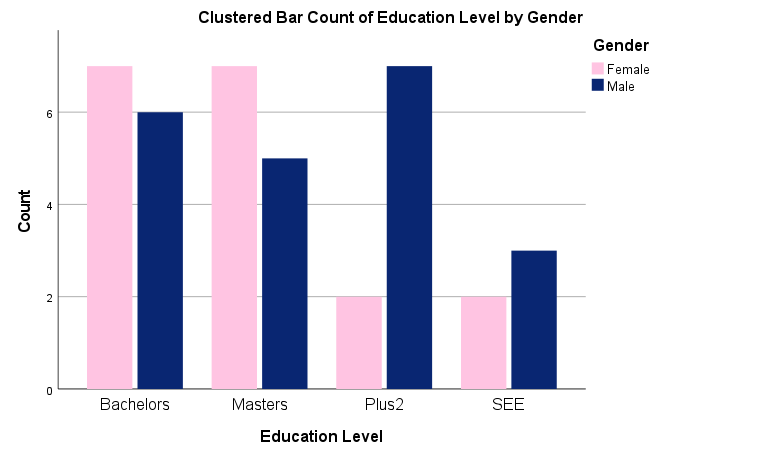
ii.

1. In the Graphs menu choose Chart Builder
2. Choose Bar from Gallery
3. Double Click simple bar
4. Drag Education Level to X-axis
5. Click Ok.



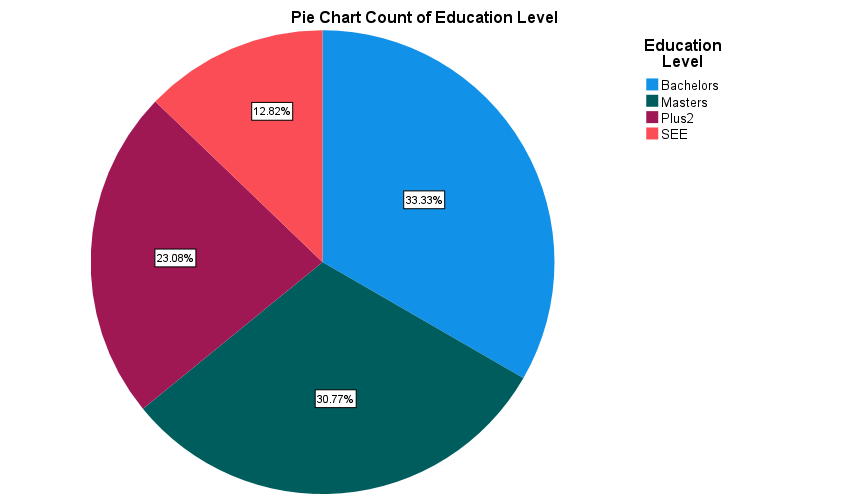
iii.

1. In the Graphs menu choose Chart Builder
2. Choose Bar from Gallery
3. Double Click simple bar
4. Drag Education Level to X-axis
5. Drag Gender to Cluster on X:set Color
6. Click Ok.



iv.

1. In the Graphs menu choose Chart Builder
2. Choose Pie/Polar from Gallery
3. Double Click Pie Chart
4. Drag Education Level to X-axis
5. Click Ok.



v.

1. In the Graphs menu choose Chart Builder
2. Choose Histogram from Gallery.
3. Double Click Simple Histogram.
4. Drag Monthly income(in Thousands)to X-axis
5. Click on set Parameters.
6. Custom the Bin size with interval width 10.
7. Click on Continue.
8. Click Ok.



# Lab 3: Descriptive Statistics

# For the data of Lab 1

1. Construct the frequency tables of gender and education level
2. For the monthly income find
3. Mean
4. Median
5. Mode
6. Standard deviation
7. Maximum
8. Minimum
9. Quartiles
10. 30th, 80th percentiles

## SPSS Procedure

i.

1. In the Analyze menu choose Descriptive Statistics
2. Choose Frequencies
3. On the dialog box send gender and education level to variables.
4. Check Display Frequency tables
5. Click OK

## In Output:

|  |  |  |  |
| --- | --- | --- | --- |
| **Statistics** | | | |
|  | | Gender | Education Level |
| N | Valid | 39 | 39 |
| Missing | 0 | 0 |

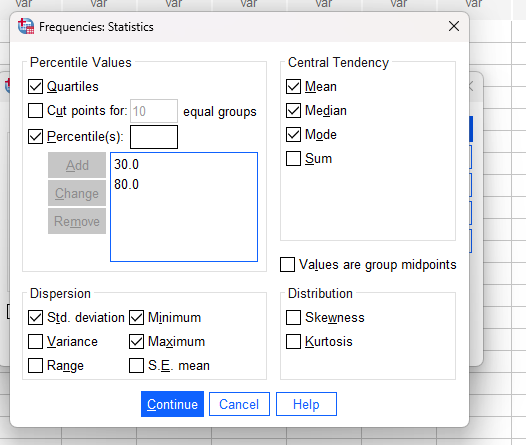
**Frequency Table**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gender** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | Female | 18 | 46.2 | 46.2 | 46.2 |
| Male | 21 | 53.8 | 53.8 | 100.0 |
| Total | 39 | 100.0 | 100.0 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Education Level** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | Bachelors | 13 | 33.3 | 33.3 | 33.3 |
| Masters | 12 | 30.8 | 30.8 | 64.1 |
| Plus2 | 9 | 23.1 | 23.1 | 87.2 |
| SEE | 5 | 12.8 | 12.8 | 100.0 |
| Total | 39 | 100.0 | 100.0 |  |

ii.

1. In the Analyze menu choose Descriptive Statistics
2. Choose Frequencies
3. On the dialog box Monthly income(in Thousands) to variables.
4. Uncheck Frequency tables.
5. Click on statistics
6. Choose required field.



1. Click on continue
2. Click on Ok.

## In Output:

|  |  |  |
| --- | --- | --- |
| **Statistics** | | |
| Monthly Income (in Thousand) | | |
| N | Valid | 39 |
| Missing | 0 |
| Mean | | 46.82 |
| Median | | 46.00 |
| Mode | | 45 |
| Std. Deviation | | 8.660 |
| Minimum | | 32 |
| Maximum | | 65 |
| Percentiles | 25 | 40.00 |
| 30 | 44.00 |
| 50 | 46.00 |
| 75 | 52.00 |
| 80 | 55.00 |

|  |
| --- |
|  |

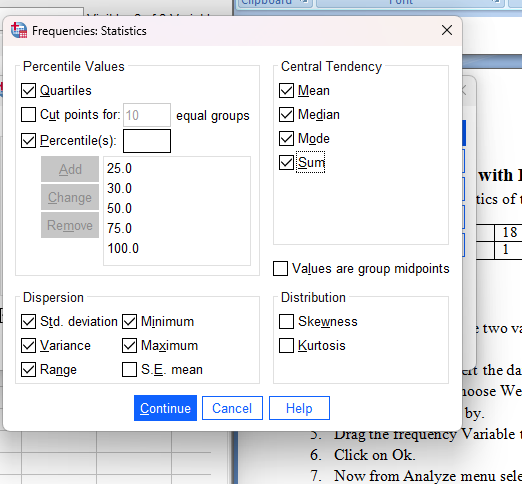
# Lab 4: Weighting the Data with Frequencies.

Find the various descriptive statistics of the following data.

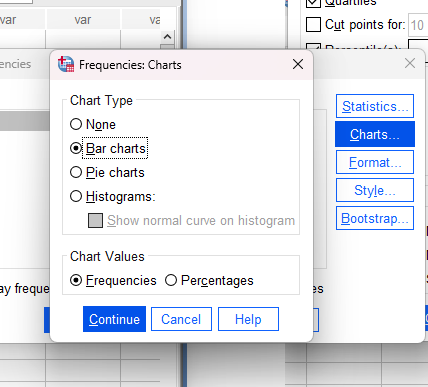
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | 10 | 12 | 15 | 16 | 18 |
| f | 3 | 6 | 9 | 2 | 1 |

## SPSS Procedure

1. In variable view create two variables with variable name ‘x’ and ‘f’ with Measure Scale.
2. In the Data view insert the data.
3. In the Data menu Choose Weight Cases.
4. Select Weight cases by.
5. Drag the frequency Variable to the Frequency Variable.
6. Click on Ok.
7. Now from Analyze menu select Descriptive Statistics.
8. Choose Frequencies.
9. From Frequencies Statistics Choose required values.



1. Click on continue.
2. From Frequencies: Charts Choose required chart type

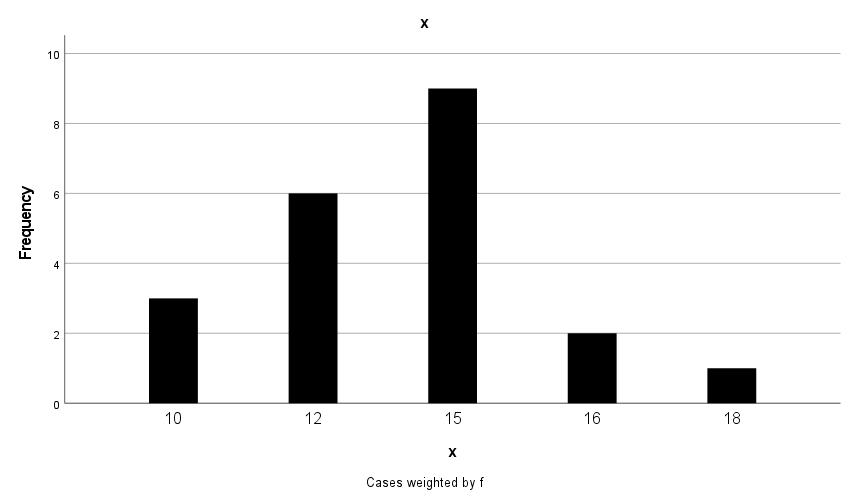


1. Click on Continue.
2. Click on OK.

## In Output:

|  |  |  |
| --- | --- | --- |
| **Statistics** | | |
| x | | |
| N | Valid | 21 |
| Missing | 0 |
| Mean | | 13.67 |
| Median | | 15.00 |
| Mode | | 15 |
| Std. Deviation | | 2.266 |
| Variance | | 5.133 |
| Range | | 8 |
| Minimum | | 10 |
| Maximum | | 18 |
| Sum | | 287 |
| Percentiles | 25 | 12.00 |
| 30 | 12.00 |
| 50 | 15.00 |
| 75 | 15.00 |
| 100 | 18.00 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **x** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 10 | 3 | 14.3 | 14.3 | 14.3 |
| 12 | 6 | 28.6 | 28.6 | 42.9 |
| 15 | 9 | 42.9 | 42.9 | 85.7 |
| 16 | 2 | 9.5 | 9.5 | 95.2 |
| 18 | 1 | 4.8 | 4.8 | 100.0 |
| Total | 21 | 100.0 | 100.0 |  |



# Lab 5: CORRELATION

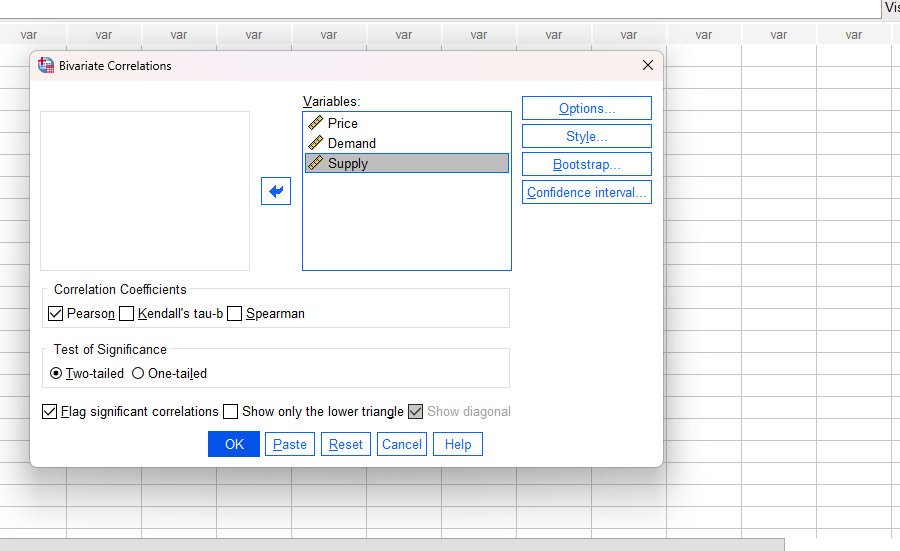
Find the pair wise correlation between the variables in the following data:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Price | 5 | 7 | 10 | 12 | 15 | 18 | 20 | 21 | 23 | 25 |
| Demand | 30 | 28 | 27 | 20 | 21 | 18 | 15 | 12 | 8 | 9 |
| Supply | 12 | 18 | 25 | 23 | 28 | 30 | 32 | 35 | 38 | 40 |

Also interpret the result.

## SPSS Procedure

1. In the variable view create three variable (Price, Demand, Supply) with Scale Measure and decimal 0.
2. Insert the data in Data View as per question.
3. From the menu bar choose Analyze and select correlate and choose Bivarient.
4. Select all the variables to the variables view as shown below:



1. Click on OK.

## In Output:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Correlations** | | | | |
|  | | Price | Demand | Supply |
| Price | Pearson Correlation | 1 | -.976\*\* | .979\*\* |
| Sig. (2-tailed) |  | .000 | .000 |
| N | 10 | 10 | 10 |
| Demand | Pearson Correlation | -.976\*\* | 1 | -.945\*\* |
| Sig. (2-tailed) | .000 |  | .000 |
| N | 10 | 10 | 10 |
| Supply | Pearson Correlation | .979\*\* | -.945\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 |  |
| N | 10 | 10 | 10 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | |

## Interpretation:

* Coefficient of Correlation between Price and Demand is -0.976. So, if the price increases, the demand decreases.
* Coefficient of Correlation between Price and Supply is 0.979.So, if the price increases the Supply will increases too.
* Coefficient of Correlation between Demand and Supply is -0.945. So, if the demand increases, the supply decreases.

# Lab 6: Regression

**Given the data :**

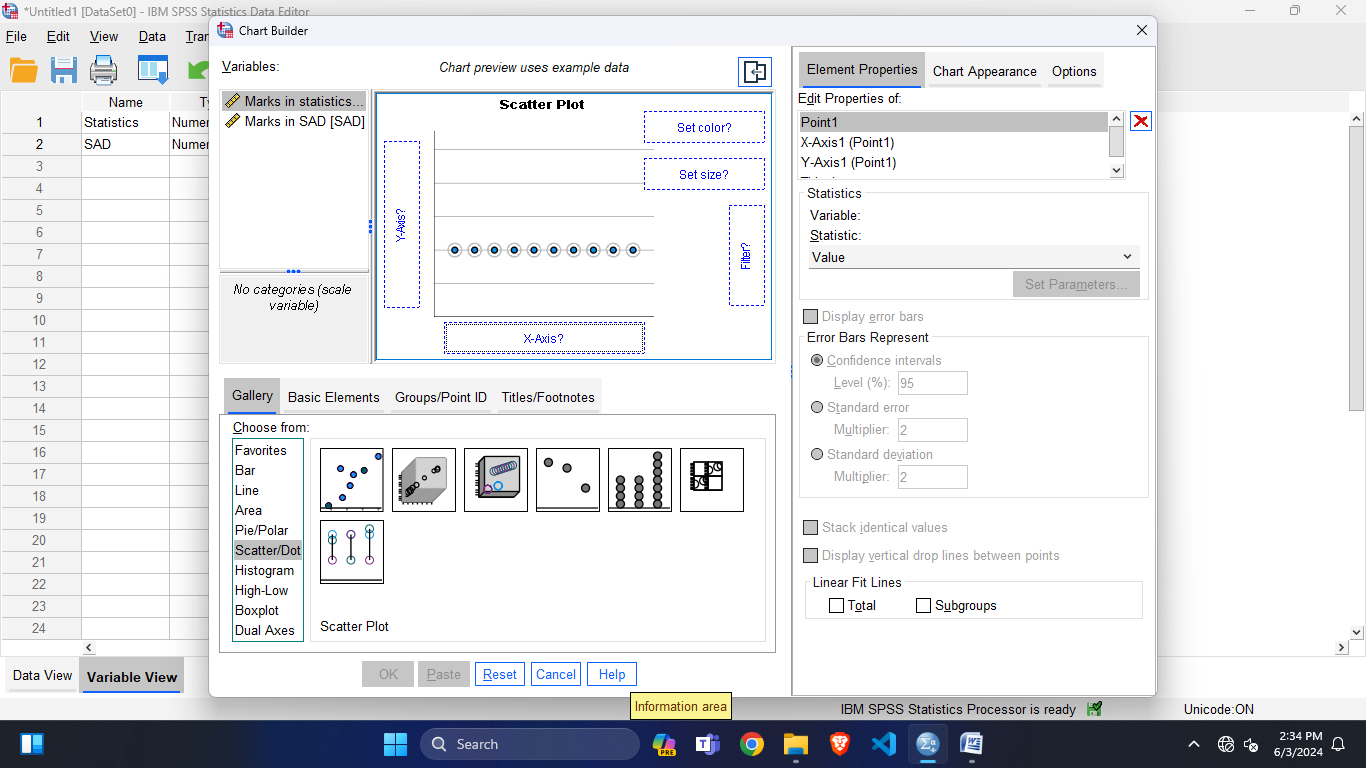
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Marks in statistics | 40 | 42 | 37 | 26 | 40 | 49 | 52 | 55 | 57 | 38 | 35 | 43 |
| Marks in  SAD | 35 | 45 | 36 | 23 | 38 | 51 | 50 | 47 | 38 | 40 | 33 | 40 |

1. Draw scatter plot with straight line trend ( marks in statistics along x-axis).
2. Run the regression analysis taking marks in statistics as independent variable.

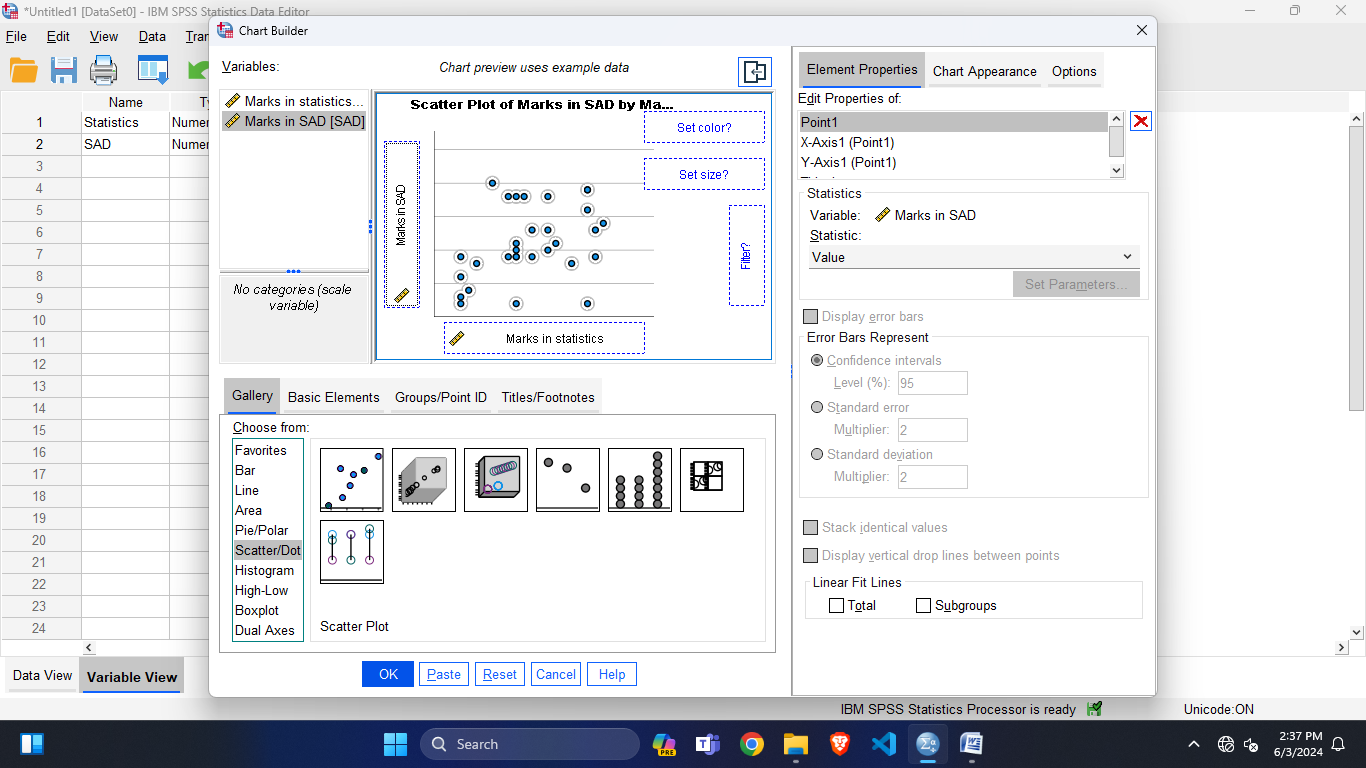
**SPSS Procedure:**

i.

1. In the variable view, create variable (Statistics, SAD) with label marks in statistics and marks in SAD along with measure scale and decimal 0.
2. Insert the data in data view as per question.
3. From the menu bar, choose graphs and select Chart Builder.
4. From the gallery in chart builder, drag and drop scatter/Dot.

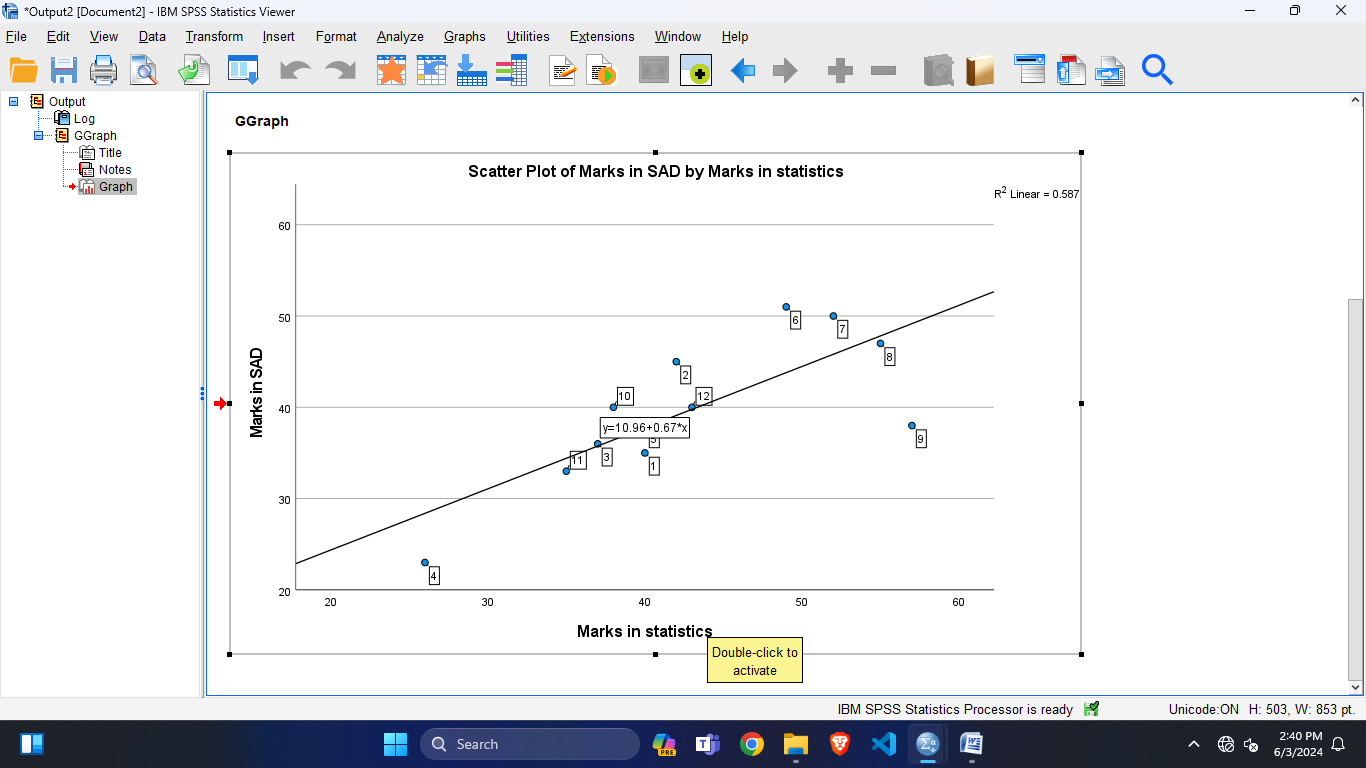


1. Then, drag and drop marks in statistics and marks in SAD from variables to X-axis and Y-axis respectively,



1. Now, click the ok button as shown.

## In Output:



ii.

1. Now, choose the Analyze from menu bar then select regression. After selecting regression, choose linear.
2. In Linear gallery ,Drag statistics to Independent(s) view and SAD to dependent view.
3. Then click on OK.

## In Output:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 10.956 | 7.764 |  | 1.411 | .189 |
| Statistics | .670 | .178 | .766 | 3.772 | .004 |
| a. Dependent Variable: SAD | | | | | | |

# Lab 7: Single Factor ANOVA

The following table shows the lifetimes under controlled conditions in hours in excess of 1000 hours, of 60W electric light bulbs of three different brands.

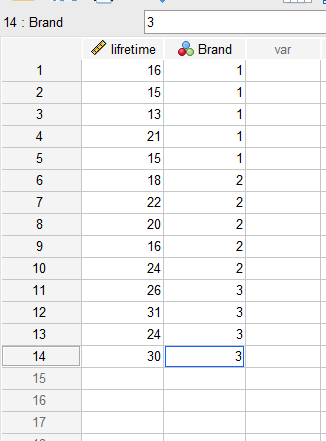
**Brands:**

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| 16 | 18 | 26 |
| 19 | 22 | 31 |
| 13 | 20 | 24 |
| 21 | 16 | 30 |
| 15 | 24 |  |

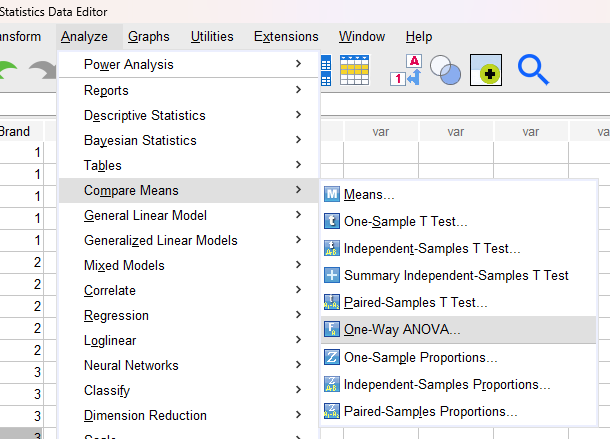
Test whether there is a significant difference between the mean life time of the three brands.

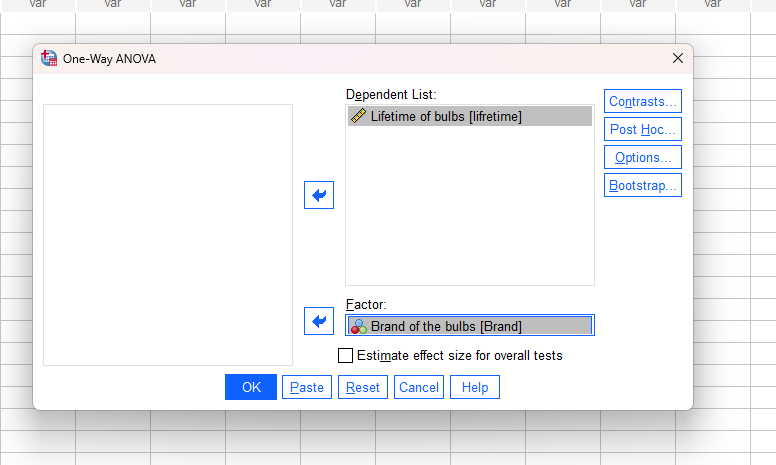
**SPSS Procedure:**

1. In the variable view create two variables (lifetime) with *label* lifetime of bulbs and *measures* Scale and (Brand) with *label* Brand of the bulbs and *measures* Nominal.
2. In data view insert the data as per the question.



1. From the menu bar choose Analyze, select Compare Means and click on One-Way ANOVA.



1. On the One-Way ANOVA gallery drag lifetime variable to the Dependent List and Brand to Factors. 
2. Click on OK.

## In Output:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ANOVA** | | | | | |
| Lifetime of bulbs | | | | | |
|  | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 311.607 | 2 | 155.804 | 15.759 | .001 |
| Within Groups | 108.750 | 11 | 9.886 |  |  |
| Total | 420.357 | 13 |  |  |  |

## Conclusion:

Since, given α=0.05 and we got (p-value) Sig=.001we should reject H0 (p-value low, H0 must go) and accept H1 that states *At least two brands differ in mean life time* (μi ≠ μj for some i, j).

# Lab 8: Two Factor ANOVA

The following are the marks obtained by students of different classes in different subjects:

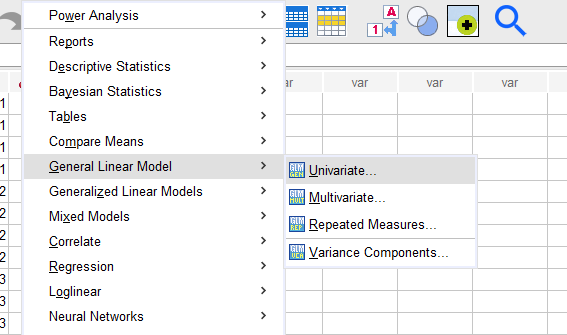
|  |  |  |  |
| --- | --- | --- | --- |
|  | BCA | BIM | BscCSIT |
| JAVA | 80 | 75 | 65 |
| Statistics | 90 | 80 | 75 |
| DSA | 70 | 65 | 85 |
| Web | 85 | 85 | 85 |

Use ANOVA to identify whether the marks differ:

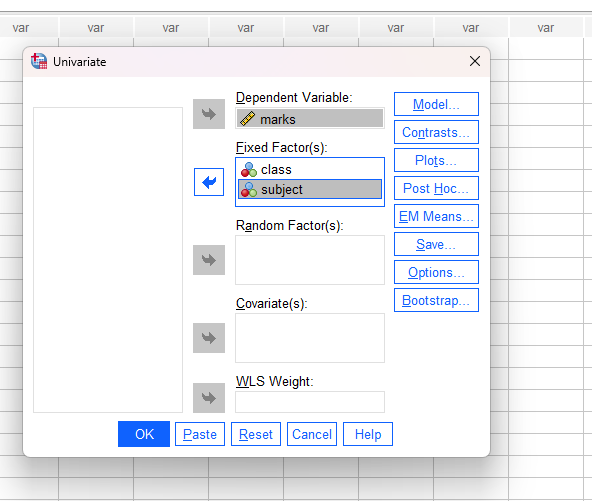
1. According to the class.
2. According to the subject

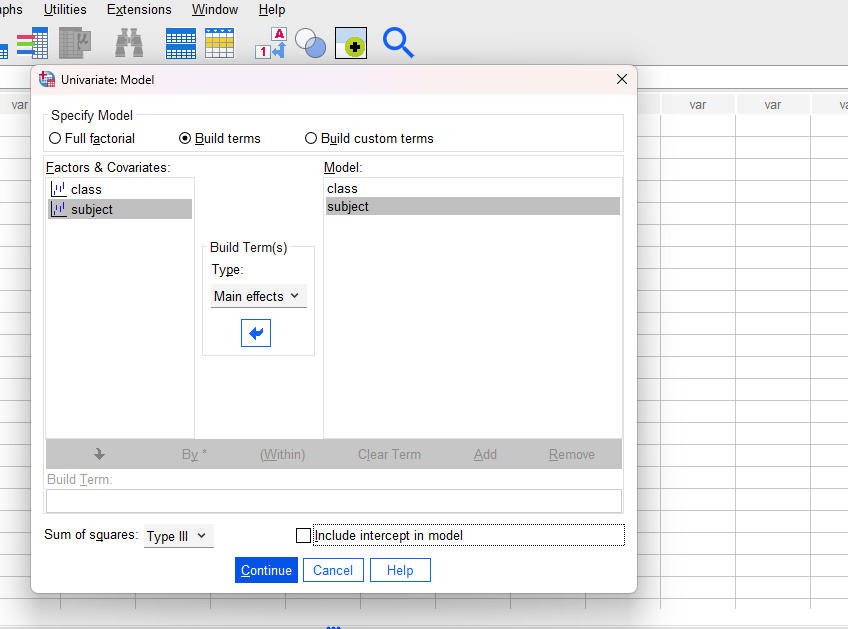
**SPSS Procedure:**

1. In the variable view of SPSS create three variable marks, class and subject with Type: Numeric, Decimals: 0, Measures: Scale, Nominal and Nominal respectively.
2. Insert the values of class:{1=BCA,2=BIM,3=BscCSIT} and subject:{1=JAVA,2=STATISTICS, 3=DSA AND 4=WEB}.
3. From the menu bar choose Analyze, select general Linear Model and select Univariate.



1. In the Univariate gallery select mark as Dependent Variable and class as well as subject as Fixed Factor(s).



1. In the right hand side click on Model.
2. In Univariate model gallery select Specify model as Build terms, Factors & Covariates to Model ,Build Term(s) Type as Main effects and Select Sum of squares as Type III.
3. Unselect include intercept in model and click on Continue. 
4. Click on OK.

## In Output:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tests of Between-Subjects Effects** | | | | | |
| Dependent Variable: marks | | | | | |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
| Model | 74004.167a | 6 | 12334.028 | 186.958 | .000 |
| class | 54.167 | 2 | 27.083 | .411 | .681 |
| subject | 316.667 | 3 | 105.556 | 1.600 | .285 |
| Error | 395.833 | 6 | 65.972 |  |  |
| Total | 74400.000 | 12 |  |  |  |
| a. R Squared = .995 (Adjusted R Squared = .989) Conclusion: Since, given α=0.05 and we got (p-value) Sig=.681 of class we should accept H0 and of Subject given α=0.05 and we got (p-value) Sig=.285 we should accept H0. i.e. There is no difference in marks neither according to class nor according to subjects. | | | | | |