SPSS TUTORIAL

MathCracker.com

Scatter Plot
Regression
ANOVA
GLM
Recoding Data
## TABLE OF CONTENTS

1. Scatter Plot ................................................................. 3
   Procedure ........................................................................ 3
   Assign Variable ................................................................ 3
   Fit .................................................................................. 4
   Spikes ........................................................................... 5
   Titles ............................................................................. 5
   Options ........................................................................... 6
   Output ............................................................................ 6

2. Linear Regression ............................................................. 7
   Simple Linear Regression .................................................. 7
   Output ............................................................................ 9
   Multiple Linear Regression ............................................. 10
   Output ............................................................................ 11

3. ANOVA ........................................................................... 13
   One Way ANOVA ............................................................ 13
   Output ............................................................................ 14

4. General Linear Model (GLM) .............................................. 16
   GLM-Univariate ............................................................... 16
   GLM Univariate-Fixed Factor(s) ....................................... 16
   Output ............................................................................ 19
   GLM Univariate-UNCOVA ............................................. 22
   Output ............................................................................ 24
   GLM Multivariate .......................................................... 26
   Output ............................................................................ 28
   GLM Repeated Measures .............................................. 30
   Output ............................................................................ 31

5. Recoding Data ................................................................. 34
   FAED/MAED ................................................................. 34
   Labeled ........................................................................... 34
   Recoding Data With Syntax .......................................... 41
1. Scatter Plot

A scatter plot may help you to understand how well linear regression fits your data. You may find that a quadratic equation would be more appropriate than a linear one.

Procedure

For example in this section we shall create a scatter plot for Employee Data.sav from SPSS data sample. Once the Employee Data.sav dataset is open, pull down the Graphs menu and point to Interactive and click on Scatterplot option.

Create Scatterplot dialog will appear. There are 5 tabs in Create Scatterplot dialog; assign variable, fit, spikes, title and option.

Assign Variable

On Assign variable you can select scatter plot coordinate between 2-D or 3-D, and then
assign variable for each axis. If you select 2-D coordinate you must choose the variables you want on the X-axis and Y-axis, and if you select 3-D coordinate you must choose the variables you want on the X-axis and Y-axis. Drag and Drop variable name into axis field. This tutorial demonstrate sample for 2-D scatterplot, we have chosen 'previous experience' vs 'salary' from employee data. sav

Fit

Select fit method, there are 4 options; None, Regression, Mean and Smoother. Select **None** for this case
Spikes

Use spikes options if you want to mark spikes data

Titles

Fill on chart title, chart sub-title and caption, as you need
Options

Select options as you need and then click **OK** to produce scatter plot diagram.

Output

Scatter plot diagram will appear on output window.
2. Linear Regression

This tutorial will explain two types of linear regression, there are simple linear regression and multiple linear regression.

Simple Linear Regression

Linear regression it is possible to output the regression coefficients necessary to predict one variable from the other. To do linear regression click on Analyze => Regression => Linear.

- Linear Regression Dialog will appear. Further, there is a need to know which variable will be used as the dependent variable and which will be used as the independent variable(s). In our current example, Revenue will be the dependent variable, and Sales Number will act as the independent variable.
• Click on **Statistics** button, and select **Estimates** and **Model fit** (as default)

• Click **Continue** button
• Click **Options** button and define confidence interval for F-test
- Click **Continue**
- Click **OK** and output will appear

### Output

Output for this case is:

#### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.998(a)</td>
<td>.997</td>
<td>.996</td>
<td>5.106</td>
</tr>
</tbody>
</table>

*a* Predictors: (Constant), Sales Number

#### ANOVA(b)

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Regression</td>
<td>107974.94</td>
<td>4</td>
<td>107974.944</td>
<td>4140.871</td>
<td>.000(a)</td>
</tr>
<tr>
<td>Residual</td>
<td>365.056</td>
<td>14</td>
<td>26.075</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>108340.00</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Predictors: (Constant), Sales Number

*b* Dependent Variable: Thousand U$

#### Coefficients(a)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>B</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>34.243</td>
<td>3.690</td>
<td>.998</td>
<td>9.281</td>
</tr>
<tr>
<td>Sales Number</td>
<td>17.821</td>
<td>.277</td>
<td>.998</td>
<td>64.350</td>
</tr>
</tbody>
</table>

*a* Dependent Variable: Thousand U$

**Linear Regression Formula Model** for this case is:

\[ Y = 34.243 + 17.821X \]
Multiple Linear Regression

- Click on Analyze => Regression => Linear.
- In this case we use revenue as dependent variable, product price and sales number as independent(s) variable.

Click on Statistics button and select Estimates, Model Fit, Colinearity diagnostics and Durbin-Watson

For further assistance in SPSS, you can contact the guys at MYGEKKTUTOR.COM
Click **Continue** button
Click **Options** button and define confidence interval for F-test

**Output**

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R$ Square</th>
<th>Adjusted $R$ Square</th>
<th>Std. Error of the Estimate</th>
<th>Durbin-Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.993(a)</td>
<td>.986</td>
<td>.983</td>
<td>5.758</td>
<td>1.910</td>
</tr>
</tbody>
</table>

a Predictors: (Constant), Sales Number, Product Price - US$
b Dependent Variable: Revenue - Thousands US$

### ANOVA(b)

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>27662.125</td>
<td>2</td>
<td>13831.063</td>
<td>417.148</td>
<td>.000(a)</td>
</tr>
<tr>
<td>Residual</td>
<td>397.875</td>
<td>12</td>
<td>33.156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>28060.000</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Predictors: (Constant), Sales Number, Product Price - US$

b Dependent Variable: Revenue - Thousands US$

### Coefficients(a)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>Tolerance</td>
<td>VIF</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>60.444</td>
<td>12.710</td>
<td>-</td>
<td>4.756</td>
<td>.000</td>
</tr>
<tr>
<td>Product Price - US$</td>
<td>-.396</td>
<td>.113</td>
<td>-.143</td>
<td>-3.491</td>
<td>.004</td>
</tr>
<tr>
<td>Sales Number</td>
<td>11.313</td>
<td>.436</td>
<td>1.064</td>
<td>25.945</td>
<td>.000</td>
</tr>
</tbody>
</table>

a Dependent Variable: Revenue - Thousands US$

**Formula Model** for this case is:

\[
Y = 30.444 + 11.313X(\text{sales number}) - 0.143X(\text{product price})
\]
3. ANOVA

Analysis of variance (ANOVA) is a collection of statistical models and their associated procedures, in which the observed variance is partitioned into components due to different explanatory variables.

Example case for this section is research about the relationship between course period and grade. There are 3 kinds of course; 3 months, 6 months, and 9 months.

One Way ANOVA

- Click on **Analyze => Compare Means => One-Way ANOVA**

<table>
<thead>
<tr>
<th>Course</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.0</td>
</tr>
<tr>
<td>2</td>
<td>8.0</td>
</tr>
<tr>
<td>3</td>
<td>7.2</td>
</tr>
<tr>
<td>4</td>
<td>7.5</td>
</tr>
<tr>
<td>5</td>
<td>7.0</td>
</tr>
<tr>
<td>6</td>
<td>7.5</td>
</tr>
<tr>
<td>7</td>
<td>7.7</td>
</tr>
<tr>
<td>8</td>
<td>5.5</td>
</tr>
<tr>
<td>9</td>
<td>7.0</td>
</tr>
<tr>
<td>10</td>
<td>7.5</td>
</tr>
<tr>
<td>11</td>
<td>8.0</td>
</tr>
<tr>
<td>12</td>
<td>8.5</td>
</tr>
<tr>
<td>13</td>
<td>8.0</td>
</tr>
<tr>
<td>14</td>
<td>8.2</td>
</tr>
<tr>
<td>15</td>
<td>7.5</td>
</tr>
</tbody>
</table>

- One-Way ANOVA dialog will appear, select **Grade** variable as dependent list and **Course** variable as factor

![One-Way ANOVA dialog](image)
• Click **Option** and select **Descriptive** and **Homogeneity of variance test**

![One-Way ANOVA: Options](image)

• Click **Continue**
• Click **Post Hoc** and select **LSD**
• Click **Continue**
• Click **Contrast** and enter **coefficients number** 0 and click **Add**, then enter coefficient number -1 and 1.
• Click **Continue**
• Click **OK**, and output will appear

### Output

#### Descriptives

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Bound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Bound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Upper Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Months</td>
<td>10</td>
<td>7.3000</td>
<td>.43780</td>
<td>.13844</td>
<td>6.9868</td>
<td>6.50</td>
<td>8.00</td>
</tr>
<tr>
<td>6 Months</td>
<td>10</td>
<td>8.0750</td>
<td>.42573</td>
<td>.13463</td>
<td>7.7704</td>
<td>7.50</td>
<td>8.75</td>
</tr>
<tr>
<td>9 Months</td>
<td>10</td>
<td>8.8750</td>
<td>.44488</td>
<td>.14068</td>
<td>8.5568</td>
<td>8.00</td>
<td>9.50</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>8.0833</td>
<td>.77774</td>
<td>.14200</td>
<td>7.7929</td>
<td>6.50</td>
<td>9.50</td>
</tr>
</tbody>
</table>

#### ANOVA

<table>
<thead>
<tr>
<th>Grade</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>12.404</td>
<td>2</td>
<td>6.202</td>
<td>32.595</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>5.138</td>
<td>27</td>
<td>.190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>17.542</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Test of Homogeneity of Variances

<table>
<thead>
<tr>
<th>Grade</th>
<th>Levene Statistic</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.006</td>
<td>2</td>
<td>27</td>
<td>.994</td>
</tr>
</tbody>
</table>

#### Multiple Comparisons

Dependent Variable: Grade
LSD

<table>
<thead>
<tr>
<th>(I) Term</th>
<th>(J) Term</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Months</td>
<td>6 Months</td>
<td>-.7750(∗)</td>
<td>.19508</td>
<td>.000</td>
<td>-1.1753 - .3747</td>
</tr>
<tr>
<td>6 Months</td>
<td>3 Months</td>
<td>.7750(∗)</td>
<td>.19508</td>
<td>.000</td>
<td>.3747 1.1753</td>
</tr>
<tr>
<td>9 Months</td>
<td>3 Months</td>
<td>1.5750(∗)</td>
<td>.19508</td>
<td>.000</td>
<td>1.1747 1.9753</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

#### Contrast Coefficients

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Term</th>
<th>3 Months</th>
<th>6 Months</th>
<th>9 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Contrast Tests

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Value of Contrast</th>
<th>Std. Error</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Assume equal variances</td>
<td>1 .8000</td>
<td>.19508</td>
<td>4.101</td>
<td>27</td>
<td>.000</td>
</tr>
<tr>
<td>Grade Does not assume equal variances</td>
<td>1 .8000</td>
<td>.19472</td>
<td>4.108</td>
<td>17.965</td>
<td>.001</td>
</tr>
</tbody>
</table>

4. General Linear Model (GLM)

This tutorial will explain four types of GLM, there are; GLM Univariate-Fixed Factor(s), GLM Univariate-UNCOVA, GLM-Multivariate and GLM-Repeates Measures.

GLM-Univariate

GLM-Univariate analysis is regression analysis and variance one dependent variable with two or more factor variable or other variables.

GLM Univariate-Fixed Factor(s)

Example case for univariate-fixed factor is to know customer shopping trend.

- Click on Analyze => General Linear Model => Univariate

- Uniavariate dialog will appear, select **shopping value** as dependent variable, **frequency** and **customer category** variable as fixed factor(s).
• Click **Plots** and Univariate: Profile Plots dialog will appear, enter *frequency* variable into horizontal axis and **customer category** (*Cust_Cat*) into separate lines and then click **Add**. *frequency*\(^{*}\)\(*\)cust\_cat* variable will move into Plots box.

• Click **Continue**

• Click **Post Hoc** and Univariate:Post Hoc dialog will appear, Select Equal Variances Assumed – *Turkey* and Equal Variance Not Assumed - *Tamhane*
- Click **Continue**
- Click **Option** and Univariate:Option dialog will appear, move frequency*Cat_Cus from Factor(s) and Factor Interactions box into Display Means for box. Select Descriptive statistic, Estimates of effect size, Homogenety test and spread vs level plot in Display groupbox.

- Click **Continue**
- Click **OK** and output will appear
### Between-Subjects Factors

<table>
<thead>
<tr>
<th>Customer Category</th>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>individu</td>
<td>337</td>
</tr>
<tr>
<td>2</td>
<td>couple</td>
<td>287</td>
</tr>
<tr>
<td>3</td>
<td>family</td>
<td>176</td>
</tr>
<tr>
<td>1</td>
<td>once-two weeks</td>
<td>187</td>
</tr>
<tr>
<td>2</td>
<td>once-a week</td>
<td>461</td>
</tr>
<tr>
<td>3</td>
<td>several - a week</td>
<td>152</td>
</tr>
</tbody>
</table>

### Descriptive Statistics

#### Dependent Variable: Shopping Value

<table>
<thead>
<tr>
<th>Customer Category</th>
<th>frequency</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>individu</td>
<td>once-two weeks</td>
<td>241549.03</td>
<td>51076.881</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>267907.85</td>
<td>47644.510</td>
<td>191</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>297827.14</td>
<td>46527.810</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>266508.14</td>
<td>51569.954</td>
<td>337</td>
</tr>
<tr>
<td>couple</td>
<td>once-two weeks</td>
<td>298406.07</td>
<td>50142.015</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>324952.65</td>
<td>47765.161</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>342457.78</td>
<td>48419.239</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>323030.95</td>
<td>50393.788</td>
<td>287</td>
</tr>
<tr>
<td>family</td>
<td>once-two weeks</td>
<td>384409.50</td>
<td>69602.060</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>400183.56</td>
<td>78776.554</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>421745.30</td>
<td>68918.993</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>400396.35</td>
<td>75637.588</td>
<td>176</td>
</tr>
<tr>
<td>Total</td>
<td>once-two weeks</td>
<td>290654.37</td>
<td>77743.025</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>318453.06</td>
<td>75860.179</td>
<td>461</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>341010.90</td>
<td>69289.815</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>316241.11</td>
<td>76812.882</td>
<td>800</td>
</tr>
</tbody>
</table>

### Tests of Between-Subjects Effects

#### Dependent Variable: Shopping Value

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>22909853751</td>
<td>8</td>
<td>28637317189</td>
<td>93.477</td>
<td>.000</td>
<td>.486</td>
</tr>
<tr>
<td>Intercept</td>
<td>27.443(a)</td>
<td>1</td>
<td>63790494928</td>
<td>20822.226</td>
<td>.000</td>
<td>.963</td>
</tr>
<tr>
<td>Cust_Cat</td>
<td>15978262869</td>
<td>2</td>
<td>79891314347</td>
<td>260.778</td>
<td>.000</td>
<td>.397</td>
</tr>
<tr>
<td>frekun</td>
<td>16112594323</td>
<td>2</td>
<td>80562971617</td>
<td>26.297</td>
<td>.000</td>
<td>.062</td>
</tr>
</tbody>
</table>
### Type III Sum of Squares

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cust_Cat * frekun</td>
<td>6899347068. 037 4</td>
<td>1724836767. 09</td>
<td>.563</td>
<td>.690</td>
<td>.003</td>
</tr>
<tr>
<td>Error</td>
<td>24232895140 93.421</td>
<td>3063577135.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84721024114 141.800</td>
<td>3063577135.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>47142748892 20.860</td>
<td>3063577135.39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* R Squared = .486 (Adjusted R Squared = .481)

---

### Estimated Marginal Means of Shopping Value

#### Customer Category * frequency

**Dependent Variable: Shopping Value**

<table>
<thead>
<tr>
<th>Customer Category</th>
<th>frequency</th>
<th>Mean</th>
<th>Std. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
</tr>
<tr>
<td>individu</td>
<td>once-two weeks</td>
<td>241549.033</td>
<td>5968.500</td>
<td>229833.061</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>267907.849</td>
<td>4004.956</td>
<td>260046.250</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>297827.138</td>
<td>7145.601</td>
<td>283800.554</td>
</tr>
<tr>
<td>couple</td>
<td>once-two weeks</td>
<td>298406.073</td>
<td>7086.789</td>
<td>284494.936</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>324952.647</td>
<td>4308.960</td>
<td>316494.299</td>
</tr>
</tbody>
</table>

---

*Estimated Marginal Means of Shopping Value*

---

*Customer Category: individu, couple, family*

---

*Estimated Marginal Means of Shopping Value*

---

*Customer Category: individu, couple, family*
### Customer Category frequency

<table>
<thead>
<tr>
<th>Customer Category</th>
<th>frequency</th>
<th>Mean</th>
<th>Std. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Lower Bound</td>
</tr>
<tr>
<td>family</td>
<td>several - a week</td>
<td>342457.775</td>
<td>7086.789</td>
<td>328546.639</td>
</tr>
<tr>
<td></td>
<td>once-two weeks</td>
<td>384409.496</td>
<td>8751.539</td>
<td>367230.510</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>400183.564</td>
<td>5401.567</td>
<td>389580.464</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>421745.298</td>
<td>9941.080</td>
<td>402231.281</td>
</tr>
</tbody>
</table>

### Multiple Comparisons

Dependent Variable: Shopping Value

<table>
<thead>
<tr>
<th>(I) frequency</th>
<th>(J) frequency</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(I-J)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>once-two weeks</td>
<td>once-a week</td>
<td>-27798.69(*)</td>
<td>6693.576</td>
<td>.000</td>
<td>-43861.00 -11736.37</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>-50356.53(*)</td>
<td>7994.172</td>
<td>.000</td>
<td>-96540.69 -31172.37</td>
</tr>
<tr>
<td>once-a week</td>
<td>once-two weeks</td>
<td>27798.69(*)</td>
<td>6693.576</td>
<td>.000</td>
<td>11736.37 43861.00</td>
</tr>
<tr>
<td></td>
<td>several - a week</td>
<td>-22557.84(*)</td>
<td>6638.469</td>
<td>.002</td>
<td>-38504.28 -6611.40</td>
</tr>
<tr>
<td>several - a week</td>
<td>once-two weeks</td>
<td>50356.53(*)</td>
<td>7994.172</td>
<td>.000</td>
<td>31172.37 69540.69</td>
</tr>
<tr>
<td></td>
<td>once-a week</td>
<td>22557.84(*)</td>
<td>6638.469</td>
<td>.002</td>
<td>6611.40 38504.28</td>
</tr>
</tbody>
</table>

Based on observed means.

* The mean difference is significant at the .05 level.
GLM Univariate-UNCova

Example case for this section is research about household income before and after participate in government program.

- Click on Analyze => General Linear Model => Univariate
- Enter result_after as dependent variable, program status variable as Fix Factor(s) and result_before as Covariate(s)

- Click Models and Univariate: Models dialog will appear, Select Custom in Specify Model. Select program variable move into Model box, select result_before and move into Model box. Select both program variable and result_before and move into Model box and then program*result_before variable will appear. Select Interaction on the Build Term(s) dropdown.
• Click **Continue**
• Click **Options** and select Estimates of effect size.

![Univariate: Options](image)

• Click **Continue**
• Click **OK** and output will appear.
• The next step is covarian analysis. Open Univariate dialog again click **Model** and select **Full Factorial** in Specify model.

![Univariate: Model](image)

• Click **Continue**
• Click **Option** and select Descriptive statistics, Estimates of effect size, Homogenety test, and Parameter Estimates.
- Click **Continue**
- Click **OK** and output will appear.

**Output**

### Between-Subjects Factors

<table>
<thead>
<tr>
<th>program status</th>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>not participate</td>
<td>293</td>
</tr>
<tr>
<td>1</td>
<td>participate</td>
<td>307</td>
</tr>
</tbody>
</table>

### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>16557688.72</td>
<td>3</td>
<td>5519229.576</td>
<td>220.962</td>
<td>.000</td>
<td>.527</td>
</tr>
<tr>
<td>Intercept</td>
<td>291519.199</td>
<td>1</td>
<td>291519.199</td>
<td>11.671</td>
<td>.001</td>
<td>.019</td>
</tr>
<tr>
<td>program</td>
<td>114931.433</td>
<td>1</td>
<td>114931.433</td>
<td>4.601</td>
<td>.032</td>
<td>.008</td>
</tr>
<tr>
<td>result_before</td>
<td>9598160.112</td>
<td>1</td>
<td>9598160.112</td>
<td>384.262</td>
<td>.000</td>
<td>.392</td>
</tr>
<tr>
<td>program * result_before</td>
<td>11836.725</td>
<td>1</td>
<td>11836.725</td>
<td>.474</td>
<td>.491</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>14886973.77</td>
<td>596</td>
<td>24978.144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>453042500.0</td>
<td>600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31444662.50</td>
<td>599</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a  R Squared = .527 (Adjusted R Squared = .524)
### Descriptive Statistics

Dependent Variable: result after program

<table>
<thead>
<tr>
<th>program status</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>not participate</td>
<td>728.84</td>
<td>194.917</td>
<td>293</td>
</tr>
<tr>
<td>participate</td>
<td>942.67</td>
<td>210.011</td>
<td>307</td>
</tr>
<tr>
<td>Total</td>
<td>838.25</td>
<td>229.118</td>
<td>600</td>
</tr>
</tbody>
</table>

### Levene’s Test of Equality of Error Variances(a)

Dependent Variable: result after program

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>.605</td>
<td>1</td>
<td>598</td>
<td>.437</td>
<td></td>
</tr>
</tbody>
</table>

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a Design: Intercept+result_before+program

### Tests of Between-Subjects Effects

Dependent Variable: result after program

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>16545852.00 (a)</td>
<td>2</td>
<td>8272926.002</td>
<td>331.499</td>
<td>.000</td>
<td>.526</td>
</tr>
<tr>
<td>result_before</td>
<td>285238.337</td>
<td>1</td>
<td>285238.337</td>
<td>11.430</td>
<td>.001</td>
<td>.019</td>
</tr>
<tr>
<td>program</td>
<td>9691004.737</td>
<td>1</td>
<td>9691004.737</td>
<td>388.322</td>
<td>.000</td>
<td>.394</td>
</tr>
<tr>
<td>Error</td>
<td>6459841.024</td>
<td>1</td>
<td>6459841.024</td>
<td>258.848</td>
<td>.000</td>
<td>.302</td>
</tr>
<tr>
<td>Total</td>
<td>14898810.49</td>
<td>6</td>
<td>24956.131</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>145304250.00</td>
<td>600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Model</td>
<td>31444662.50</td>
<td>599</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a R Squared = .526 (Adjusted R Squared = .525)

### Parameter Estimates

Dependent Variable: result after program

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>Intercept</td>
<td>227.856</td>
<td>37.378</td>
<td>6.096</td>
<td>.000</td>
<td>154.448</td>
<td>301.264</td>
</tr>
<tr>
<td>[program=0]</td>
<td>1.587</td>
<td>.081</td>
<td>19.706</td>
<td>.000</td>
<td>1.429</td>
<td>1.745</td>
</tr>
</tbody>
</table>

a This parameter is set to zero because it is redundant.
GLM Multivariate

Example case for this section is research about the impact of gender factor to expense for life style.

- Click on **Analyze => General Linear Model => Multivariate**

  ![Screenshot of SPSS interface showing the General Linear Model dialog box]

- Multivariate dialog box will appear, select **food cost** variable and **lifestyle** variable and move into Dependent Variable
- Move **Gender** variable into Fix Factor(s)
- Click **option** and then select Descriptive statistics, Estimates of effect size, and Parameter estimates.

- Click **Continue**
- Click **OK** and output will appear
Output

**Between-Subjects Factors**

<table>
<thead>
<tr>
<th></th>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>0</td>
<td>male</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>female</td>
</tr>
</tbody>
</table>

**Descriptive Statistics**

<table>
<thead>
<tr>
<th></th>
<th>gender</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>food cost</td>
<td>male</td>
<td>445.24</td>
<td>82.375</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>451.18</td>
<td>79.947</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>448.38</td>
<td>81.056</td>
<td>400</td>
</tr>
<tr>
<td>lifestyle cost</td>
<td>male</td>
<td>723.54</td>
<td>195.055</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>945.26</td>
<td>209.481</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>840.50</td>
<td>230.880</td>
<td>400</td>
</tr>
</tbody>
</table>

**Multivariate Tests(b)**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Pillai's Trace</td>
<td>.969</td>
<td>6237.900(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.969</td>
</tr>
<tr>
<td></td>
<td>Wilks' Lambda</td>
<td>.031</td>
<td>6237.900(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.969</td>
</tr>
<tr>
<td></td>
<td>Hotelling's Trace</td>
<td>31.425</td>
<td>6237.900(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.969</td>
</tr>
<tr>
<td></td>
<td>Roy's Largest Root</td>
<td>31.425</td>
<td>6237.900(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.969</td>
</tr>
<tr>
<td>gender</td>
<td>Pillai's Trace</td>
<td>.308</td>
<td>88.173(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.308</td>
</tr>
<tr>
<td></td>
<td>Wilks' Lambda</td>
<td>.692</td>
<td>88.173(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.308</td>
</tr>
<tr>
<td></td>
<td>Hotelling's Trace</td>
<td>.444</td>
<td>88.173(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.308</td>
</tr>
<tr>
<td></td>
<td>Roy's Largest Root</td>
<td>.444</td>
<td>88.173(a)</td>
<td>2.000</td>
<td>.000</td>
<td>.308</td>
</tr>
</tbody>
</table>

a  Exact statistic
b  Design: Intercept+gender

**Tests of Between-Subjects Effects**

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>food cost</td>
<td>3525.673(a)</td>
<td>1</td>
<td>3525.673</td>
<td>.536</td>
<td>.465</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>lifestyle cost</td>
<td>4900914.469(b)</td>
<td>1</td>
<td>4900914.469</td>
<td>119.169 .000</td>
<td>.230</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>food cost</td>
<td>80114325.673</td>
<td>1</td>
<td>80114325.673</td>
<td>12179.717</td>
<td>.000</td>
<td>.968</td>
</tr>
<tr>
<td></td>
<td>lifestyle cost</td>
<td>277648789.469</td>
<td>9</td>
<td>277648789.469</td>
<td>6751.241</td>
<td>.000</td>
<td>.944</td>
</tr>
<tr>
<td>gender</td>
<td>food cost</td>
<td>3525.673</td>
<td>1</td>
<td>3525.673</td>
<td>.536</td>
<td>.465</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>lifestyle cost</td>
<td>4900914.469</td>
<td>1</td>
<td>4900914.469</td>
<td>119.169 .000</td>
<td>.230</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>food cost</td>
<td>2617918.077</td>
<td>398</td>
<td>6577.684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Dependent Variable</td>
<td>Type III Sum of Squares</td>
<td>df</td>
<td>Mean Square</td>
<td>F</td>
<td>Sig.</td>
<td>Partial Eta Squared</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------</td>
<td>-------------------------</td>
<td>-----</td>
<td>-------------</td>
<td>-------</td>
<td>------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Total</td>
<td>lifestyle cost</td>
<td>16367985.531</td>
<td>398</td>
<td>41125.592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>food cost</td>
<td>83037500.000</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>lifestyle cost</td>
<td>303845000.000</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>food cost</td>
<td>2621443.750</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lifestyle cost</td>
<td>21268900.000</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a R Squared = .001 (Adjusted R Squared = -.001)
b R Squared = .230 (Adjusted R Squared = .228)

**Parameter Estimates**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Lower Bound</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>food cost</td>
<td>Intercept</td>
<td>451.185</td>
<td>5.583</td>
<td>80.809</td>
<td>.000</td>
<td>440.208</td>
<td>462.161</td>
</tr>
<tr>
<td></td>
<td>[gender=1]</td>
<td>0(a)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>lifestyle cost</td>
<td>Intercept</td>
<td>945.261</td>
<td>13.961</td>
<td>67.707</td>
<td>.000</td>
<td>917.814</td>
<td>972.707</td>
</tr>
<tr>
<td></td>
<td>[gender=0]</td>
<td>-221.716</td>
<td>20.310</td>
<td>-10.916</td>
<td>.000</td>
<td>-261.644</td>
<td>181.787</td>
</tr>
<tr>
<td></td>
<td>[gender=1]</td>
<td>0(a)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

a This parameter is set to zero because it is redundant.
GLM Repeated Measures

Example case for this section is research about performance of 4 weeks diet program between male and female.

- Click on **Analyze ⇒ General Linear Model ⇒ Repeated Measures**

Repeted Measures Define dialog will appear, write **weight** in the Within-Subject Factor Name and enter **5** on Number of Levels. Click **Add** and **weight5** will move into box.

For further assistance in SPSS, you can contact the guys at MYGEEKYTUTOR.COM
• Click **Define** and Repeated Measures dialog will appear. Enter dependent variable from `weight0`, `weight1`, `weight2`, `weight3` and `weight4` in Within-Subjects Variables (weight) and **gender** variable in Between-Subjects Factor(s) box.

• Click **Option**, select Descriptive statistics, Estimates of effect size, and Parameter estimates
• Click **Continue**
• Click **OK**, and Output will appears

**Output**
### Within-Subjects Factors

Measure: MEASURE_1

<table>
<thead>
<tr>
<th>weight</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Weight0</td>
</tr>
<tr>
<td>2</td>
<td>Weight1</td>
</tr>
<tr>
<td>3</td>
<td>Weight2</td>
</tr>
<tr>
<td>4</td>
<td>Weight3</td>
</tr>
<tr>
<td>5</td>
<td>Weight4</td>
</tr>
</tbody>
</table>

### Between-Subjects Factors

<table>
<thead>
<tr>
<th>Gender</th>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>male</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>female</td>
<td>10</td>
</tr>
</tbody>
</table>

### Descriptive Statistics

<table>
<thead>
<tr>
<th>Gender</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight before program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>84.4750</td>
<td>3.70144</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>75.5000</td>
<td>3.12027</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>79.9875</td>
<td>5.68324</td>
<td>20</td>
</tr>
<tr>
<td>Weight Weeks1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>82.0500</td>
<td>3.37021</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>73.5000</td>
<td>3.30624</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>77.7750</td>
<td>5.45912</td>
<td>20</td>
</tr>
<tr>
<td>Weight Weeks2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>78.9250</td>
<td>3.21898</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>71.4250</td>
<td>3.26609</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>75.1750</td>
<td>4.97633</td>
<td>20</td>
</tr>
<tr>
<td>Weight Weeks3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>77.0250</td>
<td>4.77617</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>70.4000</td>
<td>4.03320</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>73.7125</td>
<td>5.48279</td>
<td>20</td>
</tr>
<tr>
<td>Weight Weeks4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>74.5000</td>
<td>4.99444</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>68.1250</td>
<td>4.21843</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>71.3125</td>
<td>5.56237</td>
<td>20</td>
</tr>
</tbody>
</table>

### Multivariate Tests(b)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypotheti</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>.981</td>
<td>193.405(a)</td>
<td>4.000</td>
<td>15.000</td>
<td>.000</td>
<td>.981</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.019</td>
<td>193.405(a)</td>
<td>4.000</td>
<td>15.000</td>
<td>.000</td>
<td>.981</td>
</tr>
<tr>
<td>Hotelling's Trace</td>
<td>51.575</td>
<td>193.405(a)</td>
<td>4.000</td>
<td>15.000</td>
<td>.000</td>
<td>.981</td>
</tr>
<tr>
<td>Roy's Largest Root</td>
<td>51.575</td>
<td>193.405(a)</td>
<td>4.000</td>
<td>15.000</td>
<td>.000</td>
<td>.981</td>
</tr>
</tbody>
</table>
Tests of Within-Subjects Effects

Measure: MEASURE_1

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>Sphericity Assumed</td>
<td>922.129</td>
<td>4</td>
<td>230.532</td>
<td>73.811</td>
<td>.000</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>922.129</td>
<td>1.118</td>
<td>824.571</td>
<td>73.811</td>
<td>.000</td>
<td>.804</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>922.129</td>
<td>1.206</td>
<td>764.356</td>
<td>73.811</td>
<td>.000</td>
<td>.804</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>922.129</td>
<td>1.000</td>
<td>922.129</td>
<td>73.811</td>
<td>.000</td>
<td>.804</td>
</tr>
<tr>
<td>weight * Gender</td>
<td>Sphericity Assumed</td>
<td>26.271</td>
<td>4</td>
<td>6.568</td>
<td>2.103</td>
<td>.089</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>26.271</td>
<td>1.118</td>
<td>23.492</td>
<td>2.103</td>
<td>.162</td>
<td>.105</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>26.271</td>
<td>1.206</td>
<td>21.776</td>
<td>2.103</td>
<td>.159</td>
<td>.105</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>26.271</td>
<td>1.000</td>
<td>26.271</td>
<td>2.103</td>
<td>.164</td>
<td>.105</td>
</tr>
<tr>
<td>Error(weight)</td>
<td>Sphericity Assumed</td>
<td>224.875</td>
<td>72</td>
<td>3.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>224.875</td>
<td>20.130</td>
<td>11.171</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>224.875</td>
<td>21.715</td>
<td>10.356</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>224.875</td>
<td>18.000</td>
<td>12.493</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>571422.606</td>
<td>1</td>
<td>571422.606</td>
<td>9246.269</td>
<td>.000</td>
<td>.998</td>
</tr>
<tr>
<td>Gender</td>
<td>1445.901</td>
<td>1</td>
<td>1445.901</td>
<td>23.396</td>
<td>.000</td>
<td>.565</td>
</tr>
<tr>
<td>Error</td>
<td>1112.406</td>
<td>18</td>
<td>61.800</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Recoding Data

You can recode data into either the same variable or into a new one by going to **Transform > Recode**. This tool is especially useful for creating dummy variables, changing values from letters to numbers, increasing or decreasing the number of possible values, or for creating specialized filters that let you have fine-tuned control over which cases to exclude.

SPSS allows us to recode variables and then use the recoded variables in statistical analyses.

The values in variables **FAED** (father’s education) and **MAED** (mother’s education) range from 2 to 10 indicating 9 levels of education as:

<table>
<thead>
<tr>
<th>Labeled</th>
<th>FAED/MAED</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Less than high school</td>
</tr>
<tr>
<td>3</td>
<td>High school graduate</td>
</tr>
<tr>
<td>4</td>
<td>Less than 2 years’ vocational education</td>
</tr>
<tr>
<td>5</td>
<td>More than 2 years’ vocational education</td>
</tr>
<tr>
<td>6</td>
<td>Less than 2 years’ college education</td>
</tr>
<tr>
<td>7</td>
<td>More than 2 years’ college education</td>
</tr>
<tr>
<td>8</td>
<td>College graduate</td>
</tr>
<tr>
<td>9</td>
<td>Master’s degree</td>
</tr>
<tr>
<td>10</td>
<td>MD/PhD degree</td>
</tr>
</tbody>
</table>

Now we want to regroup (recode) the nine levels into four levels as:

<table>
<thead>
<tr>
<th>Labeled</th>
<th>FAED/MAED</th>
<th>FAEDNEW / MAEDNEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>Less than high school</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>High school graduate</td>
</tr>
<tr>
<td>4,5,6,7</td>
<td>3</td>
<td>Some post-secondary education</td>
</tr>
<tr>
<td>8,9,10</td>
<td>4</td>
<td>College graduate &amp; beyond</td>
</tr>
</tbody>
</table>

To recode the variables, please follow the steps:

- You will see the data in the SPSS Data Editor window:
• Before you recode the data, you should make a copy of original data. Make sure you save the new file into the same place as the original file.

• Recode the **FAED** (father’s education) variable into a new variable

• From **Transform** menu, choose **Rcode**, then **Into Different Variable**.

• In the “Recode into Different Variable” dialogue box, you will see a list of variables in the box on the left. Click on “**faed**”, and then click on the arrow button. You will see “**faed**” appears in the right box.
• Type “faednew” in the Output/Variable-Name box as the name of the new variable. Type “Father’s education” as the label for the new variable. Click on the Change button (see above).
• Click on the button “Old and New Values”, you will see the Old and New Values dialogue box. Under Old Value section, type 2 in the Value box, and type 1 in the Value section – this will recode the old value 2 into a new value 1 (as shown in the tables at the beginning of this module). Click on Add button.

• Check the Range radio button, then type 4 in the first box, and type 7 in the box after the word “through”. Then type 3 in the New value box. Click on Add button. You will see:

• Type 3 in Old value box, and 2 in New value box. Click on Add button. The recoding shows in the Old → New box.
• Check the Range radio button, then type 4 in the first box, and type 7 in the box after the word “through”. Then type 3 in the New value box. Click on Add button. You will see:
• Type 8 and 10 in the range boxes, and 4 in the New value box. Click on **Add**, you have recoded the nine old values into four new values:

![Image of the Recode into Same Variables dialog box]

• Click on Continue button, you will be back to the Recode into Different Variable dialogue box. Now you will recode another variable **maed**—mother’s education.

• Recode the **MAED** (mother’s education) variable

• In the Recode into Different Variable box, from the variable list, click on **maed**, and click on the arrow button to add the variable **maed** into the right box, it should be under **faed** variable.
• Make sure the variable **maed** is highlighted, type **maednew** in the Output/Variable-Name box as the name of the new variable. Type “Mother’s education” as the label for the new variable. Click on the **Change** button (see above).

• Click on **Old and New Values** button, you will see the previous recode settings:

![Image of Recode into Same Variables: Old and New Values dialog box](image)

- We will use the same recode settings. So we do not need to change. Simply click on Continue. (If you need to change the settings, click on each of the recode settings, then click on Remove. You can add new transform settings).

• Now, you are back to the original dialogue box, click on **OK**. You will see the two new variables **faednew** and **maednew**.

![Image of Recode into Different Variables dialog box](image)
• For the value of these two variables, we do not need decimals. To change the decimals, look at the bottom of the Data Editor window, you can see two tabs (Data View – which is the current window, and Variable View). Click on the Variable View tab.

![Data Editor window with Variable View selected](image)

• You will see the window changes to the Variable View mode:

![Variable View window](image)

• Click on the decimal cell of the faednew variable, the two arrows appear for you to change the decimals. Click on the down arrow to change the decimal number to 0. Repeat this step to change the decimals for the maednew variable.

![Variable View with faednew and maednew variables](image)

• Save the changes, to save the changes, from File menu, choose Save to save.
• Label the new values, click in the cell that crossing the Values column and the 12th row (faednew variable), you will see a small gray box.
Double click on the small gray box, you will see the Value Labels dialogue box as the following. Type 1 in the Value box, and “Less than High School” in the Value Label box. Then click on Add button.

Type 2 in Value box, and “High School Graduate” in the Value Label box. Click on Add. You should have:

Repeat Step above. Make the value label “Some Post-Secondary Education” for 3, and “College Graduate & Beyond” for 4. Then click on OK. You should see:
Repeat Step above to change the value labels for the variable maednew. You can repeat to add value labels to each variable.

We can add the labels for the variables with a clear description of the variable when sometimes the meaning is not clear from the variable name itself (e.g., “mathach” we can add a label “Math Achievement” as the label). To do this:

In the Variable View window, click in the Label column of the mathach variable, and type “Math Achievement” in the crossing cell.

You can repeat Step above to add variable label to each variable.

Save the changes. Make sure you save the data as SPSS (*.sav) file. Click on the Data View tab to switch to the data. Now, you are ready to use this new set of data with recoded values in faednew and maednew variable.

**Recoding Data With Syntax**

It is possible to use syntax when recoding variables. For example, if I had a variable that included the following values:

Redbird
Bluebird
Yellowbird
Elm
Butterfly

and I wanted to recode any values that included ‘bird’ into a new value ‘bird’.
To solve the problem the following syntax is an option:

DATA LIST LIST /var1(A15).
BEGIN DATA
Bluebird
Redbird
Yellowbird
Butterfly
Elm
END DATA.

STRING newVar(A15).
DO IF INDEX(UPCASE(var1),â€œ BIRDâ€œ ) > 0.
- COMPUTE newVar=â€œ BIRDâ€œ .
END IF.
EXECUTE.

Example number two is we want to recode the above variables into variables having the same name but with the last letter being replaced by x.

DATA LIST FREE /abc, sal, age, sex1, school,v1234567.
BEGIN DATA
85 95 5 87 100 1
END DATA.
LIST.
SAVE OUTFILE='c:\temp\mydata.sav'.
* suppose we want to recode the above variables into variables having the same name but with the last letter being replaced by x.

FLIP.
STRING newname(A8).
COMPUTE newname=CONCAT(SUBSTR(case_lbl,1,LENGTH(RTRIM(case_lbl))-1),"X").
WRITE OUTFILE 'c:\temp\temp.sps'
"RECODE "case_lbl" (1 THRU 87.9=1) (89 THRU 98.1=1) (ELSE=COPY) INTO "newname".""FREQ "newname".".
EXE.

For further assistance in SPSS, you can contact the guys at MYGEEKYTUTOR.COM
GET FILE='c:\temp\mydata.sav'.
INCLUDE 'c:\temp\temp.sps'.