



Factorial ANOVA analysis on BMI measurements with multiple health indicators

This dataset comes from a survey about diabetes, examining the effect of various health factors and habits. We will analyze the effect of select factors on the BMI as well as potential interactions between them.

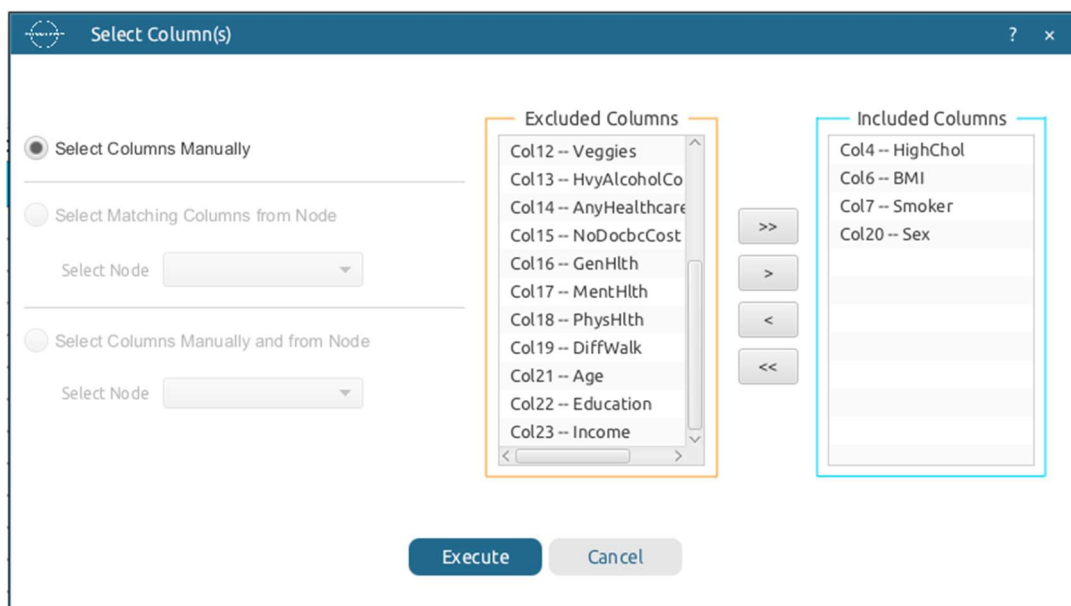
The variables examined are BMI as a dependent variable and Cholesterol, Smoking and Sex (indicators) as factors

Isalos version used: 2.1.2

Dataset: <https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset>

Step 1: Data preparation

Import dataset and rename the tab to “Data”. Go to Data Transformation → Data Manipulation → Select Column(s), transfer “Col4 -- HighCol”, “Col6 -- BMI”, “Col7 -- Smoker”, “Col20 -- Sex” to the “Included Columns” and click “Execute”



	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)
User Header	User Row ID	HighChol	BMI	Smoker	Sex
1		1.0	40.0	1.0	0.0
2		0.0	25.0	1.0	0.0
3		1.0	28.0	0.0	0.0
4		0.0	27.0	0.0	0.0
5		1.0	24.0	0.0	0.0
6		1.0	25.0	1.0	1.0
7		0.0	30.0	1.0	0.0
8		1.0	25.0	1.0	0.0
9		1.0	30.0	1.0	0.0
10		0.0	24.0	0.0	1.0
11		0.0	25.0	1.0	1.0
12		1.0	34.0	1.0	0.0
13		0.0	26.0	1.0	0.0
14		1.0	28.0	0.0	0.0
15		1.0	33.0	1.0	0.0
16		0.0	33.0	0.0	0.0
17		1.0	21.0	0.0	0.0
18		0.0	23.0	1.0	1.0
19		0.0	23.0	0.0	0.0
20		1.0	28.0	0.0	1.0
21		1.0	22.0	0.0	0.0
22		1.0	38.0	1.0	0.0
23		0.0	28.0	1.0	1.0

Step 2: Full Factorial ANOVA

Create a new tab named “Full Factorial” and import the spreadsheet from the “Data” tab. To perform the analysis, go to **Statistics** → **Analysis of (Co)Variance** → **Factorial ANOVA** and enter “95” for **Confidence Level (%)**, “BMI” as **Dependent Variable**, move all columns to **Factors** and select “Full Factorial” as **Formula Type**. Then, click **Execute**

Factorial ANOVA

Confidence Level (%) Double (0,100), Default: -

Dependent Variable Col3 -- BMI

Sum of Squares for Tests Adjusted (Type III)

Coding for Factors (-1, 0, +1)

Excluded Columns

Factors

Col2 -- HighChol
Col4 -- Smoker
Col5 -- Sex

Specify Reference Le...

Formula Type: ☐ Custom ☐ Main Effects ☒ Full Factorial

Specify Custom Model Formula

Formula: HighChol + Smoker + Sex + HighChol:Smoker + HighChol:Sex + Smoker:Sex

Execute Cancel

The results are shown in the right spreadsheet:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (S)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant
1		HighChol	1	112683.2347913	112683.2347913	2615.1102945	0.0	Yes
2		Smoker	1	0.6246393	0.6246393	0.0144964	0.9041656	No
3		Sex	1	14005.7690741	14005.7690741	325.0406412	0.0	Yes
4		HighChol*Smoker	1	58.5864854	58.5864854	1.3596532	0.2435987	No
5		HighChol*Sex	1	4754.3852249	4754.3852249	110.3379910	0.0	Yes
6		Smoker*Sex	1	13.6788648	13.6788648	0.3174540	0.5731422	No
7		HighChol*Smoker*Sex	1	116.7219060	116.7219060	2.7088383	0.0997949	No
8		Error	253672	10930545.2989466	43.0892858			
9		Total	253679	11079389.4947757				

Step 3: Reduced Factorial ANOVA

Create a new tab named “Reduced” and import the spreadsheet from the “Data” tab. To perform the analysis, go to **Statistics** → **Analysis of (Co)Variance** → **Factorial ANOVA** and enter “95” for **Confidence Level (%)**, “BMI” as **Dependent Variable**, move all columns to **Factors** and select “Custom” as **Formula Type**. Click **Specify Custom Model Formula**, select all columns from **Included Columns** and click **Add interactions up to order 2**, then, click **Save**. Click **Execute**

Custom Formula Creation

Included Columns

- HighChol
- Smoker
- Sex

Add interactions
up to order **2**

Add terms
up to order **2**

Add Custom Interaction

Delete Selected

Reset Selections

Included Terms

- HighChol
- Smoker
- Sex
- HighChol:Smoker
- HighChol:Sex
- Smoker:Sex

Move up **Move down**

HighChol + Smoker + Sex + HighChol:Smoker + HighChol:Sex + Smoker:Sex

Save **Cancel**

Factorial ANOVA

Confidence Level (%)

Dependent Variable

Sum of Squares for Tests

Coding for Factors

Excluded Columns

Factors

Col2 -- HighChol
Col4 -- Smoker
Col5 -- Sex

Specify Reference Le...

Formula Type: ☒ Custom ☐ Main Effects ☐ Full Factorial

Specify Custom Model Formula

Formula:

The results are shown in right spreadsheet:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (S)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant
1		HighChol	1	114073.1910108	114073.1910108	2647.3500441	0.0	Yes
2		Smoker	1	0.6651405	0.6651405	0.0154362	0.9011233	No
3		Sex	1	14291.3538676	14291.3538676	331.6661519	0.0	Yes
4		HighChol*Smoker	1	76.3305060	76.3305060	1.7714379	0.1832054	No
5		HighChol*Sex	1	4880.6666751	4880.6666751	113.2679206	0.0	Yes
6		Smoker*Sex	1	5.3030310	5.3030310	0.1230699	0.7257288	No
7		Error	253673	10930662.0208526	43.0895760			
8		Total	253679	11079389.4947757				

Step 4: Main Effects Factorial ANOVA

Create a new tab named “Main Effects” and import the spreadsheet from the “Data” tab. To perform the analysis, go to **Statistics** → **Analysis of (Co)Variance** → **Factorial ANOVA** and enter “95” for **Confidence Level (%)**, “BMI” as **Dependent Variable**, move all columns to **Factors** and select “Main Effects” as **Formula Type**. Click **Execute**

Factorial ANOVA

Confidence Level (%)

Dependent Variable

Sum of Squares for Tests

Coding for Factors

Excluded Columns

Factors

>>
 Custom ☒ Main Effects ☐ Full Factorial

Specify Custom Model Formula

Formula:

The results are shown in the right spreadsheet:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (S)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant
1		HighChol	1	122093.0594546	122093.0594546	2832.1761100	0.0	Yes
2		Smoker	1	2.3764677	2.3764677	0.0551266	0.8143715	No
3		Sex	1	17226.9063006	17226.9063006	399.6102046	0.0	Yes
4		Error	253676	10935788.5058910	43.1092752			
5		Total	253679	11079389.4947757				

Final Isalos Workflow:

```

graph TD
    Data[Data] --> FullFactorial[Full Factorial]
    Data --> Reduced[Reduced]
    Data --> MainEffects[Main Effects]
  
```

5