KURDISTAN REGIONAL GOVERNMENT

MINISTRY OF HIGHER EDUCATION & SCIENTIFIC RESEARCH

SALAHADDINUNIVERSITY – HAWLER (SUH)

COLLEGE OF ADM. & ECO.

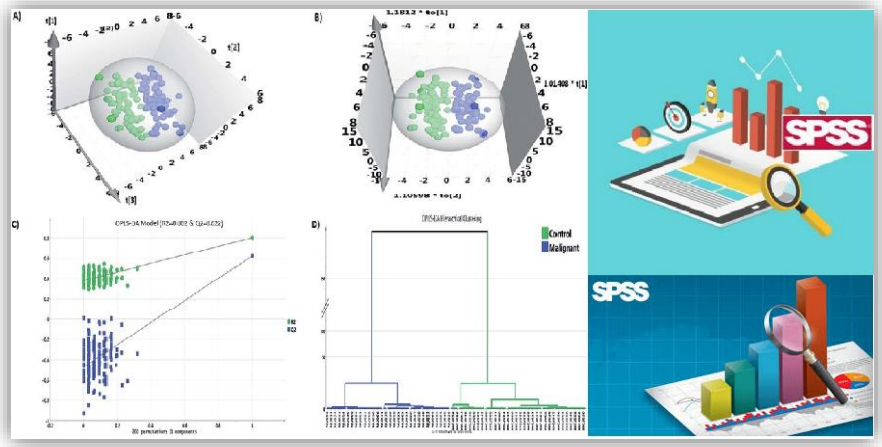
STATISTICS & INFORMATICS

DEPARTMENT

Syllabus of

MULTIVARIATE STATISTICAL ANALYSIS with SPSS

Undergraduate Students (Senior)



by

Dr. Delshad Shaker Ismael Botani & Mr. Sami Ali Obed

E-mail: delshd.botani@su.edu.krd , sami.obed@su.edu.krd

2022 – 2023, Erbil

Chapter One

SPSS Syntax

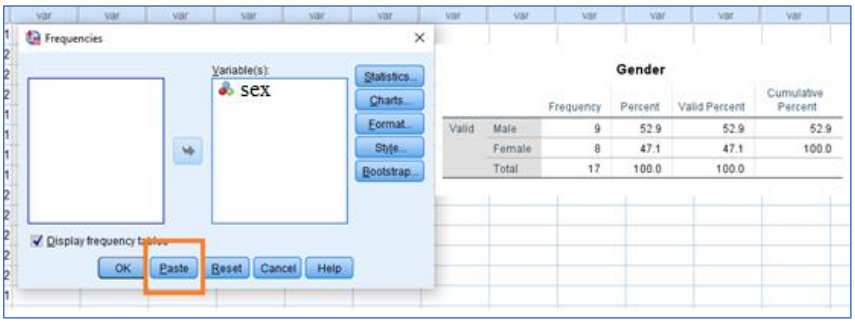
SPSS *syntax*

It is a programming language that is unique to SPSS. It allows you to write commands that run SPSS procedures, rather than using the graphical user interface (GUI).

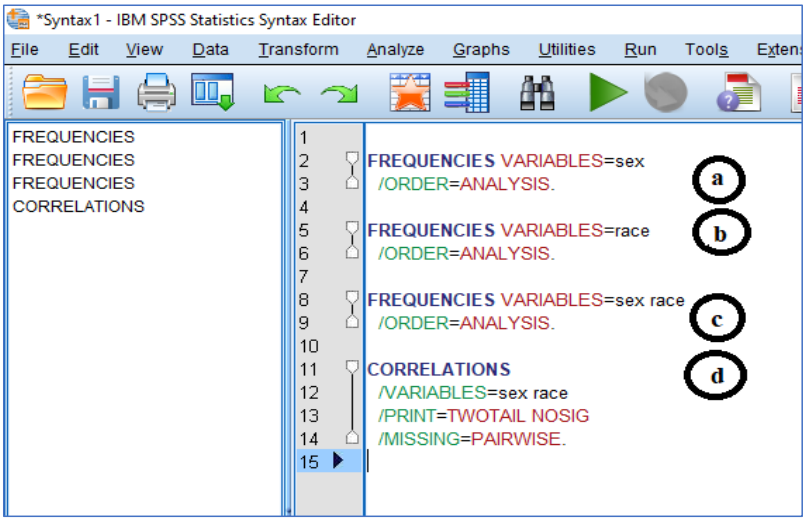
Example1: For the following data (survey_sample.sav), find the following using GUI and syntax (paste) methods.

a. Frequency distribution for SEX variable;
b. Frequency distribution for RACE variable;
c. Frequency distribution for SEX and RACE variables;
d. Correlation between SEX and RACE variables.

Solution: Using GUI method



Using syntax (paste) method:



Example2: For the following data (survey_sample.sav), find frequency distribution for SEX variable using syntax method.

Solution:

**********Frequency**********.**

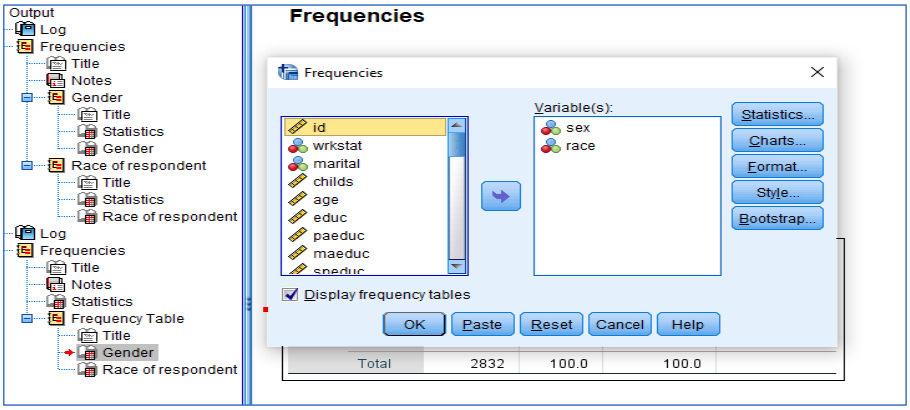
FREQUENCIES VARIABLES = sex.

*The * symbol means what?*

The . (dot) symbol means what?

Example3: For the following data (survey_sample.sav), find frequency distribution for SEX and RACE variables using GUI and syntax methods.

Solution: Using GUI method



Using syntax method:

```
frequencies variables=sex race
  /order=analysis.
```

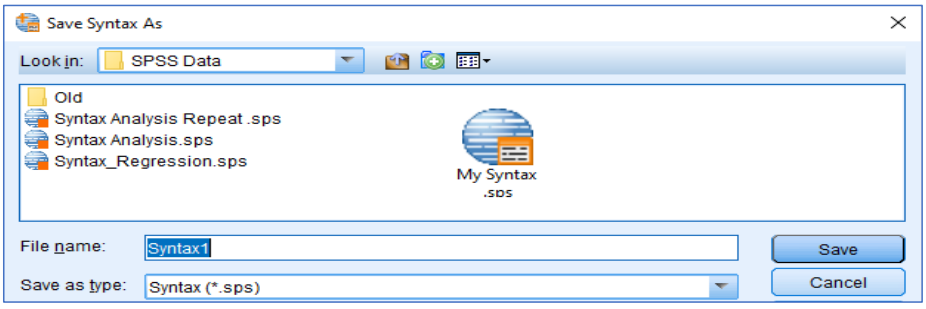
Or

```
frequencies variables=sex race
  /order=variable.
```

The command ORDER refers to what?

The / symbol means what?

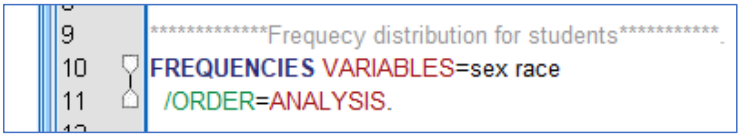
Remark: When saving a syntax file, its extension and icon are different (see below figure).



Color-Coding

SPSS uses color and bolding to indicate the roles of the words in the syntax. The colors are as follows:

Dark blue/purple	Procedure names; execution statements
Green	Statements associated with the given procedure
Dark red/orange	Option keywords
Gray	Comments
Black	Variable names; other text



Variable Labels and Value Labels

The labeling of one or more variables will be accomplished as follows:

VARIABLE LABELS
sex "Gender of Respondent"
age “Age of my family”
race “Race of people in Kurdistan’.

The value labeling of one or more variables will be accomplished as follows:

VALUE LABELS
/sex
0 "Female"
1 "Male"
/race
1 “White”
2 “Black”
3 “Brown“
4 “Other”
/Marital
1 “Married”
2 “Widowed”
3 “Divorced”.

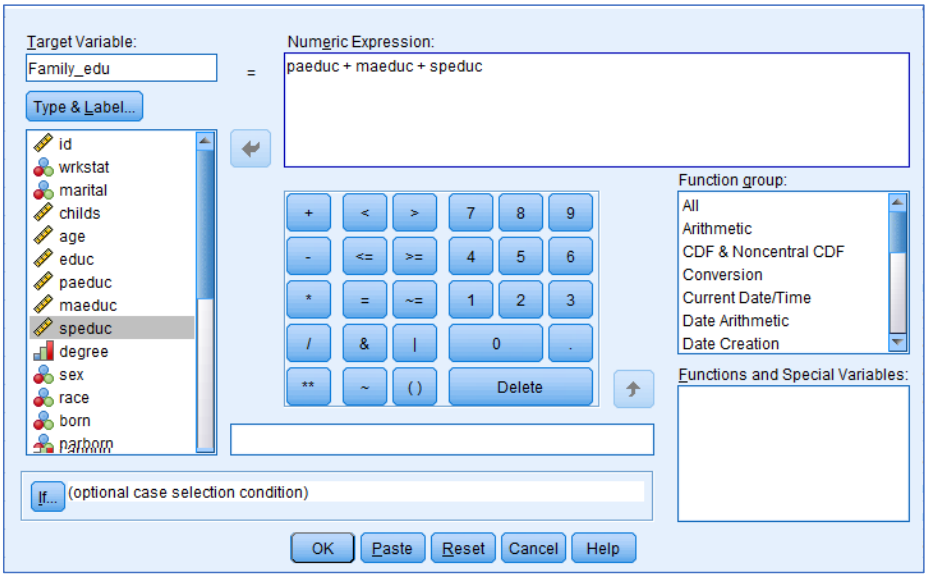
Computing Variables using Syntax

The general form of the syntax for computing a new (numeric) variable is:

Compute y = (x1+x2) / 5.
Execute.

Example4: For the following data (survey_sample.sav), compute new variable from adding the values of the highest year school completed by father, mother,and spouse (**paeduc, maeduc, speduc**) using GUI and syntax methods.

Solution: Using GUI method



Using Syntax method:

```
COMPUTE Family_edu=paeduc + maeduc + speduc.
EXECUTE.
```

Remark: The following table describes the arithmetic operators, arithmetic functions, and logical operations.

Arithmetic operations	Arithmetic functions	Comparisons and logical operations
+ Addition	ABS (expr)	EQ or = Equal to
- Subtraction	RND (expr)	NE or ~= Not equal to
* Multiplication	TRUNC (expr)	GE or >= Greater than or equal to
/ Division	MOD (expr, divid)	GT or > Greater than
** Exponentiation	SQRT (expr)	LE or <= Less than or equal to
	LN (expr)	LT or < Less than

Homework1: Find the mean, variance, log, square root, and cubic of the new variable (Family_edu) in example 4.

Chapter Two

Matrices and Linear Algebra.

Matrix Operations:

To start a Matrix session, the first command must be: **Matrix.**
To end a Matrix session, the final command must be: **End Matrix.**
At any time, to display a vector or matrix or results of some computation in the output:
Print vector_matrix_name.
To create a vector or matrix, use the compute command. Vectors and matrices are enclosed in braces {}. The elements of each row are separated by commas, and rows are separated by semicolons. For example: **Compute name= {a1, a2, a3, a4, a5} or**
Compute name= {a1; a2; a3; a4; a5}

Example5: Write the following Matrices in SPSS syntax

$$X = [1 \quad 2 \quad 3 \quad 4], Y = \begin{bmatrix} 7 \\ 8 \\ 9 \\ 10 \\ 11 \end{bmatrix}, A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{bmatrix}$$

Solution:

***** X is a 1 X 4 matrix********.
Matrix.
compute x = {1, 2, 3, 4}.
print x
/ title “Matrix X”.
End Matrix.

***** Y is a 5 X 1 matrix********.
Matrix.
compute y = {7; 8;9;10;11}.
print y
/ title "Matrix Y".
End Matrix.

***** A is a 3 X 3 matrix********.
Matrix.
compute A = {1,2,3;2,4,6;3,6,1}.
print A
/ title "Matrix A".
End Matrix.

Example6: Write the following Matrix operations for matrix A and B in SPSS syntax

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

a.) A + B b.) A – B c.) A*B d.) A^2

Solution:

a.)
***** A + B ********.
Matrix.
compute A = {1,2,3;2,4,6;3,6,1}.
compute B = {1,0,0;0,2,0;0,0,3}.
compute Z = A + B.
print A / title "Matrix A".
print B / title "Matrix B".
print Z / title "Matrix A + B".
End Matrix.

b.) & c.) & d.)

***** A - B & A*B & A^2********.
Matrix.
compute A = {1,2,3;2,4,6;3,6,1}.
compute B = {1,0,0;0,2,0;0,0,3}.
compute Z = A - B.
compute M = A * B.
compute K = A * A.
print A / title "Matrix A".
print B / title "Matrix B".
print Z / title "Matrix A - B".
print M / title "Matrix A * B".
print K / title "Matrix A^2".
End Matrix.

Example7: Using SPSS syntax, find A', |A|, and A^{-1} .

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix}$$

Solution:

***** Transpose, Determinant, & Inverse********.
Matrix.
compute A = {1,2,3;0,4,0;3,6,1}.
compute Trans_A =transpos (A).
compute det_A = det(A).
compute halg_A = inv (A).
print Trans_A / title "Transpose of Matrix A".
print det_A / title " Determinant of Matrix A".
print halg_A / title "Inverse of Matrix A ".
End Matrix.

Homework2: Find for the following matrix: A', det(A), A^{-1}, $\sqrt{A}$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 3 & 6 & 1 \end{bmatrix}$$

Opening SPSS Data Files and Variables:

When we select File⇨Open⇨Data, SPSS issues a "GET" command to open an SPSS-formatted file and load it into SPSS. For example, the following opens and loads the file named survey_sample.sav:

```
GET FILE = "C:\Program Files\IBM\SPSS\Statistics\26\Samples\English\survey_sample.sav".
```

Example8: Open survey_sample.sav data, find a matrix (X) consists of respondent, father, mother, and spouse education (**educ, paeduc, maeduc, speduc**) using syntax methods. Then find $X'X$.

Solution:

```
GET FILE = "C:\Program Files\IBM\SPSS\Statistics\26\Samples\English\survey_sample.sav".
Matrix.
 /* the following creates a matrix with the n rows and  p=4 columns.
get x
 /variables = educ, paeduc, maeduc, speduc
 /missing=accept
 /sysmis=omit.
* /sysmis=15.
compute z=transpos(x)*x.
print x / title "Matrix X".
print z / title "Matrix X'X".
End Matrix.
```

Remark: in case we have missing values, they must be **omitted or accepted**, while when we have system missing values, they must be **omitted** or **changing** them to a number such as 15 or 20 or any other numbers.

Example9: Using survey_sample.sav data, find a matrix (X) consists of father, mother, and spouse education (**paeduc, maeduc, speduc**), and a matrix (Y) consists of respondents' education (**educ**) using syntax methods. Then find $(X'X)^{-1}X'Y$.

Solution:

```
Matrix.
 /* the following creates a matrix with the n rows and  p=3 columns.
get x
 /variables = paeduc, maeduc, speduc
 /missing=accept
 /sysmis=omit.
 /* the following creates a matrix with the n rows and  p=1 columns.
get y
/variables = educ
 /missing=accept
 /sysmis=omit.
compute z=transpos(x)*x.
compute inv_z=inv(z).
compute xtrans_y=transpos(x)*y.
compute B=inv_z*xtrans_y.
print x / title "Matrix X".
print z / title "Matrix X'X".
```

```
print xtrans_y / title "Matrix X'Y".
print inv_z / title "Inverse of Matrix (X'X)".
print B / title "Matrix (X'X)-1X'Y".
End Matrix.
```

Question: Is there any mistake in the above example?

Commented [s1]: Yes, in X matrix we did not add a colu or variable with values 1s.

Homework3: Using car_sales.sav data, find a matrix (X) consists of Engine size, Horsepower, Fuel capacity, and Fuel efficiency (**engine_s, horsepow, fuel_cap, mpg**), and a matrix (Y) consists of Price in thousands (**price**) using syntax methods. Then find $(X'X)^{-1}X'Y$.

Finding Mean and Variance-Covariance matrix

In order to find Mean matrix and Variance-Covariance matrix, the following equations will be used:

$$\bar{\boldsymbol{X}} = \mathbf{1}\mathbf{1}'\boldsymbol{X}\left(\frac{\mathbf{1}}{\mathbf{n}}\right)$$, where **1** is a matrix of ones (n×1)

$$\boldsymbol{S} = \frac{\left(\left(\boldsymbol{X}-\mathbf{1}\mathbf{1}'\boldsymbol{X}\left(\frac{\mathbf{1}}{\mathbf{n}}\right)\right)'\left(\boldsymbol{X}-\mathbf{1}\mathbf{1}'\boldsymbol{X}\left(\frac{\mathbf{1}}{\mathbf{n}}\right)\right)\right)}{n-1}$$

Example10: Open survey_sample.sav data, find a matrix (X) consists of father, mother, and spouse education (**paeduc, maeduc, speduc**) using syntax methods. Then find mean and var-covar matrix of X.

Solution:

```
Matrix.
get one
 /variables = one.
get x
 /variables = paeduc, maeduc, speduc
 /missing=accept
 /sysmis=omit.
get y
/variables = educ
 /missing=accept
 /sysmis=omit.
compute n=nrow (x).
compute xbar=transpos(one)*x/n.
compute x_xbar=x-one*xbar.
compute var_x=(transpos(x_xbar)*x_xbar)/n-1.
print n / title "Number of Rows".
print xbar / title "X bar".
print x_xbar / title "X - X Bar".
print var_x/ title "Variance of X Matrix".
End Matrix.
```

Homework4: Using car_sales.sav data, find a matrix (X) consists of Engine size, Horsepower, Fuel capacity, and Fuel efficiency (**engine_s, horsepow, fuel_cap, mpg**) using syntax methods. Then find mean, var-covar, and correlation matrix of X.

Finding Correlation matrix

In order to find the correlation matrix, the following equations will be used:

$$\boldsymbol{R = D^{-1/2} S D^{-1/2}} \quad ,$$

Where D is the diagonal matrix of var-covar matrix. In other words, $\boldsymbol{D^{-1/2}}$ is a diagonal matrix with 1/sd for each variable as the diagonal elements.

To use SPSS syntax for finding correlation matrix, the following commands should be used:

diag: Is the diagonal elements of a matrix;

mdiag: Is the diagonal matrix of the diagonal elements.

Example11: Open survey_sample.sav data, find correlation matrix father, mother, and spouse education (**paeduc, maeduc, speduc**) using syntax methods.

Solution:

```
compute one=1.
execute.
Matrix.
 get one
 /variables = one.
get x
 /variables = paeduc, maeduc, speduc
 /missing=accept
 /sysmis=omit.
compute n=nrow (x).
compute xbar=transpos(one)*x/n.
compute x_xbar=x-one*xbar.
compute var_x=(transpos(x_xbar)*x_xbar)/n-1.
compute diag_var_x=diag(var_x).
compute sqrt_diag_var_x=sqrt(diag(var_x)).
compute mat_sqrt_diag_var_x=mdiag(sqrt_diag_var_x).
compute corr_x=inv(mat_sqrt_diag_var_x)*var_x*inv(mat_sqrt_diag_var_x).
print n / title "Number of Rows".
print xbar / title "X bar".
print x_xbar / title "X - X Bar".
print var_x/ title "Variance of X Matrix".
print sqrt_diag_var_x/ title "Sqrt Diagonal of Variance X Matrix".
print mat_sqrt_diag_var_x/ title "Matrix of Sqrt Diagonal of Variance X Matrix".
print corr_x  /title="R, the correlation matrix".
End Matrix.
```