



LINEAR REGRESSION AND MULTIPLE LINEAR REGRESSION USING SPSS

REGRESSION ANALYSIS

- Predict the value of a dependent variable based on the value of at least one independent variable
- Explain the impact of changes in an independent variable on the dependent variable

LINEAR REGRESSION

- Only one independent variable, X
- Relationship between X and Y is described by a linear function
- Changes in Y are assumed to be caused by changes in X

$$Y = B_0 + B_1X$$

Where, B_0 = constant

B_1 = regression coefficient

X = value of the independent variable

Y = value of the dependent variable.

MULTIPLE LINEAR REGRESSION

- Used to predict a relation between more than one independent variables and a dependent variable

$$Y = b_0 + b_1x_1 + b_2x_2 + \dots + b_{k-1}x_{k-1} + b_kx_k$$

where b_0 = constant

b_i = regression coefficients

x_i = value of the independent variable

Y = value of the dependent variable.



PROCEDURE

LINEAR REGRESSION USING SPSS

Question: A salesperson for a large car brand wants to determine whether there is a relation between an individual's income and the price they pay for a car.

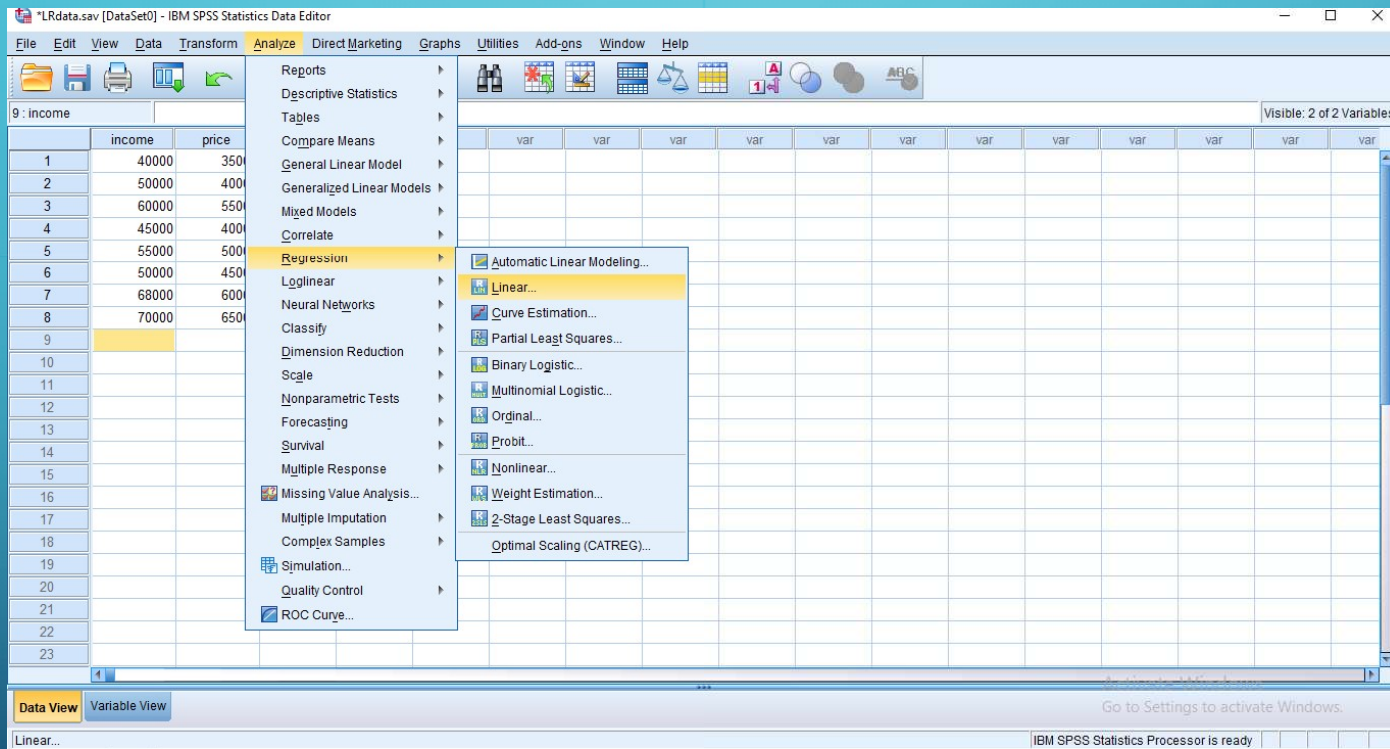
STEP1: ENTER THE INCOME AND PRICE DATA IN SPSS

The screenshot displays the IBM SPSS Statistics Data Editor interface. The window title is "*LRdata.sav [DataSet0] - IBM SPSS Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations, data manipulation, and analysis. The main data grid shows two columns: 'income' and 'price'. The data is as follows:

	income	price	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	40000	35000															
2	50000	40000															
3	60000	55000															
4	45000	40000															
5	55000	50000															
6	50000	45000															
7	68000	60000															
8	70000	65000															
9																	
10																	
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	

The interface also shows a status bar at the bottom with the text "IBM SPSS Statistics Processor is ready" and a message "Go to Settings to activate Windows." The "Data View" tab is selected.

STEP 2: CLICK ANALYZE AND SELECT REGRESSION & THEN CLICK LINEAR



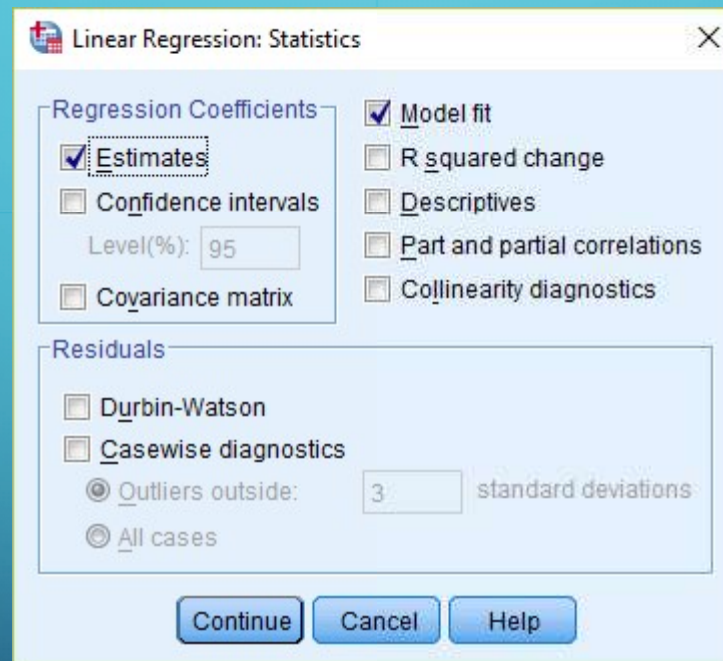
STEP 3: TRANSFER THE INDEPENDENT VARIABLE INCOME INTO THE INDEPENDENT BOX & TRANSFER THE DEPENDENT VARIABLE PRICE INTO THE DEPENDENT BOX

The image shows a screenshot of the 'Linear Regression' dialog box in SPSS. The dialog box is titled 'Linear Regression' and has a close button (X) in the top right corner. On the left side, there is a list of variables, with 'Income of the pers...' selected. In the center, there are several sections for configuring the regression model:

- Dependent:** A field containing 'price of the car [price]'.
- Block 1 of 1:** A section containing 'Previous' and 'Next' buttons.
- Independent(s):** A field containing 'Income of the person [inco...]'.
- Method:** A dropdown menu set to 'Enter'.
- Selection Variable:** An empty field with a 'Rule...' button.
- Case Labels:** An empty field.
- WLS Weight:** An empty field.

On the right side of the dialog box, there are several buttons: 'Statistics...', 'Plots...', 'Save...', 'Options...', and 'Bootstrap...'. At the bottom of the dialog box, there are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

STEP 4: IN STATISTICS FEATURE SELECT ESTIMATE AND MODEL FIT IN REGRESSION COEFFICIENTS AND THEN CLICK CONTINUE TO GET THE OUTPUT.



OUTPUT

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help

7 METHOD ENTER INCOME

Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.984 ^a	.967	.962	2068.710

a. Predictors: (Constant), Income of the person

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	761822621.3	1	761822621.3	178.014	.000 ^b
	Residual	25677378.70	6	4279563.117		
	Total	787500000.0	7			

a. Dependent Variable: price of the car
b. Predictors: (Constant), Income of the person

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4896.030	4086.762		-1.198	.276
	Income of the person	.980	.073	.984	13.342	.000

a. Dependent Variable: price of the car

Activate Windows
Go to Settings to activate Windows.

IBM SPSS Statistics Processor is ready

INTERPRETATION

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.984 ^a	.967	.962	2068.710

a. Predictors: (Constant), Income of the person

- The R value represent the simple correlation and is 0.984 which indicates high degree of correlation.
- The R² indicates how much of the total variation in the dependent variable can be explained by the independent variable, In this case 96.7 which is large.

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4896.030	4086.762		-1.198	.276
	Income of the person	.980	.073	.984	13.342	.000

a. Dependent Variable: price of the car

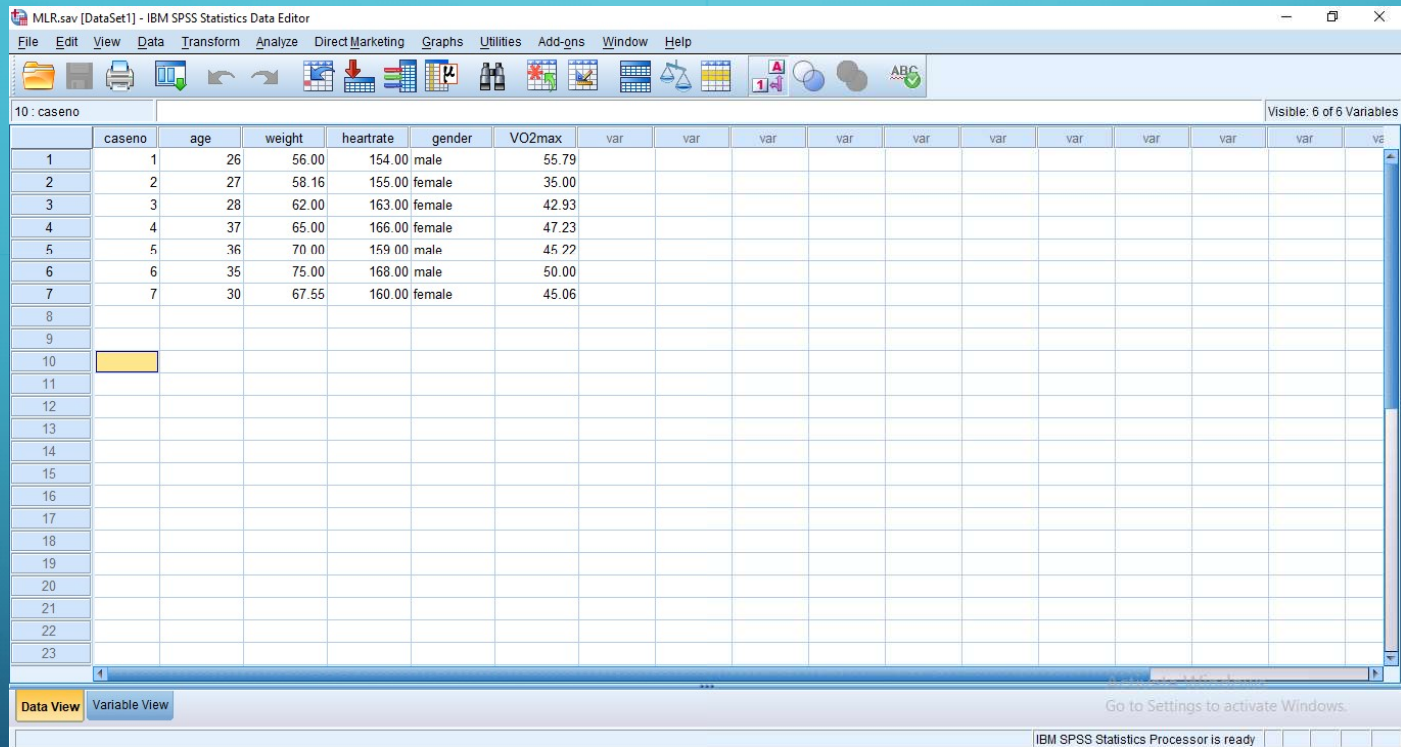
**To present the regression equation as
Price = 4896 + 0.980 (Income)**

PROCEDURE

MULTIPLE LINEAR REGRESSION USING SPSS

- Question: A health researcher wants to predict VO₂ max, an indicator of fitness and health.
- The researchers goal is to predict vo₂ max based on these four attributes: age, weight, heartrate, gender.

STEP 1: ENTER THE ATTRIBUTES IN SPSS

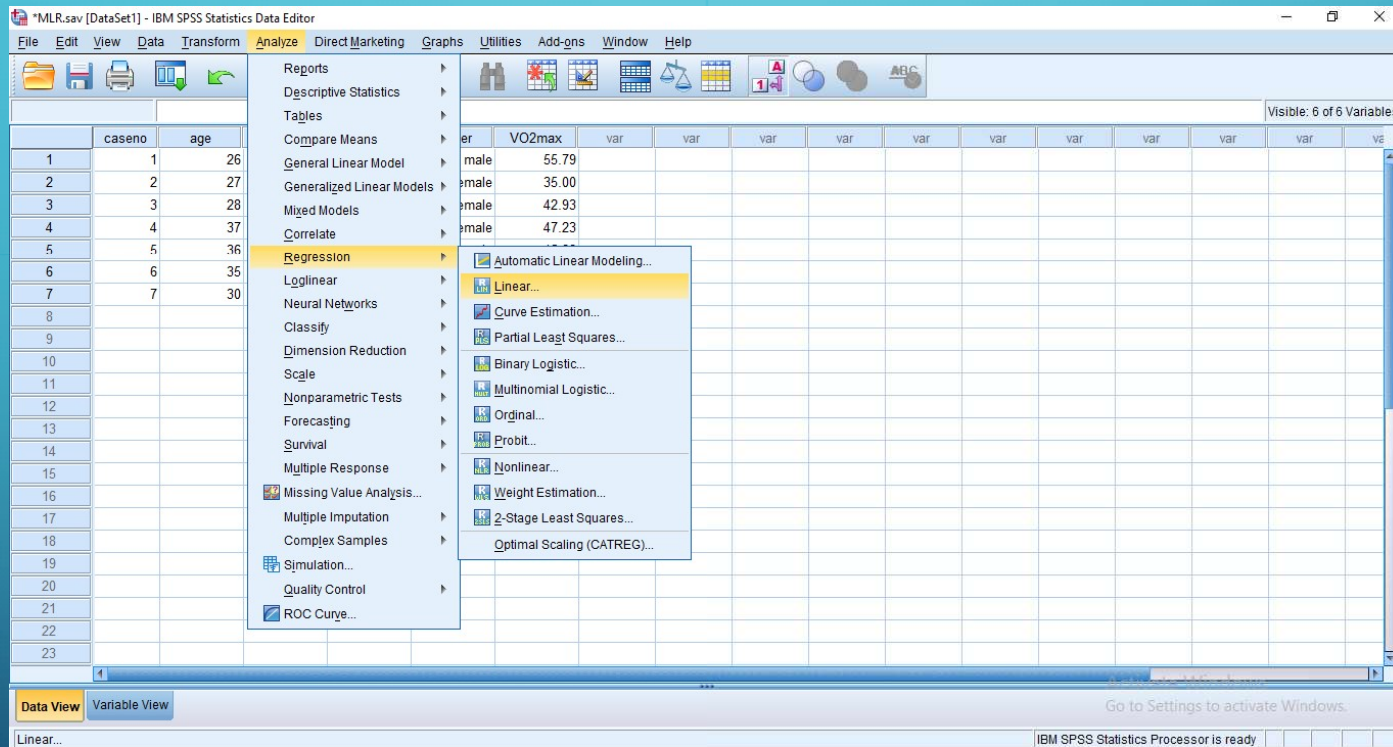


The screenshot displays the IBM SPSS Statistics Data Editor window. The title bar reads "MLR.sav [DataSet1] - IBM SPSS Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations, editing, and analysis. The main data grid shows 10 cases and 6 variables. The first six variables are: caseno, age, weight, heartrate, gender, and VO2max. The remaining six variables are labeled "var". The data is as follows:

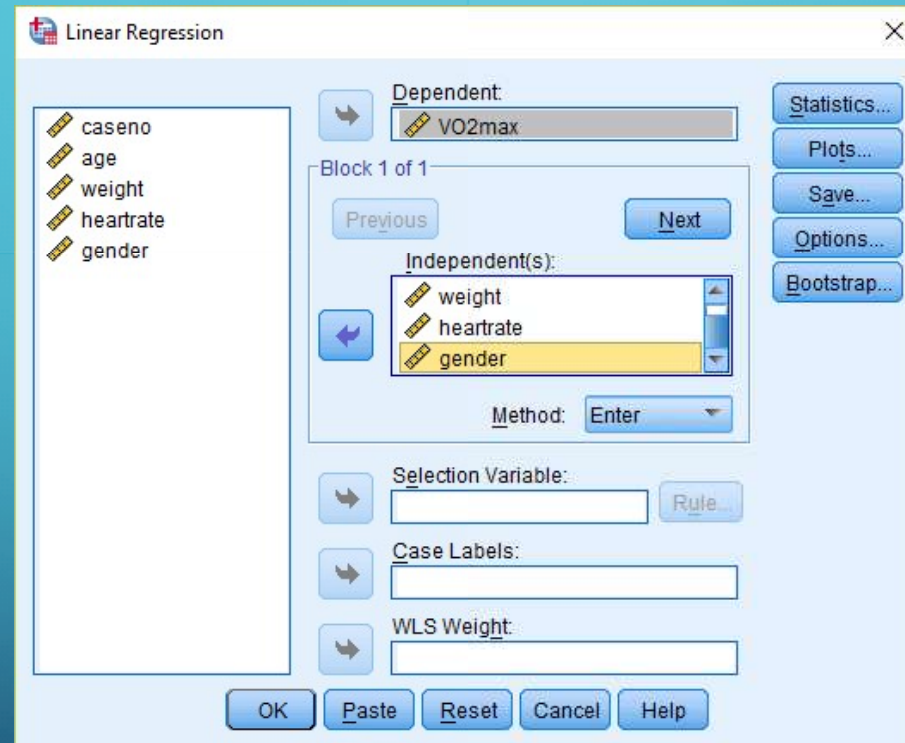
	caseno	age	weight	heartrate	gender	VO2max	var	var	var	var	var	var	var	var	var
1	1	26	56.00	154.00	male	55.79									
2	2	27	58.16	155.00	female	35.00									
3	3	28	62.00	163.00	female	42.93									
4	4	37	65.00	166.00	female	47.23									
5	5	36	70.00	159.00	male	45.22									
6	6	35	75.00	168.00	male	50.00									
7	7	30	67.55	160.00	female	45.06									
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															

The status bar at the bottom indicates "Data View" is selected, "Variable View" is also available, and "IBM SPSS Statistics Processor is ready".

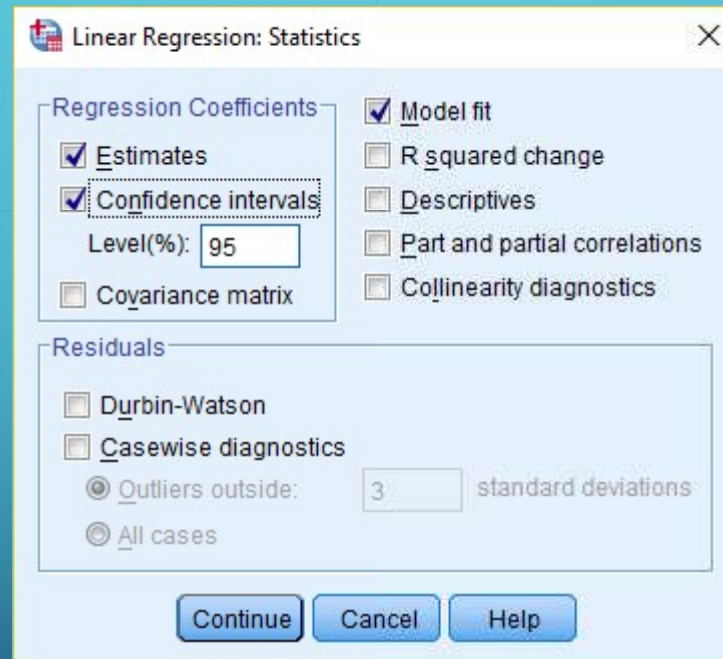
STEP 2: CLICK ANALYSE AND SELECT REGRESSION & THEN SELECT LINEAR.



STEP 3: TRANSFER THE DEPENDENT VARIABLE
VO2 MAX INTO THE DEPENDENT BOX
TRANSFER THE INDEPENDENT VARIABLES AGE,
GENDER, HEARTRATE INTO INDEPENDENT BOX



STEP 4: IN STATISTICS FEATURE SELECT ESTIMATE MODEL FIT & CONFIDENCE INTERVALS IN REGRESSION COEFFICIENTS AND THEN CLICK CONTINUE TO GET THE OUTPUT



OUTPUT

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Output
 Log
 Regression
 Title
 Notes
 Active Dataset
 Variables Entered/Removed
 Model Summary
 ANOVA
 Coefficients

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.774 ^a	.599	-.202	7.00910

a. Predictors: (Constant), gender, heartrate, age, weight

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	146.934	4	36.734	.748	.641 ^b
	Residual	98.255	2	49.128		
	Total	245.189	6			

a. Dependent Variable: VO2max
 b. Predictors: (Constant), gender, heartrate, age, weight

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-30.167	112.955		-.267	.814	-516.172	455.837
	age	-.060	1.050	-.043	-.057	.959	-4.576	4.455
	weight	-.600	.826	-.627	-.727	.543	-4.153	2.952
	heartrate	.832	.941	.688	.885	.470	-3.215	4.879
	gender	-10.741	6.333	-.898	-1.696	.232	-37.988	16.506

a. Dependent Variable: VO2max

Activate Windows
 Go to Settings to activate Windows.

Double click to edit Pivot Table

IBM SPSS Statistics Processor is ready

JH:111 W: 390 pt