Regression in SPSS

Example regression analysis in SPSS

Example from the 1974 cars data set to predict fuel efficiency in miles per gallon (mpg) of cars given their mass (wt) in kilograms.

▶ Our **response** or **dependent** variable is miles per gallon
▶ Our **predictor** or **independent** variable is car mass

Before fitting our regression model, we should check our assumptions of linear regression.

Checking assumptions of linear regression

1. **The independent variable X is measured without error.**
   Independent variable (car mass) should be fixed (i.e., measured without error).
2. **The relationship between X and Y is linear.** We can check in SPSS; select 'Graph > Legacy Dialogs > Scatter/Dot', then select 'Simple Scatter' from the available options.
3. **For any value of X, Y is normally distributed.** Values of the dependent variable, miles per gallon, should be normally distributed. We can check this in SPSS using 'Analyse > Descriptive Statistics > Explore', then putting 'mpg' in the Dependent List.
4. **For all values of X, the variance of the residuals is identical**. We can check by plotting the data in SPSS in the same scatter plot that we used to check that the relationship was linear.

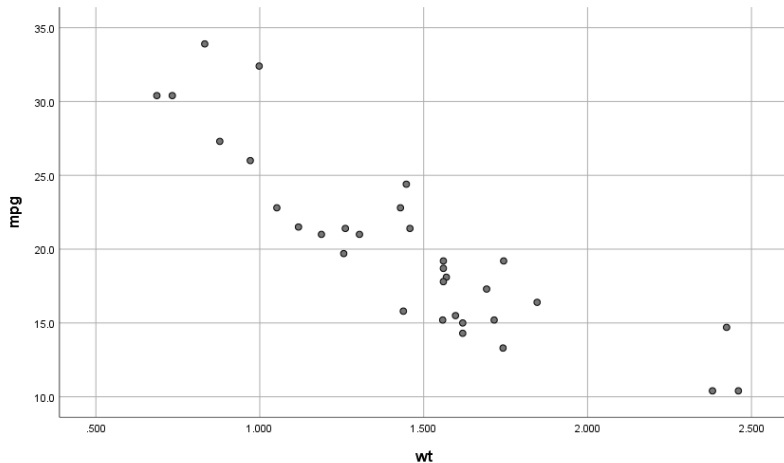# Checking assumptions of linear regression

Figure 1: Scatterplot showing SPSS output for car mass (wt) and car miles per gallon (mpg)

Fitting the linear regression model
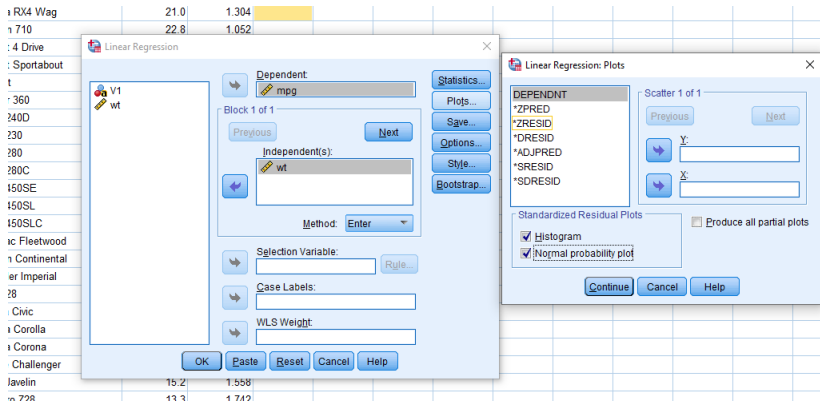
'Analyse > Regression > Linear'

Figure 2: Options in SPSS for creating linear regression plots

Every regression model in SPSS will produce four tables (we can ignore the first for now).

Linear regression: Model summary output

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.868[a]	.753	.745	3.0457

a. Predictors: (Constant), wt

b. Dependent Variable: mpg

Figure 3: Model summary output table in SPSS

Linear regression: ANOVA output

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	847.754	1	847.754	91.388	.000[b]
	Residual	278.293	30	9.276		
	Total	1126.047	31			

a. Dependent Variable: mpg

b. Predictors: (Constant), wt

Figure 4: ANOVA summary output table in SPSS

Linear regression: Coefficients output

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	37.287	1.878		19.858	.000
	wt	-11.785	1.233	-.868	-9.560	.000

a. Dependent Variable: mpg

Figure 5: Coefficients summary output table in SPSS

Assumption of normally distributed residuals

Next, we can look at the output of the residual diagnostics to check our assumption that the residuals are normally distributed.

We can see the Histogram showing the Regression Standardised Residual output.
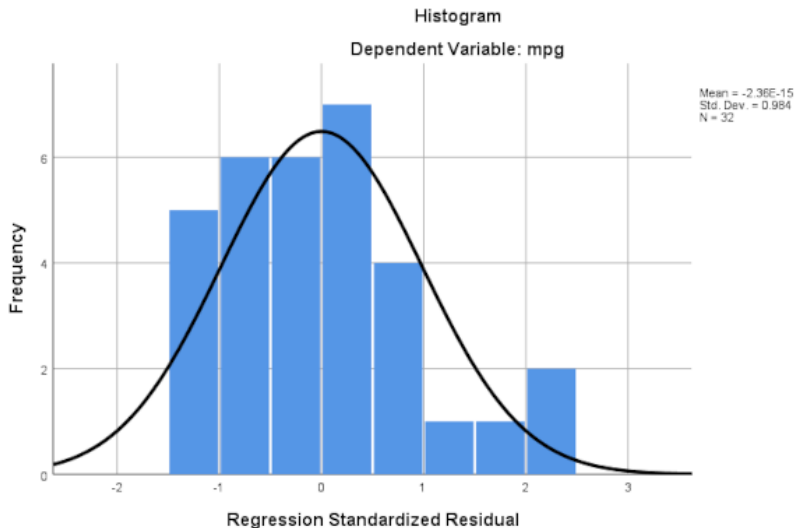
Assumption of normally distributed residuals

Figure 6: SPSS output showing regression residuals.