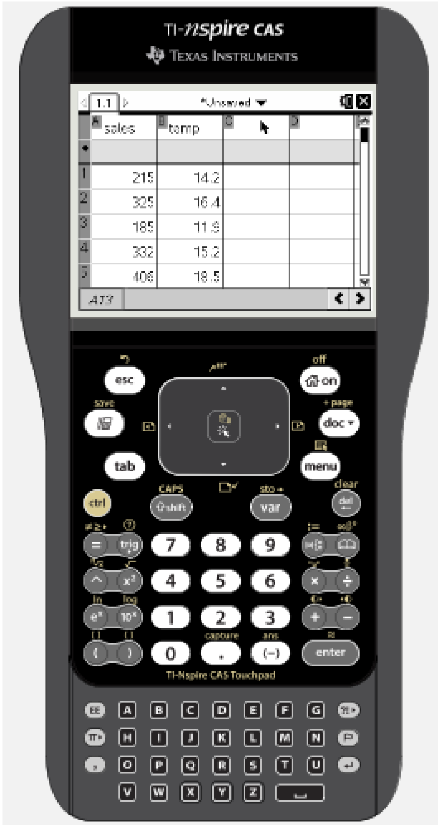
Residual Analysis

You would recall the earlier example of daily ice creams sales versus daily maximum temperature. To date we know how to:

Skill.1 Enter data into TI-nspire

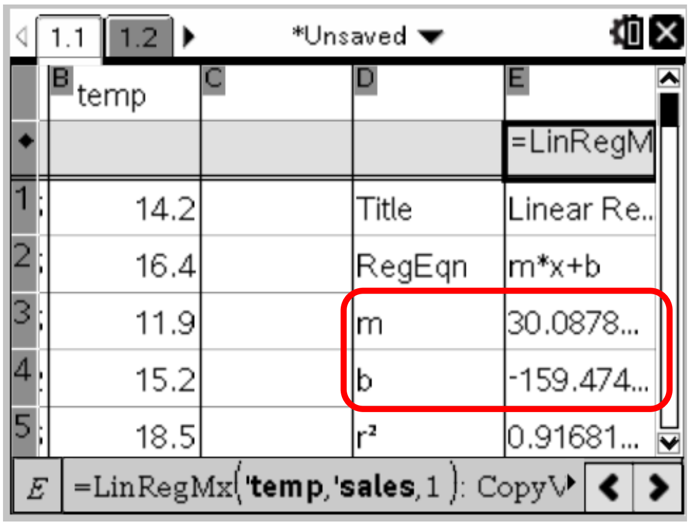
Temperature °C	Ice Cream Sales
14.2	$215
16.4	$325
11.9	$185
15.2	$332
18.5	$406
22.1	$522
19.4	$412
25.1	$614
23.4	$544
18.1	$421
22.6	$445
17.2	$408



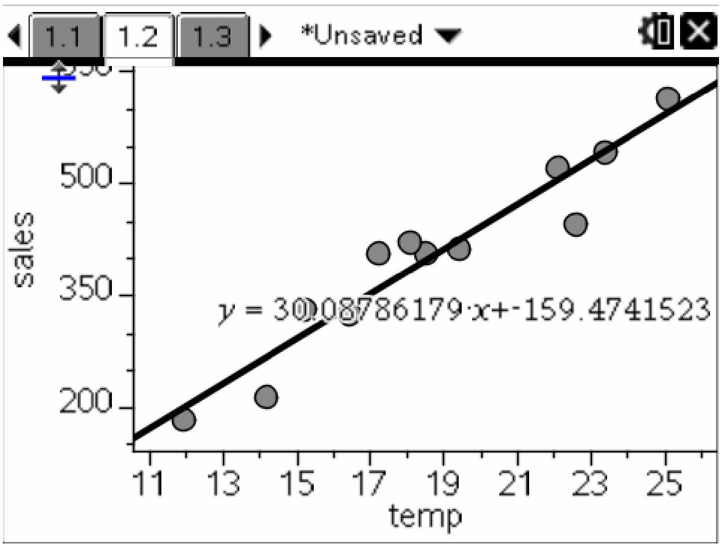
.

Skill.2 Perform a linear regression on the TI-Nspire CAS calculator

Ice cream sales = 30.1 x Daily Temp - 159.5
(Accuracy to 1 decimal place)



Skill.3 Construct a scatterplot using the TI-Nspire CAS calculator



Residuals

Upon the above scatterplot you will notice that not all of the original points (x,y) are located upon the least squares regression line. For this to be true we would need a correlation coefficient (r) = 1, this is not the case.

The difference between each y-value (actual data) and the corresponding predicted y-value (linear regression line) is what is called the **residual**.

Residual = (actual y-value) – (predicted y-value)

Let's consider the residuals in this example and what they tell us about the **linearity of the relationship** between the two variables

Residual table

Temp (°C)	Sales ($)	Sales Pred ($) (Sales = 30.1T – 159.5)	Residual (Sales- Sales Pred)
14.2	215	267.92	-52.92
16.4	325	334.14	-9.14
11.9	185	198.69	-13.69
15.2	332	298.02	**33.98**
18.5	406	397.35	**8.65**
22.1	522	505.71	**16.29**
19.4	412	424.44	-12.44
25.1	614	596.01	**17.99**
23.4	544	544.84	-0.84
18.1	421	385.31	**35.69**
22.6	445	520.76	-75.76
17.2	408	358.22	**49.78**

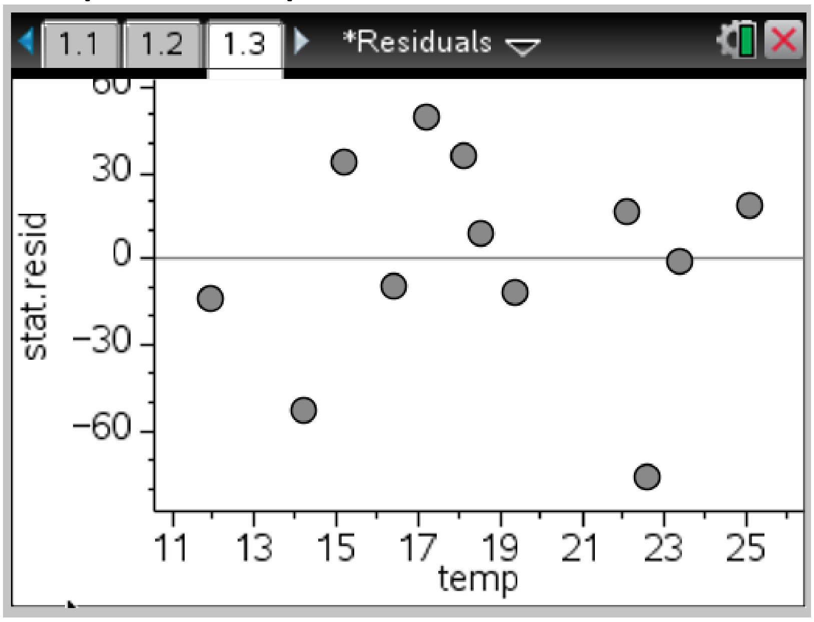
NB: A **positive residual** indicates that the plotted sales point is **above** the regression lines predicted value.
A **negative residual** indicates that the plotted sales point is **below** the regression lines predicted value.

Residual calculations

* Predicted y-value calculations	*Residual (y_{plot}-y_{pred})
(**x=14.2**); Pred y = 30.1 x (**14.2**) - 159.5 = 267.92	215-267.92=-52.92
(**x=16.4**); Pred y = 30.1 x (**16.4**) - 159.5 = 334.14	325-334.14=-9.14
(**x=11.9**); Pred y = 30.1 x (**11.9**) - 159.5 = 198.69	185-198.69=-13.69
(**x=15.2**); Pred y = 30.1 x (**15.2**) - 159.5 = 298.02	332-298.02=33.98
(**x=18.5**); Pred y = 30.1 x (**18.5**) - 159.5 = 397.35	406-397.35=8.65
(**x=22.1**); Pred y = 30.1 x (**22.1**) - 159.5 = 505.71	552-505.71=46.29
(**x=19.4**); Pred y = 30.1 x (**19.4**) - 159.5 = 424.44	412-424.44=-12.44
(**x=25.1**); Pred y = 30.1 x (**25.1**) - 159.5 = 596.01	614-596.01=17.99
(**x=23.4**); Pred y = 30.1 x (**23.4**) - 159.5 = 544.84	544-544.84=-0.84
(**x=18.1**); Pred y = 30.1 x (**18.1**) - 159.5 = 385.31	421-385.31=35.69
(**x=22.6**); Pred y = 30.1 x (**22.6**) - 159.5 = 520.76	445-520.76=-75.76
(**x=17.2**); Pred y = 30.1 x (**17.2**) - 159.5 = 358.22	408-358.22=49.78

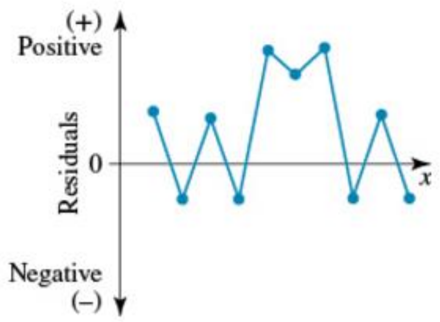
The best way to analyse the residual pattern is to generate a residual plot using your TI-Nspire calculator.

TI-Nspire residual plot for the Ice cream Sales V Max Temp

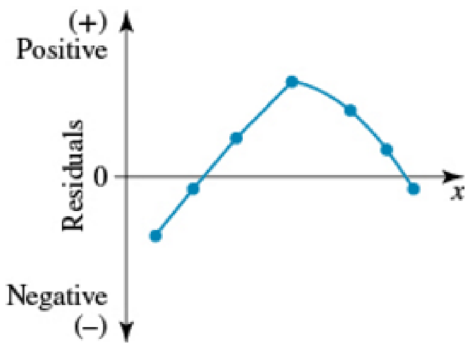


This plot can be used to examine whether a scatter plot has a **linear** or **non-linear** relationship.

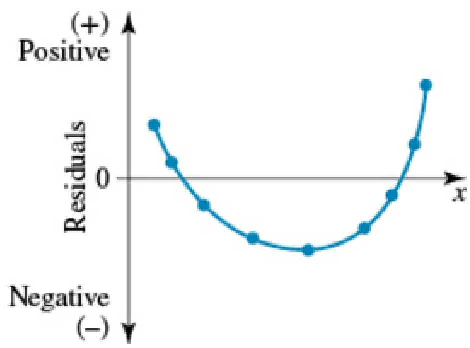
The following classifications can be used when analyzing residuals graphically:



The points of the residuals are randomly scattered above and below the x-axis. The original data probably have a **linear relationship.**



The points of the residuals show a curved pattern (∩), with a series of negative, then positive and back to negative residuals along the x-axis. The original data probably have a **non-linear relationship**. Transformation of the data may be required



The points of the residuals show a curved pattern (∪), with a series of positive, then negative and back to positive residuals along the x-axis. The original data probably have a **non-linear relationship**.
Transformation of the data may be required.

In our above example of Ice cream sales ($) V maximum daily temperature (°C), the residual plot are **randomly scattered above and below the x-axis**. Therefore we can conclude that the original data **probably have a linear relationship**.