Unit 4
Functions and Models

Relation

- Relation is the relationship between two set of numbers
- It can be represented as ordered pairs
- When we do correlation, we find the relationship between two set of numbers(i.e, variables)
- The correlation coefficient explains about the relation.
- It gives an idea to expand it as function.

Functions

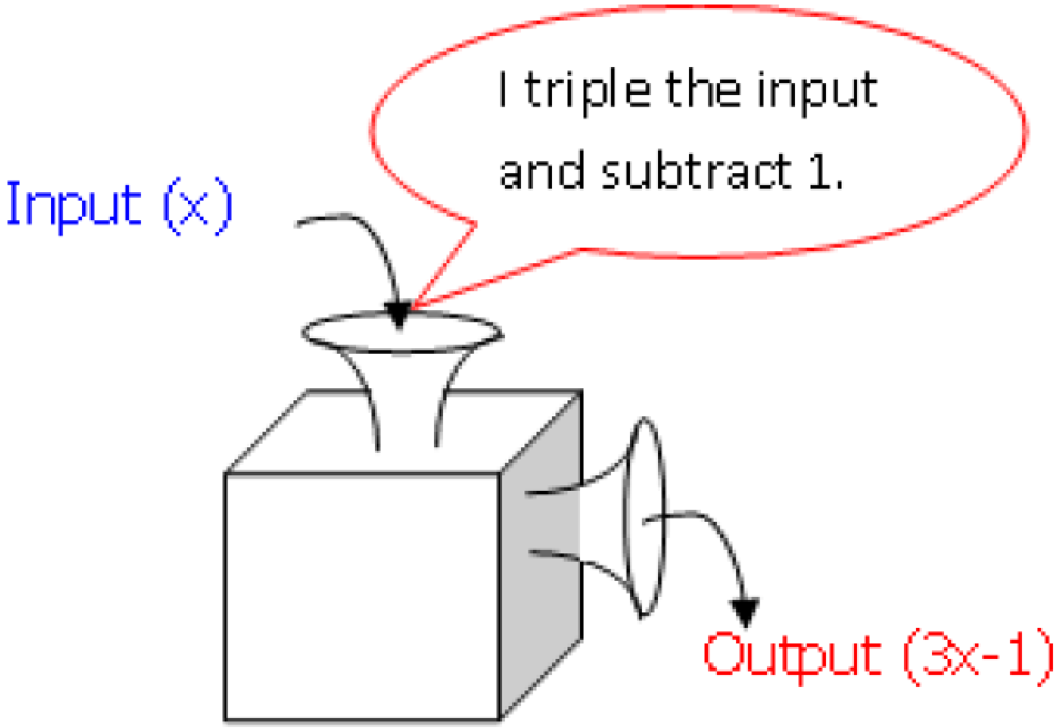
- A function is a set of ordered pairs in which the first coordinate, usually x, matches with exactly one second coordinate, y.
- The y coordinate represents the dependent variable
- A function can be expressed as an equation

$$\underbrace{\overset{\substack{input \\ \downarrow}}{f(x)}}_{\substack{function \\ box}} = y \leftarrow output$$

- f(x) – f is the name of the function and x is the name of the independent variable

Function

- A function is a relationship where each input number corresponds to one and only one output number

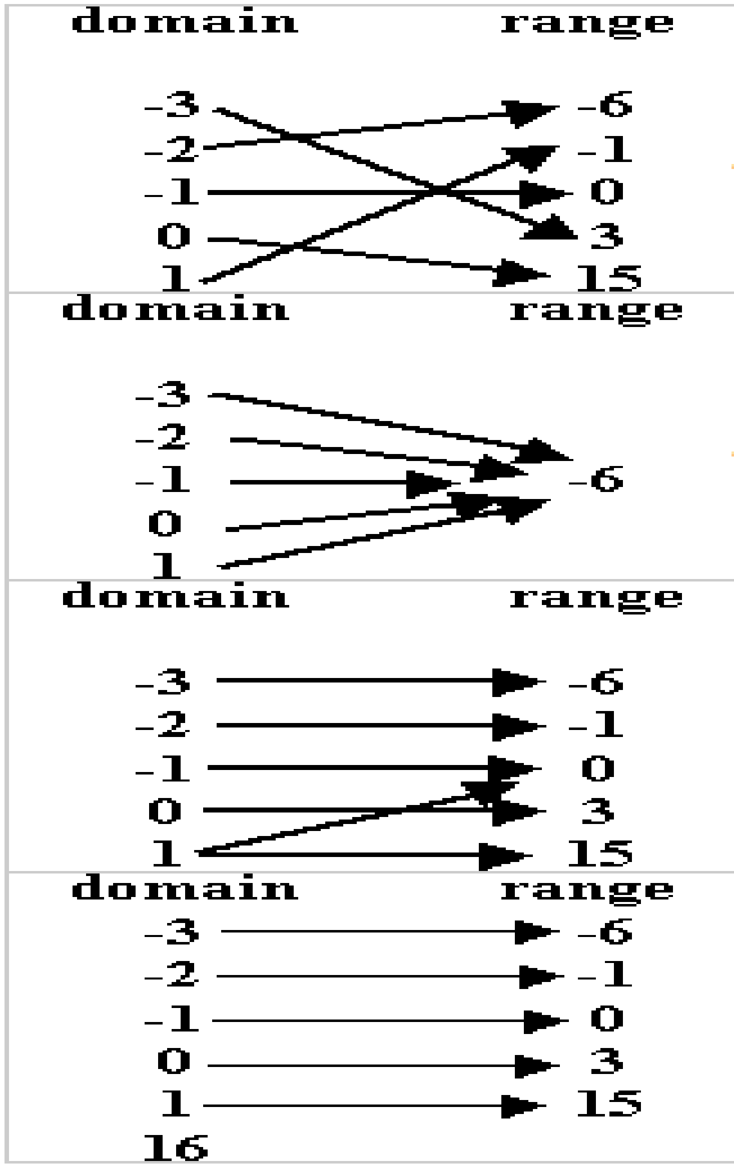


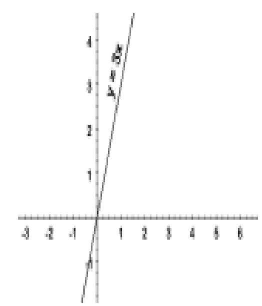
Functions as Relations

- Situation 1: You are selling candy bars for a school fundraiser. Each candy bar costs Rs.3
- Situation 2: You collect data from several students in your class on their ages and their heights

  (18, 65"), (17,64"), (18,67"), (18, 68"), (17,66")
- In Situation 1, for each different number of candy bar sales you input, there is one and only one output number representing your profit.
- In Situation 2, if you input "18 years", there are multiple outputs, so you can’t identify a specific relationship between age and height.

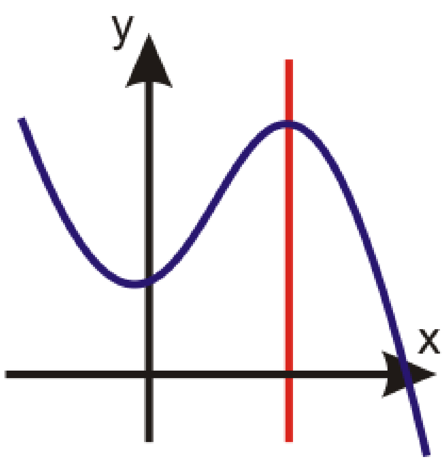
- Relation - simply a relationship between two sets of numbers or data.
- Function - every x is paired with only one y. Functions are well-behaved relation
- Representation of function
  - a graph
  - ordered pairs
  - an equation
  - a table of values
  - an arrow or mapping diagram


domain
range
-3
-2
-1
0
1
-6
-1
0
3
15
domain
range
-3
-2
-1
0
1
-6
domain
range
-3
-2
-1
0
1
-6
-1
0
3
15
domain
range
-3
-2
-1
0
1
16
-6
-1
0
3
15

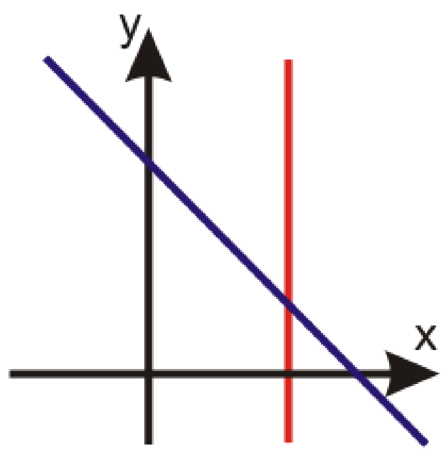
Representation	**Example**
Set of ordered pairs	(1,3), (2,6), (3,9), (4,12) (a subset of the ordered pairs for this function)
Equation	$y = 3x$
Graph	

Way to identify the Function from graph

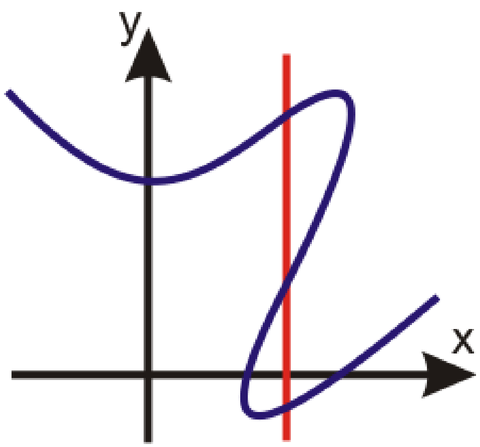
- Draw a vertical line on the graph.
- Function : If the line crosses the relation at one point
- Relation: If it is passes at multiple points



This is a function



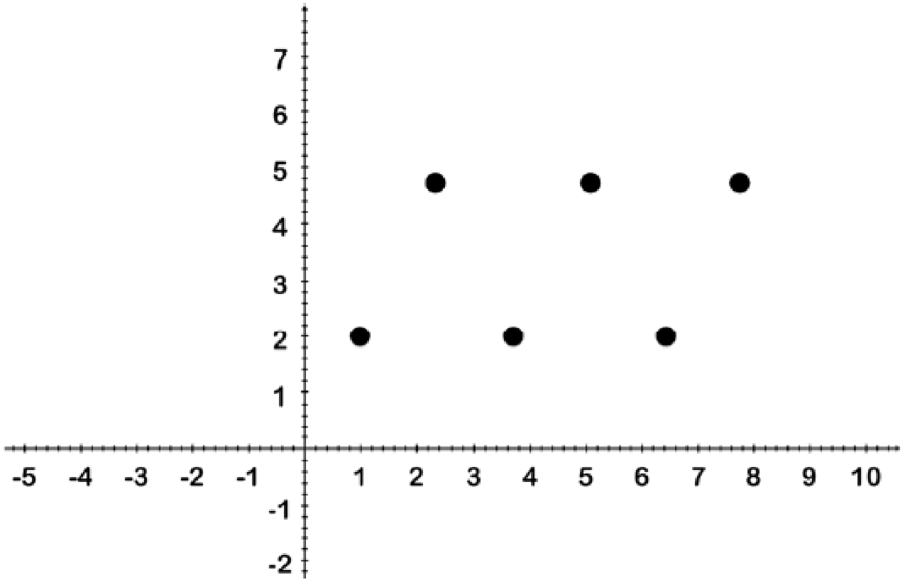
This is a function



This is NOT a function

Quiz

- Determine if each relation is a function:
  a. (2, 4), (3, 9), (5, 11), (5, 12)
  b. See the graph below

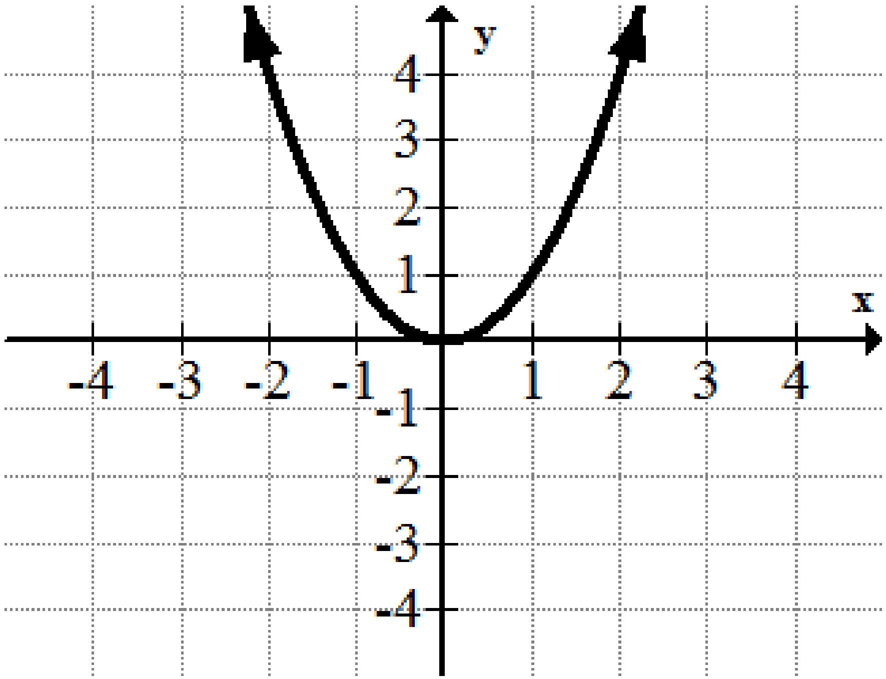


Answer

a. This relation is not a function because 5 is paired with 11 and with 12.
b. This relation is a function because every x is paired with only one y. A vertical line through the graph will always only encounter a single point.

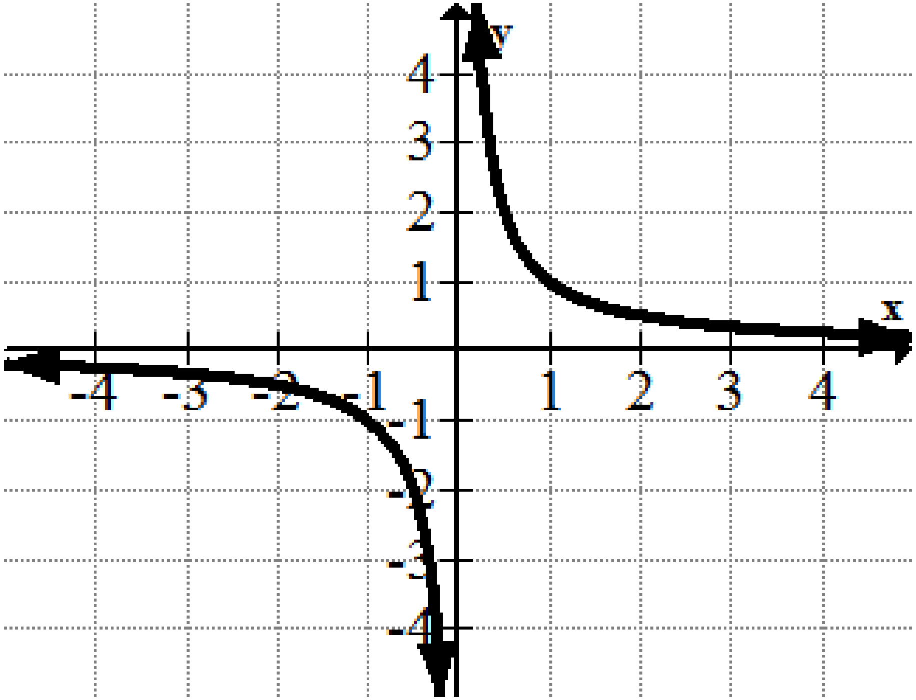
Modelling Data with functions

- Squaring Function: $f(x) = x^2$



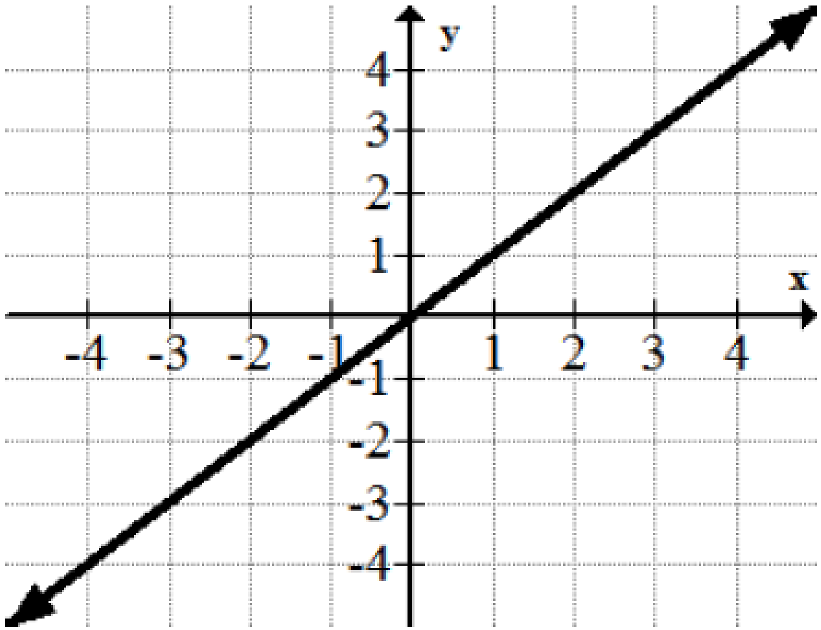
Modelling Data with functions

- The Reciprocal Function: $f(x) = 1/x = x^{-1}$



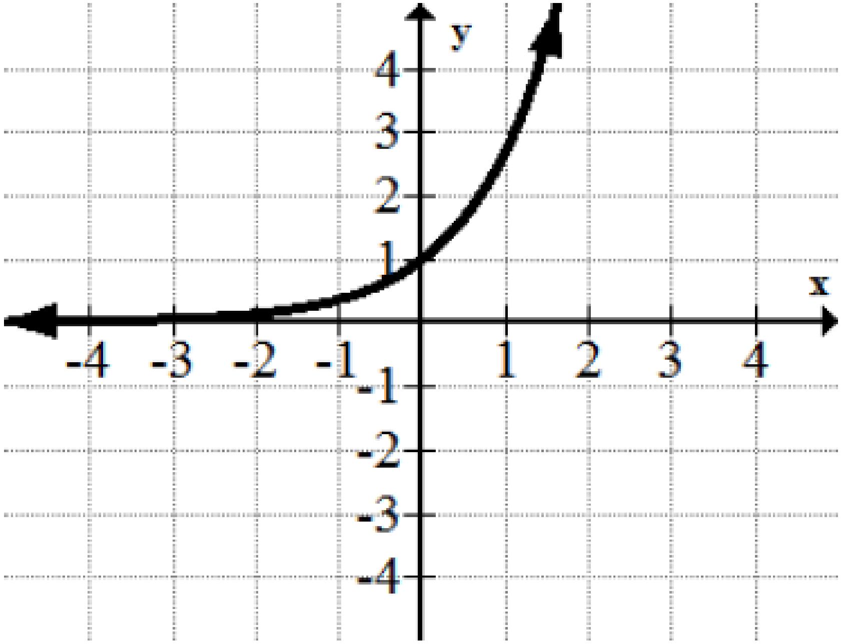
Modelling Data with functions

- Linear Function: $f(x) = x$

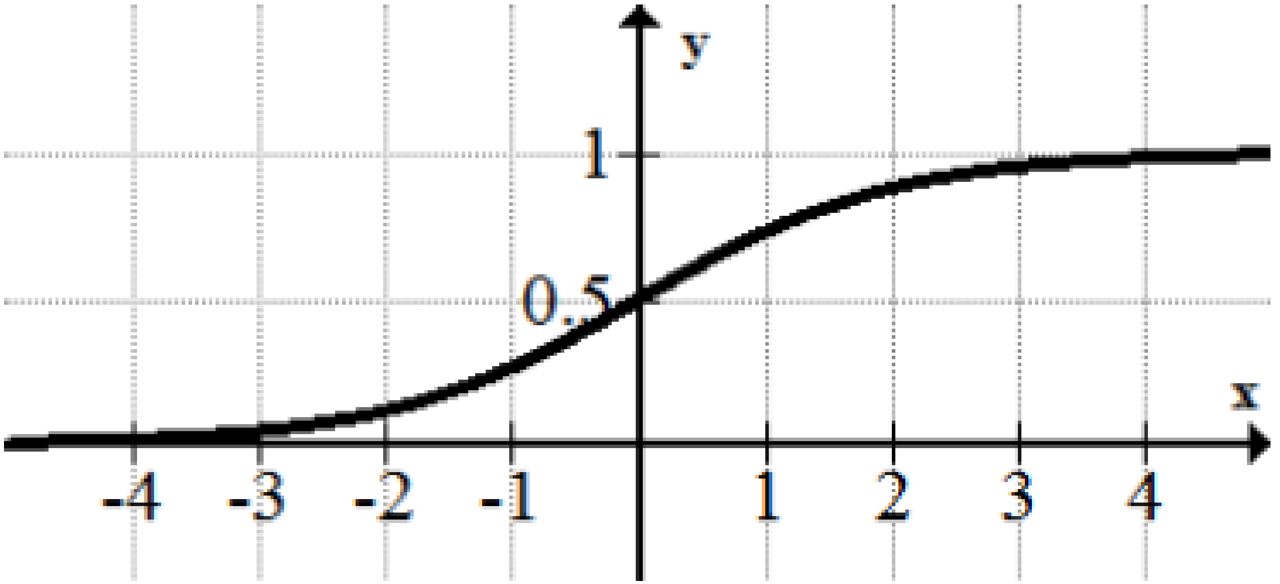


Modelling Data with functions

- Exponential Function: f (x) = e^x



The Logistic Function: $f(t) = \frac{C}{1+ab^{-t}} = \frac{C}{1+ae^{-kt}}$



Recap

- A relation is a comparison of two or more sets of values.
- A function is a relation of two or more sets of values in which each input number corresponds to one and only one output number.
- A function family is a group of functions that all have the same basic shape.

Mathematical modeling

- Mathematical modeling involves creating a set of mathematical equations that describes a situation, solving those equations, and using them to understand the real-life problem.
- Models can also be used to predict what a system will do for different values of the independent variable. Lastly, a model can be used to estimate quantities that are difficult to evaluate exactly.
- The goal is not to produce an exact copy of the "real" object but rather to give a representation of some aspect of the real thing.

Linear Models

- Linear regression involves using data to calculate a line that best fits that data, and then using that line to predict scores on one variable from another.
- Prediction is simply the process of estimating scores of the outcome (or dependent) variable based on the scores of the predictor (or independent) variable.
- How to predict???

Line of best fit – Regression line

- Find the mean of x and mean of y.
- Any line we decide, should pass through this coordinate
- Draw all possible lines
- Notice for every point on the graph, our line is wrong by some distance
- **This amount wrong is called a residual**
- Difference between the observed value of the dependent variable (y) and the predicted value ($\hat{y}$) is called the **residual** (e).
- Each data point has one **residual**.
- **Residual** = Observed value - Predicted value.
  - $e = y - \hat{y}$

Line of Best Fit – Regression line

- Both the sum and the mean of the **residuals** are equal to zero.
- Find the squares and sum it it, sum of squared residuals
- Find the least squares to fit the regression line
- **Least squares** is a statistical **method** used to determine a line of best fit by minimizing the sum of **squares** created by a mathematical function. A "**square**" is determined by squaring the distance between a data point and the regression line.
- **Line of Best Fit = Lowest SS residuals**

Interpreting the Linear Model

- Slope – the coefficient of x
- It is the one whole unit of change in x variable, how much difference in y it makes
- Rate of change = f(b)-f(a)/(b-a)
- Linear Model – constant rate of change
- Intercept -Our intercept is the predicted value of our outcome variable **when our x-axis variable is ZERO.**
- Be sure that the value of zero make sense
- If the intercept is not making sense, make z=x-min(x)
- Fit for the y~z (instead of x), which will give a +ve intercept. But the slope will not change.

Exponential Model

- Linear models are straight whereas exponential models are "Not Straight"
- Linear Model – the change in one unit of x will give a constant change in the outcome of y ie, constant rate of change
- Exponential model – Constant % of change
- In general, the exponential function takes the form:
  - $y = A * b^x$
  - Where A is the initial value and b is the growth factor
- The growth factor is expressed as $b=1+r$
  - where r is the change in % each unit of increase in x
- $b > 1$ - Growth
- $b < 1$ - Decay

Logistic Growth Model

- Logistic growth model will be best explainable when the data has reached the upper limit and there after the growth is consistent, logistic growth model can be used.
- Initially, it gives a illusion of using a exponential model and later, the growth will stop and shows only consistent value
- The logistic model can be expressed as

$$f(t) = \frac{C}{1+ab^{-t}}$$

- C – Carrying capacity. Determines the upper limit
- Log(a)/log(b)
- a – where our input variable is at zero, b is the growth factor.
- In logistic growth model, b should be greater than 1.

Finding the Best Model

- Do decisions based on the context of the data. If the contenxt of the data is not known then look for residuals.
- Residuals - comes from total sums of squares and model sums of squares
- R square = total sum of squares/model sum of squares
- It represents the proportion of variance accounted.
- The best model is one which has higher R square value.
- To check graphically use tripleFit function in the SDSFoundation Package.

Model fit in R

- linFit(independent,dependent)
- expFit(independent,dependent)
- logisticFit(independent,dependent)
- To compare with all three models
- tripleFit(independent,dependent)
- To predict
- expFitPred(inde,dep,12)
- logisticFitPred(inde,dep,12)

Example

You have a cylinder that is filled with water to a height of 50 centimeters. The cylinder has a hole at the bottom which is covered with a stopper. The stopper is released at time t = 0 seconds and allowed to empty. The following data shows the height of the water in the cylinder at different times.

Time(sec)	0	2	4	6	8	10	12	14	16	18	20	22	24
Height(cm)	50	42.5	35.7	29.5	23.8	18.8	14.3	10.5	7.2	4.6	2.5	1.1	0.2

a. Find the height (in centimeters) of water in the cylinder as a function of time in seconds.
b. Find the height of the water when $t = 5$ seconds.
c. Find the height of the water when $t = 13$ seconds.