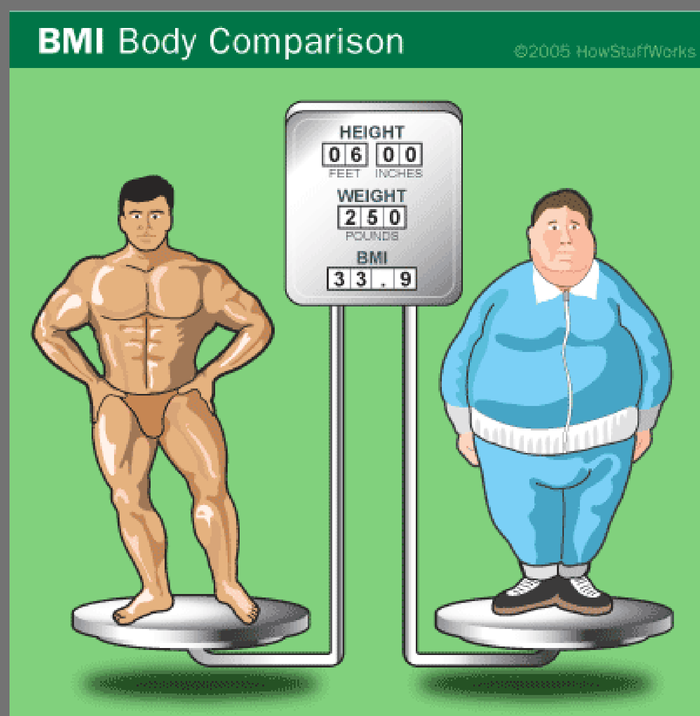
2008

BMI of females in the United States

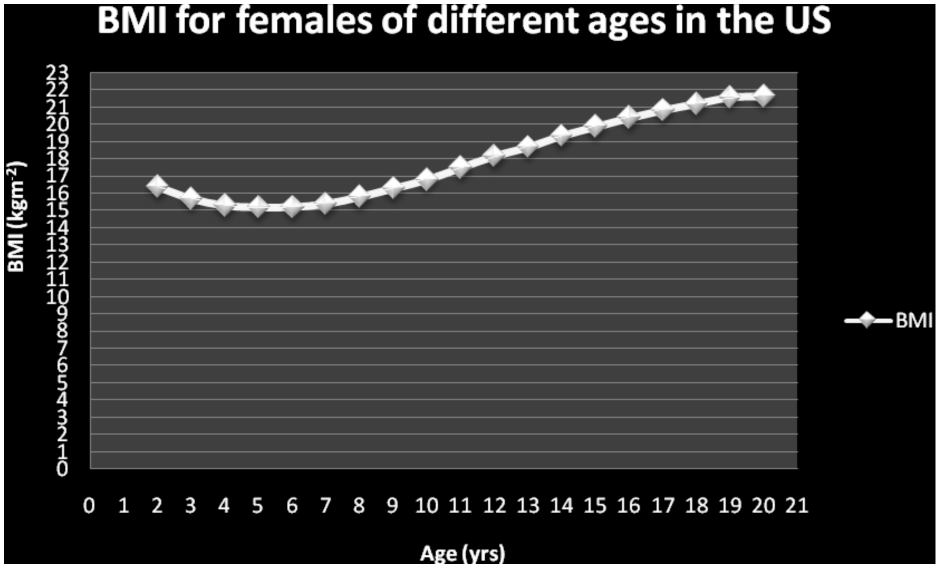


Anas Mansour
John Paul College
7/31/2008

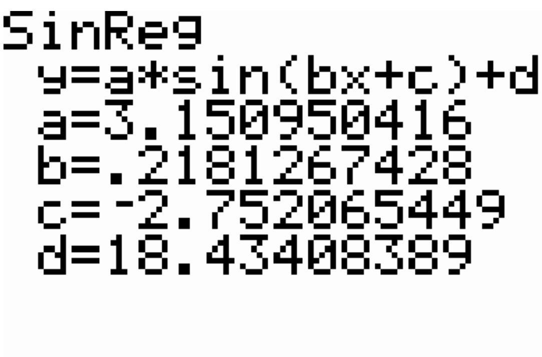
The table below shows the median body mass index (BMI) for females in the United States of America (USA) for the year 2000 (source www.cdc.gov).

Age (yrs)	*BMI*
2	16.4
3	15.7
4	15.3
5	15.2
6	15.21
7	15.4
8	15.8
9	16.3
10	16.8
11	17.5
12	18.18
13	18.7
14	19.36
15	19.88
16	20.4
17	20.85
18	21.22
19	21.6
20	21.65

Using the two variables above in the table, the independent variable (age), and the dependent variable BMI which has the units $kg.m^{-2}$. The graph below was generated using Microsoft Excel 2007.



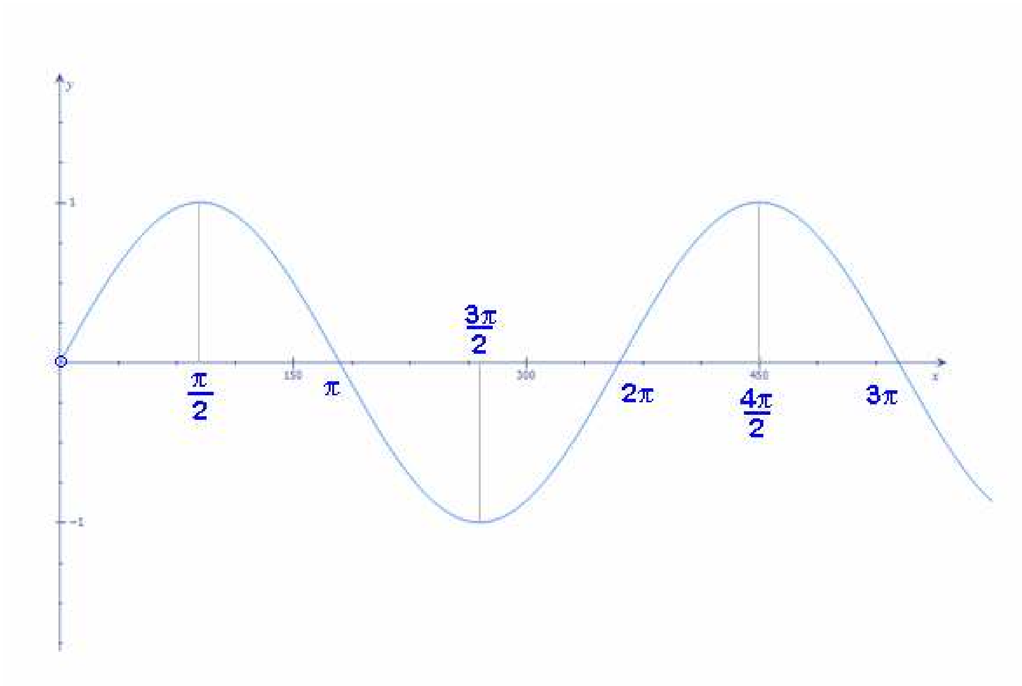
Using a GDC, the data was fitted to a set of parameters, a screen shot of these parameters can be seen below:



It can be seen from the plot that one possible mathematical fit for the graph is a sine function. This is apparent because it has the same periodic nature (it appears to be ¾ of a period of a sine curve) as a sine function. The next step in modelling a generic sine function to the data in the plot above, and applying a number of changes to f(x) = sin(x) curve. There are quite a number of calculations that are needed to acquire an accurate model of the data. Each aspect of the graph is investigated separately, for example vertical shift.

The graph below shows you the general function of a sine equ ation.

y=Sin(x)

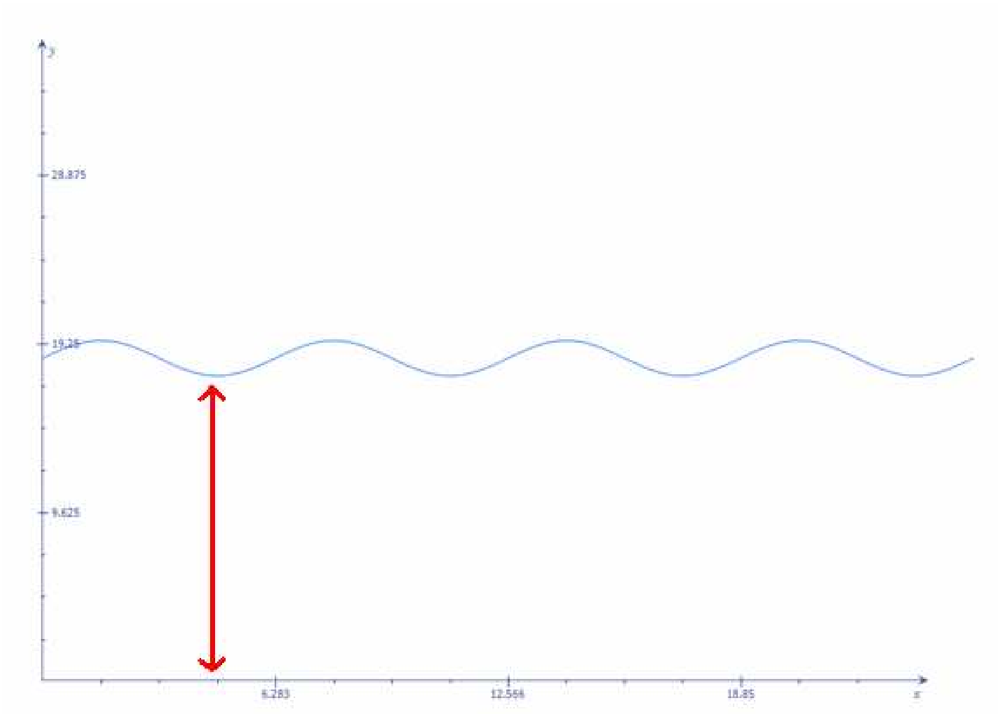


The midpoint in the data presented below is the maximum value plus the minim um value divided by 2. The mid point for a standard sine curve is zero. So the first step is to change from the standard f(x) = sin (x), to f(x) = sin (x) + d where'd' is the vertical shift.

$$\frac{\min + max}{2} = d$$

$$\frac{21.65 + 15.2}{2} = 18.43$$

$$d = 18.43$$



So at this stage the model for the data is f(x) = sin (x) + 18.425

However this does not accurately representing the data , there are further changes that take place. The next step is to increase the amplitude to the same amplitude of the data.

The coefficient 'a' determines the amplitude, $\boldsymbol{a}\,sin(x) + d$

Amplitude = distance from the midpoint to the minimum/ maximum, or similarly the maximum point minus the minimum point divided by two.

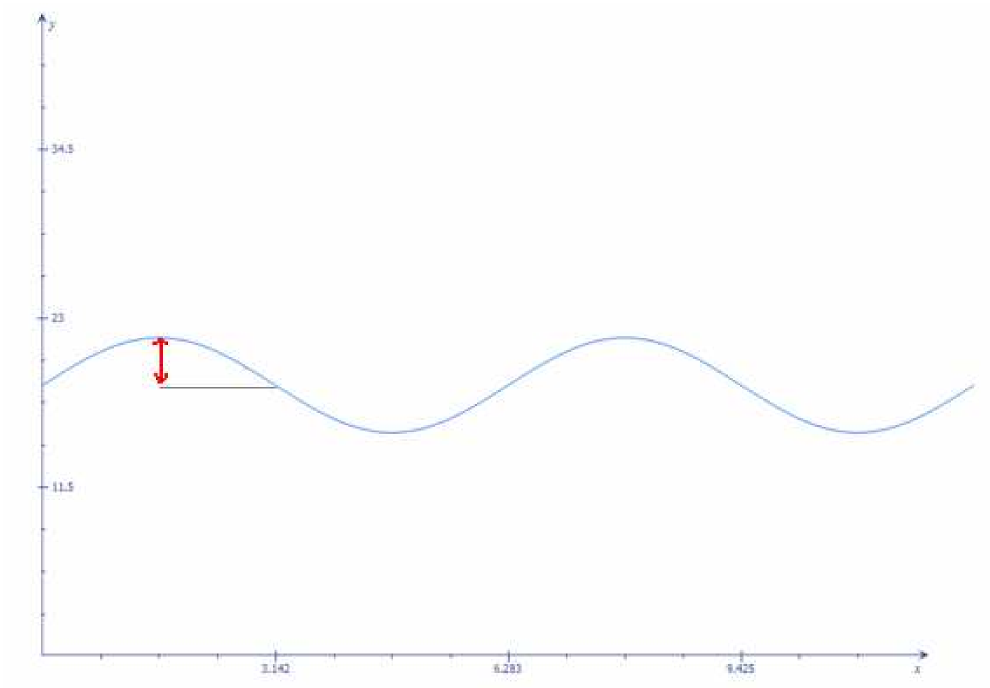
Maximum = 21.65

Minimum = 15.2

$\frac{\max - min}{2} = amplitude$

$\therefore \frac{21.65 - 15.2}{2} = 3.225 = a$

Hence f(x) becomes, f(x) = 3.225 sin(x) + 18.425



The model now has the correct amplitude and the right vertical shift; however it has the incorrect period and horizontal shift. To account for this two parameters' introduced, 'b' and 'c'. The model then becomes f(x) = 3.225 sin (bx +c) + 18.425

The coefficient of 'b' determines the period, the data provided only models half a period, which is 15 (this can be seen as the minimum is 5 and the maximum is 20 years??) hence the period is 30.

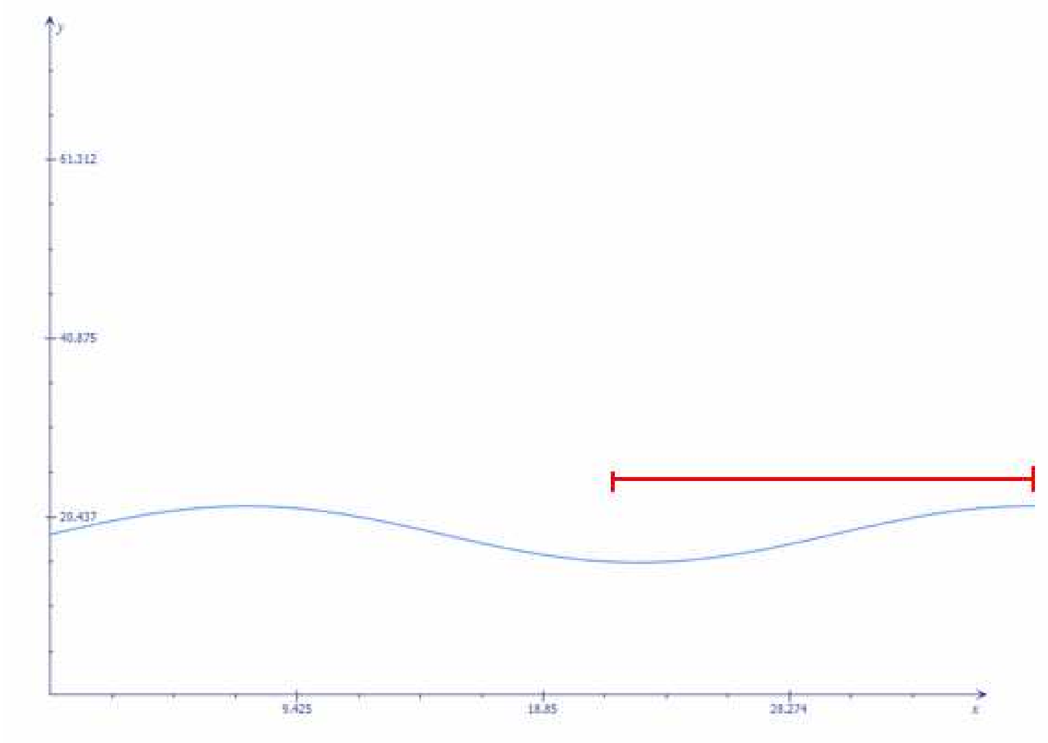
The period of a general sine curve is $\frac{2\pi}{b}$

$$\frac{2\pi}{b} = period$$

$$\frac{2\pi}{b} = 30$$

$$b = \frac{\pi}{15}$$

Hence the model now becomes f(x) = 3.225 sin ($\frac{\pi}{15}x$) + 18.425



The model now is close to fitting the data, the last step is to add the horizontal shift. This is the letter 'c' in the following general form of the sine function : f(x) = 3.225 sin ($\frac{\pi}{15}x - 2.61$) + 18.425

The coefficient 'c' determines the h orizontal shift:

$$\frac{-c}{B} = horizontal\ shift$$

The horizontal shift in the data can be found by adding the minimum va lue and the maximum for the x values, and dividing by two.

$$\frac{5 + 20}{2} = 12.5$$

$$\frac{-c}{b} = 12.5$$

is already known to be $\frac{\pi}{15}$

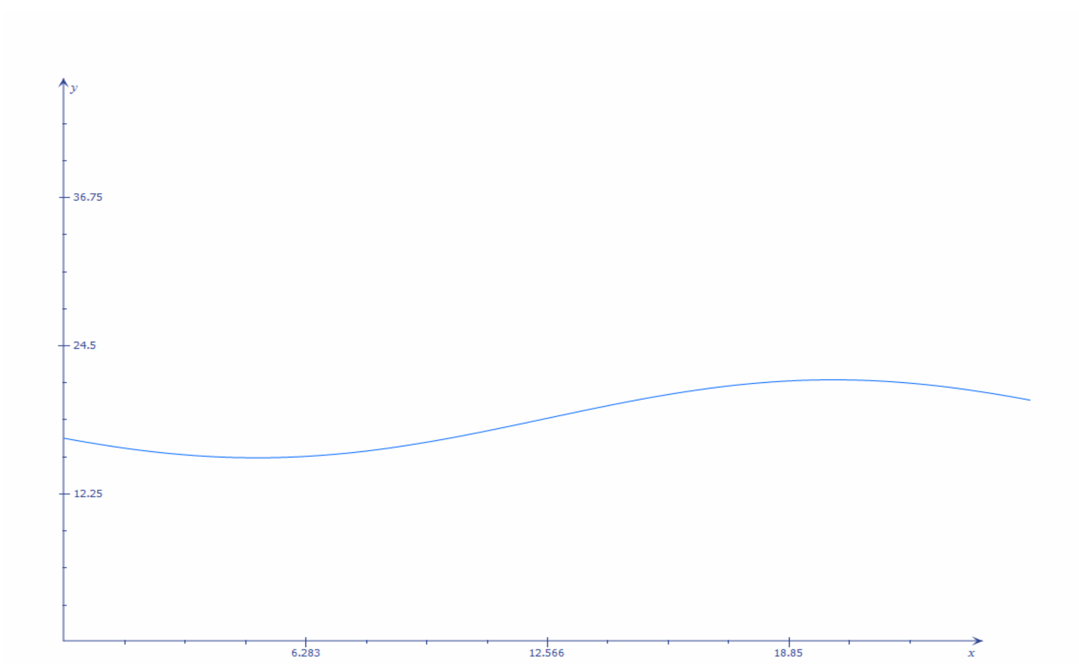
$$\therefore \frac{-c}{\frac{\pi}{15}} = 12.5$$

$$-c = 12.5 \, x \, \frac{\pi}{15}$$

$$-c = 2.61$$

$$c = -2.61$$

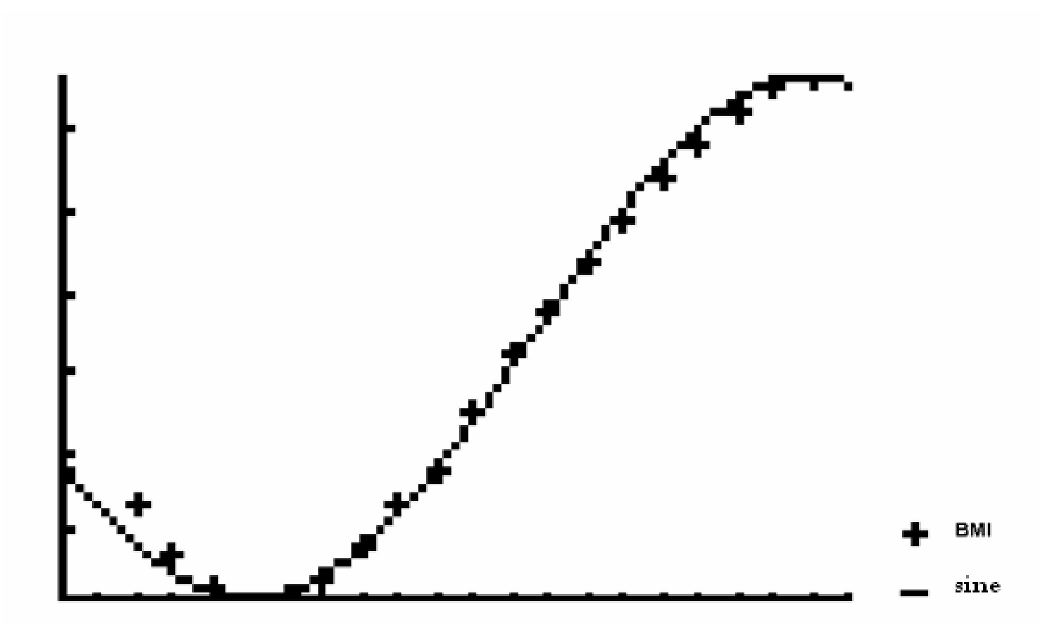
Now the model becomes, f(x) = $3.225 \, sin\left({}^{\pi}/_{15} \, x - 2.62\right) + 18.43$



The model function $y = 3.225 \, sin\left({}^{\pi}/_{15} x - 2.62\right) + 18.43$ is now compared to the plot of the original data table (BMI).

The graph below shows the model compared to the actual data using the GDC.

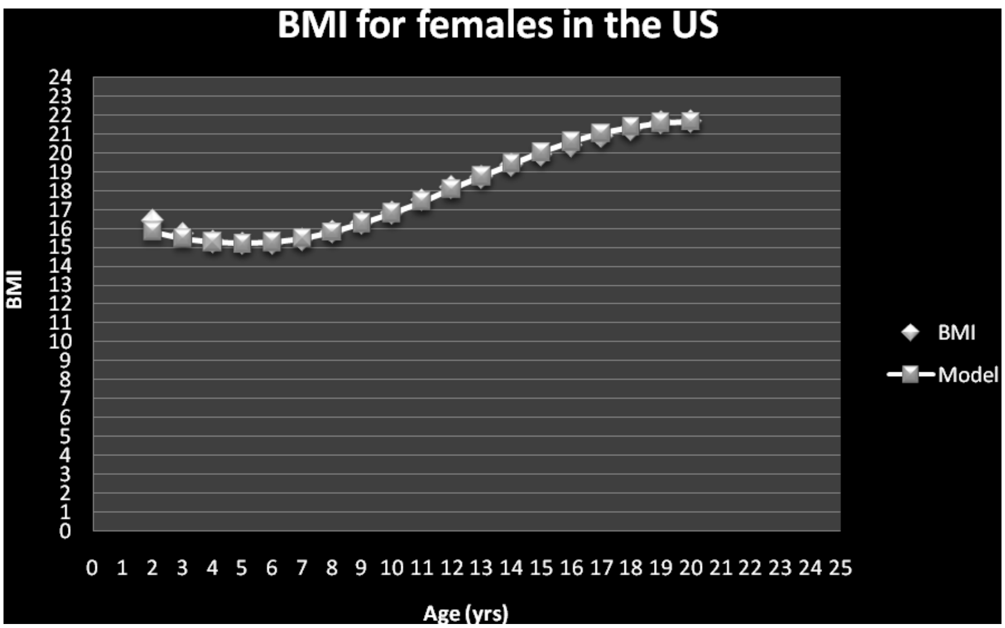
L1	L2
2	16.4
3	15.7
4	15.3
5	15.2
6	15.21
7	15.4
8	15.8
9	16.3
10	16.8
11	17.5
12	18.18
13	18.7
14	19.36
15	19.88
16	20.4
17	20.85
18	21.22
19	21.6
20	21.65



This can also be seen using excel, the model was applied to the same independent variable values (ages), the results can be seen in the following diagram.

Age (yrs)	BMI	Model
2	16.4	15.82473
3	15.7	15.48645

4	15.3	15.27683
5	15.2	15.20501
6	15.21	15.27413
7	15.4	15.48119
8	15.8	15.81712
9	16.3	16.26725
10	16.8	16.8119
11	17.5	17.42726
12	18.18	18.08646
13	18.7	18.76066
14	19.36	19.42042
15	19.88	20.03689
16	20.4	20.58313
17	20.85	21.03527
18	21.22	21.37354
19	21.6	21.58317
20	21.65	21.65499

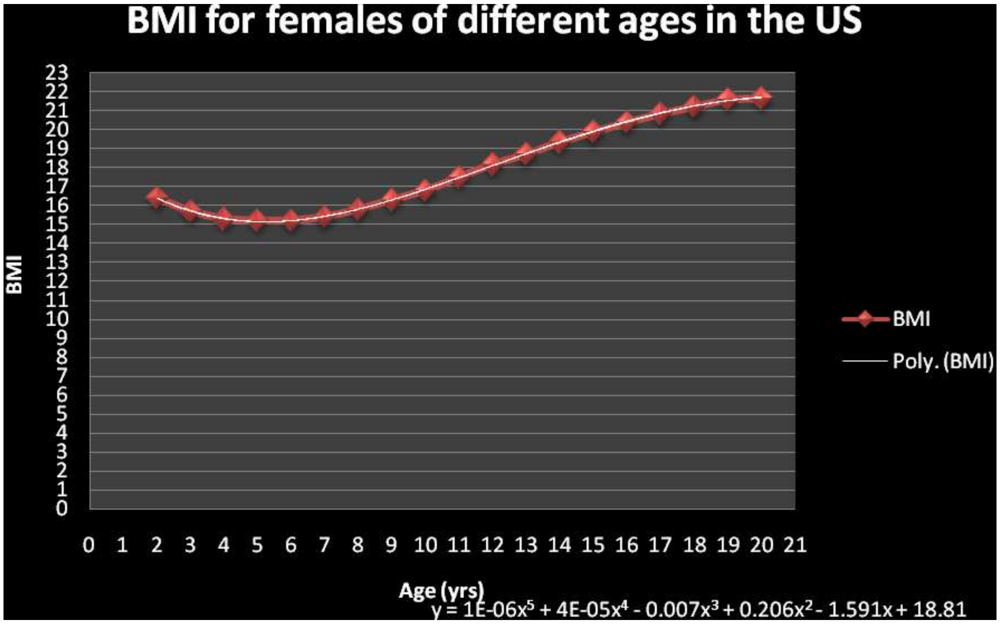


As you can see there is a slight difference between the model graph and the original data graph.

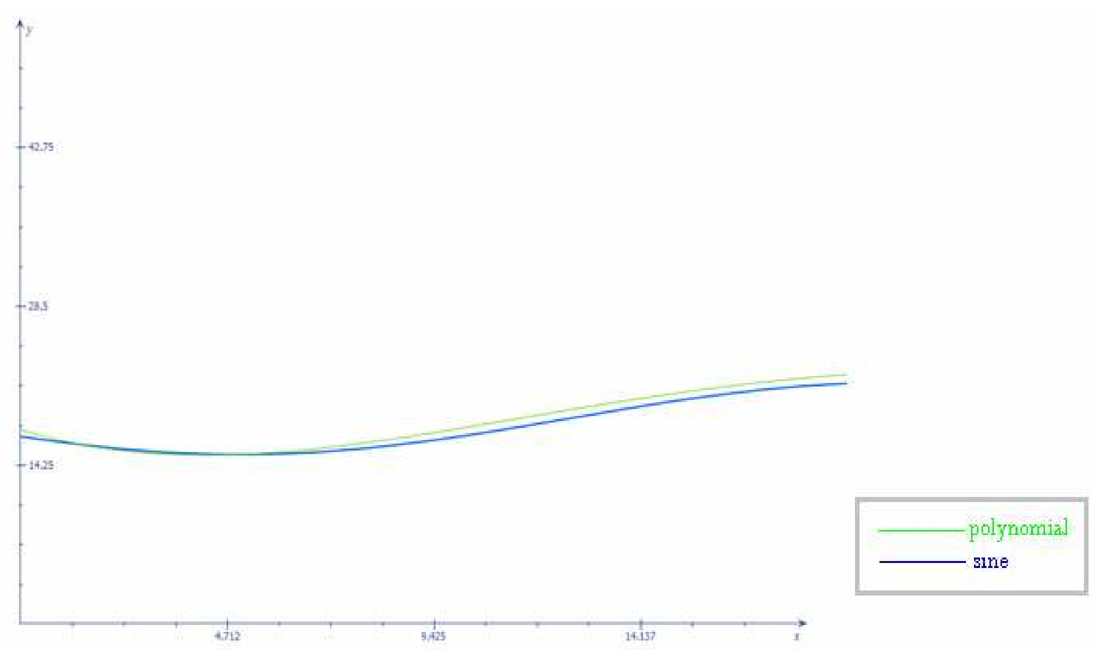
They both have similar curve nature. The model function has a constant up and down curve while the original data graph has one curve and then it remains constant.

By using excel we were able to find another graph that fits the data. This is a polynomial function.

The equation for the function is $X^5+X^4+X^3+X^2+X^1+X$ by adding your data in the graph we were able to find that the polynomial function has the best fit to the graph.



The graph below shows the function compared to the model:

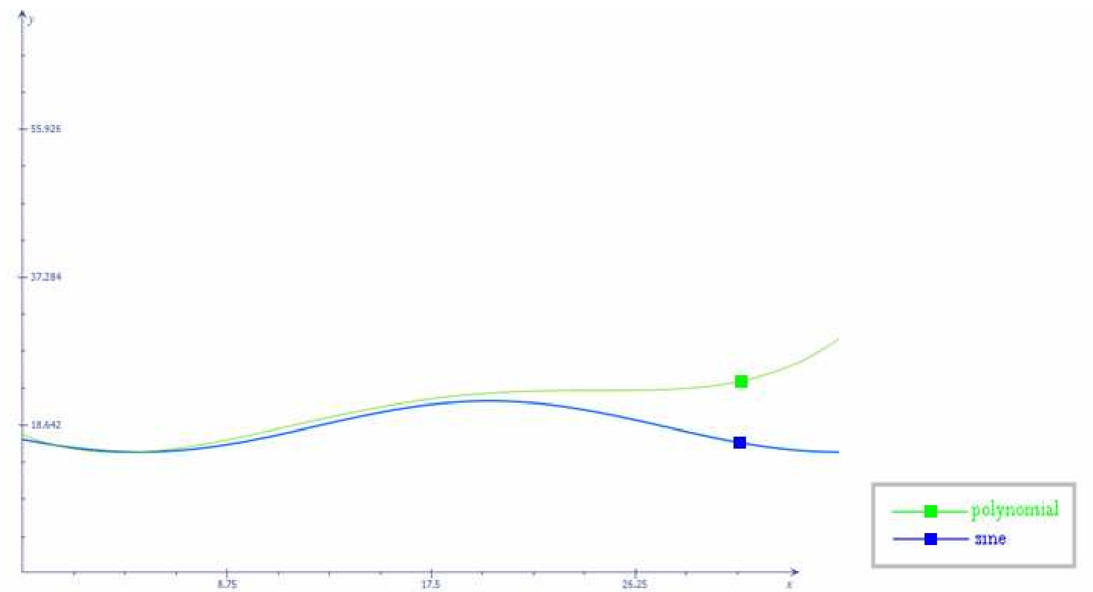


The difference between the sin function and the polynomial function was that the sin function has constant curve movement; however the polynomial function had the best fit.

-The polynomial is a closer fit then the sine curve because it takes into account the different rates of curvature in the actual data.

-The sine is periodic which will become important if trying to estimate BMI's outside the range of the actual data. Whereas the polynomial gives reasonable answers until it starts to increase rapidly at a value of approximately.

Using the model of BMI of a women aged 30 years old in the USA is estimated:



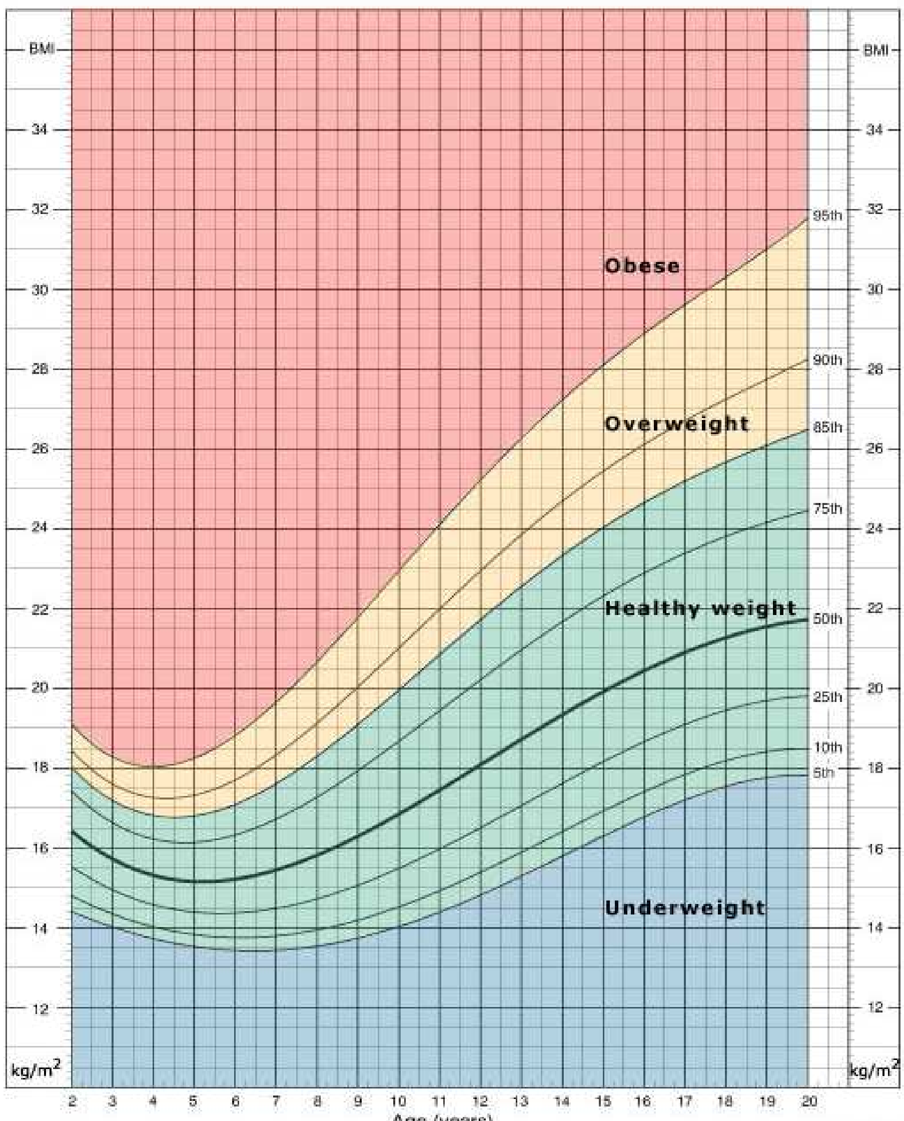
-As you can see the curve of the sine function starts to decrease around 20 age(yrs) . The sine function has a pattern of increase and decrease in curve every 10 age(yrs), therefore giving us a less accurate measure of the BMI of a 30 year old woman. We obtained the result to be ***[illegible]4.[illegible]7 BMI***

- The polynomial function maintains constant to around 27 age (yrs) after that the model increases rapidly. Therefore it cannot be used to measure the BMI after around 33 age (yrs) however it gives a more accurate result of the BMI of a 30 year old woman. The results were ***23.3 BMI***

- Neither of these two models is an accurate fit when rather large ages are used. However there may exist a function which is periodic to begin with, but then follows a more steady value.

-The average BMI for a normal woman is 22.5, hence this states that the BMI should not change with age after 20 http://www.massgeneral.org/children/adolescenthealth/articles/aa_body_mass_index.aspx)

From the following source http://www.betterhealth.vic.gov.au/bhcv2/bhcsite.nsf/pages/bmi4child and by assuming from the graph below that the average BMI is the healthy curve in the data further results were obtained.



SAFER • HEALTHIER • PEOPLE

The following table is produced by extrapolating the values for each year.

Age (yrs)	BMI
2	16.5
3	15.6
4	15.4
5	15.2
6	15.4
7	15.5
8	15.75
9	16.4
10	16.8
11	17.5
12	18.1
13	18.6
14	19.3
15	20
16	20.4
17	20.9
18	21.3
19	21.5
20	21.6

Accuracy is approximate (Because the data in the table above was generated by extrapolating the value from the curve, the accuracy is limited to the scale of the y-axis in the graph. This scale was .2, so the values in the table are only shown to one significant figure.

Calculating the different aspects of the sine function will help to obtain best fit for the data above:

Vertical shift

$$\frac{\min + \max}{2} = \text{vertical shift}$$

$$\frac{15.2 + 21.6}{2} = 18.4$$

d = 18.4

f(x)= sin(x)+18.4

Amplitude

$$\frac{\min - \max}{2} = \text{amplitude}$$

$$\frac{21.6 - 15.2}{2} = 3.2$$

a = 3.2

f(x)=3.2sin(x)+18.4

Period

$$\frac{2\pi}{b} = period$$

$$\frac{2\pi}{b} = 30$$

$$B = \frac{\pi}{15}$$

F(x)=3.2sin($\frac{\pi}{15}$x)+18.4

Horizontal shift

$$\frac{-\text{c}}{\text{b}} = \text{horizontal shift}$$

$$\frac{5 + 20}{2} = 12.5$$

$$\frac{-\text{c}}{\text{b}} = 12.5$$

b is known to be $\frac{\pi}{15}$

$$\frac{-c}{\frac{\pi}{15}} = 12.5$$

-c = 12.5 x $\frac{\pi}{15}$

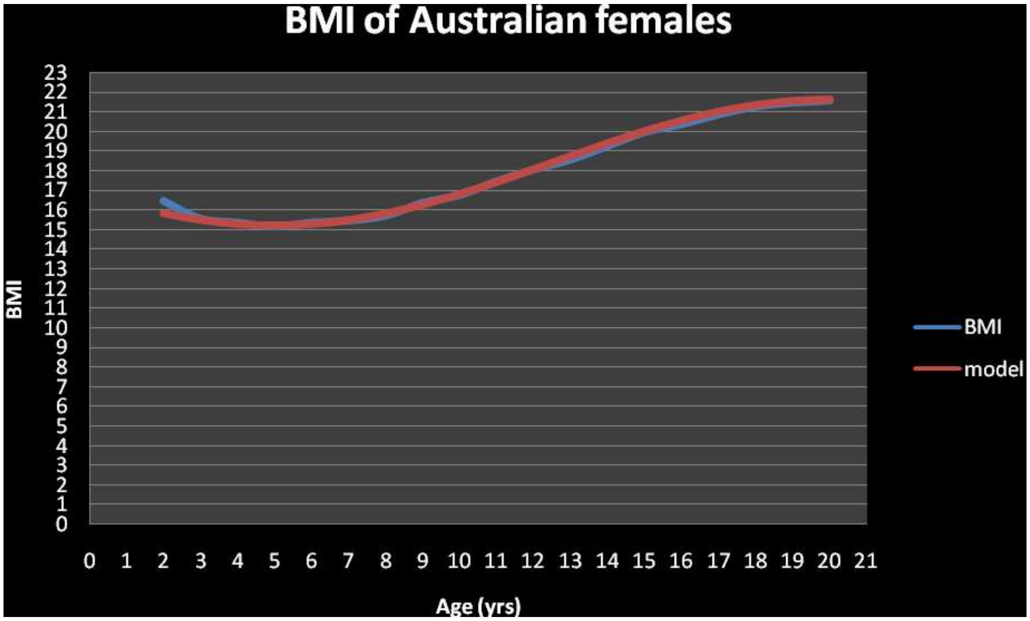
-C =2.61

C= -2.61

$F(x)=3.2\sin(\frac{\pi}{15}x-2.61)+18.4$

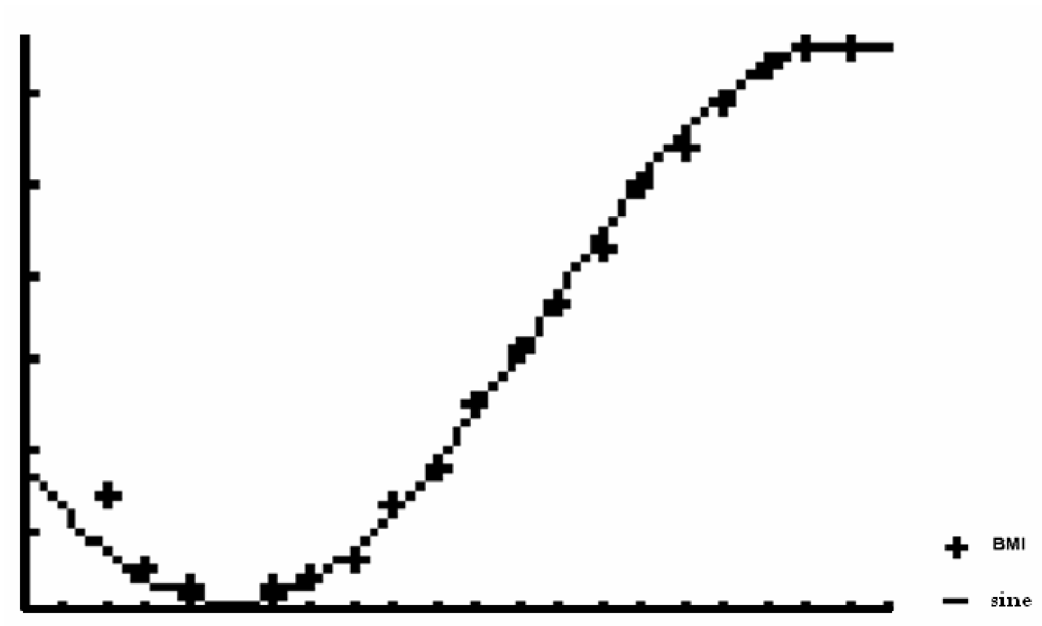
Using excel the model was applied to the same independent variable values (ages), the results can be seen in the following diagram.

Age (yrs)	BMI	model
2	16.5	15.82473
3	15.6	15.48645
4	15.4	15.27683
5	15.2	15.20501
6	15.4	15.27413
7	15.5	15.48119
8	15.75	15.81712
9	16.4	16.26725
10	16.8	16.8119
11	17.5	17.42726
12	18.1	18.08646
13	18.6	18.76066
14	19.3	19.42042
15	20	20.03689
16	20.4	20.58313
17	20.9	21.03527
18	21.3	21.37354
19	21.5	21.58317
20	21.6	21.65499



The graph below shows the model compared to the actual data using the GDC.

L1	L2
2	16.5
3	15.6
4	15.4
5	15.2
6	15.4
7	15.5
8	15.75
9	16.4
10	16.8
11	17.5
12	18.1
13	18.6
14	19.3
15	20
16	20.4
17	20.9
18	21.3
19	21.5
20	21.6



-There is a slight difference between the model graph and body mass index (BMI) of young Australian woman. They have a similar curve nature for a specific period before the sine function changes.

-The sine function has a constant up and down curve while the original data graph has one curve and then it remains constant.

-The BMI of the Australian females is very similar to the BMI of US females from the age of 2 -20.

-Until this point the sine function has been able to calculate the BMI of different woman in different countries, however if you apply more tests to the function with different countries and ages then you will see slight limitations to the graph which would require you to either recalculate the equation or use other functions.