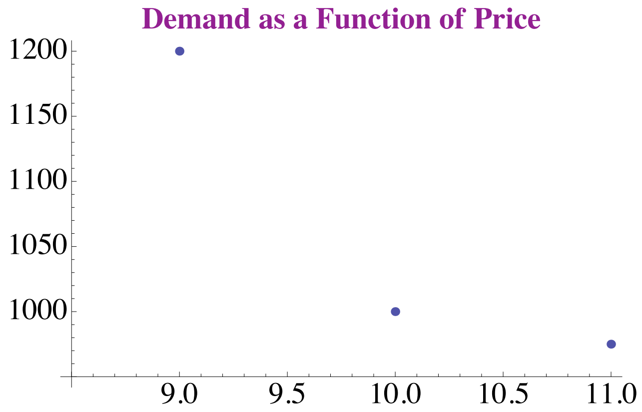
Price – Demand Curve (p. 111–114)

Example. A company is trying to determine how demand for a new product depends on its price and collect the following data:

price p	\$9	\$10	\$11
demand d	1200/mo.	1000/mo.	975/mo.

The company has reason to believe that price and demand are **inversely proportional**, that is, $d = \frac{c}{p}$ for some constant c.

$\longrightarrow$ Use the method of least squares to determine this constant c.



Price – Demand Curve (p. 111–114)

Solution. Since $f(p) = \dfrac{c}{p}$, then the sum $S = \sum_{(p_i, d_i)} \left[d_i - \left(\dfrac{c}{p_i} \right) \right]^2$.

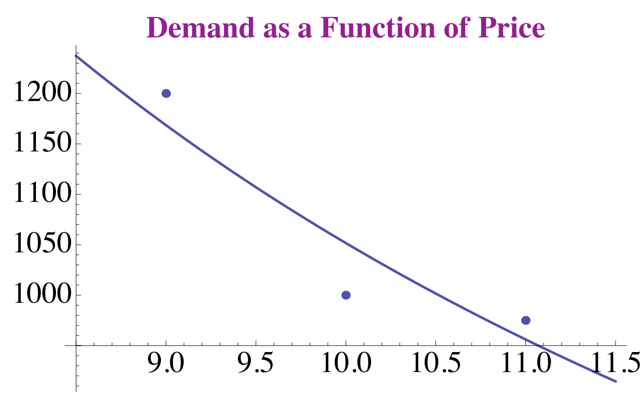
Specifying datapoints gives

$$S = \left[1200 - \frac{c}{9}\right]^2 + \left[1000 - \frac{c}{10}\right]^2 + \left[975 - \frac{c}{11}\right]^2$$

Setting the derivative equal to zero gives

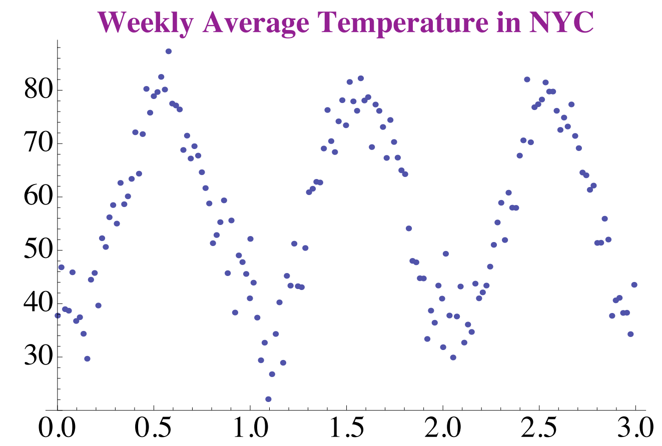
$$\frac{dS}{dc} = \frac{-2}{9}\left[1200 - \frac{c}{9}\right] + \frac{-2}{10}\left[1000 - \frac{c}{10}\right] + \frac{-2}{11}\left[975 - \frac{c}{11}\right] = 0$$

Solving for c gives $c \approx 10517$.



New York City Temperature (similar to p. 158)

The graph of average weekly temperature in New York City from Jan. 2006 to Dec. 2008 gives the distinct impression of a ________.

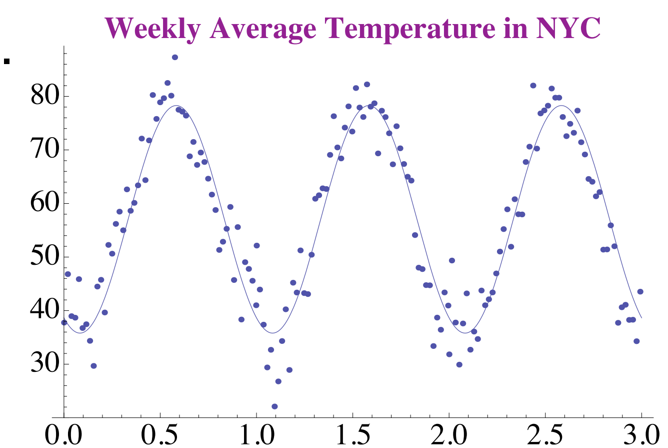


We need to determine the constants in:

$Temp(t) = A + B\sin(C(t - D))$.

Let's simplify our model to only determine amplitude B and vertical shift A. We must make assumptions about ________ C and ______________ D. We can assume that $C =$ ________.

For D, find when the sine passes through zero. Since January is cold and July is hot, the zero should occur in April; guess $D \approx \frac{4}{12}$.



Fitting to $Temp(t) = A + B\sin[t - \frac{4}{12}]$

gives: $Temp(t) = 13.9 + 11.8\sin[t - \frac{4}{12}]$

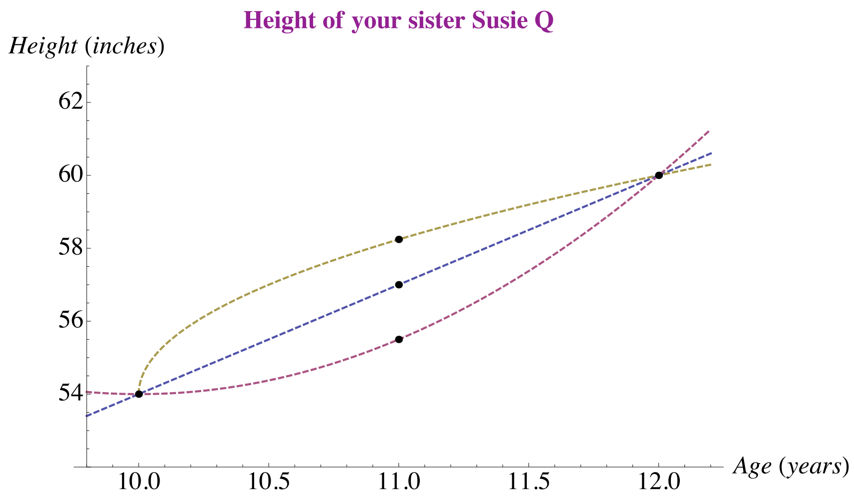
Interpolation vs. Extrapolation

Suppose you have collected a set of *known* data points (x_i, y_i), and you would like to estimate the y-value for an *unknown* x-value.

The name for such an estimation depends on the placement of the x-value relative to the known x-values.

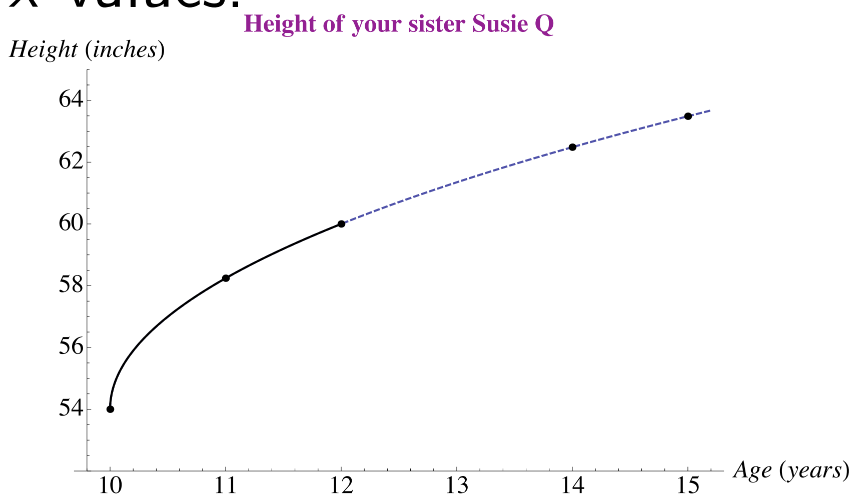
Interpolation

Inserting one or more x-values between known x-values.



Extrapolation

Inserting one or more x-values outside of the range of known x-values.



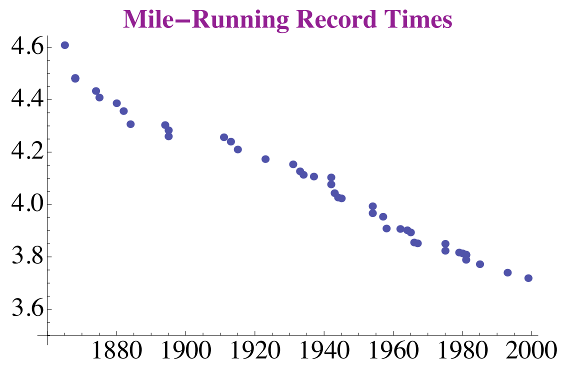
Interpolation vs. Extrapolation

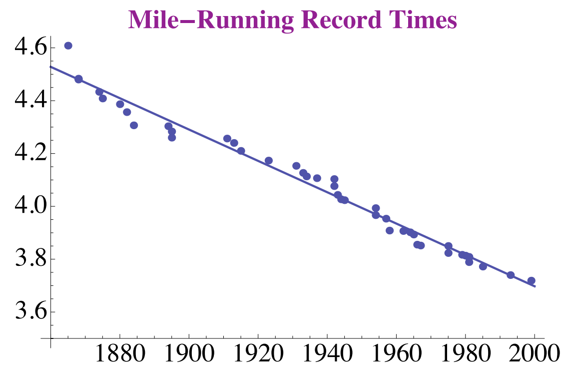
- The most common method for interpolation is taking a weighted average of the two nearest data points; suppose $x_1 < x < x_2$, then,
$$f(x) \approx y_1 + \frac{y_2 - y_1}{x_2 - x_2}(x - x_1).$$
- In both interpolation and extrapolation, when you have a function f that is a good fit to the data, simply plug in $y = f(x)$.
- Confidence in approximated values depend on confidence in your data and your model.
- Confidence in extrapolated data higher when closer to the range of known x-values.

Extrapolation: Running the Mile (p. 162)

Below is a plot of the years in which a record was broken for running a mile and the record-breaking time.
The data appears to fit a line; running least-squares gives

$$T(t) = 15.5639 - 0.00593323t$$





Solving for T(t)=0 gives $t \approx 2623$.

Conclusion: in the year 2623, the record will be zero minutes!

- Note the lack of descriptive realism.
- Always be careful when you extrapolate!