2.3 & 2.4 Scatter Plots and Trends in Data

Independent Variable: a variable that affects the value of another variable

Dependent Variable: a variable that is affected by some other variable

Example:

Independent - time spent practicing free throws

Dependent - free throw percentage in games

Your free throw percentage ***depends on*** the amount of time you spend practicing free throws

# of Hours John Studies	John's Test Score
0	75
.5	80
1	85
1.5	90
2	95
2.5	100

Independent Variable: # of Hours John Studies

Dependent Variable: John's Test Score

How are they related?

The more you study, the higher your test score will be.

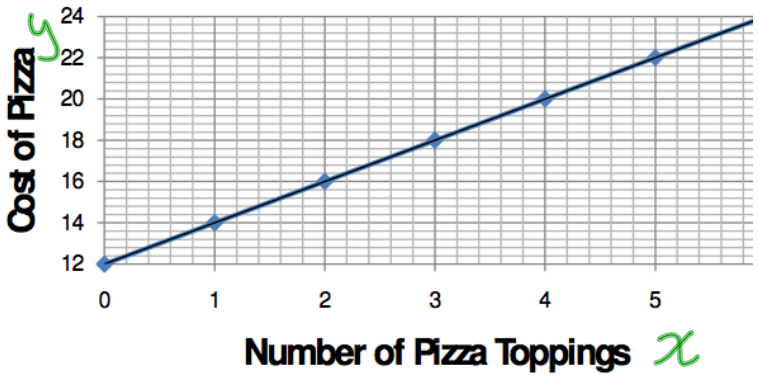
Number of Guests	Meal Preparation Time (min)
3	25
4	33
5	41
6	49
7	57
8	65

Independent Variable: Number of Guests

Dependent Variable: Meal Prep Time

How are they related?

The more guests you have attending dinner, the longer it will take to prepare the meal.



Independent Variable:

Number of Pizza Toppings

Dependent Variable:

Cost of Pizza

How are they related?

The more toppings you put on your pizza, the more it will cost

Now fill in the following the chart using your understanding of each type of variable:

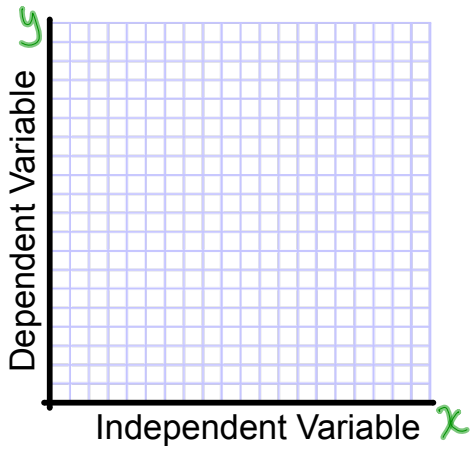
x y

Independent Variable	Dependent Variable
Number of gallons in your gas tank	How far you can drive
How much you read	Your IQ
Number of calories you eat each day	How much you weigh
How physically active you are	Your level of happiness
Number of hours you study for a test	Your test mark

Scatter Plots

A Scatter plot is a graph that shows the correlation between two variables.

The Independent variable goes on the horizontal (x) axis, and the dependent variable goes on the vertical (y) axis.

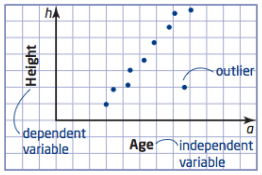


Types of correlations:

	A scatter plot shows a positive correlation when the pattern rises up to the right. *This means that the two quantities increase together.*
	A scatter plot shows a negative correlation when the pattern falls down to the right. *This means that as one quantity increases the other decreases.*
	A scatter plot shows no correlation when no pattern appears. *Hint:* *If the points are roughly enclosed by a circle, then there is no correlation.*

Correlations can also be strong or weak depending on how close or spread out the points on the scatter plot are.

Define an outlier:



measurement that differs significantly from the rest of the data.

When should you include an outlier in your data set?

If you can't show that it is inaccurate or unrepresentative

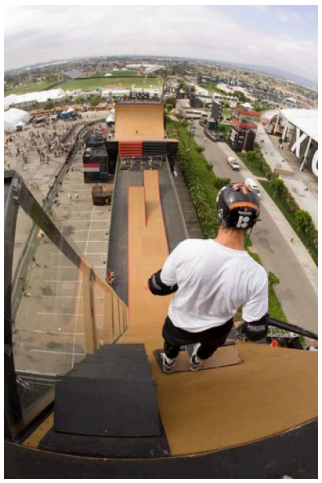
When shouldn't you?

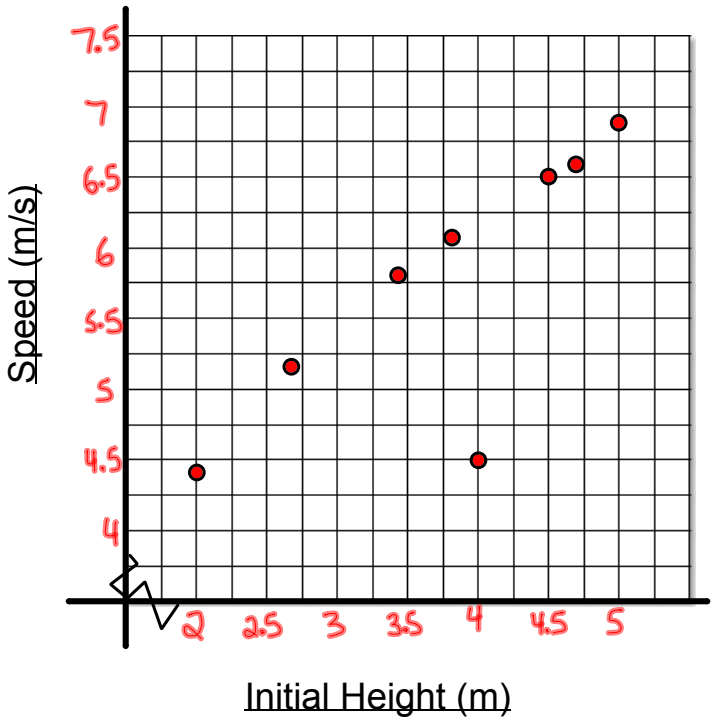
If you can show that it is inaccurate or unrepresentative

Make a Scatter Plot

A skateboarder starts from various points along a steep ramp and coasts to the bottom. This table lists the initial height and his speed at the bottom of the ramp.

Initial Height (m)	2.0	2.7	3.4	3.8	4.0	4.5	4.7	5.0
Speed (m/s)	4.4	5.2	5.8	6.1	4.5	6.5	6.6	6.9





Independent Variable: Initial Height

Dependent Variable: Speed

Describe the relationship:

The higher the initial height, the faster the speed.

There is a strong positive linear correlation

Are there any outliers? If so what are possible reasons for the outlier?

Yes, (4.0, 4.5) is an outlier. Maybe the skateboarder fell...

Line of Best Fit

A line of best fit can help you see the relationship between variables and also to make interpolations and extrapolations

Properties of a line of best fit:

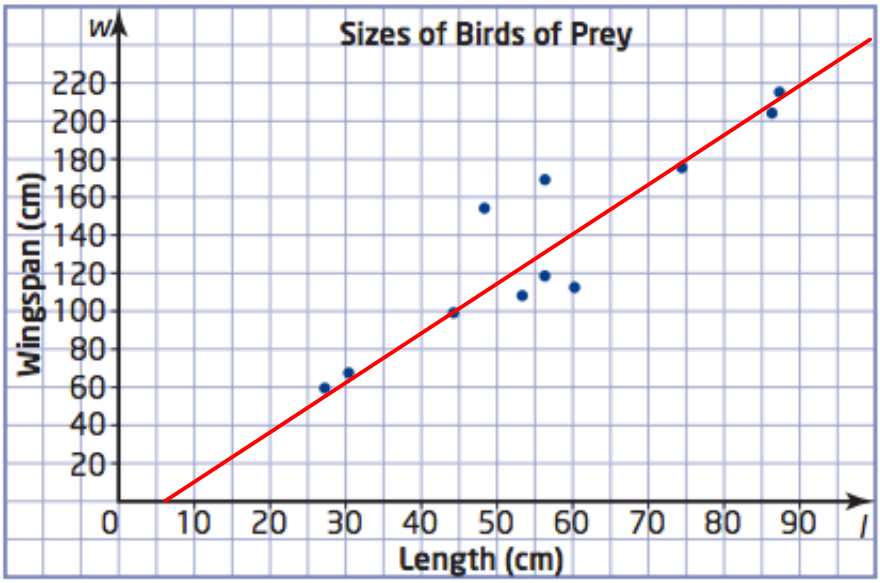
1. Straight line that passes through or close to as many points as possible

2. Any points that are not on the line should be evenly distributed above and below it

Interpolation: estimate a value between two measurements in a set of data

Extrapolation: estimate a value beyond the range of a set of data

Practice drawing a line of best fit:



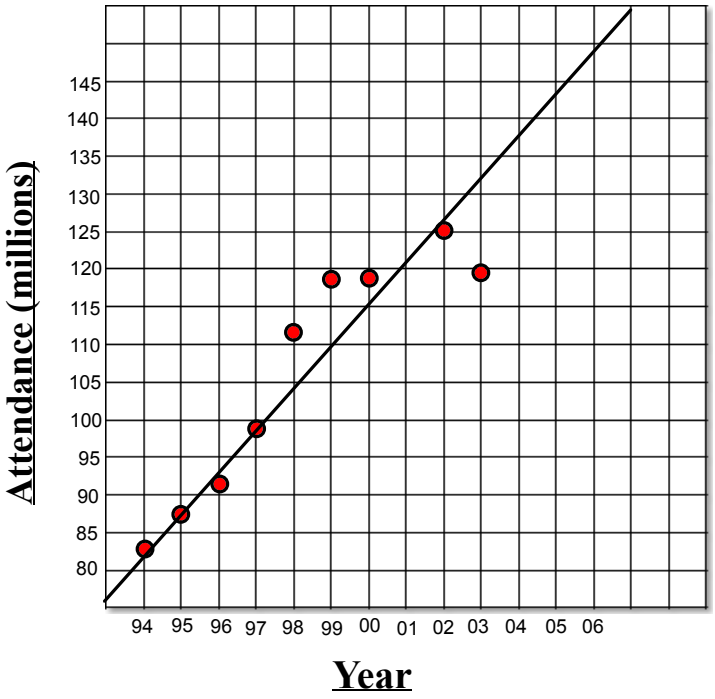
This table shows the number of paid movie admissions in Canada for 12 month periods

Fiscal Year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Attendance (millions)	83.8	87.3	91.3	99.1	111.6	119.3	119.3	no data	125.4	119.6

Independent Variable: Fiscal Year

Dependent Variable: Attendance

Graph the data and draw a line of best fit:



Describe the correlation:

There is a strong positive linear correlation between the year and movie attendance.

Movie attendance increases as the year increases.

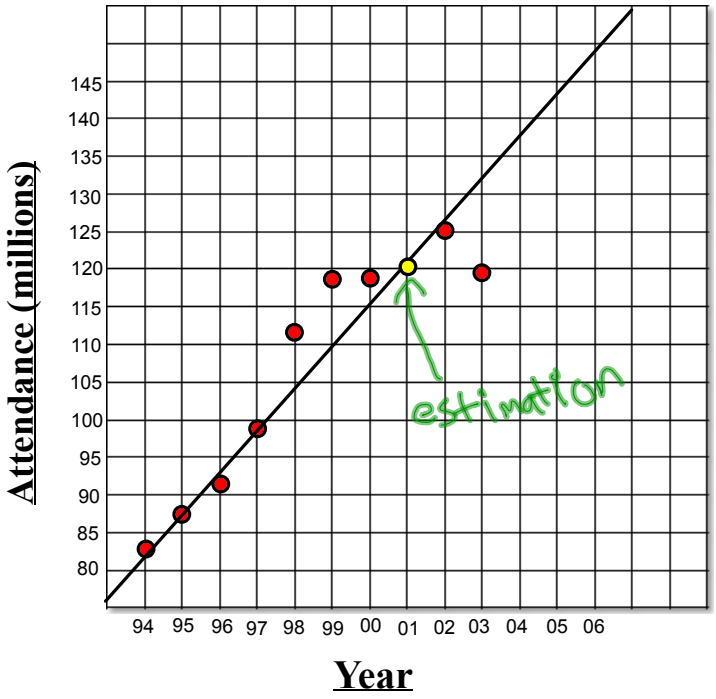
There is no data for 2001, estimate the movie attendance for this year using your line of best fit?

120 million

Did you use interpolation or extrapolation to estimate this data?

Interpolation

Graph the data and draw a line of best fit:



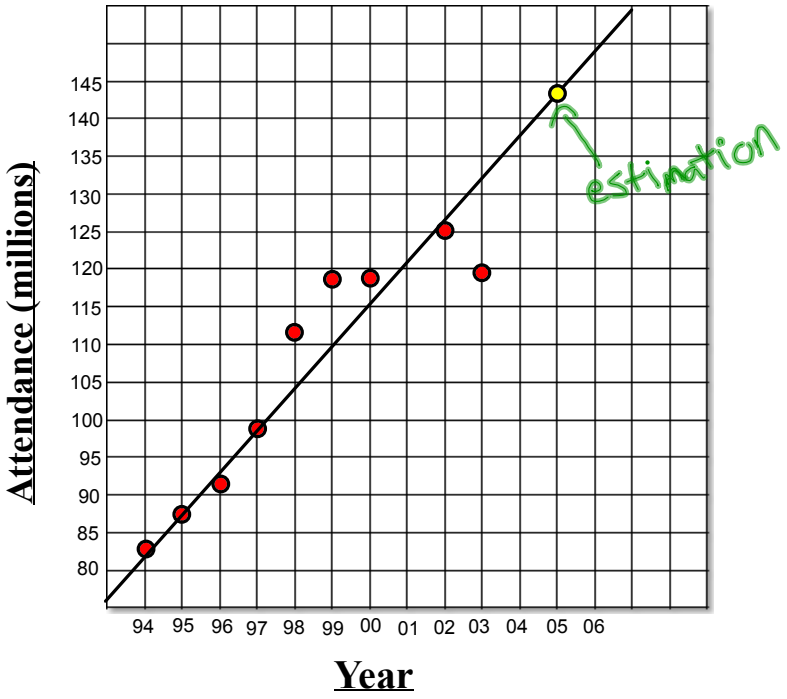
Estimate the movie attendance for 2005 by extending your line of best fit:

144 million

Did you use interpolation or extrapolation to estimate this data?

Extrapolation

Graph the data and draw a line of best fit:



Homework

Complete Worksheet