**Chapter 3.1– Scatterplots & Correlation**

p. 142-157

*Quick Summary*

* A scatterplot displays the relationship between two quantitative variables measured on the same individuals. Each individual’s data point is plotted on the coordinate grid.
* Explanatory variables (x-axis) help explain, predict, or cause changes in response variables (y-axis).
* In examining a scatterplot, look for overall pattern using DOFS:
	+ Direction
		- positive association (as x increases, so does y)
		- negative association (as x increases, y decreases)
	+ Outlier – any point outside of the general trend of the rest of the data
	+ Form – linear relationships, curved relationships, and clusters are important forms to look for
	+ Strength – determined by how close points in a scatterplot lie to a simple form such as a line
* Correlation, “r,” measures strength and direction of a linear relationship between two variables.
	+ A correlation of r = 0 represents no linear relationship
	+ Correlation values near zero have a very weak linear relationship
	+ Correlation values near 1 indicate a very strong positive linear relationship
	+ Correlation values near – 1 indicate a very strong negative linear relationship
	+ Correlation value of r = $\pm 1$ indicate a perfect linear relationship
* Facts about correlation:
	+ Ignores distinction between explanatory & response variables
	+ The value of *r* is not affected by changes in units of measurement
	+ Correlation itself has no unit of measurement
	+ Correlation requires both variables to be quantitative
	+ Correlation only describes strength and direction of linear relationships
	+ Correlation is not resistant to outliers
	+ Correlation is not a complete summary of two-variable data – you still need to calculate statistics like mean, standard deviation, etc. of both variables.