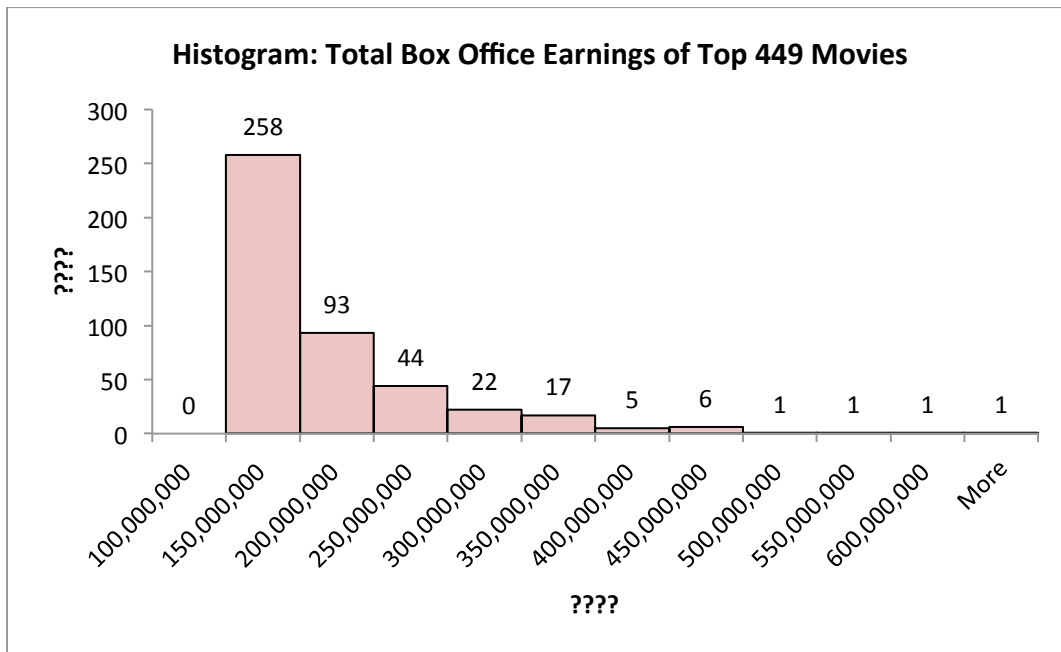


**Math 263: Review for Exam 1: Deb Hughes Hallett**

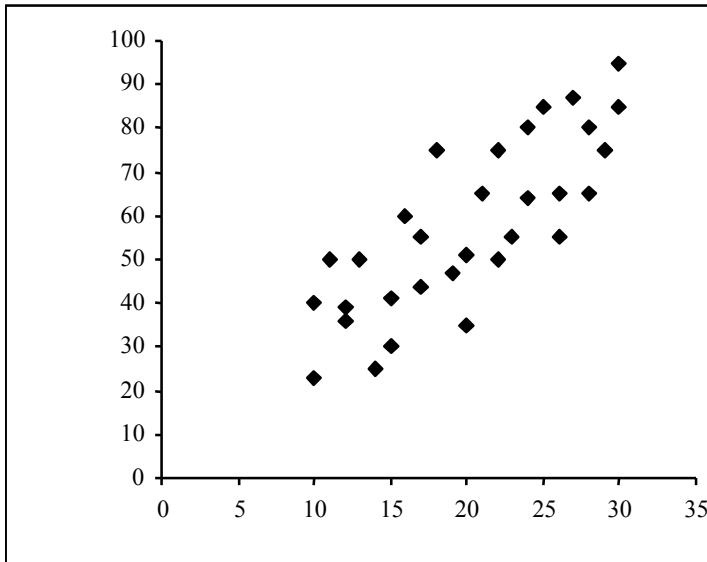
1. The histogram shows the distribution of total box office receipts of all time highest grossing movies.<sup>1</sup> (Those that made over \$100 million.)
  - a) What is on each of the axes? Give units.
  - b) What is the relationship between the mean and median? Give a reason.
  - c) From the histogram, estimate (*very* roughly) the mean and median.
  - d) How would you explain to someone who does not know statistics what the mean, median and standard deviation represent in this example? Which are resistant?
  - e) What proportion of the movies represented here made over 500 million?



2. The weights of fish of type A are normally distributed with mean 3.24 pounds and standard deviation 0.41 pounds. The weights of fish of type B are normally distributed with mean 5.73 pounds and standard deviation 0.81 pounds.
  - (a) Find the standardized weight for a 3.82 pound fish of type A.
  - (b) Find the standardized weight for a 6.80 pound fish of type B.
  - (c) Which fish is the heaviest, relative to its population?
  - (d) Find the weight of a fish of type A that is heavier than 90% of its population.
  - (e) Find the weight of a fish of type B that is lighter than 90% of its population.

<sup>1</sup> <http://www.the-numbers.com/movies/records/100million.php>

3. In the scatterplot below, scores on a foreign language exam are being predicted by scores on a standard test of verbal ability.



- What do the numbers on the horizontal axis represent?
- What do the numbers on the vertical axis represent?
- Roughly sketch a least squares regression line on the scatterplot.
- Estimate the slope of your line. Show your calculation.
- What is the approximate equation of your line?
- Use your equation to predict the foreign language score of someone who got 10 on the verbal ability test. (Show the calculation.)
- Approximately what the maximum error in the predictions given by x-values of data points on the line?
- (Optional) Mark the point  $(\bar{x}, \bar{y})$  on the graph. (Don't do any calculations; just pick a reasonable point.)
- (Optional) Explain as precisely as you can what knowing the value of the value of  $r^2$  would tell you about the regression line.

4. The NCES (National Center for Educational Statistics) publishes data US engineers and scientists that gives the following joint probability distribution.<sup>2</sup>

	Scientist	Engineer	
BA	0.289	0.343	0.632
MA	0.146	0.098	0.244
PhD	0.091	0.017	0.108
Other	0.003	0.013	0.016
	0.529	0.471	1

In the population described by this table, find

- (a) The proportion of engineers
- (b) The proportion with PhDs.
- (c) The proportion of people who are engineers with PhDs.
- (d) The proportion of scientists who have PhDs.
- (e) The proportion of scientists who do not have PhDs.
- (f) The proportion of BAs who are scientists.

---

<sup>2</sup> Reported in *Intro Statistics*, 9<sup>th</sup> ed, N. Weiss (Pearson, 2012).