Math 263 Solutions to Review for Exam 1: Deb Hughes Hallett

1.

(a) Horizontal: Box office receipts in dollars. Vertical: Number of movies
(b) Mean larger because skewed right.
(c) There are 449 movies.
To find the **median**, count from the left till you find the 225^{th} movie; it is in the first bar. Thus the median is less than \$150 million, but not a whole lots less Any estimate a bit less that \$150 is reasonable.
For the **mean**, look for the balance point. Since the distribution is skewed right, the mean will be large that the median, and it looks like it will be larger than \$150 million, but not much.
From the original data, we have:

Mean	\$ 166,518,750.32
Median	\$ 139,605,150.00
Standard deviation	\$ 76,027,162.98

(d) Mean: Balance point; Not resistant
Median: Middle value: Resistant
Standard deviation: "Average distance from the average": Not resistant
(e) Since three out of the 449 moves make over \$500 million

$$\text{Proportion} = \frac{3}{449} = 0.0068 = 0.68\%.$$

2.

(a) Normalizing gives

$$z = \frac{3.82 - 3.24}{0.41} = 1.415.$$

.

(b) Normalizing gives

$$z = \frac{6.80 - 5.73}{0.81} = 1.321.$$

.

(c) Since the *z*-score of the fish of type A is larger, this fish is heavier relative to its population.

(d) From the table or calculator, the *z*-score of a fish that is heavier than 90% of its population is 1.28.
Let *x* be the weight of a fish of type A. If

$$\frac{x - 3.24}{0.41} = 1.28$$

$$x = 1.28(0.41) + 3.24 = 3.76.$$

Thus, a 3.76 pound fish of type A is heavier than 90% of the type A fish.

(e) Since the fish is lighter than 90% of its population, it is heavier than 10% of its population. From the table or calculator, the *z*-score of a fish that is heavier than 10% of its population is – 1.28.
Let *y* be the weight of a fish of type B. If

$$\frac{y - 5.73}{0.81} = -1.28$$
$$y = -1.28(0.81) + 5.73 = 4.69.$$

Thus, a 4.69 pound fish of type B is lighter than 90% of the type B fish.

3. We have
(a) Scores on verbal test.
(b) Scores on foreign language exam.
(c) Draw a line following trend of data.
(d) The points (10, 30) and (30, 90) are approximately on the line.

$$\text{Slope} = \frac{90 - 30}{30 - 10} = 3.$$

(e) The line goes approximately through the origin, so equation is approximately $y = 3x$. Both points we used to find the slope are on this line.
(f) Foreign language score $= 3(10) = 30$.
(g) Read the largest vertical distance of a data point from the line. Perhaps about 25.
(h) The point must lie on the line sketched in part c). The coordinates are the mean values of each variable, so should be roughly in the middle. In general, the point $(\bar{x}, \bar{y})$ is not one of the original data points.
(i) The value of r^2 tells us the fraction of variation in the *y*-values from the mean that is explained by the regression line. If y_i is an observed value, and $\hat{y}_i$ is a predicted value, and $\bar{y}$ is the mean, we have

$$R^2 = \frac{\text{Variation of points on line from } \bar{y}}{\text{Variation of orginal data points from } \bar{y}} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

More roughly, the value of r^2 tells us how closely the data points fit the regression line.

4. We have:
(a) Proportion of engineers is 0.471 or 47.1%
(b) Proportion of PhDs is 0.108 or 10.8%.
(c) Proportion of people who are engineers with PhDs is 0.017 or 1.7%.

Restricting to scientists, we look at the proportions as a fraction of 0.529, the proportion of scientists.
(d) Proportion of scientists who have PhDs is

$$\frac{0.091}{0.529} = 0.172 = 17.2\%.$$

(e) Proportion of scientists who do not have PhDs is $1 - 0.172 = 0.828 = 82.8\%$.

Restricting to those with Bas, we look at proportion as a fraction of 0.632:
(f) The proportion of those with BAs who are scientists is $0.289/0.632 = 0.457 = 45.7\%$.