**Name:** *Caleb McWhorter — Solutions*
**MAT 121**
**Summer 2019**
**Homework 2**

*"I understand now that boundaries between noise & sound are conventions. All boundaries are conventions, waiting to be transcended. One may transcend any convention, if only one can first conceive of doing so."*

*–Robert Frobisher, Cloud Atlas*

**Problem 1:** Construct a frequency distribution, relative frequency distribution, and cumulative frequency distribution table for the following data using 6 classes. Be sure to list your lower class limits, upper class limits, class boundaries, and class widths. Also, sketch a histogram for the frequency distribution of the data and describe the distribution (symmetric, left/right skewed, neither, etc.).

| | | | | |
|---|---|---|---|---|
| 8.0 | 36.1 | 40.3 | 25.7 | 21.9 |
| 16.8 | 40.4 | 39.0 | 8.1 | 27.8 |
| 36.1 | 31.4 | 33.7 | 29.3 | 34.5 |

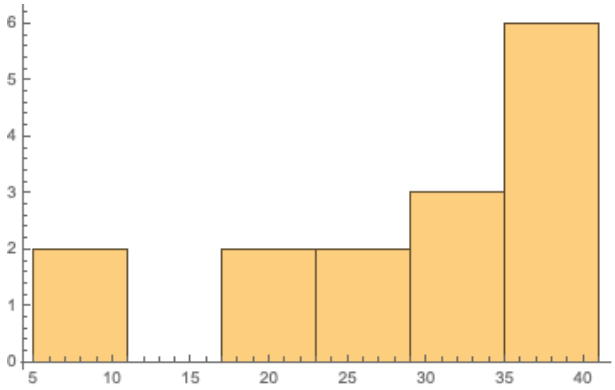
***Solution.*** *We have* $15$ *values, the largest value is 40.4 and the smallest value is 8. We want 6 classes so we need a class width of approximately* $\frac{40.4-8}{6} = 5.4$*. Given the dataset, it may be useful to treat the numbers as being between* $5$ *and* $45$ *so that we should use class width approximately* $\frac{45-5}{6} = 6.6$*. So we want a class width between 5.5 and 6.6, so that a class width of 6 seems appropriate (starting at 5). This yields a frequency distribution of. . .*

| Class | Frequency | Relative Frequency (%) | Cumulative Frequency (%) |
|---|---|---|---|
| 5–10 | 2 | 13.3 | 13.3 |
| 11–16 | 0 | 0.0 | 13.3 |
| 17–22 | 2 | 13.3 | 26.6 |
| 23–28 | 2 | 13.3 | 39.9 |
| 29–34 | 3 | 20.0 | 59.9 |
| 35–40 | 6 | 40.0 | 99.9 |

*This frequency distribution has the following table values:*

| | |
|---|---|
| Lower Class Limits: | 5, 11, 17, 23, 29, 35 |
| Upper Class Limits: | 10, 16, 22, 28, 34, 41 |
| Class Boundaries: | 4.5, 10.5, 16.5, 22.5, 28.5, 34.5, 40.5 |
| Class Widths: | 6 |

*A histogram for the dataset can be found in the figure to the right. Observe that the dataset is clearly left skewed.*



**Problem 2:** Construct a leaf-and-stem plot of the following data. Also, say whether the distribution is left/right skewed, symmetric, or neither.

$$100, \quad 125, \quad 114, \quad 130, \quad 155, \quad 166, \quad 125, \quad 108, \quad 129, \quad 136, \quad 110, \quad 115, \quad 119$$

**Solution.** The distribution is also clearly right skewed.

| Stem | Leaf |
|---|---|
| 10 | 08 |
| 11 | 0459 |
| 12 | 559 |
| 13 | 06 |
| 14 | |
| 15 | 5 |
| 16 | 6 |

Stem Unit: 10

**Problem 3:** The following table is a count of 'Mexican' food preferences in a bridge club, broken down by age.

| Location/Age | 18 – 25 | 26 – 50 | 51 – 100 |
|---|---|---|---|
| Taco Bell | 5 | 6 | 6 |
| Chipotle | 61 | 44 | 30 |
| Moes | 3 | 3 | 8 |
| Del Taco | 27 | 31 | 19 |

(a) How many people between the ages of 26 and 50 were surveyed?

$$6 + 44 + 3 + 31 = 84$$

(b) How many people preferred Moes?

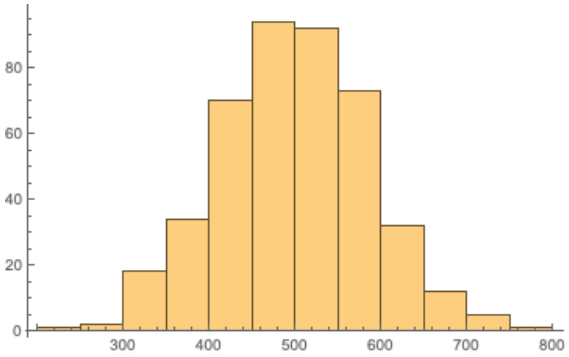
$$3 + 3 + 8 = 14$$

(c) How many people preferred Taco Bell or were aged 18 to 25?

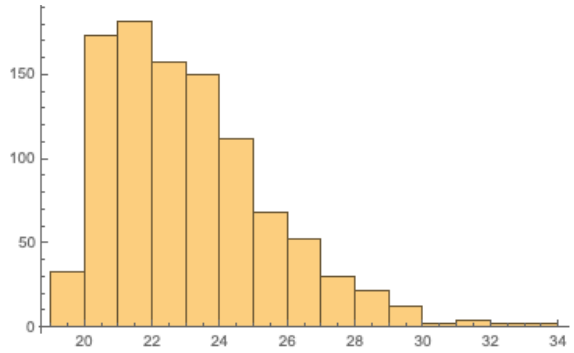
$$5 + 6 + 6 + 61 + 3 + 27 = 108$$

(d) How many people surveyed were aged 51 to 100 and prefered Del Taco?

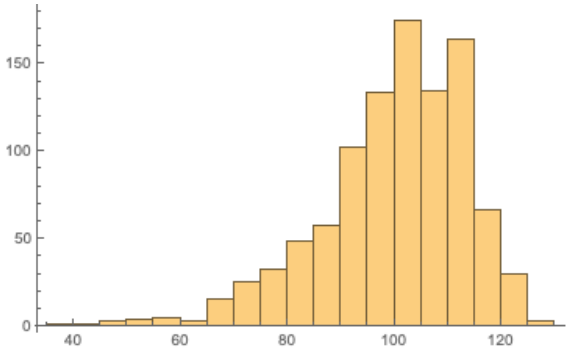
**Problem 4:** Determine if the following distributions are symmetric or skewed. If it is skewed, indicate whether the distribution is right skewed or left skewed.



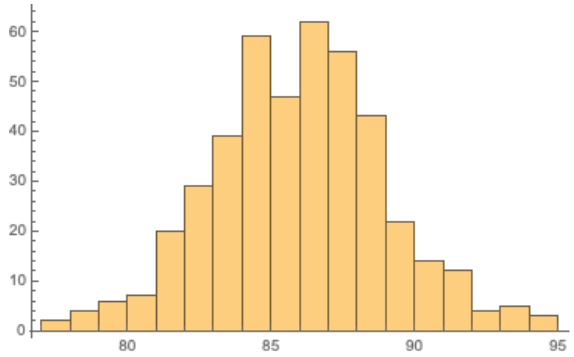

(a)




(b)




(c)




(d)

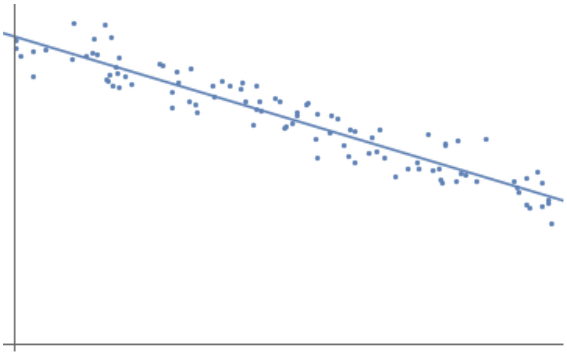
(a) *Symmetric*

(b) *Right Skewed*

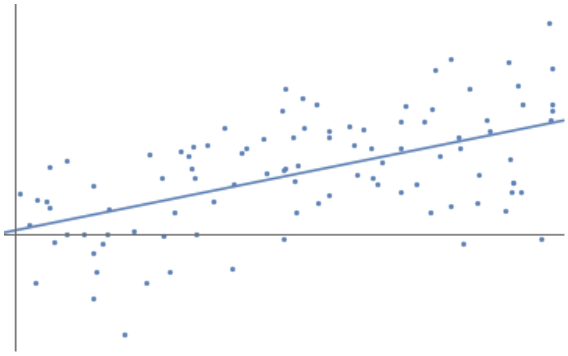
(c) *Left Skewed*

(d) *Symmetric*

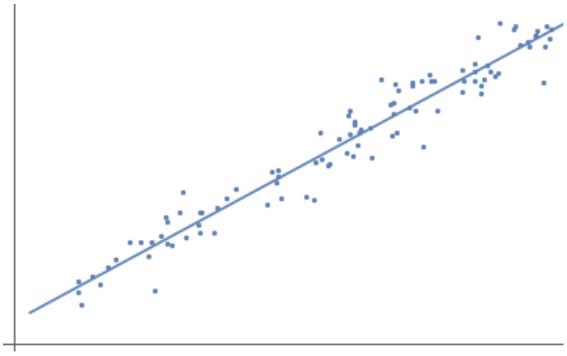
**Problem 5:** Match each of the following scatterplots with its regression equation and correlation coefficients. Note that the scale on each axes is the same.



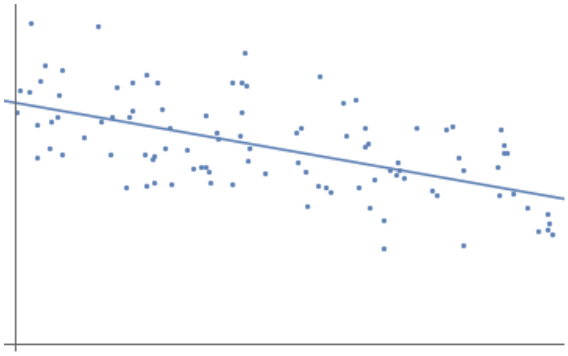
(a)



(b)



(c)



(d)

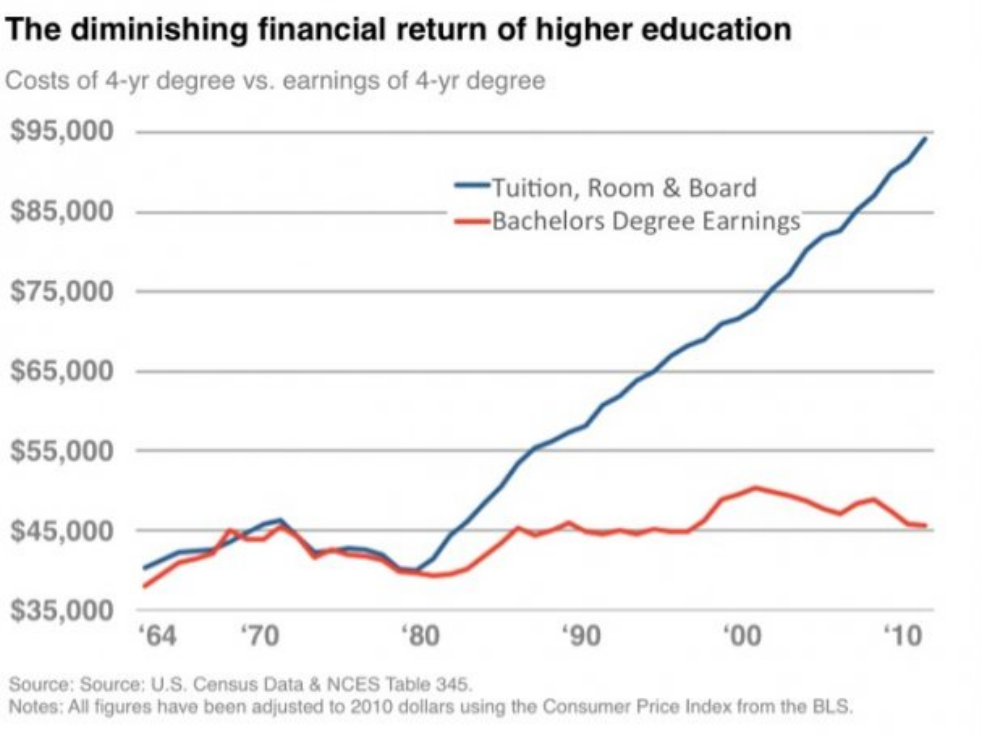
_____*b*_____ : $r = 0.556$, $\hat{y} = 1.7\hat{x} + 2.1$

_____*c*_____ : $r = 0.972$, $\hat{y} = 4.27\hat{x} + 26.2$

_____*d*_____ : $r = -0.641$, $\hat{y} = -113.2\hat{x} + 2478.0$

_____*a*_____ : $r = -0.946$, $\hat{y} = -97.8\hat{x} + 1877.6$

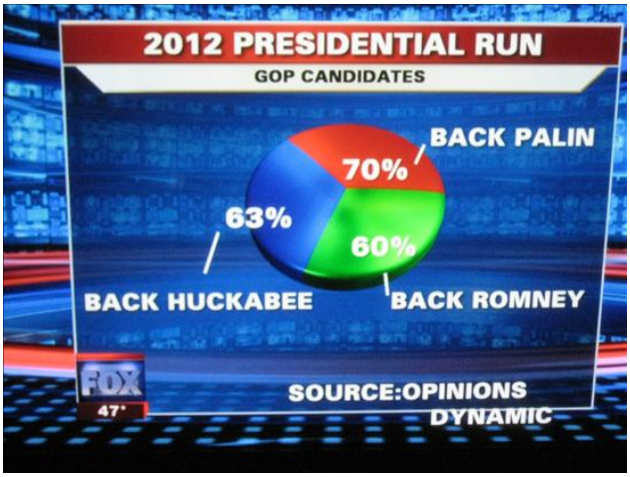
**Problem 6:** Identify problems (if any) in the presentation of data in the following chart. Identify how you might correct the figure.




***Solution.*** *The graph is just fine but as a stand-alone figure, the title could be misleading. The graph shows that the cost of college is increasing. However, the figure does not indicate or take into account the income difference between a college graduate job and a non-college graduate job. This is a concept one often sees in an introductory microeconomics cost, e.g. opportunity cost. For example, a cost of college is tuition, while a cost of not going to college is the possible 'loss' of income from obtaining a lower paying job.*

*To understand this better in the context of the graph, consider this possibility: a person spends $100,000 on a college degree and obtains a job that earns $50,000 a year. If one had not gone to college, they save $100,000 (though this is not money in pocket) but may only get a job that earns $40,000. Due to the difference in income, if the person making $50,000 a year paid $10,000 a year on their loan (ignoring interest), the college debt can be paid in 10 years. Both would then effectively be earning $40,000 a year. Then at the end of the 10 years, they would both have the same amount of money. But in another 10 years, the person with the $50,000 a year job would have $100,000 more in their bank account. [All assuming both spent no money.]*

**Problem 7:** Identify problems (if any) in the presentation of data in the following chart. Identify how you might correct the figure.




***Solution.*** *The percentages do not add up to 100%. In fact, they add up to much more than 100%. This could be an error or the percentages are not being measured from the total $n$, e.g. people were asked if they would vote for candidate $x$, allowing to say 'yes' to more than one candidate. Such a scenario would explain the over 100% total. Furthermore, the title perspective of the pie chart makes the 60% green portion look larger than the 70% red portion, despite the 70% being larger.*

**Problem 8:** Identify problems (if any) in the presentation of data in the following chart. Identify how you might correct the figure.


***Solution.*** *There is no $y$-axis. There is only an approximately 6% difference between the amounts but one bar is 4–5 times larger in size. The graph should begin at a true 0.*

Image: http://flowingdata.com/2009/11/26/fox-news-makes-the-best-pie-chart-ever/

Image: Media Matters