HW1 - Probability - Christopher Munoz

1.2) Are some cities more windy than others? Does Chicago deserve to be nicknamed "The Windy City"? Given below are the average wind speeds (in miles per hour) for 45 selected U.S. cities:

8.9 12.4 8.6 11.3 9.2 8.8 35.1 6.2 7.0
7.1 11.8 10.7 7.6 9.1 9.2 8.2 9.0 8.7
9.1 10.9 10.3 9.6 7.8 11.5 9.3 7.9 8.8 8.8
12.7 8.4 7.8 5.7 10.5 10.5 9.6 8.9 10.2
10.3 7.7 10.6 8.3 8.8 9.5 8.8 9.4
Source: The World Almanac and Book of Facts, 2004.

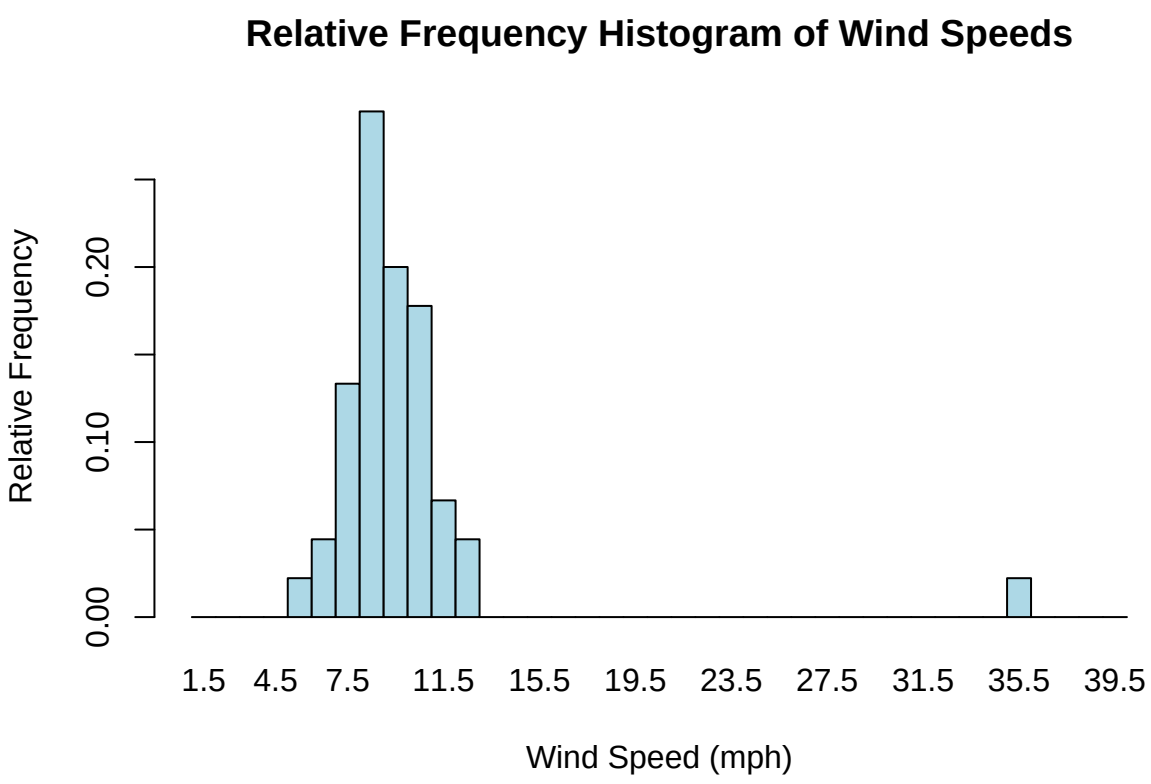
a. Construct a relative frequency histogram for these data. (Choose the class boundaries without including the value 35.1 in the range of values.)

```
wind_speeds_full <- c(8.9, 12.4, 8.6, 11.3, 9.2, 8.8, 35.1, 6.2, 7.0,
                      7.1, 11.8, 10.7, 7.6, 9.1, 9.2, 8.2, 9.0, 8.7,
                      9.1, 10.9, 10.3, 9.6, 7.8, 11.5, 9.3, 7.9, 8.8, 8.8,
                      12.7, 8.4, 7.8, 5.7, 10.5, 10.5, 9.6, 8.9, 10.2,
                      10.3, 7.7, 10.6, 8.3, 8.8, 9.5, 8.8, 9.4)

breaks <- seq(1, 40, by=1)

hist_data <- hist(wind_speeds_full, breaks=breaks, plot=FALSE)  #
rel_freq <- hist_data$counts / length(wind_speeds_full)  # Calculate

barplot(rel_freq, names.arg=hist_data$mids, space=0,
        main="Relative Frequency Histogram of Wind Speeds",
        xlab="Wind Speed (mph)", ylab="Relative Frequency", col="lightblue")
```



b. The value 35.1 was recorded at Mt. Washington, New Hampshire. Does the geography of that city explain the magnitude of its average wind speed?

Yes, apparently located between major storm tracks and thanks to its topography. But also on a more pedantic note, Mt Washington is not a "city". It may be the case that it shouldn't even be in this dataset since it doesn't belong in our population.

c. The average wind speed for Chicago is 10.3 miles per hour. What percentage of the cities have average wind speeds in excess of Chicago's?

```
count_above_chicago <- sum(wind_speeds_full > 10.3)
percentage_above_chicago <- (count_above_chicago / length(wind_speeds_full))
print(percentage_above_chicago)
```

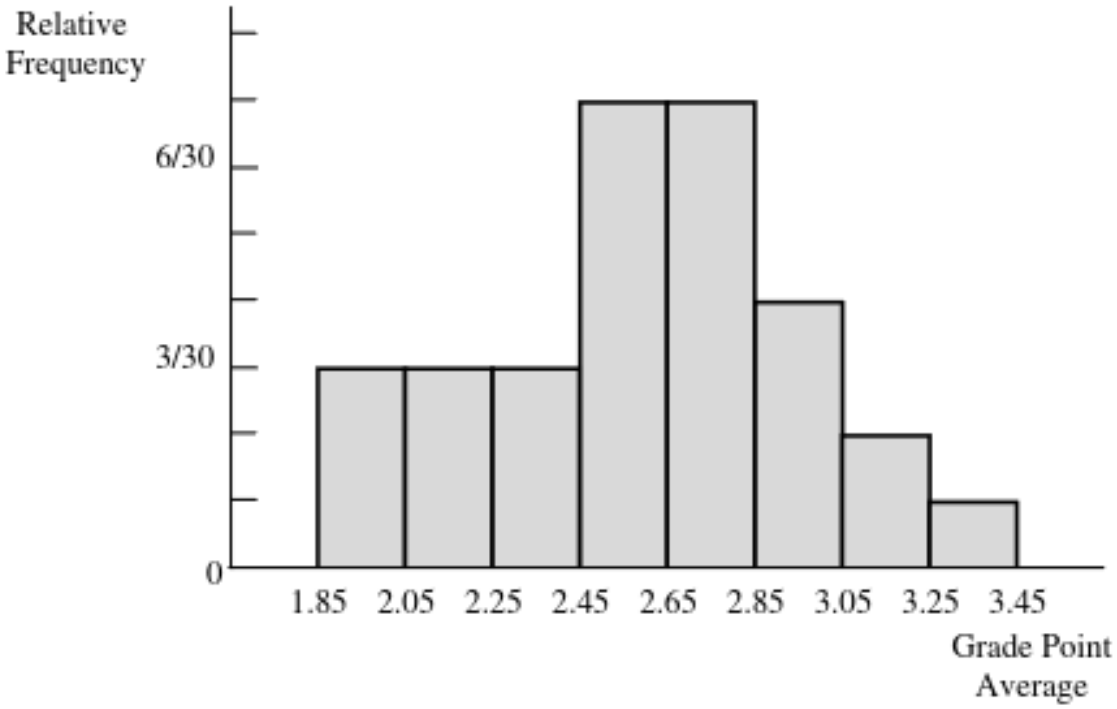
```
## [1] 0.2444444
```

The percentage of cities that have average wind speeds above Chicago is approximately 24.4444%

d. Do you think that Chicago is unusually windy?

Not especially in our data-set.

1.5) Given here is the relative frequency histogram associated with grade point averages (GPAs) of a sample of 30 students



a. Which of the GPA categories identified on the horizontal axis are associated with the largest proportion of students?

The categories ranging from 2.45 to 2.65 and 2.65 to 2.85 with 7 students in both of these ranges. 14 out of 30 students lie in this range.

b. What proportion of students had GPAs in each of the categories that you identified?

There are 7/30 students in each of these categories or "ranges".

c. What proportion of the students had GPAs less than 2.65

There are 16/30 students that had a GPA less than 2.65.

1.9)Resting breathing rates for college-age students are approximately normally distributed with mean 12 and standard deviation 2.3 breaths per minute. What fraction of all college-age students have breathing rates in the following intervals?

a. 9.7 to 14.3 breaths per minute
Since these values are 1 standard deviation away from the mean, 68%

b. 7.4 to 16.6 breaths per minute
Since these values are 2 standard deviations away from the mean, 95%

c. 9.7 to 16.6 breaths per minute

~50%	47.5%	34%	mean	34%	47.5%	~50%
5.1	**7.4**	**9.7**	**12**	**14.3**	**16.7**	**19**

34% + 47% = 81.4%

d. Less than 5.1 or more than 18.9 breaths per minute

This would be approximately 0%

1.13)Refer to Exercise 1.2.

a Calculate $\bar{y}$ and s for the data given.

```
print(mean(wind_speeds_full))
```

```
## [1] 9.791111
```

```
print(sd(wind_speeds_full))
```

```
## [1] 4.138007
```

The mean for the wind speeds is approximately 9.791111
The standard deviation is approximately 4.138007.

b Calculate the interval $\bar{y}\acute{s}ks$ for $k = 1, 2$, and 3. Count the number of measurements that fall within each interval and compare this result with the number that you would expect according to the empirical rule

k	interval	frequency	Exp.frequency
1	**5.65, 13.93**	**44**	**30.6**
2	**1.51, 18.07**	**44**	**42.75**
3	**-2.63, 22.21**	**44**	**45**

1.20) Weekly maintenance costs for a factory, recorded over a long period of time and adjusted for inflation, tend to have an approximately normal distribution with an average of $420 and a standard deviation of $30. If $450 is budgeted for next week, what is an approximate probability that this budgeted figure will be exceeded?

The appropriate calculation for this would be 100% - 68% and then dividing this by 2, so 16%.