

Types of Data

1. Which of the questions from the class survey have categorical/qualitative answers?

Solution: Gender, Birth Month (?), Country, State, Major, Employment Status.

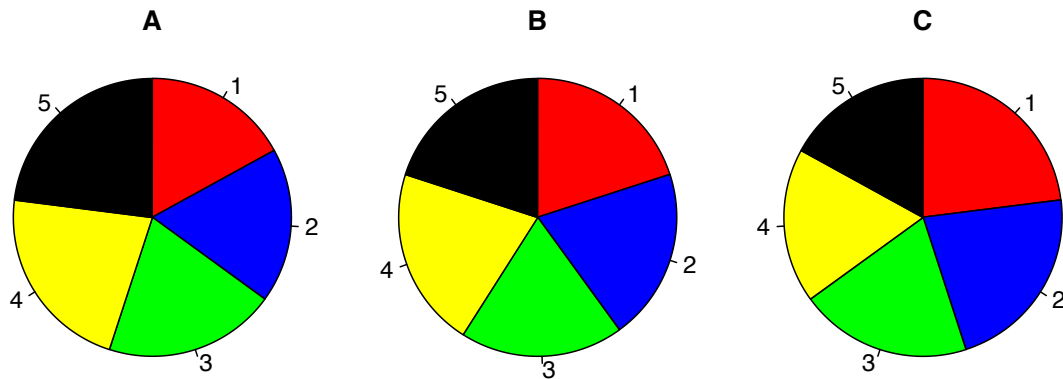
2. Which of the questions from the class survey have numerical/quantitative answers?

Solution:

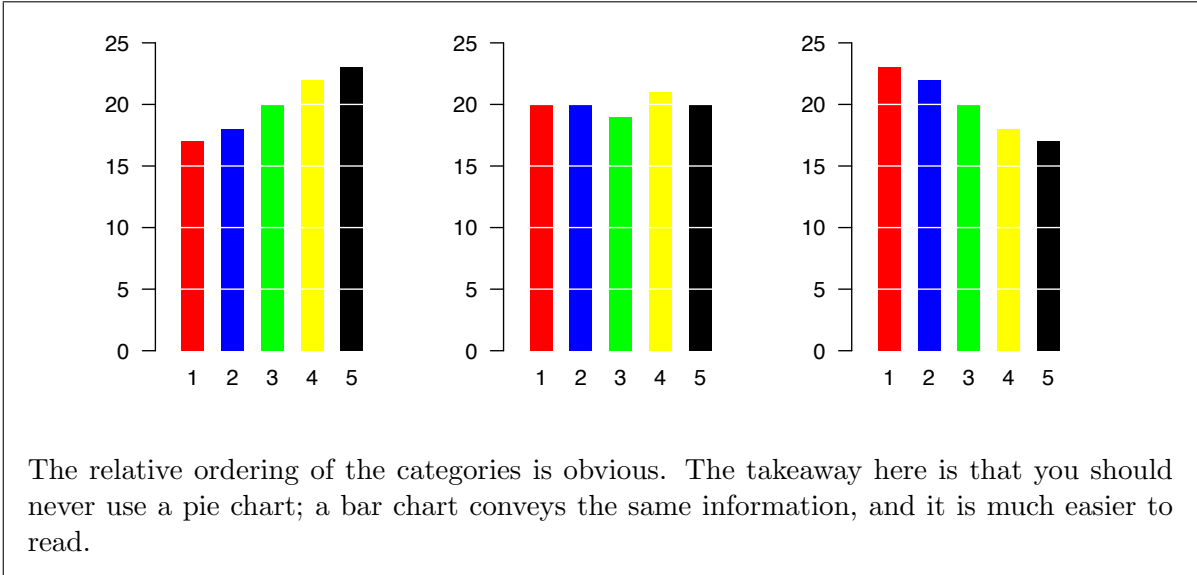
Birth Month (?), Study Hours per Week, GPA, SAT Score, Work Hours per Week, Expected Starting Salary, Dinners per Month, Pairs of Shoes, Social Media Hours per Week.

Describing Categorical (Qualitative) Data

3. Use the following pie charts to rank the categories (1–5) by size.



Solution: This is much easier if we have bar charts instead:



4. Draw what you think the bar chart for “Birth Month” will look like.

Solution:

5. Draw what you think the bar chart for “Major” will look like.

Solution:

Describing Numerical (Quantitative) Data

6. Draw what you think the histogram for “Expected Salary” will look like.

Solution:

7. Draw what you think the histogram for “Dinners per Month” will look like.

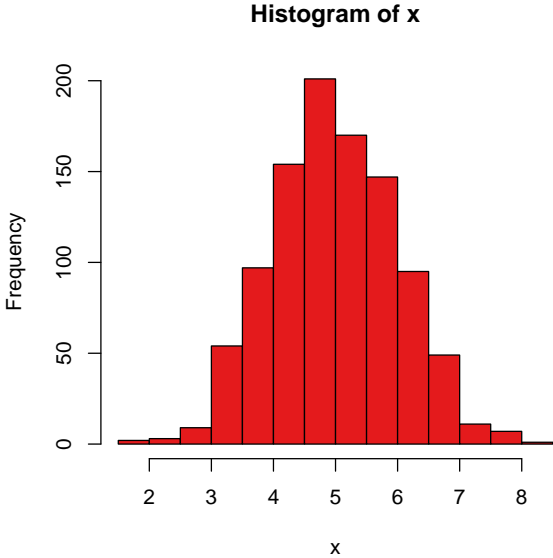
Solution:

8. Draw what you think the histogram for “Pairs of Shoes” will look like.

Solution:

Measures of Central Tendency

9. Here are some histograms. Estimate the mean and median of the data.
 (a) Symmetric and mound-shaped data.



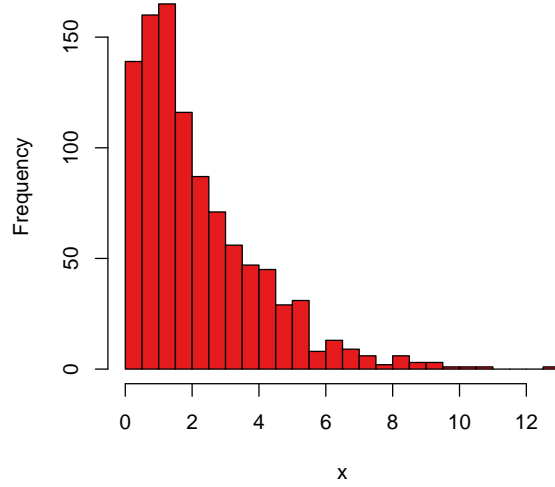
Solution: The median (solid) is roughly in the same place as the mean (dashed).

Histogram of x

x	Frequency
2	2
3	10
4	55
5	100
6	155
7	200
8	170
9	145
10	95
11	50
12	15
13	10
14	5
15	2

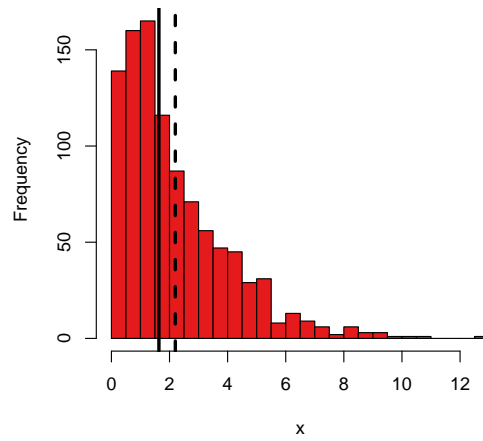
- (b) Skewed data.

Histogram of x

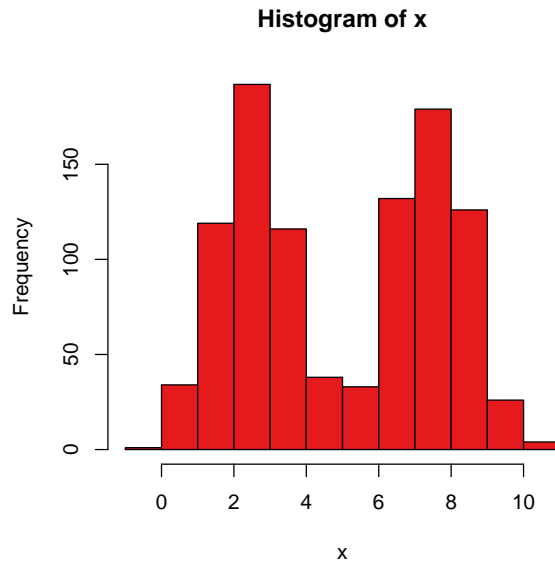


Solution: The mean is pulled to the right by the long tail.

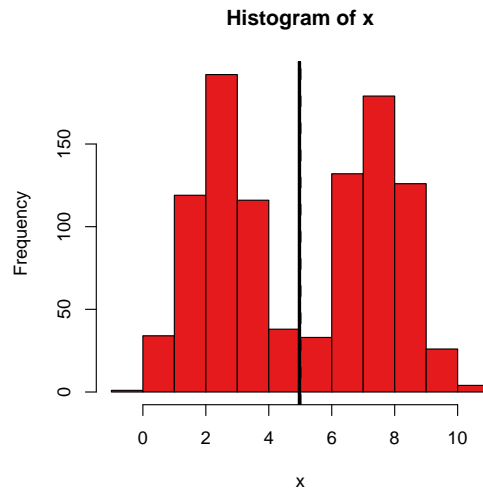
Histogram of x



(c) Bimodal data.



Solution: The median and the mean are roughly in the center. Note that neither number conveys much information about the distribution.



10. For the examples (a)–(c) of the previous problem, which is appropriate, the mean or the median?

Solution: (a) Both are appropriate; (b) the median is more appropriate; (c) neither is appropriate.

Percentiles

11. Find the quartiles of the following datasets:

(a) 10, 10, 11, 11, 12, 13, 14, 14, 15, 17, 17, 18, 18, 18, 22, 23

Solution: $Q_1 = 11.5$, $M = 14.5$, $Q_3 = 18$.

(b) 5, 7, 8, 8, 8, 8, 9, 10, 11, 13, 13, 19, 19

Solution: $Q_1 = 8$, $M = 9$, $Q_3 = 13$.