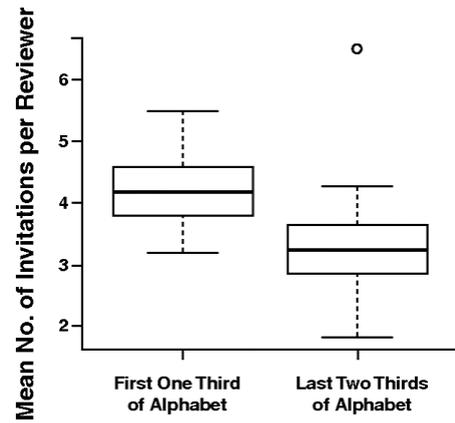


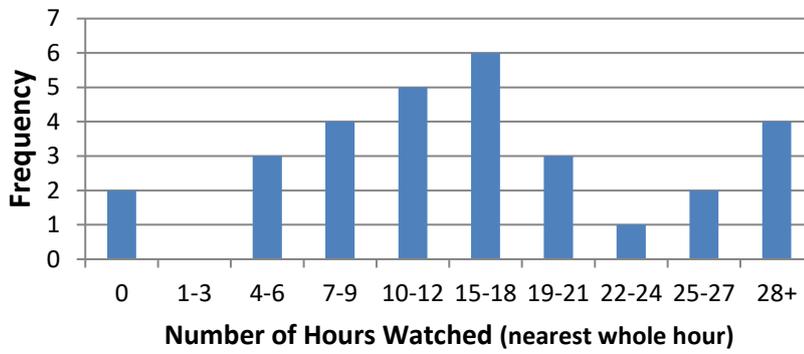
Routine Statistics Practice #3

1. It is claimed that the number of invitations a scientist gets to review papers is increased by having a name early in alphabetical order:
 - a) Does the graph on the right show the claim is correct?
 - b) The dot at the top right is an outlier. What does it represent in the real world, and why is it shown separately like that?



2. A class records how many hours of television they watch in one week.

Hours of TV watched in a Week



- a) What is wrong with this graph?
 - b) How many students were surveyed?
 - c) What is another way the data could be more graphed for analysis? How would that be better?
3. A company sells two types of devices. The table lists the sales values for recent years.

Year	Widgets (\$)	Gadgets (\$)
2014	8,000	4,000
2015	6,500	5,200
2016	7,400	5,000
2017	7,900	5,700
2018	8,100	7,000
2019	7,000	6,600
2020	8,300	7,500
2021	7,500	8,000

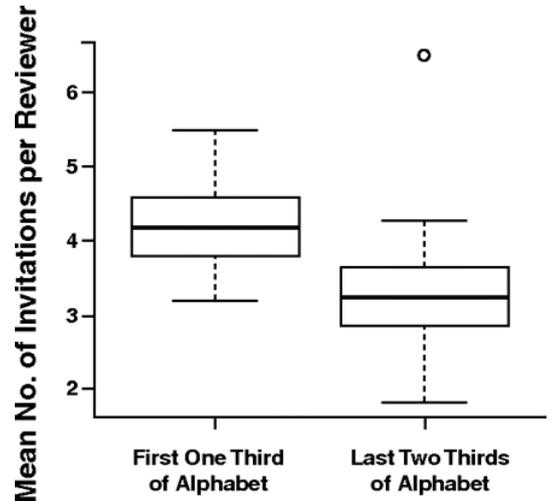


- a) Plot the appropriate graph to show the trend in sales.
- b) What does the graph show that the company might find useful?

Answers: Routine Statistics Practice #3

1. It is claimed that the number of invitations a scientist gets to review is increased by having a name early in alphabetical order:

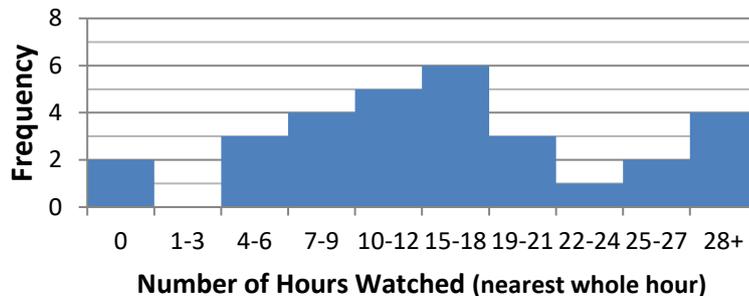
- a) Does the graph on the right show the claim is correct? **Yes. Each quartile is higher for the first third of the alphabet. The medians are outside the IQRs, so it is not likely a result of sample variation.**
- b) The dot at the top right is an outlier. What does it represent in the real world, and why is it shown separately like that?



It represents one person with 7 invitations. It is shown separately to stop a single value making the whisker very long, which could lead to the impression that many more values were in that whisker.

2. A class records how many hours of television they watch in one week.

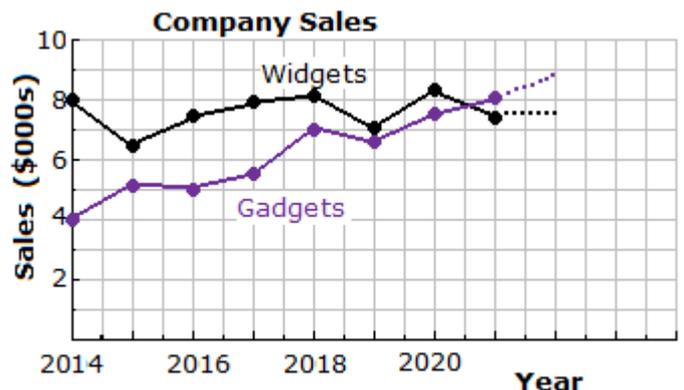
Hours of TV watched in a Week



- a) The graph should be a **histogram** (shown above) as the data is not categorical. Also the **"28+hours"** group is far **too vague**, and the values should keep going up.
- b) **30 students** were sampled.
- c) A **box-and-whisker** graph could be used. It would make comparing the class to other groups easier. It would clearly mark the extremes. It would not have to group the data, and would not have to round to the nearest hour.

3. A company sells two types of devices. The table lists the sales figures for recent years.

Year	Widgets	Gadgets
2014	8,000	4,000
2015	6,500	5,200
2016	7,400	5,000
2017	7,900	5,700
2018	8,100	7,000
2019	7,000	6,600
2020	8,300	7,500
2021	7,500	8,000



- a) It should be a **time series**, as shown. A bar chart is poor, and a histogram worse.
- b) Sales of widgets are not increasing although they vary a lot. **2022 is likely to be about \$8,000 of widgets sold.** Gadgets are showing a trend of a fairly constant increase in sales. **Next year is likely to have about \$9,000 gadgets sold.**