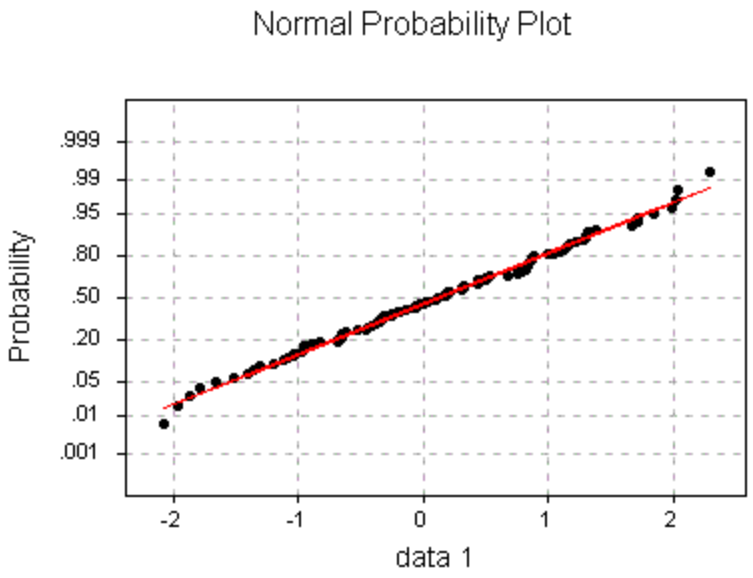
Math 361

Normal Probability Plots (Inv. 2.1)

Boxplots (Inv. 2.2)

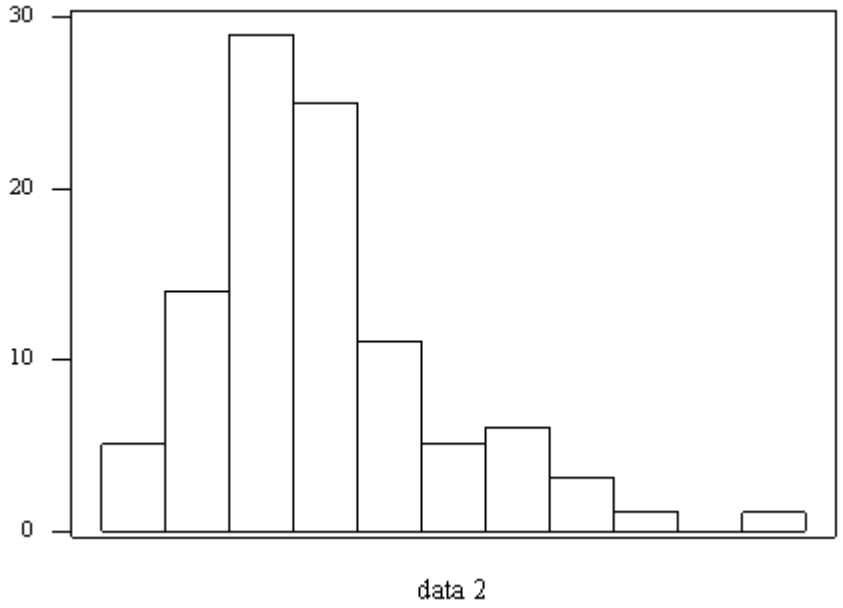
Normal probability plots

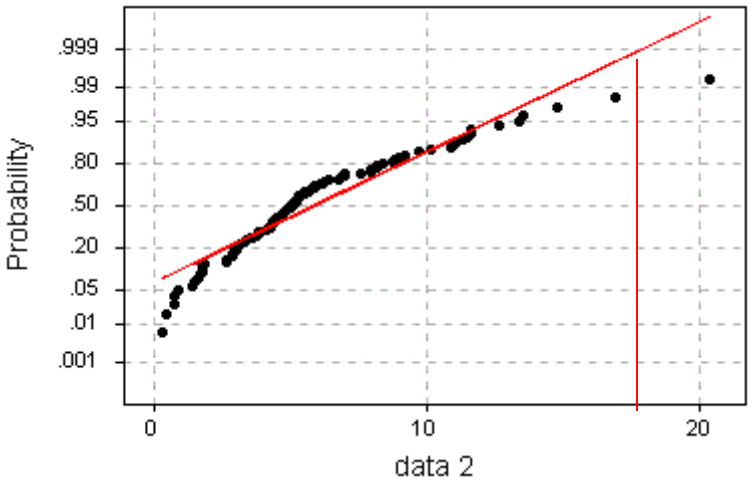
- A straight line means the data behave like observations from a normal distribution



Normal probability plots

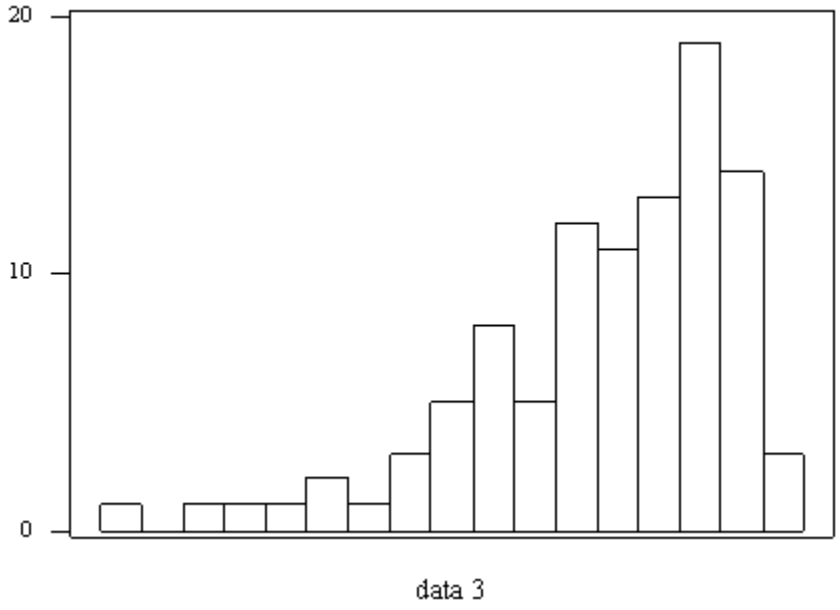
- Skewed to the right

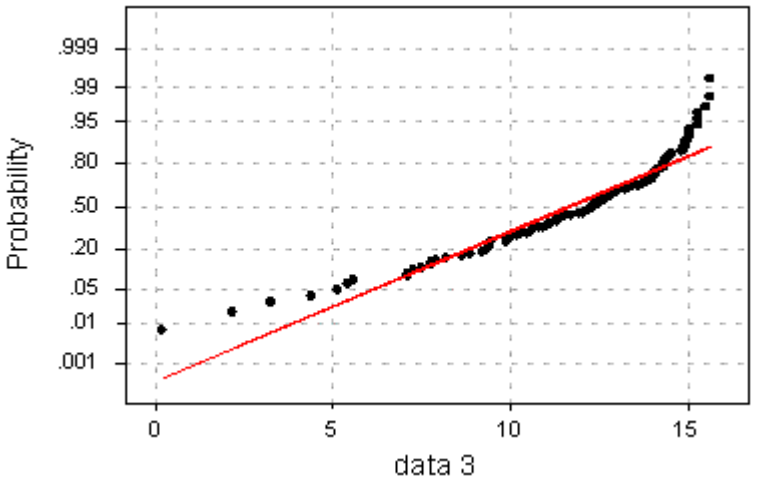




Normal probability plots

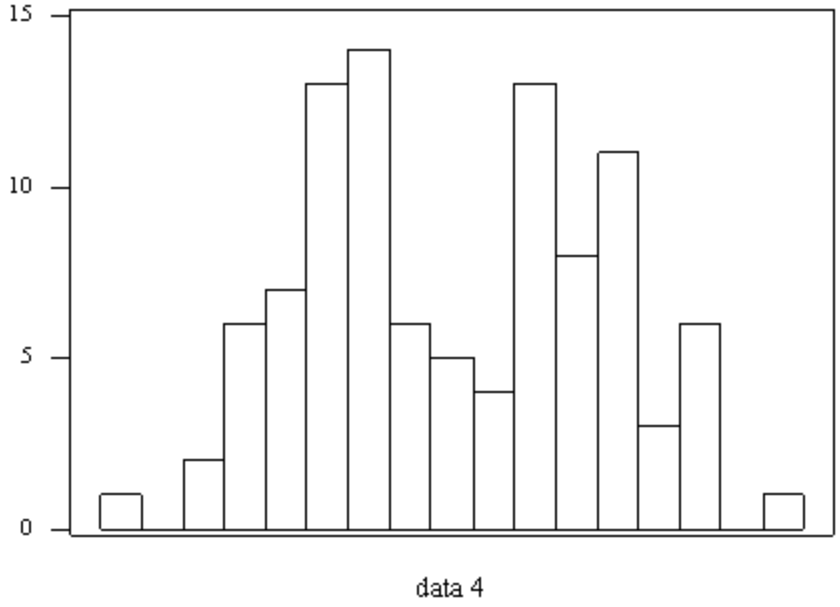
- Skewed to the left

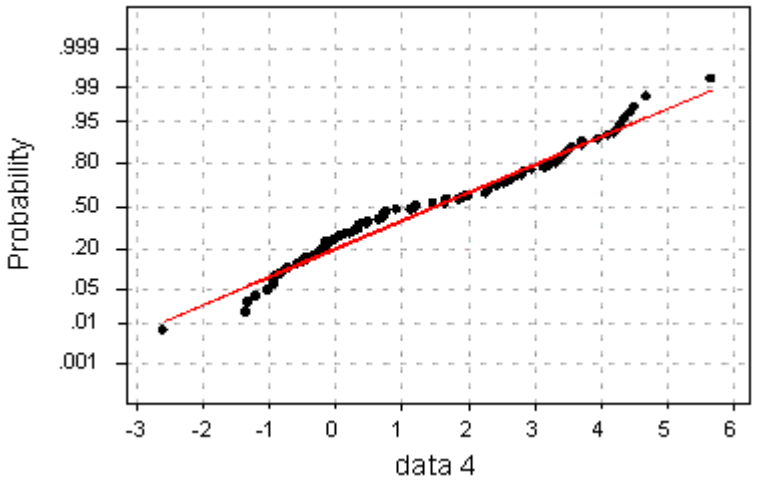




Normal probability plots

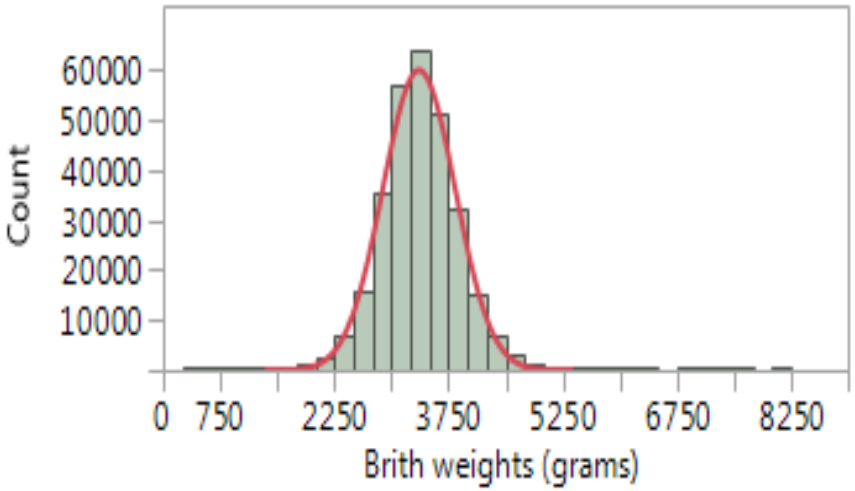
- Bimodal distribution

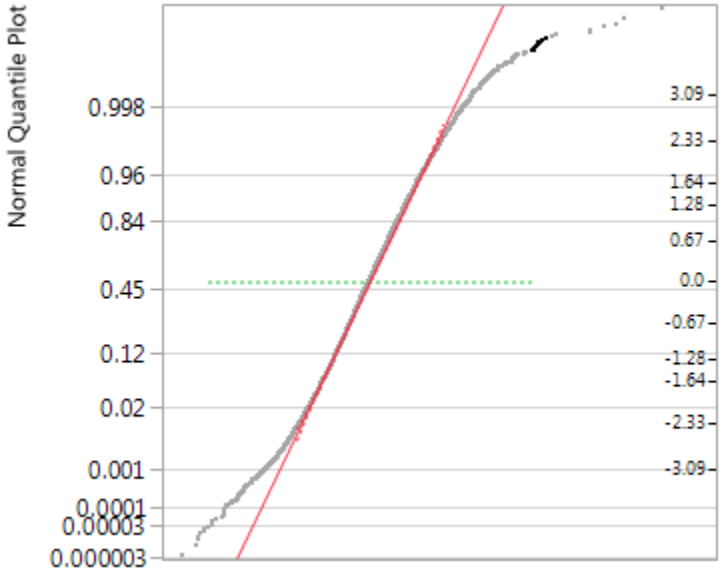




Normal probability plots

- Heavy tails





Applications

What can we do if a variable is **normally** distributed?

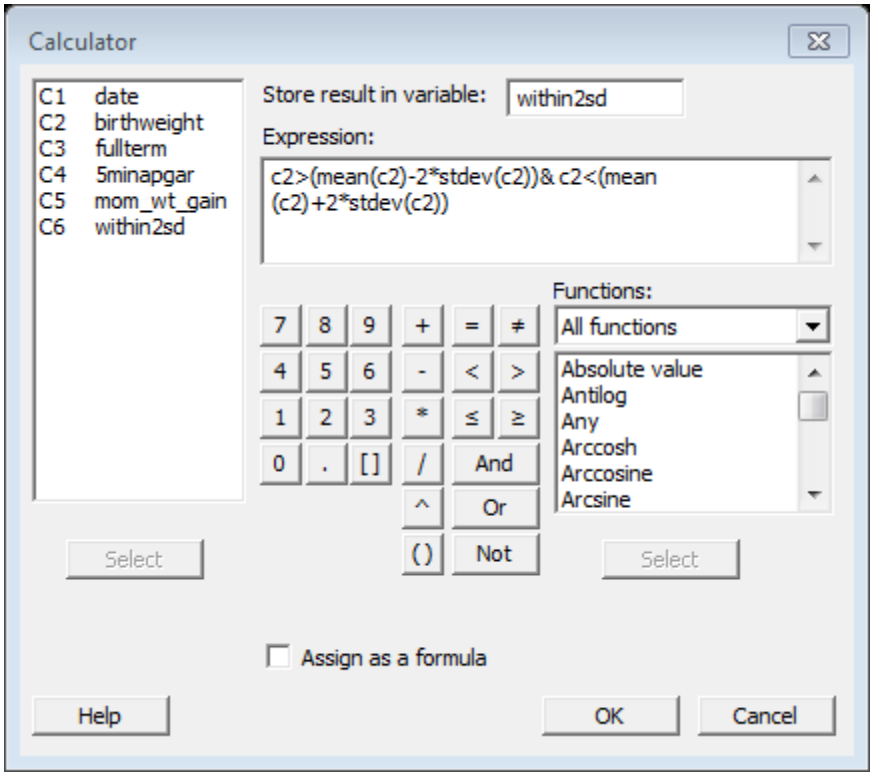
- Use the empirical rule to say where 68%, 95% or 99% of the data falls (i.e., within 1, 2, or 3 SDs of the mean)

- Predict the probability of being above or below a certain value.

Inv. 2.1 part (n)

Recall that the Empirical Rule states that about 95% of observations in a Normal Distribution are within 2 SDs of the mean.

Let's compute the proportion of Birthweights that are within 2 SDs of the mean…



Inv. 2.1 part (n)

What proportion of birth weights are within 2 SDs of the mean?

Tally for Discrete Variables: within2sd

Tally

within2sd	Count	Percent
0	13720	4.80
1	272187	95.20
N=	285907	

Inv. 2.1, part (q)

Normal Probability Calculator

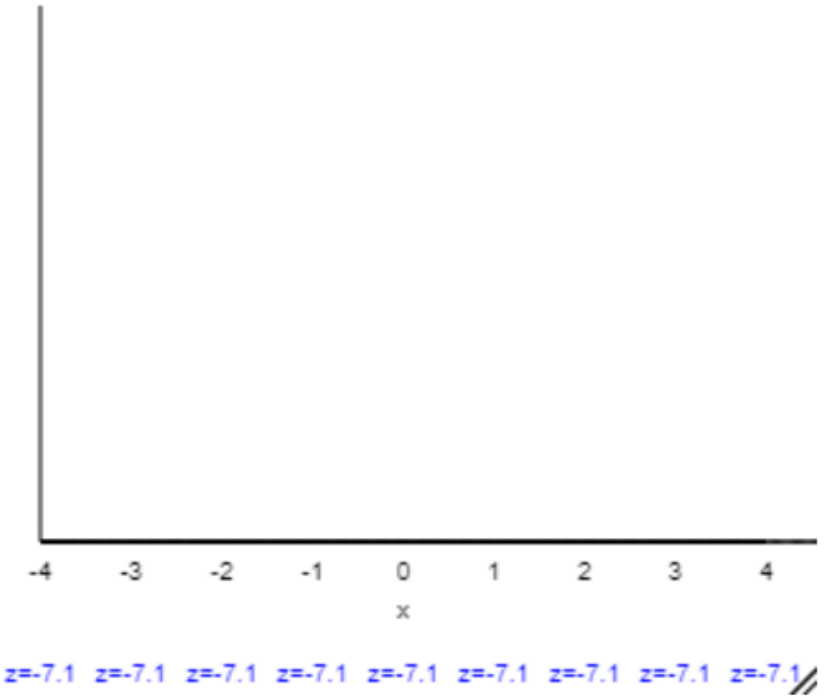
Variable: x

Mean: 3361.6 SD: 474.6

Mean: 0 SD: 2

Scale to Fit

		x	z	Probability
☑	<	2500	-1.815	0.0347
☐	<			



Using mean of 3361.6 and SD of 474.6 from part (k)

Inv. 2.1, part (r)

Actually, 3.3% of fullterm babies had low birthweights.

How many pairs of shoes do you own?

In the initial course survey, you answered:

6, 12, 22, 13, 20, 10, 6, 15, 10, 5, 17, 14, 6, 35, 15, 15, 22, 11, 6, 25, 5

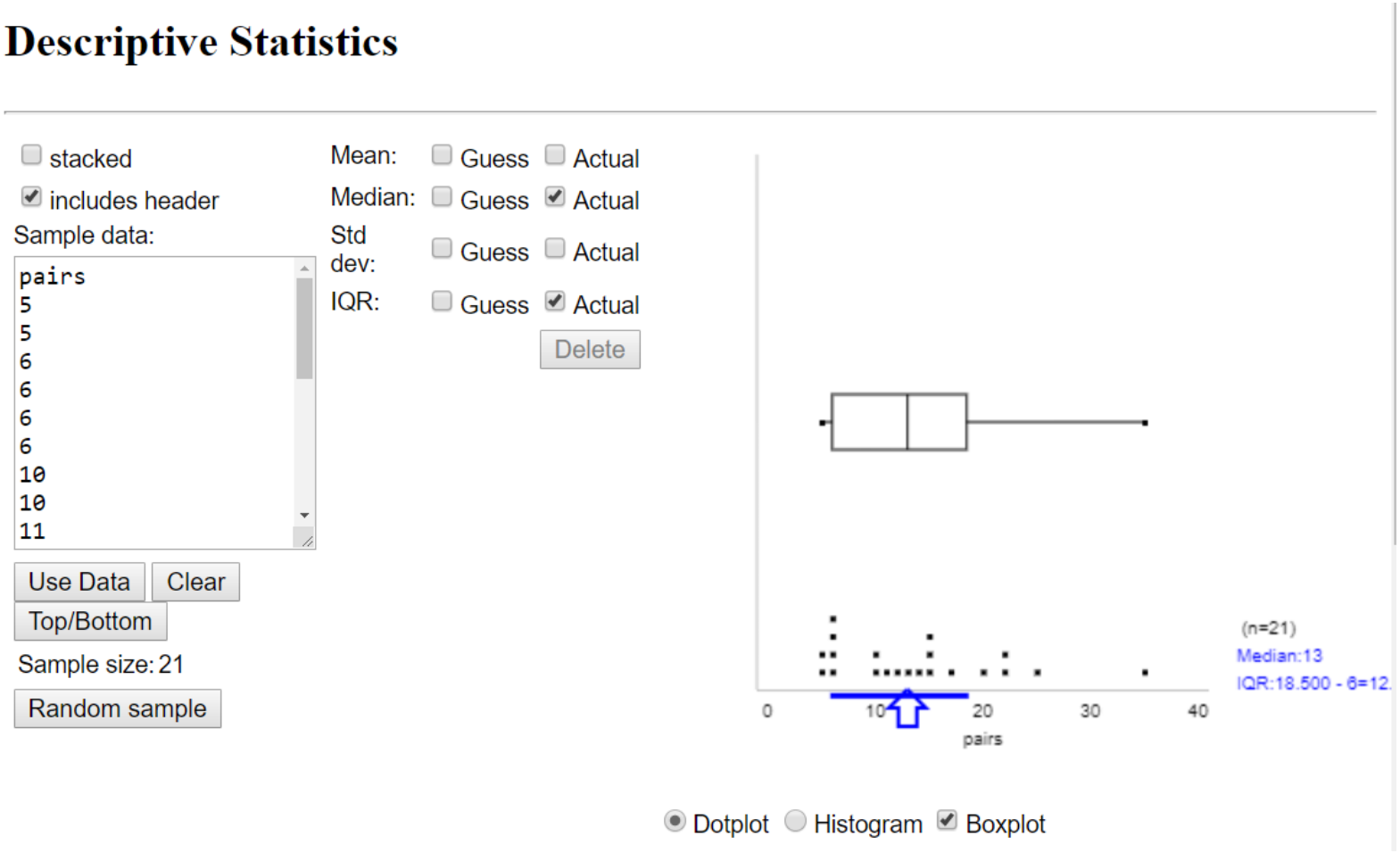
Find:

- median (Q2),
- first quartile (Q1),
- third quartile (Q3)
- Interquartile range (Q3-Q1)
- Minimum
- Maximum
- Any outliers (values more than 1.5xIQR from the median)

Draw a boxplot by hand

Five number summary *= min, Q1, median, Q3 and max*

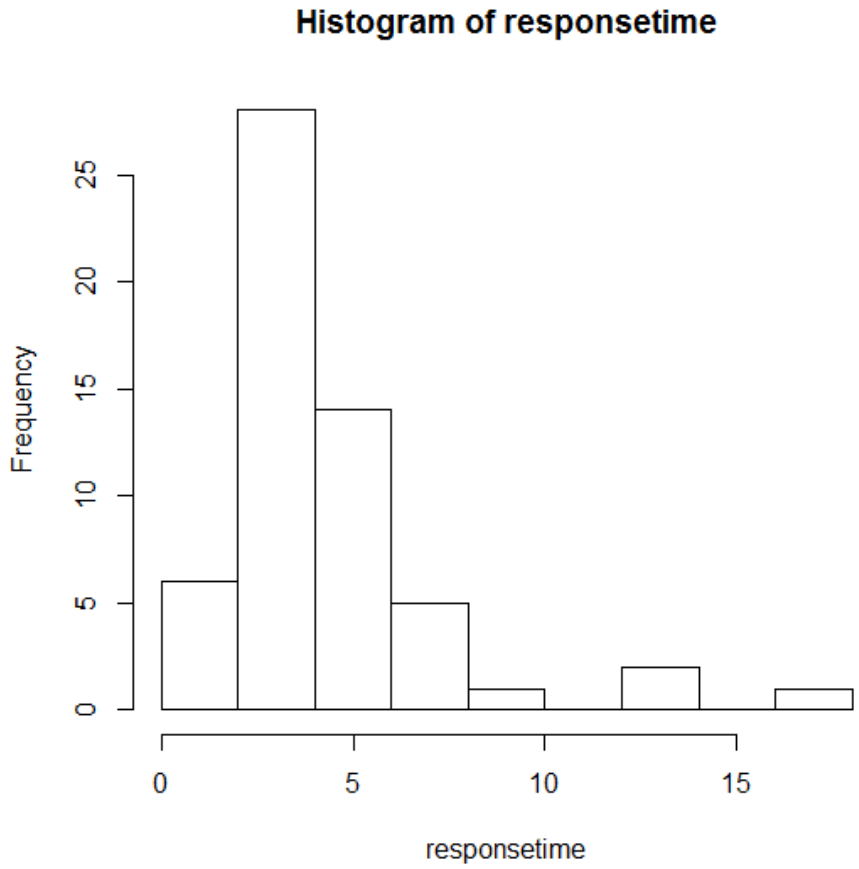
Boxplot using applet



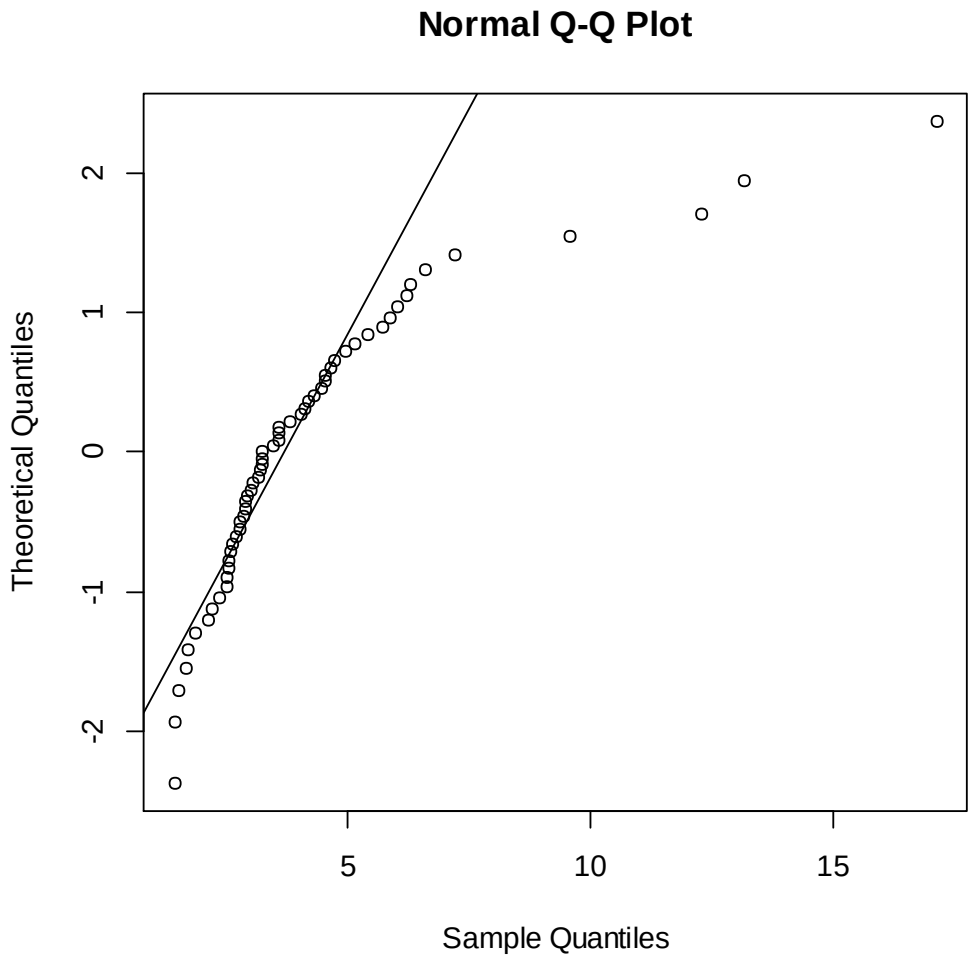
Inv. 2.2: Honking

Researchers blocked the intersection and recorded the length of time it took for the driver behind them to honk…

Inv. 2.2 part (b)


Histogram of responsetime
Frequency
0
5
10
15
20
25
0
5
10
15
responsetime

Normal Probability Plot

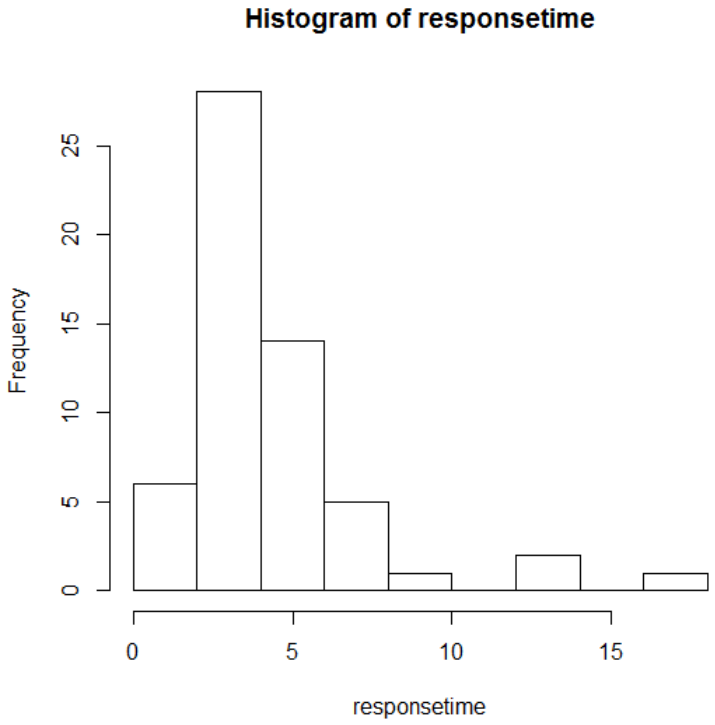


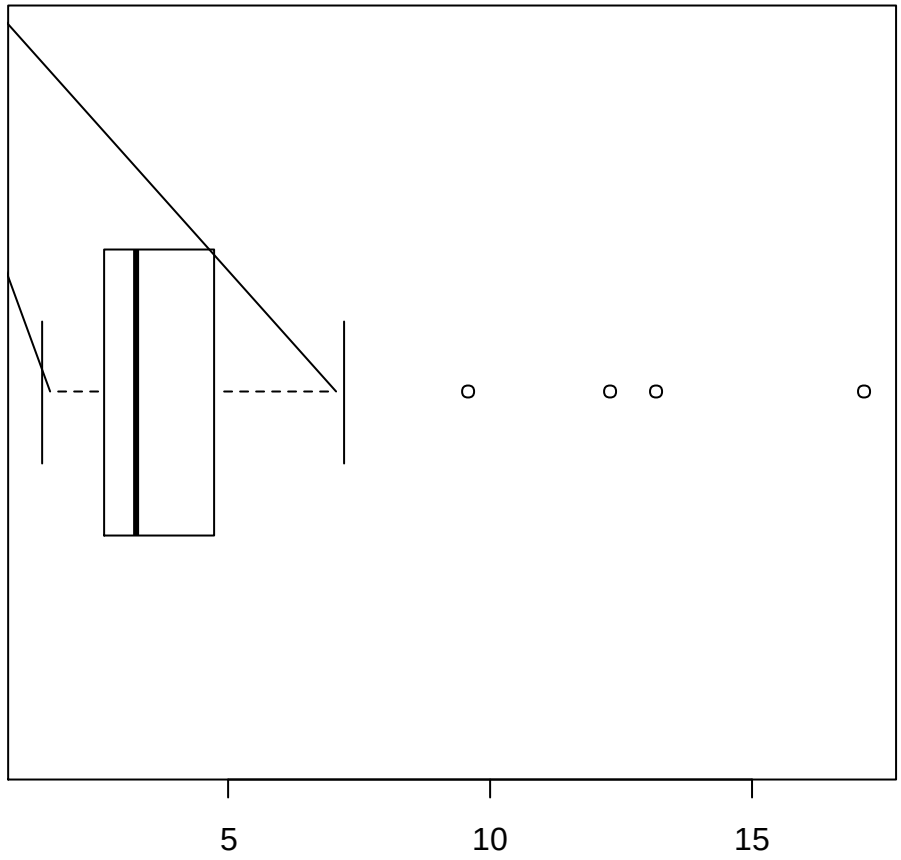
5 number summary

n	Min	Q1	Median	Q3	Max	Mean	SD
57.000	0.344	0.963	1.180	1.550	2.840	1.290	0.528

Use this information to draw a boxplot by hand

Modified Boxplot





Statistical Inference for 1 Quantitative Variable

Suppose we want to make *inferences* beyond the sample data

- Need random sample from population/process
- Need to how about the behavior of sample means from different random samples from the same population

Next time: Investigation 2.4 (p. 143)

- Wikipedia

The ***Ethan Allen*** was a 40-foot, glass-enclosed tour boat operated by Shoreline Cruises on Lake George in upstate New York. On October 2, 2005, at 2:55 p.m., with 47 passengers–all from Michigan and Ohio and mostly seniors–aboard, the *Ethan Allen* capsized and sank just south of Cramer Point in the Town of Lake George. Twenty passengers died. The accident caused government regulators to consider new laws on passenger boat capacity.

Contents [hide]

The *Ethan Allen* is raised to the surface of Lake George the day after it capsized