

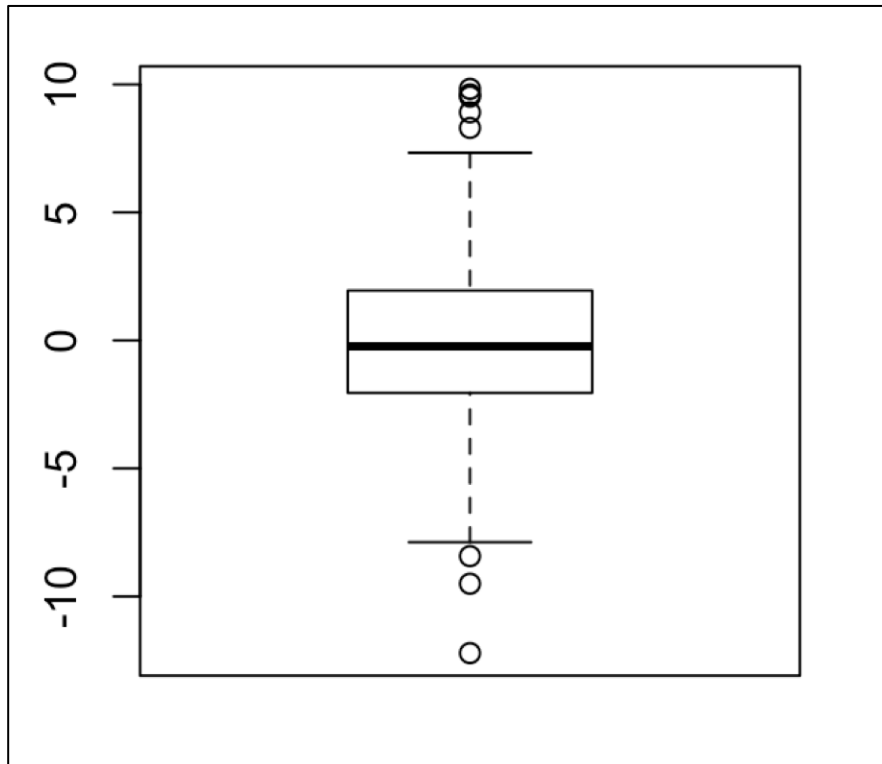
# Quiz Review

Stanford OLI Exploratory Data  
Analysis

True or False: A Boxplot visually depicts the standard deviation of a group of numbers.

- What is standard deviation of a group of numbers?
  - Write the equation for standard deviation from memory (on the board or on a piece of paper)
  - Explain what it means, in words.
  - Why it matters, what is it, what it means
- What is a boxplot?
  - Be able to point out and explain graphical elements of a boxplot.
  - Plots are just drawings with graphical elements - circles, lines, curves, etc - that signify something about a data set.

# Boxplot



```
> x = rnorm(1000,sd=3,mean=0)  
> boxplot(x)
```

outliers

Q1

Q3

median

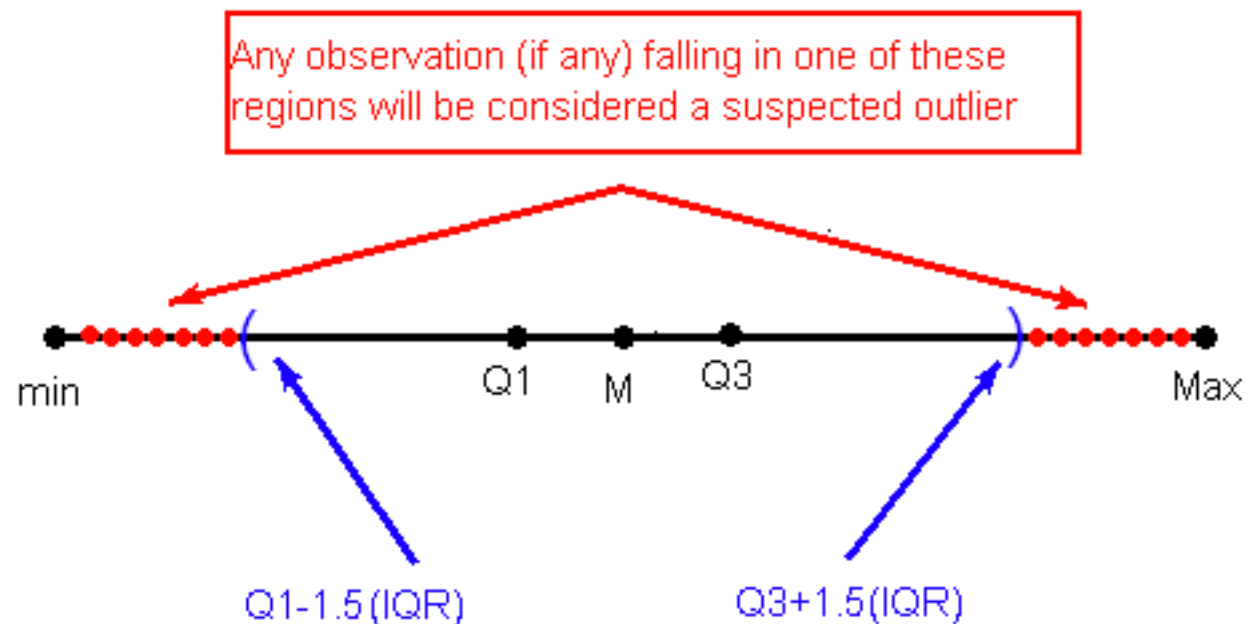
What is the y-coordinate for the whiskers?

## The 1.5(IQR) Criterion for Outliers

An observation is considered a suspected outlier if it is:

- below  $Q1 - 1.5(IQR)$  or
- above  $Q3 + 1.5(IQR)$

The following picture illustrates this rule:



A survey taken in a large statistics class contained the question: "What's the fastest you have driven a car (in miles per hour)?" The five-number summary for the 87 males surveyed is:

min = 55, Q1 = 95, Median = 110, Q3 = 120, Max = 155

Should the largest observation in this data set be classified as an outlier?

Try it in R ... related to today's homework due Friday.

A survey taken in a large statistics class contained the question: "What's the fastest you have driven a car (in miles per hour)?" The five-number summary for the 87 males surveyed is:

min = 55, Q1 = 95, Median = 110, Q3 = 120, Max = 155

Should the largest observation in this data set be classified as an outlier?

```
> Q1 = 95  
> Q3 = 120  
> Max = 155  
> iqr=Q3-Q1  
> upper_whisker=Q3+iqr*1.5  
> Max>upper_whisker
```

Which of the following plot types give information regarding the center, range and outliers for a single variable? Select all that apply.

Step one: Make sure you know what each plot looks like.

Step two: Ask yourself if the plot has a graphical element that shows **center**, **range**, **outliers**.

