**Topic 10**

**More Summary Measures and Graphs**

**Five Number Summary (FNS)**

1. Minimum
2. 1st Quartile
3. Median (2nd Quartile)
4. 3nd Quartile
5. Maximum

Example {1,3,5,4,3,8,7,6,0,6,8,5,3,5,3,4,8,7}

Ordered, {0,1,3,3,3,3,4,4,5,5,5,6,6,7,7,8,8}

Min = 0, Max=8

Next find the median. n=17 is odd, so find median using (n+1)/2 method

(n+1)/2=9, and the 9th value of the ordered set = 5.

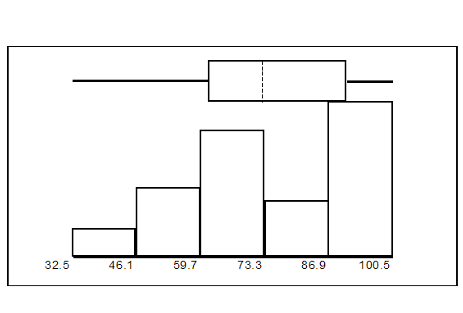
Median is 5.

On either side of 5 are eight values. The median of those values are the average of values 4 and 5 in each set. The median of the first set is Q1, and that turns out to be (3+3)/2=3, and the median of the second set is (6+7)/2=6.5, which is Q3. We can now build our FNS table

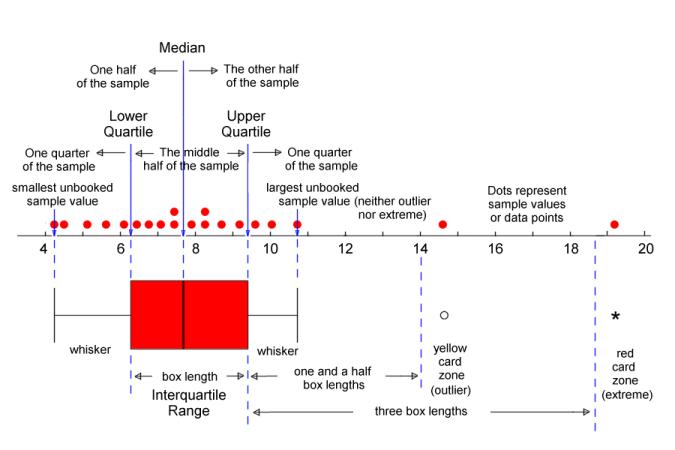
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Min** | **Q1** | **Med** | **Q3** | **Max** |
| **0** | **3** | **5** | **6.5** | **8** |

**Boxplot** (Box and Whisker)

* “Box” is middle 50% of data (Q1 to Q3) with median line
* Whiskers extend to min and max
* No outliers



**(Modified) Box and Whisker Plot** (or modified boxplot)



**Modified Boxplot**

* Box is the same
* Whiskers only extend to the last data point
  + <= Q1 - 1.5\*IRQ
  + >= Q1 + 1.5\*IRQ
* After that, they are labeled as outliers
* Sometimes, label outliers > 3\*IRQ differently

The boxplot is an extremely useful graphic for viewing *2 dimensional data* where

* One dimension is **categorical,** the other is **quantitative**
* We don’t know the parameters or arithmetic measures of the data
* We want a graphic representation of FNS
* The data distribution is asymmetric or funky in some other way

Recall that Ordinal can be treated as either quantitative or categorical, depending on what you want to show.

