

### 3.3 percentiles and Box-and-Whisker Plots

Percentile

## What are Percentiles?

- Quantitative data
- Remember standardized tests...
- Example: If you test at the 77<sup>th</sup> percentile, it means you did better than 77% of the people taking the test.
- If 100 people took the test, you'd have done better than 77 of them.



## Percentile Definition

Percentiles can be between 1 and 99

- You can't have a -2<sup>nd</sup> percentile, or a 105<sup>th</sup> percentile

Whatever number you pick:

- That % of values fall below the number
- And 100 minus that % of values fall above the number

Example: 20 people take a test.

- Let's say there is a maximum score of 5 on the test.
- The 25<sup>th</sup> percentile means 25% of the scores fall below this score, and 75% fall above that score.

Let's say it is an easy test, and 12 people get a 4, and the remaining 8 get a 5. The 25<sup>th</sup> percentile, or the score the cuts off the bottom 5 tests scores, will be 4. (Even the 50<sup>th</sup> percentile will be 4.)

This would come out very different if it were a hard test, and most people got below a score of 3.

# Quartiles

- Quartiles is a specific set of percentiles
  - 1<sup>st</sup> quartile: 25<sup>th</sup> percentile
  - 2<sup>nd</sup> quartile: 50<sup>th</sup> percentile (also median!)
  - 3<sup>rd</sup> quartile: 75<sup>th</sup> percentile
- These can be calculated by hand.



## Computing Quartiles

1. Order the data from smallest to largest.
  2. Find the median.
    - 2<sup>nd</sup> quartile
    - 50<sup>th</sup> percentile
  3. Find the median of the lower half of the data.
    - 1<sup>st</sup> quartile
    - 25<sup>th</sup> percentile
  4. Find the median of the upper half of the data.
    - 3<sup>rd</sup> quartile
    - 75<sup>th</sup> percentile
- Remember the range?
  - **New!** Interquartile range
  - Once you have 3<sup>rd</sup> quartile and 1<sup>st</sup> quartile you can calculate interquartile range (IQR)
  - 3<sup>rd</sup> quartile minus 1<sup>st</sup> quartile = IQR

# Quartile Example

1. Order the data from smallest to largest.
2. Find the median. That's Quartile 2/50<sup>th</sup> percentile.
3. Find the median of the lower half of the data for Quartile 1/25<sup>th</sup> percentile.
4. Find the median of the upper half of the data for Quartile 3/75<sup>th</sup> percentile.

$$\text{IQR} = 254 - 90 = 164$$

The following are a sample of numbers of beds from 11 Mass. hospitals (ordered).

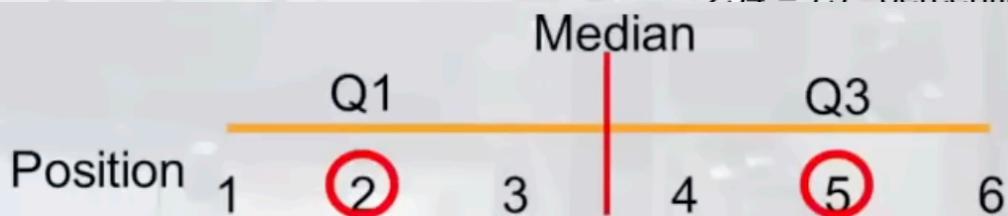
From ahd.com

41 74 90 97 121  
126 142 155 254 318  
364

126 = 50<sup>th</sup> percentile

90 = 25<sup>th</sup> percentile

254 = 75<sup>th</sup> percentile



Box-And-Whisker plot

# Box Plot Ingredients

Units: Hospital Beds

Minimum = 41

Q1 = 90

Median (Q2) = 126

Q3 = 254

Maximum = 364

Let's make a box plot!

The following are a sample of numbers of beds from 11 Mass. hospitals (ordered)

41 74 90 97 121  
126 142 155 254 318  
364

# Box Plot Ingredients

Units: Hospital Beds

Minimum = 41

Q1 = 90

Median (Q2) = 126

Q3 = 254

Maximum = 364



# Why Box-and-Whisker Plots?

- Help see the distribution in the data
- Another way to look at distribution
  - Histogram
  - Stem-and-leaf
- Also can see spread of data

