



CATS SEE THINGS WE CAN'T SEE

- ✓ Range ↙ max
↘ min
- ✓ Median or Q_2
- ✓ Lower Quartile or Q_1
- ✓ Upper Quartile or Q_3

Box and Whiskers Plots

Homework:

Page 70-71-2J: odds

Page 72-2K: odds

REVIEW-Median + Quartiles

How do you get values of median + Quartiles?

It depends on how the data given

- Discrete-List $(\frac{n+1}{2})$ for median, then split the data into two equal halves
 $Q_1 = \text{median of 1st half}$ $Q_2 = \text{median of 2nd half}$
- Discrete - frequency table
Use $\frac{n+1}{4}$ for Median, then $\frac{n+1}{4} \Rightarrow$ for Q_1 and $\frac{3(n+1)}{4}$ for Q_3
- Discrete-grouped
Be careful w/ upper endpoint } you must draw cum. freq curve
- Continuous-grouped

Drawing the Box & Whisker Plot:



- First STEP: Gather the 5 clues
 1. Minimum Value of the Data
 2. Lower Quartile or Q_1
 3. Median or Q_2
 4. Upper Quartile or Q_3
 5. Maximum Value of the Data

Be careful with outliers (thinking of your project...)

They are represented with an asterisk * and the whiskers

Do not reach all the way to the outliers.

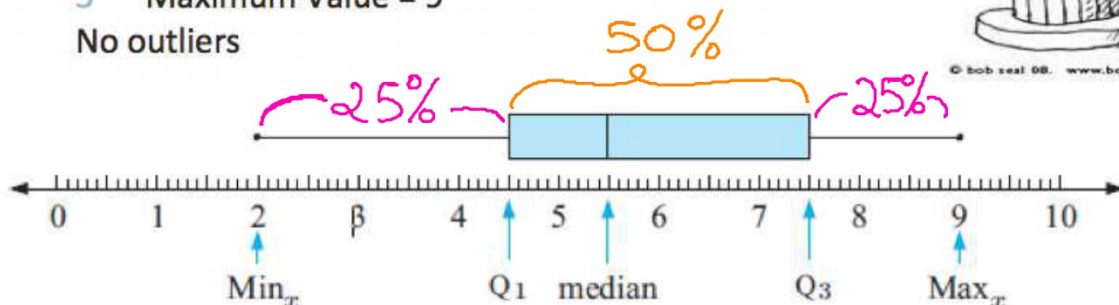


Drawing the Box & Whisker Plot:



- Example:
 1. Minimum Value = 2
 2. Lower Quartile = 4.5
 3. Median = 5.5
 4. Upper Quartile = 7.5
 5. Maximum Value = 9

No outliers



• What it represents:

1. Whiskers: 25% of the population
It does not matter how long the whisker is, it is still 25% of n
2. Rectangular Box: 50% of the population



Reading from the B & W plot...

100 boys

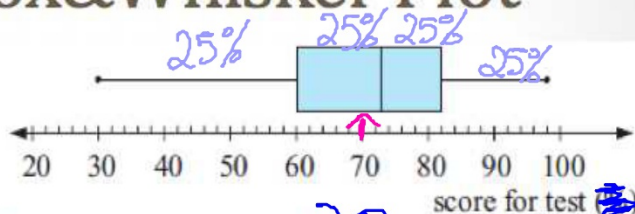


100 girls



Reading from a Box&Whisker Plot

A boxplot has been drawn to show the distribution of marks (out of 100) in a test for a particular class.



- a What was the *i* highest mark *ii* lowest mark scored? **98** **30**
- b What was the median test score for this class? **73**
- c What was the range of marks scored for this test? **$98 - 30 = 68$**
- d What percentage of students scored 60 or more for the test? **75%**
- e What was the interquartile range for this test? **$IQR = 82 - 60 = 22$**
- f The top 25% of students scored a mark between **82** and **98**
- g If you scored 70 for this test, would you be in the top 50% of students in this class?
- h Comment on the symmetry of the distribution of marks. **No, bottom 50%**

The bottom 50% has values more spread than the top 50%.

Parallel BoxPlots

Parallel boxplots enable us to make a *visual comparison* of the distribution of the data and the descriptive statistics (median, range and interquartile range). Parallel boxplots could be horizontal or vertical. For example:

