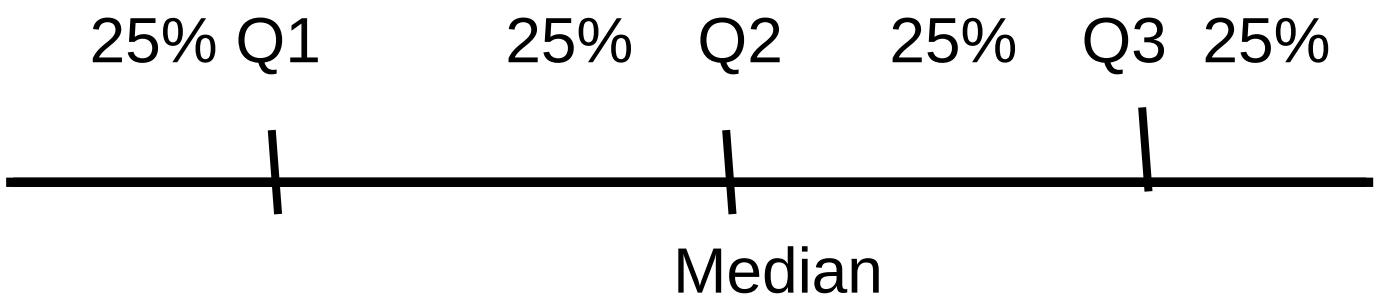
Box and Whisker Plots and the 5 number summary

Quantiles

- Types of quantiles
  - *quartiles*
  - *deciles*
  - *percentiles*

Quartiles

- Quartiles approximately divide an **ordered** data set into four equal parts.
- First quartile, Q1: About one quarter of the data fall on or below Q1.
- Second quartile, Q2: About one half of the data fall on or below Q2 (median).
- Third quartile, Q3: About three quarters of the data fall on or below Q3
- 

QUARTILES

GROUP DATA:

Quartile i=Qi= l + h/fi [i * Σ f /**4** – c.f<]

i= 1,2,3

- **UNGROUP DATA:**

Q1= (n +1)/4 th value

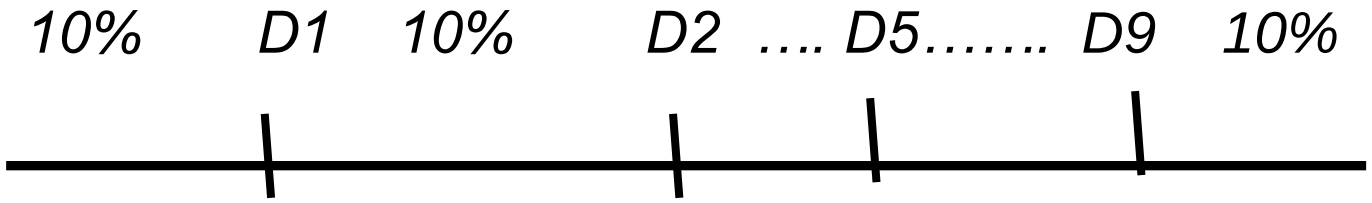
Q2=2*(n+1)/4 th value = (n+1)/2 th value

Q3= 3*(n+1)/4 th value

Qi = i * (n+1)/4 th value in the ordered dataset

Deciles

- Deciles approximately divide an **ordered** data set into TEN equal parts.
- First Decile, D1: About 10% of the data fall on or below D1.
- Second quartile, D2: About 20% half of the data fall on or below D2.
- …….
- NINTH quartile, D9: About 90% of the data fall on or below D9



DECILES

- **UNGROUP DATA:**

D1= (n +1)/10 th value

D2=2*(n+1)/10 th value

……

D9= 9*(n+1)/10 th value

Di = i * (n+1)/10 th value

i= 1,2,3,4,5,6,7,8,9

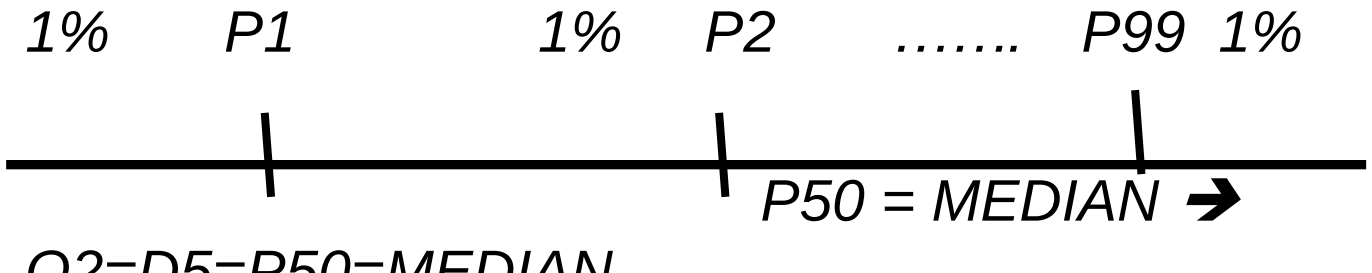
GROUP DATA:

Decile i=Di= l + h/fi [i*Σ f /10 – c.f<]

i= 1,2,3,4,5,6,7,8,9

Percentiles

- Percentiles approximately divide an **ordered** data set into HUNDRED equal parts.
- First Percentile, P1: About one Percentile of the data fall on or below P1.
- Second Percentile, P2: About one half of the data fall on or below P2 (median).
- …….
- Ninety Nine quartile, P99: About NINty nine Percentile of the data fall on or below P99



PERCENTILES

- **UNGROUP DATA:**

P1= (n +1)/100 th value

P2=2*(n+1)/100 th value

……

PP99= 99*(n+1)/100 th value

Pi = i * (n+1)/100 th value

i= 1,2,3,4,5,6,7,8,9,……,99

GROUP DATA:

Percentile i=Pi= l + h/fi [i * Σ f /100 – c.

i= 1,2,3,4,5,6,7,8,9,……,99

Some statistical concepts

- Five No Summary
  - *Min*
  - *Q1*
  - *Q2*
  - *Q3*
  - *Max*

. Inter quartile Range (I.Q.R)

.Pottential outlier (P.O)

Interquartile Range

- Interquartile Range (IQR)•The difference between the third and first quartiles
- I.Q.R = Q3 – Q1

Potential Outliers

- Obs that fall well outside the pattern of the due
  - Measurement error/ recording error
  - An unusal extreme Obs

CALCULATION:

We first define Lower limit and Upper Limit in the Data:

Lower Limit=L.L=Q1- 1.5 (I.Q.R)

Upper Limit= U.L=Q3+ 1.5(I.Q.R)

So the obs in the data below L.L and above Upper can be listed in the category of P.O:

Box and Whisker Plots

A **box plot** summarizes data using the median, upper(First) and lower Third (THIRD QUARTILE), and the extreme least(MIN) and greatest(values. It allows you to see important characteristics of the data at a glance.

The 5 Number Summary

- The five number summary is another name for the visual representation of the **box and whisker plot.**

- The five number summary consist of :
  - *The median (2nd quartile)*
  - *The 1st quartile*
  - *The 3rd quartile*
  - *The maximum value in a data set*
  - *The minimum value in a data set*

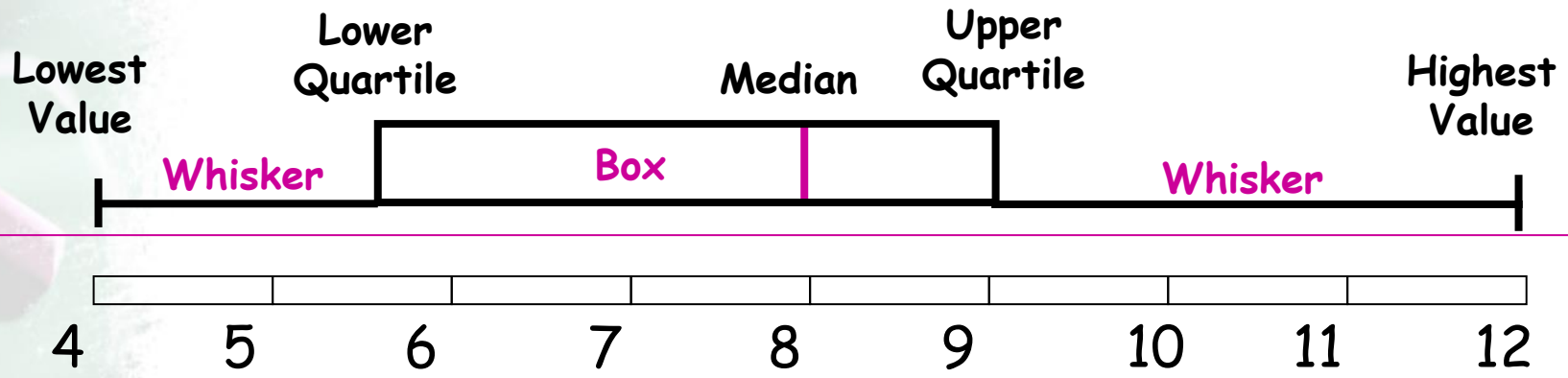
Median(revise formula)

UNGROUP DATA

- Arrange data in some order
- **ODD when Total no of obs is oDD:**
- Median= (n+1)/2 th value

Box and Whisker Diagrams.

Anatomy of a Box and Whisker Diagram.



Constructing a box and whisker plot

Step 1 - take the set of numbers given…

34, 18, 100, 27, 54, 52, 93, 59, 61, 87, 68, 85, 78, 82, 91

Place the numbers in order from least to greatest:

18, 27, 34, 52, 54, 59, 61, 68, 78, 82, 85, 87, 91, 93, 100

Constructing a box and whisker plot

- **Step 2** - Find the median.
- Remember, the median is the middle value in a data set.

18, 27, 34, 52, 54, 59, 61, 68, 78, 82, 85, 87, 91, 93, 100

68 is the median of this data set.

Constructing a box and whisker plot

- **Step 3** – Find the lower quartile.
- The lower quartile is the median of the data set to the left of 68.

(18, 27, 34, 52, 54, 59, 61,) 68, 78, 82, 85, 87, 91, 93, 100

52 is the lower quartile

Constructing a box and whisker plot

- **Step 4** – Find the upper quartile.
- The upper quartile is the median of the data set to the right of 68.

18, 27, 34, 52, 54, 59, 61, 68, (78, 82, 85, 87, 91, 93, 100)

87 is the upper quartile

Constructing a box and whisker plot

- Step 5 – Find the maximum and minimum values in the set.
- The maximum is the greatest value in the data set.
- The minimum is the least value in the data set.

18, 27, 34, 52, 54, 59, 61, 68, 78, 82, 85, 87, 91, 93, 100

18 is the minimum and 100 is the maximum.

Constructing a box and whisker plot

- **Step 5** – Find the inter-quartile range (IQR).
- The inter-quartile (IQR) range is the difference between the upper and lower quartiles.
  - Upper Quartile = 87
  - Lower Quartile = 52
  - 87 – 52 = 35
  - 35 = IQR

POTENTIAL OUTLIER:

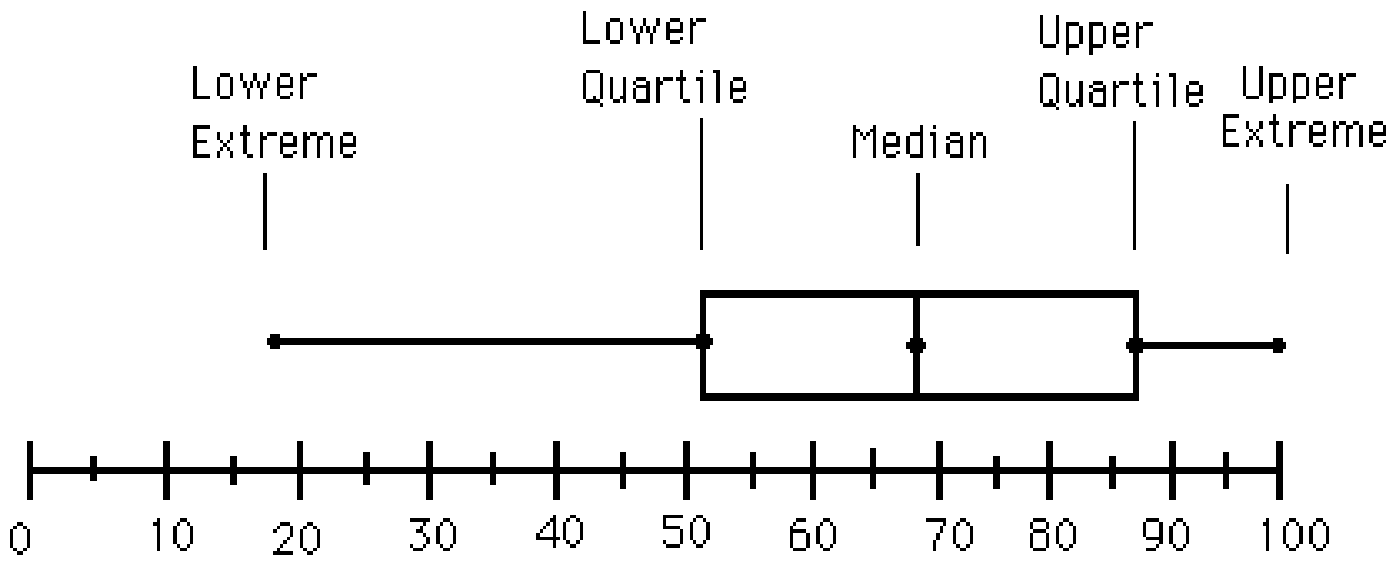
L.L= 52 – (1.5)(35)=0.5

U.L= 87 + (1.5)(35) =139.5

P.O:

Graphing The Data

- Notice, the Box includes the lower quartile, median, and upper quartile.
- The Whiskers extend from the Box to the max and min.



Interpreting the Box Plot:

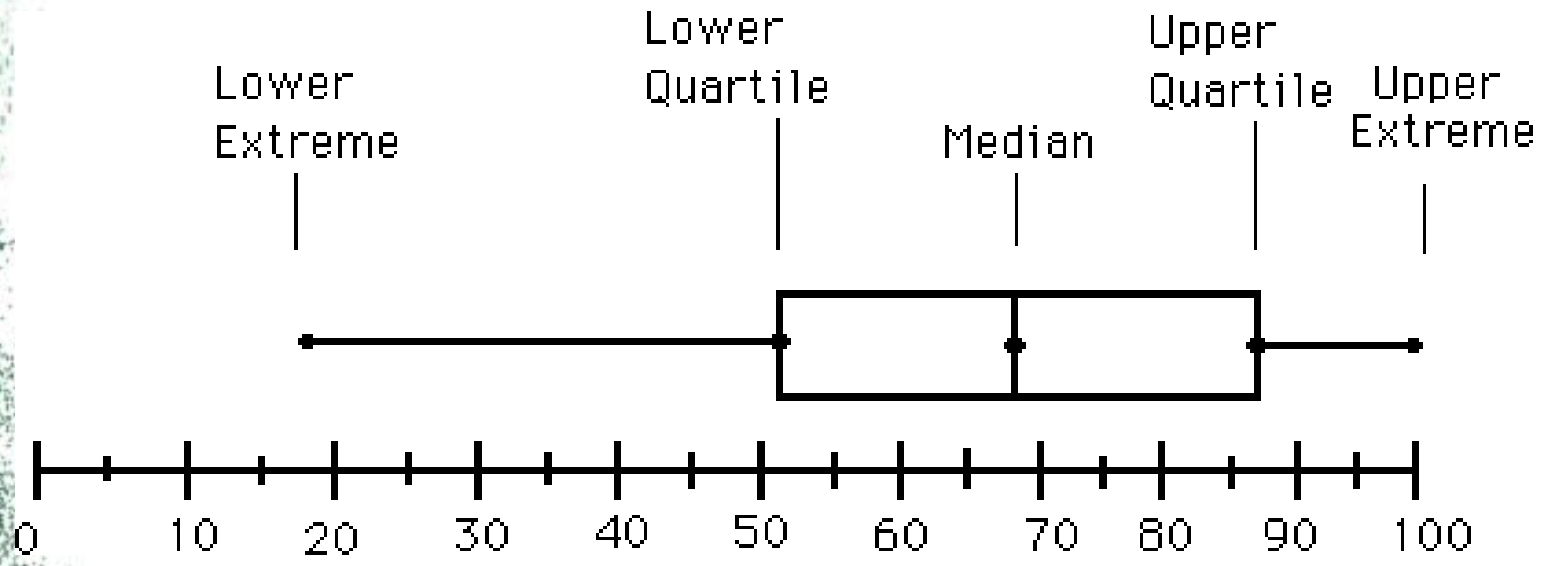
Study your Box and Whisker Plot to determine what it is telling you. Make a statement about what it is saying, then support the statement with facts from your graph.

You should include the following in your interpretation:

- **Range** or spread of the data and what it means to your graph
- **Quartiles**—compare them. What are they telling you about the data?
- **Median**- this is an important part of the graph, and should be an important part of the interpretation.
- **Percentages** should be used to interpret the data, where relevant.

Analyzing The Graph

- The data values found inside the box represent the middle half (50%) of the data.
- The line segment inside the box represents the median



Practice

- Use the following set of data to create the 5 number summary.

3, 7, 11, 11, 15, 21, 23, 39, 41, 45, 50, 61, 87, 99, 220

Median

- What is the median or 2nd quartile?

  3, 7, 11, 11, 15, 21, 23, 39, 41, 45, 50, 61, 87, 99, 220

- The median is 39

Lower Quartile (1st Quartile)

- What is the lower or 1st quartile?

(3, 7, 11, 11, 15, 21, 23), 39, 41, 45, 50, 61, 87, 99, 220

- The lower quartile is 11

Upper Quartile (3rd Quartile)

- What is the upper or 3rd quartile?

  3, 7, 11, 11, 15, 21, 23, 39, (41, 45, 50, 61, 87, 99, 220)

- The upper quartile is 61

Maximum

- What is the maximum?

  3, 7, 11, 11, 15, 21, 23, 39, 41, 45, 50, 61, 87, 99, 220

- The max is 220

Minimum

- What is the minimum?

  3, 7, 11, 11, 15, 21, 23, 39, 41, 45, 50, 61, 87, 99, 220

- The min is 3

The 5 Number Summary

- Median - 39
- Lower Quartile - 11
- Upper Quartile - 61
- Max - 220
- Min - 3

Graphing The Data

Take out your graph paper so we can practice graphing the data.

Box and Whisker Plots and the 5 number summary