

3.4 Five Number Summary; BoxPlots

GOALS:

1. Find the 5 Number Summary
2. Understand Quartiles
3. Create Box Plots, using Quartiles

Study Ch. 3.4, # 149-153, 159 - 165,
167, 169, 177

Class Notes: Prof. G. Battaly, Westchester Community College, NY

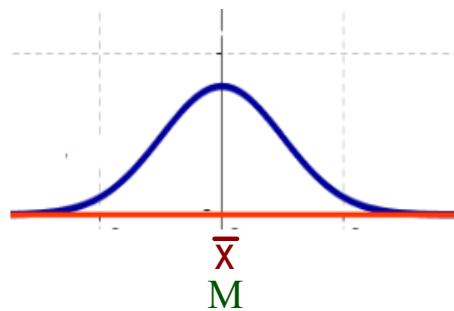
[Statistics Home Page](#)

[Homework](#)

3.4 Five Number Summary; BoxPlots

Median: midpoint:
50th percentile
50% of data is lower
50% of data is higher

Resistant to extremes



Are there other Resistant measures?

Yes. Percentiles (100ths), and QUARTILES

Q 1 25% lower, or 25th percentile

Q 2 50% median

Q 3 75%, lower, or 75th percentile

can be actual value, or mean of two values
(like the median)

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

[Homework](#)

3.4 Five Number Summary; BoxPlots

Percentiles & QUARTILES

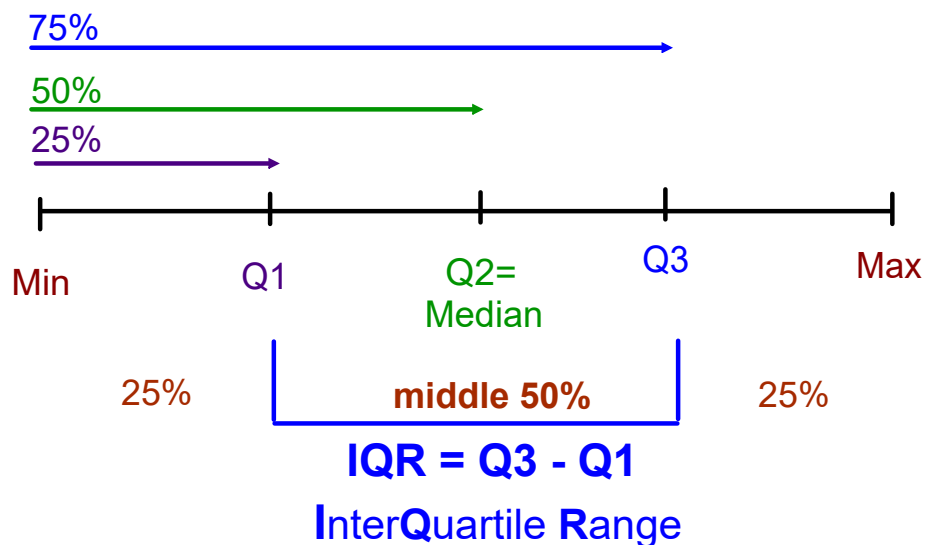
1. Arrange data in numerical order.
2. Quartiles are based on **position** in ordered data
 - Q 1 25%,
 - Q 2 50% median
 - Q 3 75%
3. Provide information about the distribution of data.

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)[Homework](#)

3.4 Five Number Summary; BoxPlots

For sequenced (ordered) data



Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)[Homework](#)

3.4 Five Number Summary; BoxPlots

Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: a) Q 1, Q 2, Q3
 b) IQR: InterQuartile Range = Q3 - Q1
 c) 5 # Sum: Min, Q1, Median, Q3, Max
 d) LL: Lower Limit
 UL: Upper Limit
 e) Box Plot

F: a) Q 1, Q 2, Q3
 Need **position**, so **sort first**
 and look for 25%, 50%, and 75%

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: a) Q 1, Q 2, Q3
 b) IQR: InterQuartile Range = Q3 - Q1
 c) 5 # Sum: Min, Q1, Median, Q3, Max
 d) LL: Lower Limit
 UL: Upper Limit
 e) Box Plot

F: a) Q 1, Q 2, Q3

1. Calculator
2. STAT/EDIT/ enter data
3. STAT/EDIT/SortA(L1),copy to paper
4. Find Median first
5. Then find half of 1st 50%
and half of the 2nd 50%

4
4.1
4.4
4.4
4.5
4.9
4.9
5.0
5.1
5.1
5.7
5.8
5.9
6.1
6.5

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- F: a) Q 1, Q 2, Q 3
 b) IQR: InterQuartile Range = $Q3 - Q1$
 c) 5 # Sum: Min, Q1, Median, Q3, Max
 d) LL: Lower Limit
 UL: Upper Limit
 e) Box Plot

F: a) Q 1, Q 2, Q 3

1. Calculator
2. STAT/EDIT/
3. STAT/EDIT/SortA(L1), copy to paper
4. Find Median first
5. Then find half of 1st 50%
and half of the 2nd 50%

Q1=4.4, Q2=5.0, Q3=5.8

5.0

4
4.1
4.4
4.4
4.5
4.9
4.9
5.1
5.1
5.7
5.8
5.9
6.1
6.5

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)[Homework](#)

3.4 Five Number Summary; BoxPlots

#94. Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: b) IQR, c) 5 # Sum

InterQuartile Range = $Q3 - Q1$
 range of **middle 50%**

Five Number Summary:
 Min, Q 1, Q 2, Q 3, Max

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)[Homework](#)

3.4 Five Number Summary; BoxPlots

Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: a) Q 1, Q 2, Q3
 ✓ b) IQR: InterQuartile Range = Q3 - Q1
 ✓ c) 5 # Sum: Min, Q1, Median, Q3, Max
 d) LL: Lower Limit
 UL: Upper Limit
 e) Box Plot

a) $n=14$
 $Q_2 = \text{med} = \text{bet. } 6^{\text{th}} \text{ \& } 7^{\text{th}}$
 $= \frac{4.9+5.1}{2} = 5.0$
 $Q_1 = 4.4, Q_3 = 5.8$

F: b) IQR
 $IQR = Q3 - Q1 = 5.8 - 4.4 = 1.4$

F: c) 5 Number Sum:
 4 4.4 5.0 5.8 6.5

F: e) Box Plot

	5 # SUM
4	4.0
4.1	
4.4	
4.4	4.4
4.5	
4.9	
4.9	5.0
5.1	5.0
5.1	
5.7	
5.8	5.8
5.9	
6.1	
6.5	6.5

Class Notes: Prof. G. Battaly, Westchester Community College, NY
[Statistics Home Page](#) [Homework](#)

3.4 Five Number Summary; BoxPlots

Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

5 # Sum: 4, 4.4, 5.0, 5.8, 6.5

A **boxplot** displays the distribution as **quartiles**.

This one is **not** symmetrical.

Axis is **uniform**.

5 Number Sum:
 4 4.4 5.0 5.8 6.5

F: a) Q 1=4.4, Q 2=5.0, Q3=5.8

Class Notes: Prof. G. Battaly, Westchester Community College, NY
[Statistics Home Page](#) [Homework](#)

3.4 Five Number Summary; BoxPlots

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: a) Q 1, Q 2, Q3
 b) IQR: InterQuartile Range = Q3 - Q1
 c) 5 # Sum: Min, Q1, Median, Q3, Max
 d) LL: Lower Limit
 UL: Upper Limit
 e) Box Plot

So, what are LL and UL?

LL and UL: 1. Boundaries beyond which data is considered an **outlier**
 2. Computed; usually Not a data item

LL = Q 1 - 1.5 (IQR)
 UL = Q 3 + 1.5 (IQR)

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

5 Number Sum:

#94. Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: d) LL and UL

IQR $\frac{1.4}{0.7} = 2.1$

LL = Q1 - 1.5 (IQR) = ___ - 1.5 (___) = ___ - ___ = 2.3

UL = Q3 + 1.5 (IQR) = ___ + 1.5 (___) = ___ + ___ = 7.9

Any data item outside the limits is an **outlier**.

Any outliers?

min= 4.0 ? 2.3 = LL any outlier to left ?

max=6.5 ? 7.9 = UL any outlier to right ?

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

5 Number Sum:

#94. Given the following:

4.0 4.4 5.0 5.8 6.5

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: d) LL and UL

$$LL = Q1 - 1.5(IQR) = 4.4 - 1.5(1.4) = 4.4 - 2.1 = 2.3$$

$$UL = Q3 + 1.5(IQR) = 5.8 + 1.5(1.4) = 5.8 + 2.1 = 7.9$$

Any data item outside the limits is an outlier.

Any outliers? **NO**

$$\begin{array}{r} 1.4 \text{ IQR} \\ 0.7 \\ \hline 2.1 \end{array}$$

$$\text{min} = 4 > 2.3 = LL \text{ no outlier to left}$$

$$\text{max} = 6.5 < 7.9 = UL \text{ no outlier to right}$$

Or: all our data is between the LL and UL NO OUTLIERS

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

#94. Given the following:

5.9	4	5.7	5.1	4.1	4.4	6.5	4.4	5.8	5.1	6.1	4.5	4.9	4.9
-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

F: e) Box Plot or **Box-and-Whisker Diagram**

BOX: bounded by Q1 and Q3, with Q2(median) noted

WHISKERS: lines from BOX to **adjacent values**,
actual data at or within the LL and UL

(ie: if LL and UL are actual values of the data, then LL and UL are used; if not, then use the data closest to LL and UL towards the BOX)

$$LL \leq \text{adjacent values} \leq UL$$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)
[Homework](#)

3.4 Five Number Summary; BoxPlots

Medieval Cremation Burials. The frequency of cremation burials found in 17 archaeological sites in eastern England are:

83 64 46 48 21 265 385 523
86 46 51 34 35 429 258 119 2484



Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

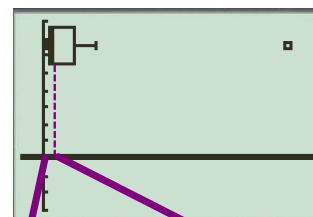
[Homework](#)

3.4 Five Number Summary; BoxPlots

Medieval Cremation Burials. The frequency of cremation burials found in 17 archaeological sites in eastern England are:

83 64 46 48 21 265 385 523
86 46 51 34 35 429 258 119 2484

21 MIN
34
35
46 < Q1=46
46
48
51 IQR=325-46=279
64 1.5*IQR=418.5
83 Med UL=325+418.5=743.5
86
119
258
265 < Q3=325
385
429
523
2484 > UL=743.5; 2484 is outlier



Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

[Homework](#)

3.121 The Great Gretzky - 5 NUMBER SUMMARY

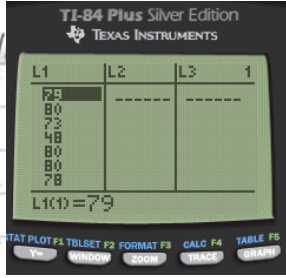
1. Enter data into a list on the calculator.

Different data

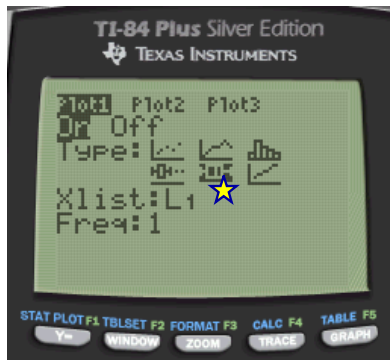
STAT/EDIT

6. Gretzky: # games pl

79	80	80	80	74
80	80	79	64	78
78	78	74	45	81
48	80	82	82	70



2. Turn ON STAT PLOT, Plot1 and enter settings for a 5 Number Summary BOX PLOT. Select the 5th graph (line 2, middle) under Type:



Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

[Homework](#)

3. Using TRACE button, arrow through the 5 Number Summary from left to right to get all 5 numbers.

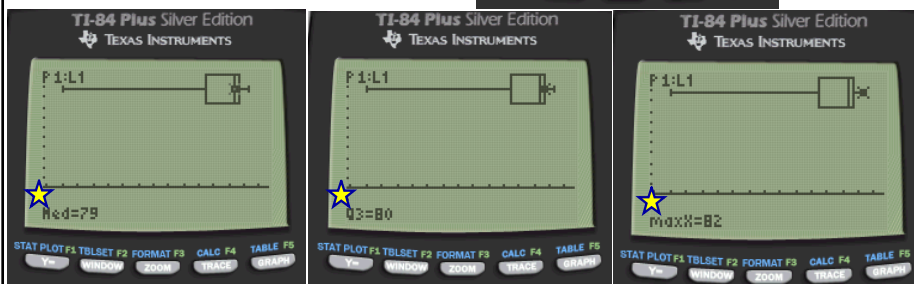
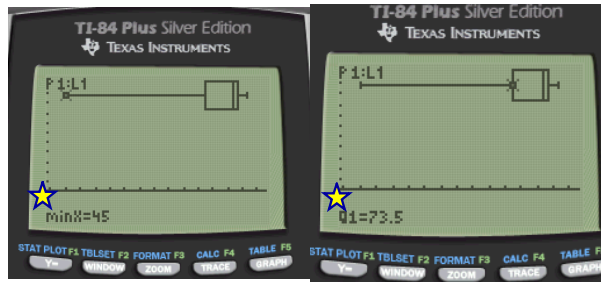
6. Gretzky: # games played in 20 seasons

79	80	80	80	74	F:
80	80	79	64	78	c)
78	78	74	45	81	d)
48	80	82	82	70	e) boxplot

2 Boxplots, Calculator

5 Number Sum

Outliers



5 # SUM: 45 73.5 79 80 82

Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

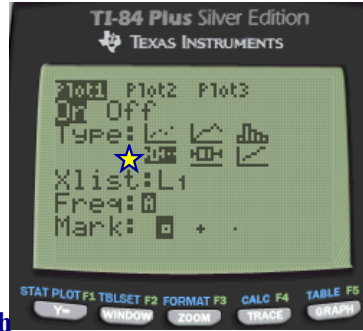
[Homework](#)

3.121 The Great Gretzky - BOX PLOT

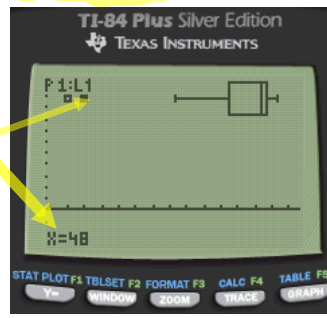
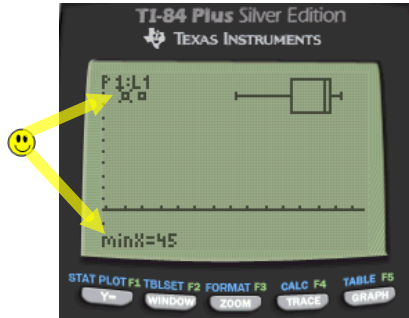
- In STAT PLOT, select the 4th graph (line 2, left) under Type: This is a **BOX PLOT w. Outliers**

6: Gretzky :# games played in 20 seasons

79	80	80	80	74	F: a) q_1, q_2, q_3 b) IQR
80	80	79	64	78	c) 5#Sum
78	78	74	45	81	d) outliers
48	80	82	82	70	e) boxplot



- Using TRACE button, arrow through the **Box Plot** from left to right to get: adjacent $x_i > LL$, Q1, M, Q3, adjacent $x_i < UL$, along with any potential outliers. 😊



5 # SUM: 45 73.5 79 80 82 $LL=73.5-6.5*1.5=63.75$
45, 48 both < 63.75 ; outliers

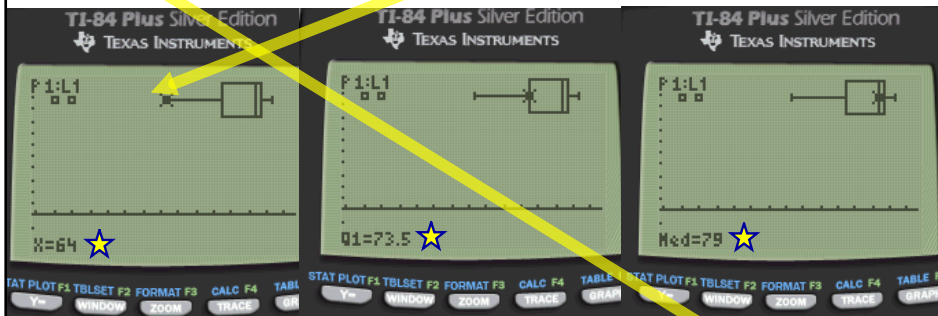


Class Notes: Prof. G. Battaly, Westchester Community College, NY

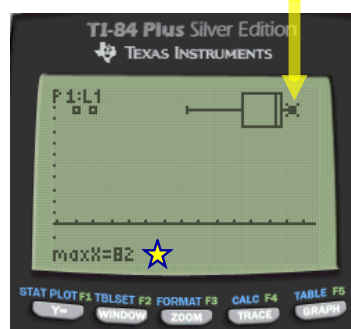
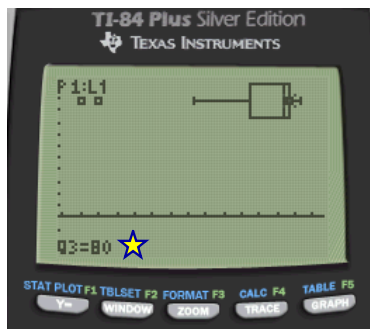
[Statistics Home Page](#)

[Homework](#)

- Using TRACE button, arrow through the **Box Plot** from left to right to get: adjacent $x_i > LL$, Q1, M, Q3, adjacent $x_i < UL$, along with any potential outliers.



5 # SUM: 45 73.5 79 80 82



Class Notes: Prof. G. Battaly, Westchester Community College, NY

[Statistics Home Page](#)

[Homework](#)

