Statistics

Lecture 4

Feb 19-8:47 AM

Class QZ 1

Consider the Sample below

2 4 6 **6** 8 9 10

1) n = 7 ✓

2) Range = 10 − 2 = 8 ✓

3) Midrange = $\frac{10+2}{2}$ = 6 ✓

4) Mode = 6 ✓

5) Median = 6 ✓

Jan 8-7:09 PM

Class Qz 2

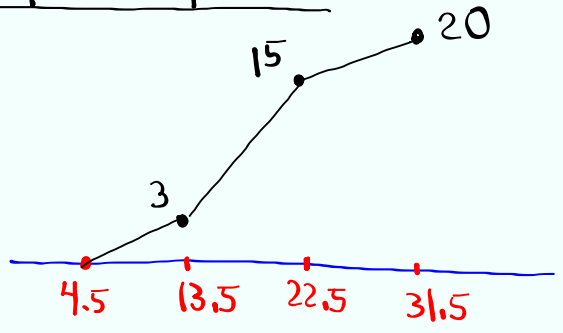
Complete the chart below:

class limits	class BNDRS	class F	Cum. F
5 – 13	4.5 – 13.5	3	3
14 – 22	13.5 – 22.5	12	15
23 – 31	22.5 – 31.5	5	20 ← n = 20

×9 ×9

then draw ogive.

CW = 9



Jan 13-3:58 PM

Consider the Sample below

28	32	30	18	19
25	28	24	40	45
33	24	20	18	27

1) n = 15

2) Make Stemplot

Data must be Sorted

Clear all lists [2nd] [+] [4: Clear All lists] [Enter]

Reset all lists [STAT] Edit [5: SetupEditor] [Enter]

Store in L1

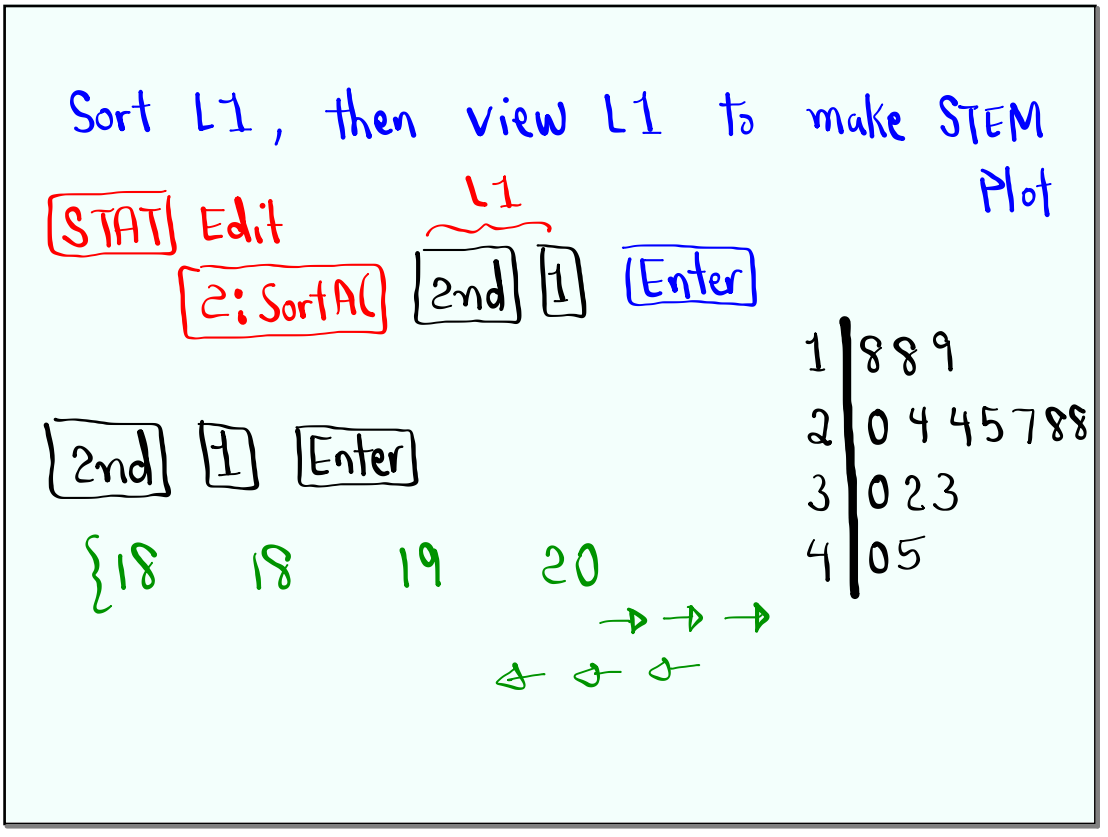
[STAT] Edit [1: Edit]

L1
28
32
⋮
27

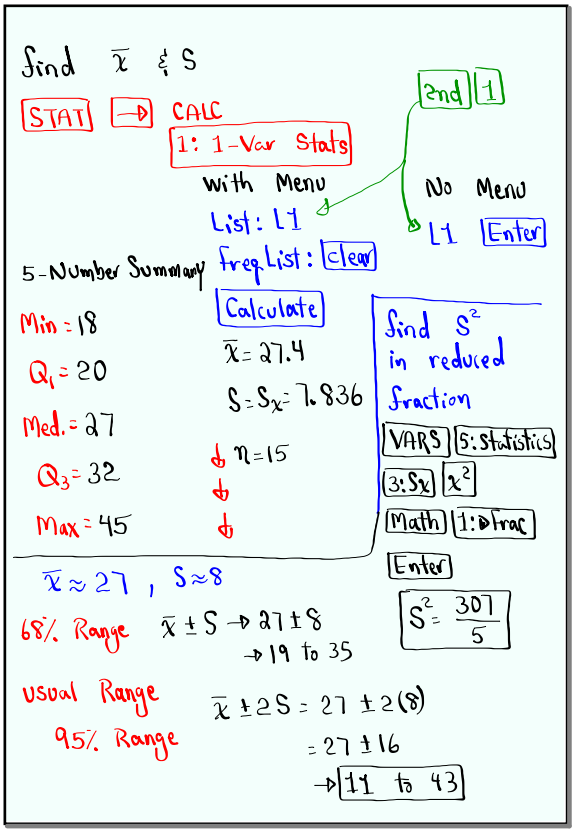
quit & clear Screen

[2nd] [MODE] [clear]

Jan 13-4:44 PM



Jan 13-4:51 PM

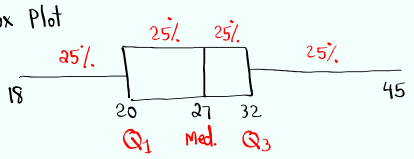


Jan 13-4:55 PM

5 - Number Summary

Min	Q_1	Med	Q_3	Max
18	20	27	32	45

Box Plot



IQR (Inter - Quartile - Range) $= Q_3 - Q_1 = 12$

Upper Fence $= Q_3 + 1.5(IQR)$
$= 32 + 1.5(12) = \boxed{50}$

Lower Fence $= Q_1 - 1.5(IQR)$
$= 20 - 1.5(12) = \boxed{2}$

Let's discuss outliers

below Lower Fence OR above upper Fence

NO outliers

Range = Max - Min = 45 - 18 = 27

Estimate $S \approx \frac{\text{Range}}{4} = \frac{27}{4} = \boxed{6.75}$

Jan 13-5:08 PM

Ages of 120 students have the following 5 - number Summary

15 20 28 36 60

Min. = 15

$Q_1 = 20$

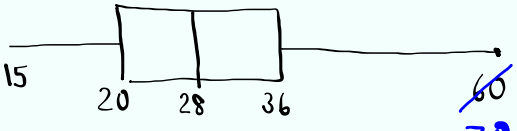
Med. = 28

$Q_3 = 36$

Max = 60

NO outliers

Box Plot



$IQR = Q_3 - Q_1 = 16$

Upper Fence
$= Q_3 + 1.5(IQR)$
$= 36 + 1.5(16) = \boxed{60}$

60-70 outliers

Lower Fence $= Q_1 - 1.5(IQR) = 20 - 1.5(16)$
$= \boxed{-4}$

Jan 13-5:18 PM

Complete the chart below

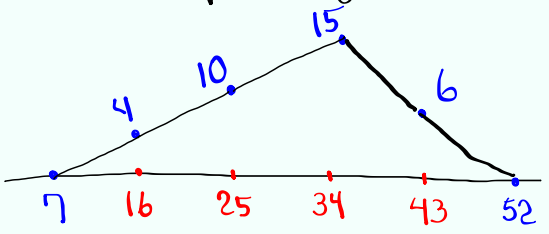
Class limits	Class MP	Class F
12 - 20	16	4
21 - 29	25	10
30 - 38	34	15
39 - 47	43	6

1) 4 classes

2) CW = 9

3) $n = 4+10+15+6 = 35$

Draw Freq. Polygon



Jan 13-5:24 PM

Find $\bar{x}$ & s of the freq. table.

Clear all lists

Class MP → L1

Class F → L2

STAT → CALC

1:1-Var Stats

$\bar{x} = 30.914$

$S = S_x = 8.151$

$n = 35$

With Menu
List: L1
FreqList: L2
Calculate

No Menu
L1 , L2
1
Enter

Find S^2 in reduced fraction

VARS → 5: Statistics → 3: S_x → χ^2 → Math → 1: ▶Frac → Enter

$S^2 = \frac{39528}{595}$

Jan 13-5:30 PM

I randomly selected 40 exams, here are the Scores

5	5 8 9
6	0 3 7 8 8 9
7	0 2 4 5 5 5 8 9 9 9
8	2 3 6 6 6 8 8 8 9 9
9	0 2 3 5 5 7 7
10	0 0 0 0

1) $n = 40$

2) Range $= 100 - 55 = 45$

3) Estimate S

$S \approx \frac{\text{Range}}{4} = \frac{45}{4} = 11.25$

4) Find P_{10}

$L = \frac{10}{100} \cdot 40 = 4$

$P_{10} = \frac{4\text{th value} + \text{Next value}}{2}$

$= \frac{60 + 63}{2} = \boxed{61.5}$

5) Find P_{72}

$L = \frac{72}{100} \cdot 40 = 28.8 \rightarrow \boxed{L = 29}$

$P_{72} = 29\text{th element}$

$= \boxed{89}$

Jan 13-5:37 PM

5	5 8 9
6	0 3 7 8 8 9
7	0 2 4 5 5 5 8 9 9 9
8	2 3 6 6 6 8 8 8 9 9
9	0 2 3 5 5 7 7
10	0 0 0 0

Find K such that

$P_K = 75$

$K = \frac{B}{n} \cdot 100$ (B: Below)

Round to whole %

$K = \frac{12}{40} \cdot 100 = \boxed{30}$

30% | 70%

$P_{30} = 75$

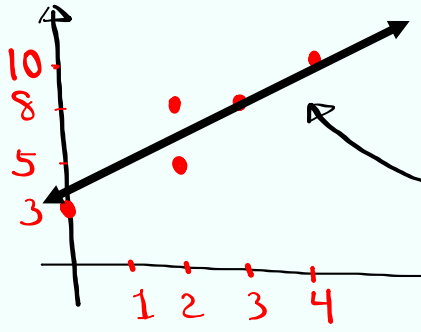
$\boxed{P_{30} = 75}$

Jan 13-5:46 PM

Working with ordered-Pairs (x, y)

SG9

Plot (0,3), (2,5), (2,8), (3,8), (4,10)



Scatter Plot

Regression line

$y = a + bx$

Jan 13-6:07 PM

How to use TI

x	y
0	3
2	5
2	8
3	8
4	10

Clear all lists

$x \to L1$, $y \to L2$

quit & clear Screen

STAT → CALC

2: 2-Var Stats

$\sum x = 11$

$\sum x^2 = 33$

$n = 5$

$\sum y = 34$

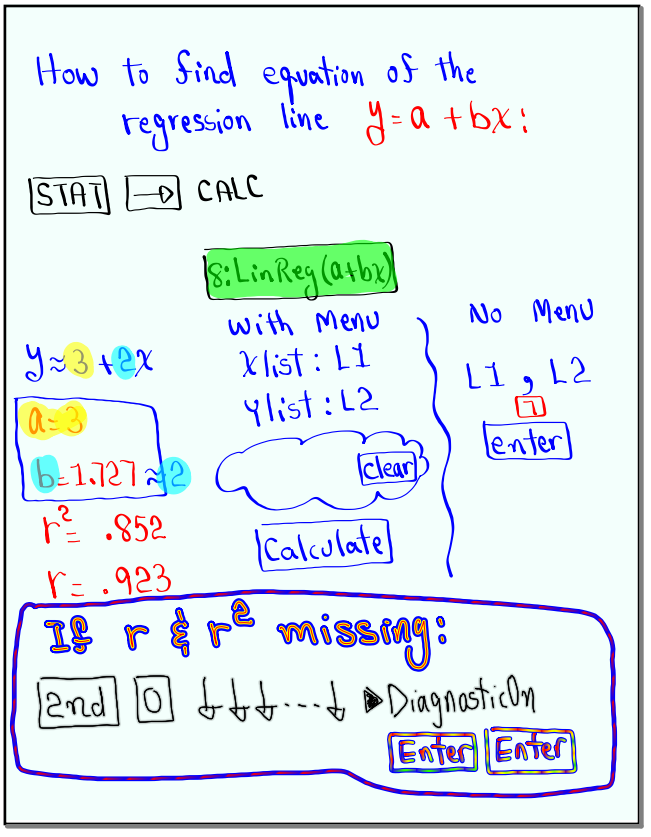
$\sum y^2 = 262$

$\sum xy = 90$

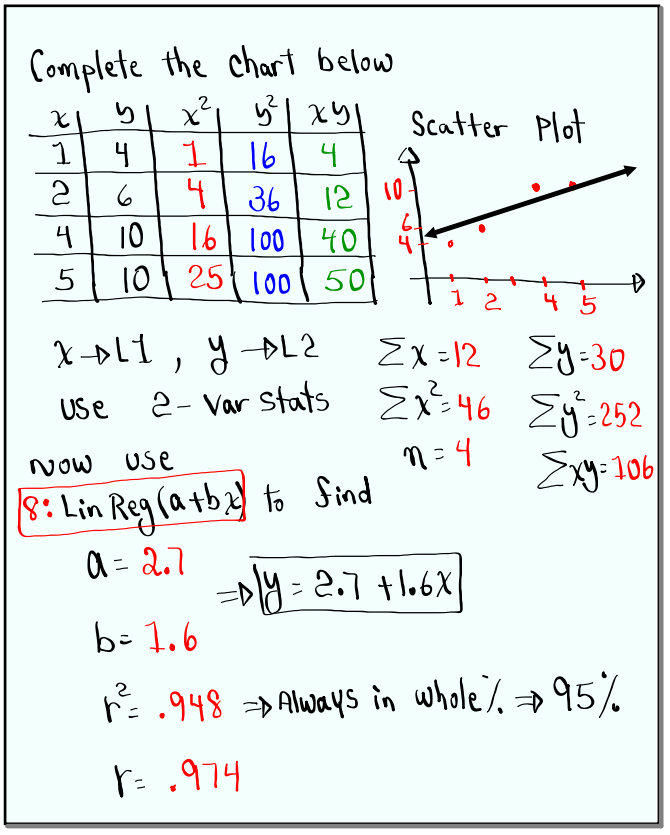
With Menu
- x list: L1
- y list: L2
- Freq List: clear
- Calculate

No Menu
- L1, L2
- ,
- Enter

Jan 13-6:11 PM



Jan 13-6:17 PM

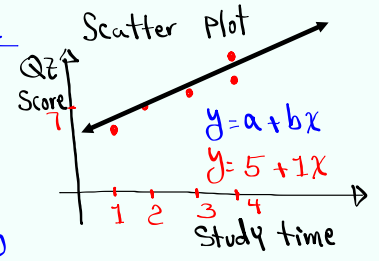


x	y	x^2	y^2	xy
1	4	1	16	4
2	6	4	36	12
4	10	16	100	40
5	10	25	100	50

Jan 13-6:25 PM

I randomly selected 5 students, chart below show study time and Quiz Scores.

Time	Score
2	7
3	8
1	6
4	8
4	10



Study time → x → L1

Quiz Score → y → L2

use 2-Var Stats

$\sum x = 14$ $\sum y = 39$

$\sum x^2 = 46$ $\sum y^2 = 313$

$n = 5$ $\sum xy = 116$

option 8:

use LinReg(a+bx)

$a = 5$

$b = 1$

$r^2 = (.)\approx 77\%$,

$r = .879$

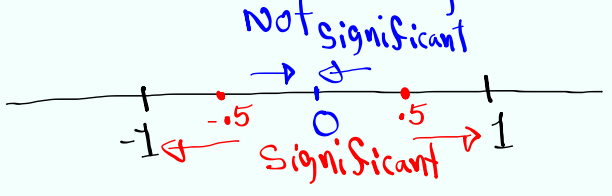
Jan 13-6:37 PM

$r \rightarrow$ Linear Correlation Coef.

$$-1 \le r \le 1$$

when r is close to 1 or -1, the linear Correlation is Significant.

when r is close to 0, The linear Correlation is not Significant.



Jan 13-6:47 PM

what about r^2?

r^2 Coef. of determination

Always in whole %

r^2 tells us what % of y-values are explained by x-values.

From Last example $r^2 = 77\%$.

77% of Quiz Scores are explained by Study time.

23% are unexplained.

Jan 13-6:52 PM

How do we use these?

Two Braches in Statistics:

1) Descriptive

2) Inferential (Make predictions)

How to make predictions:

If r is Significant,
use the regression line

If r is not Significant,
use $\bar{y}$, $\bar{y} = \frac{\Sigma y}{n}$

Jan 13-6:57 PM

Using last example

Ryan Studies 4 hrs, Predict his Score.

1) Assume r is Significant

$y = 5 + 1x$

$y = 5 + 1(4) = \boxed{9}$

2) Assume r is not Significant

use $\bar{y}$

$\bar{y} = 7.8 \approx \boxed{8}$

VARS

5: Statistics

5: $\bar{y}$

Enter

Jan 13-7:00 PM

Class QZ 3

Consider the Sample below

20	23	18	25
30	25	19	24
32	28		

Find

1) $\bar{x} = 24.4 \approx \boxed{24}$

2) $S = 4.648 \approx \boxed{5}$

} Round to whole #

3) $S^2 = \boxed{\frac{108}{5}}$

} Reduced Fraction

Jan 13-7:04 PM