Statistics

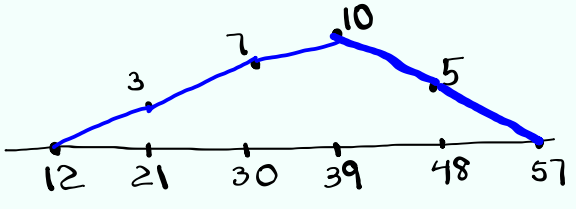
Lecture 7

Feb 19-8:47 AM

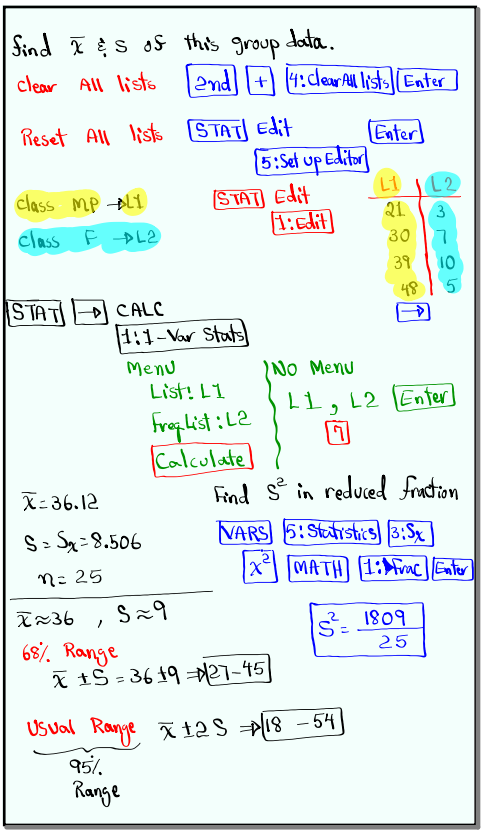
Complete the freq. table below

Class limits	Class BNDRS	Class MP	Class F	Cum. F	Rel. F	% F
17 – 25	16.5 – 25.5	21	3	3	.12	12%
26 – 34	25.5 – 34.5	30	7	10	.28	28%
35 – 43	34.5 – 43.5	39	10	20	.40	40%
44 – 52	43.5 – 52.5	48	5	25	.20	20%

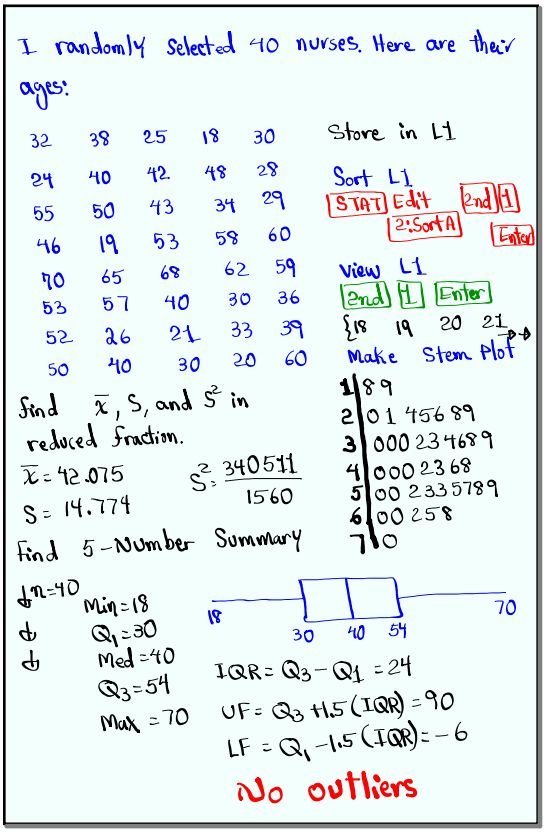
4 classes , CW = 9 , $n = 25$, Rel. F $= \frac{f}{n} = \frac{f}{25}$



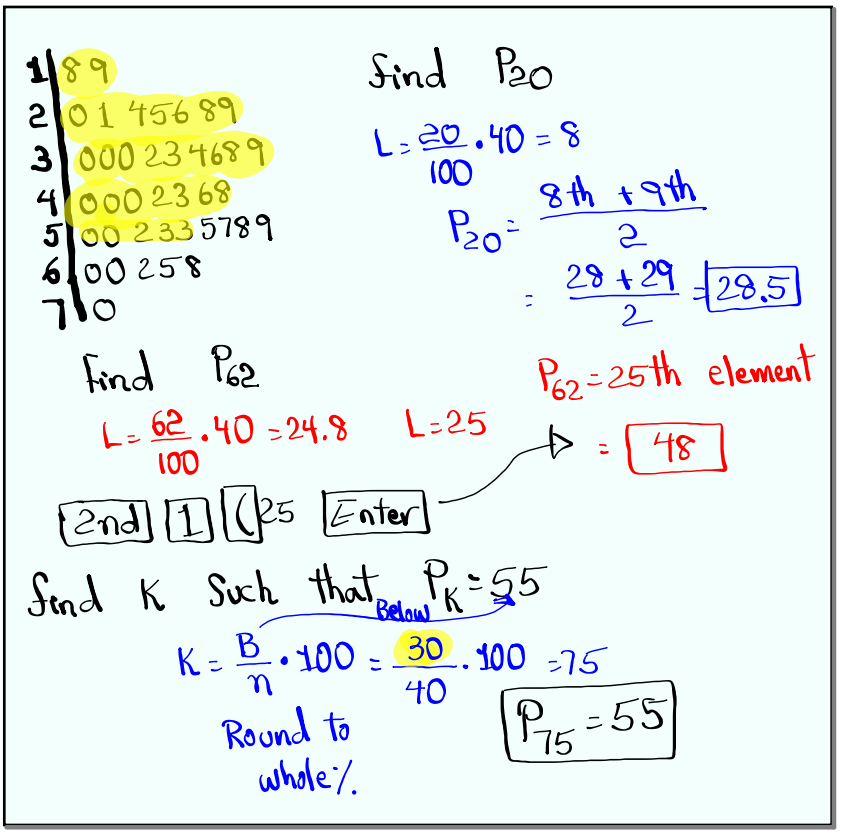
Mar 5-1:50 PM



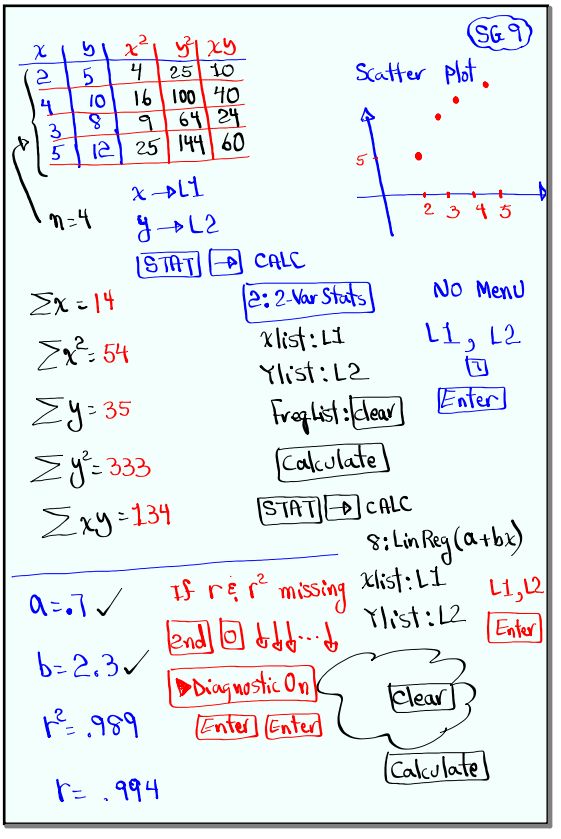
Mar 5-2:00 PM



Mar 5-2:13 PM



Mar 5-2:32 PM



Mar 5-2:39 PM

$\Sigma x = 14$

$\Sigma x^2 = 54$

$\Sigma y = 35$

$\Sigma y^2 = 333$

$\Sigma xy = 134$

Regression Line $y = a + bx$

$$a = \frac{\Sigma y \Sigma x^2 - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{35 \cdot 54 - 14 \cdot 134}{4 \cdot 54 - 14^2} = \frac{14}{20} = .7$$

$$b = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{4 \cdot 134 - 14 \cdot 35}{4 \cdot 54 - 14^2} = \frac{46}{20} = 2.3$$

Mar 5-3:02 PM

Linear Correlation Coefficient r

$-1 \le r \le 1$

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$\Sigma x = 14$

$\Sigma x^2 = 54$

$\Sigma y = 35$

$\Sigma y^2 = 333$

$\Sigma xy = 134$

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{4 \cdot 134 - 14 \cdot 35}{\sqrt{4 \cdot 54 - 14^2}\sqrt{4 \cdot 333 - 35^2}}$$

$$= \frac{46}{\sqrt{20}\sqrt{101}} = \frac{46}{\sqrt{2140}}$$

46 ÷ 2nd x^2 2140 Enter .994

when r is close to ± 1, Linear Correlation is Significant

when r is close to 0, Linear Correlation is not Significant

Mar 5-3:08 PM

What about r^2?

Always write r^2 in %.

Coef. of determination r^2

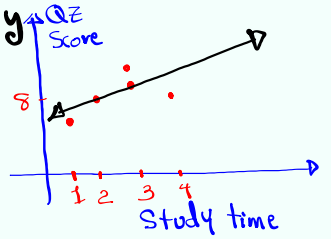
r^2 in % tells us what % of Y-values are explained by x-values.

$r^2 \approx 98.9\% \approx 99\%$

99% of Y-values are explained by x-values

Mar 5-3:18 PM

Study time	Qz Score
2	8
3	9
3	10
4	8
1	6



Study time → x → L1

Qz Score → y → L2

STAT → CALC

8: LinReg(a+bx) with L1, L2

Regression line

$\hat{y} \approx 6 + .8x$

Coef. of determination

$r^2 \approx 42\%$

42% of Qz Scores are explained by study time.

58% unexplained.

$a = 6$

$b = .846$

$r^2 = .423$

$r = .650$

Linear Correlation Coef. $r = .650$

it is closer to 1 than 0. Somewhat significant.

Mar 5-3:23 PM

How to make Predictions:

When r is significant

use the regression line

when r is not significant

use $\bar{y}$ $\bar{y} = \frac{\Sigma y}{n}$

VARS

5: Statistics

5: $\bar{y}$ Enter

Mar 5-3:32 PM

Predict QZ Score when Study time is 4hrs

1) Assume r is Significant

$\hat{y} = 6 + .8x$

$= 6 + .8(4) = 6 + 3.2 = \boxed{9.2} \approx 9$

2) Assume r is not Significant.

use $\bar{y} = 8.2 \approx \boxed{8}$

Mar 5-3:36 PM

Qz Score	Exam Score
8	85
7	80
9	92
10	90
6	75

Qz Score → X → L1

Exam Score → y → L2

Use LinReg(a + bx)

$a = 50.8 \approx 51$

$b = 4.2 \approx 4$

$r^2 = .895 \approx 89\%$

$r = .946$

r is close to 1

Linear Correlation is Significant.

Regression line

$y \approx 51 + 4x$

89% of exam Scores are explained by Qz Scores.

Predict exam Score for Qz Score 8.

1) Assume r is Significant

$y \approx 51 + 4x = 51 + 4(8) = 51 + 32 = \boxed{83}$

2) Assume r is not Significant

Use $\bar{y} = 84.4 \approx \boxed{84}$

VARS

5: Statistics | 5: $\bar{y}$ | Enter

Mar 5-3:39 PM