Pick Up the Solutions to the HW

assuming that you did do the assignment with fidelity.

WarmUp in your notes

HW Tally

Convert

$$y - 3.47 = 0.68(x - 4.36)$$

$$y - y_1 = m(x - x_1)$$

which is in **POINT-SLOPE** form to **slope intercept** form and then state the **gradient** and **y-intercept**.

$y = mx + c$

Convert $y - 3.47 = 0.68(x - 4.36)$ to slope intercept form and then state the **gradient** and y-intercept.

$$y - 3.47 = 0.68x - 2.9648$$

$$y = 0.68x + 0.5052$$

gradient is 0.68
y-intercept (0, 0.5052)

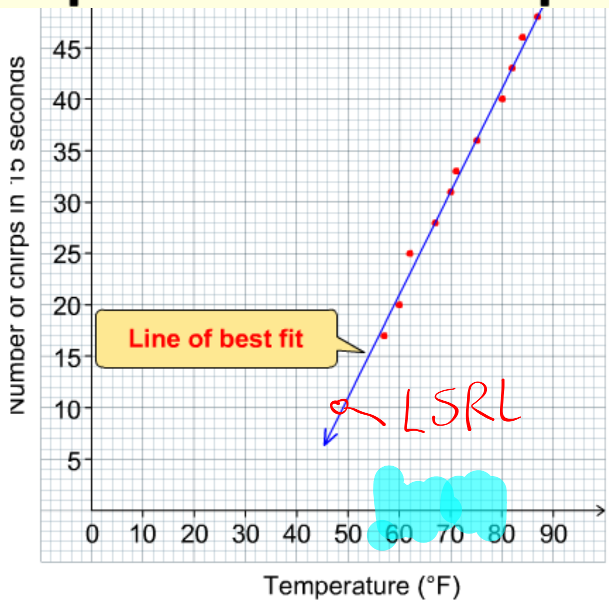
$y - y_1 = m(x - x_1)$

Write the equation, In POINT-SLOPE form, of the line with a slope of 3/7 that passes through (-4, 6)

leave your answer in POINT-SLOPE form

$$y - 6 = \frac{3}{7}(x + 4)$$

Interpolation / Extrapolation



Check your HW solutions.

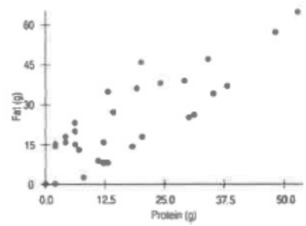
Let me know if you have any questions.

Statistical Applications

Name ______________________________

Assignment 2

The following is a scatterplot of total *fat* versus *protein* for 30 items on the Burger King menu:



1. Estimate a reasonable value for the linear correlation coefficient $r =$ ____
2. Interpret the correlation *(remember, there may be two things you need to do)*

Do heavier cars really use more gasoline?

3. Create a scatter plot given the two-variable data. Be sure to put the dependent variable, the response variable, on the x-axis. <u>Always</u> label each axis fully.

Weight of car in hundreds of pounds (x)	27	44	32	47	23	40	34	52
Miles per gallon (y)	30	19	24	13	29	17	21	14

4. Just by viewing the scatter plot, interpret the correlation.

5. Calculate the linear correlation coefficient to confirm your interpretation. $r =$ ____

6. Now calculate r by "hand" showing the complete formula, followed by the formula with the three critical totals shown, followed by the answer.

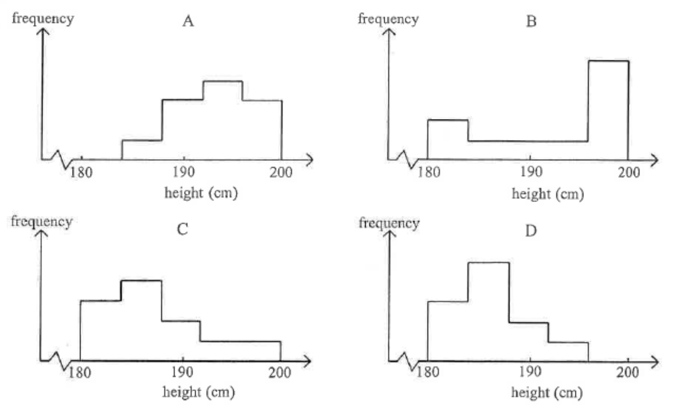
7. Calculate the LSRL (least squares regression line which is a commonly accepted line of best fit). Use the calculator basics reference sheet if needed.

   y = ______________________

8. Use the LSRL equation to estimate the gas mileage of a car that weighs 2000 pounds.

   Do you feel this estimate is trustworthy?

9. The heights in cm of the members of 4 volleyball teams A, B, C and D were taken and represented the frequency histograms given below.

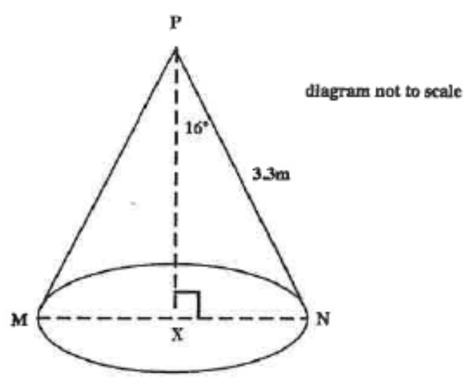


$\overline{x}$ and σ	Team
I	
II	
III	
IV	

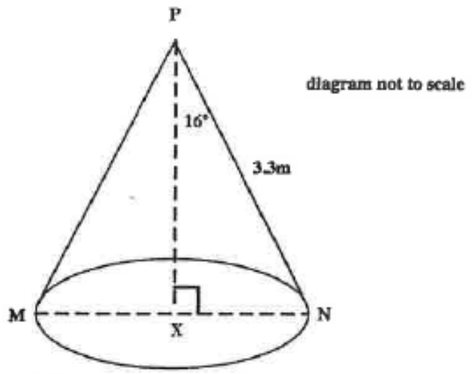
The mean $\overline{x}$ and standard deviation σ of each team are shown in the following table.

	I	II	III	IV
$\overline{x}$	194	189	188	195
σ	6.50	4.91	3.60	3.74

low is the diagram of a cone shaped tent. Angle NPX is 16°, the slant height of the ne is 3.3m.



A right circular cone



A right circular cone

(a) Find the radius of the cone.

(b) Find the vertical height of the cone.

(c) Find the volume of the cone.

Random HW Check

Turn in your HW

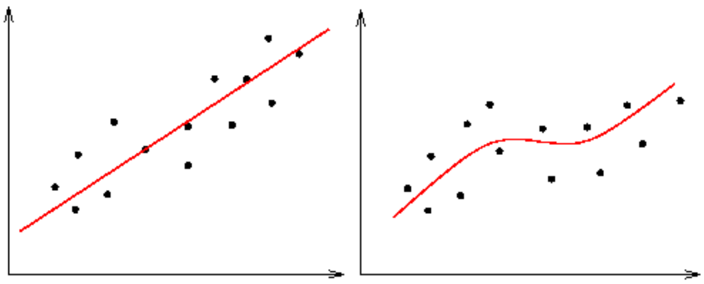
If you end up doing something with correlation on your project, you may want to read the section in the Ch. 11 packet on

The Coefficient of Determination

R^2

$R^2 = r^2$

Curve Fitting



Once we have the equation of the curve, we can use the curve to predict values of y for other values of x. The equation of the curve can also be very helpful for understanding phenomena:

AIM TODAY:

notes will be given

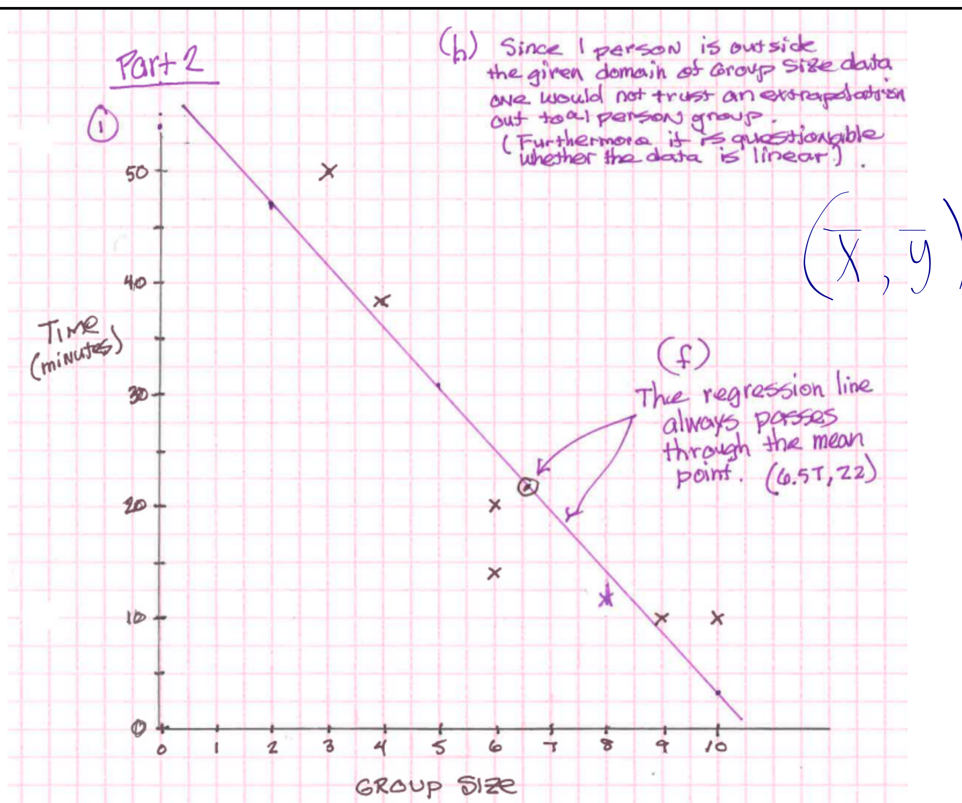
Calculate the LSRL equation "by hand"

LSRL

The line of best fit will always pass through the **mean point** of the two variables

$$(\bar{x}, \bar{y})$$

This mean point is also called the center of gravity of the data.



b) $\bar{x} = 6.57$ people
$\bar{y} = 22$ minutes

c) $r = -0.903$ is the correlation coefficient.

d) There is a strong negative correlation, but not necessarily linear. With this little quantity of data, it might be risky to trust a linear model derived from this data.

e) $y = -5.53x + 58.4$ is the equation of the regression line (a.k.a. LSRL)

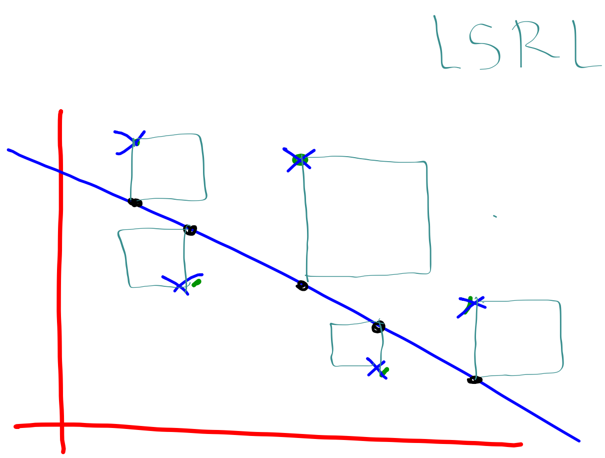
(g) $y = -5.53(5) + 58.4 \doteq 30.8$ minutes for 5 people to do the job.

the differences are called

errors

or

residuals



Applet

Method of Least Squares

Applet

NOTES

Finding the

Least Squares Regression Line

by hand

a.k.a. by First Principles

$$y - \overline{y} = \frac{s_{xy}}{(s_x)^2} (x - \overline{x})$$

in Point-Slope form

the line passes through the mean point, $(\overline{x}, \overline{y})$

required stats

$\bar{x} = 66.444$

$\bar{y} = 62.889$

$S_x = 24.409$

mean point

$(66.444, 62.889)$

① Covariance

$$S_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{1979.444}{9}$$

$$= 219.9382$$

$$y - \bar{y}_1 = \frac{S_{xy}}{(S_x)^2}(x - \bar{x})$$

$$y - 62.889 = \frac{219.9382}{(24.409)^2}(x - 66.444)$$

$$\boxed{y - 62.889 = 0.369118\,(x - 66.444)}$$

LSRL in point-slope form

$$\boxed{y = 0.369x + 38.4}$$

LSRL in gradient-intercept form

$(s_x)^2$ **is the standard deviation, squared**

s_{xy} **is the covariance and will be given to you on the IB test, but you might need to calculate this on a project if you really wanted to.**

Lit Rate vs Population

required stats

$\bar{x} = 66.444$

$\bar{y} = 62.889$

$S_x = 24.410$ ← Use σ on GDC

If using on on a project round beyond one or two decimal places if your data is in the hundreds.

on a project you would want to briefly explain

① Covariance $S_{xy} = \dfrac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n}$

$= \dfrac{1979.444\ldots}{9}$

$= 219.9383$

$(66.444, 62.889)$

(2) Point-Slope

$$y - \bar{y} = \frac{s_{xy}}{(s_x)^2}(x - \bar{x}) \qquad y - 62.889 = \frac{219.9\ldots}{(24.\ldots}$$

$$\boxed{y - 62.889 = 0.3691\ldots}$$

if on an exam an answer on blank . $y - 62.9 = 0.369(x\ldots$

Slope-intercept form

$$\begin{aligned} y - 62.889 &= .369118x - 24.5257 \\ +62.889 &\qquad\qquad\qquad +62.889 \end{aligned}$$

$$y = .369x + 38.4$$

Way of Communicating Understanding

$$S_{xy} = \frac{(36 - 66.444)(55 - 62.889) + (93 - 66.444)(72 - 62.889) + \ldots}{9}$$

$$= \frac{1979.444\ldots}{9}$$

$$= 219.9383$$

With a partner

do the back side

TV / GPA example

STATS Needed : $\bar{x} = 12.4444$ $\bar{y} = 3.0222$ $s_x = 5.4997$

(i) covariance

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{-22.7889}{9}$$

$$= -2.5321$$

② Equation $y - \bar{y} = \frac{S_{xy}}{(S_x)^2}(x - \bar{x})$

mean point (12.4444, 3.0222)

$y - 3.0222 = \frac{-2.5321}{(5.4997)^2}(x - 12.4444)$

simplify slope

Point-slope

$\boxed{y - 3.0222 = -.08371(x - 12.4444)}$

$y - 3.0222 = -.08371x + 1.0418$

$+ 3.0222 \qquad + 3.0222$

slope-intercept

$y = -.08371x + 4.0640$

$\boxed{y = -.0837x + 4.06}$

Gradient = Slope

Quiz on Normal Distrib.

Assignment

"Day 3 Statistical Applications"

Fascinating Furball Fluffies

Why don't statisticians like to model new clothes ?

Lack of Fit

Next......

Intriguing Inert Igloos