

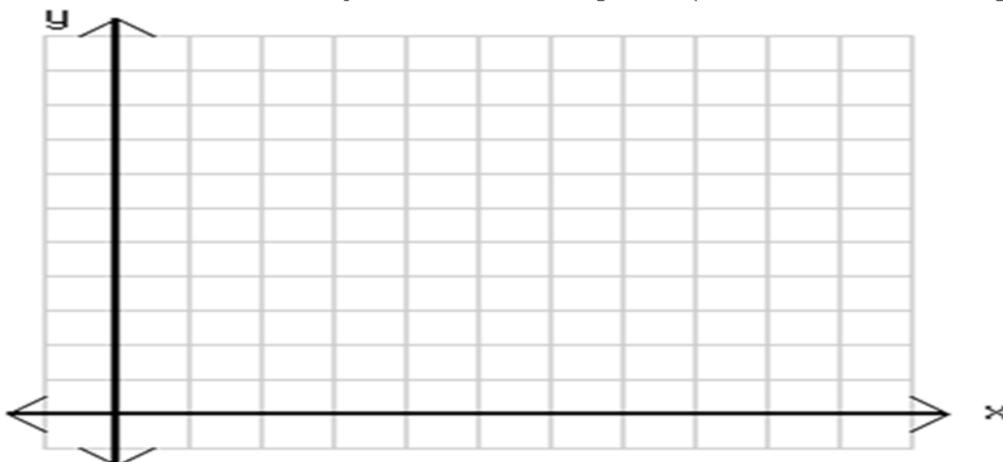
Help can be found in class lecture, topics review or related PowerPoints

1. What is the first step in linear regression?
2. What are the reasons for doing the scatter plot?
3. If scatter plot does not exhibit a linear pattern, can we continue with linear regression?
4. What are different names for x and y variables?
5. What is a lurking variable? Google me to find out!!
6. What a linear correlation coefficient tries to measure?
7. What letter is used for a linear correlation coefficient?
8. What is the range for a linear correlation coefficient?
9. If the value of a linear correlation coefficient is close to +1 what does that suggest?
10. If the value of a linear correlation coefficient is close to -1 what does that suggest?
11. If the value of a linear correlation coefficient is close to 0 what does that suggest?
12. Draw the range of correlation coefficient and label the parts that considered no correlation, strong positive and strong negative correlation.
13. What x or y values for the purpose of estimation can be used in regression equation?
14. In using the regression equation to estimate y and x, what notation is used?

The city's transportation department is interested in studying the relationship between the temperature and the number of passengers that ride the main bus line in order to better serve their customers. The manager recorded the temperature at the beginning of the hour, and then had a bus driver record the number of passengers that boarded the bus throughout the hour. Their findings are listed below.

X= temperature	42	37	46	30	50	43	43	46	46	49
Y= passengers	173	149	185	123	201	174	175	188	186	198

1. **Plot the data** as a scatter diagram, which one is explanatory and which one is the response variable?



- 2) **Comment** on the pattern of the points.

- 3) Compute correlation coefficient (**Round in 2 decimal**) and comment on that 3) $r =$
- 4) Compute the slope and y-intercept (**Round in 2 decimal**) 4) _____
- 5) Write the equation of regression line. 5) _____
- 6) Compute average for temperature and number of passengers (**Round in 2 decimal**) 6) _____
7. Standard deviation for temperature and number of passengers (**Round in 2 decimal**) 7) _____
8. If temperature increases to 47, how many passengers are expected to ride? (**Round in 2 decimal**) 8) _____
9. If 160 passengers are riding the bus what can be the estimated temperature? 9) _____
10. Explain the slope based in the regression equation and in the context of relationship between x and y variables.

For every additional _____ in _____, the _____ changes by _____

Use the data from the following table and find the summation for each column, then use the required formula to find regression coefficient, slope and intercept for the regression equation. Also use TI- calculator to check the answers.

	$x = \text{Hours Study/week}$	$y = \text{Test Score}$	x^2	y^2	$x y$
1	6	72			
2	9	84			
3	4	56			
4	11	94			
5	8	79			
	$\Sigma x =$	$\Sigma y =$	$\Sigma x^2 =$	$\Sigma y^2 =$	$\Sigma xy =$

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

$$\text{Slope} = a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{\quad}{\quad}$$

$$y\text{-itc} = b = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} = \frac{\quad}{\quad}$$

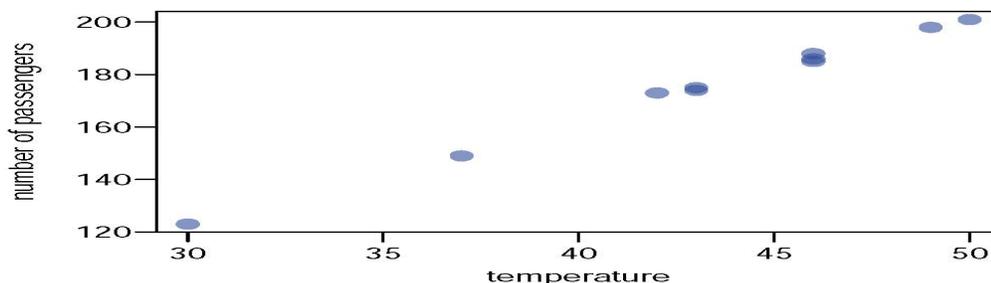
All answers can be found in topics review or related PowerPoints

1. What is the first step in linear regression?
2. What are the reasons for doing the scatter plot?
3. If scatter plot does not exhibit a linear pattern, can we continue with linear regression?
4. What are different names for x and y variables?
5. What is a lurking variable? Google me to find out!!
6. What a linear correlation coefficient tries to measure?
7. What letter is used for a linear correlation coefficient?
8. What is the range for a linear correlation coefficient?
9. If the value of a linear correlation coefficient is close to +1 what does that suggest?
10. If the value of a linear correlation coefficient is close to -1 what does that suggest?
11. If the value of a linear correlation coefficient is close to 0 what does that suggest?
12. Draw the range of correlation coefficient and label the parts that considered no correlation, strong positive and strong negative correlation.
13. What x or y values for the purpose of estimation can be used in regression equation?
14. In using the regression equation to estimate y and x , what notation is used?

The city's transportation department is interested in studying the relationship between the temperature and the number of passengers that ride the main bus line in order to better serve their customers. The manager recorded the temperature at the beginning of the hour, and then had a bus driver record the number of passengers that boarded the bus throughout the hour. Their findings are listed below.

X= temperature	42	37	46	30	50	43	43	46	46	49
Y= passengers	173	149	185	123	201	174	175	188	186	198

1. Plot the data as a scatter diagram, which one is explanatory and which one is the response variable?
Temperature as the explanatory and number of passengers as the response variable



- 2) **Comment** on the pattern of the points. **A very strong positive linear correlation.**

- 3) Compute correlation coefficient (**Round in 2 decimal**) and comment on that 3) $r = 0.9984$
Very strong Positive correlation
- 4) Compute the slope and y-intercept (**Round in 2 decimal**) 4) 3.95 and 4.43
- 5) Write the equation of regression line. 5) $y = 3.953x + 4.3$
passengers = 3.953 temperature + 4.43
- 6) Compute average for temperature and number of passengers (**Round in 2 decimal**) 6) 43.2 and 175.2
- 7) Compute standard deviation for temperature and number of passengers (**Round in 2 decimal**) 7) 5.94 and 23.52
8. If temperature increases to 47, how many passengers are expected to ride?
(**Round in 2 decimal**) 8) $y' = 190$
9. If 160 passengers are riding the bus what can be the estimated temperature? 9) $x' = 39.35$
10. Explain the slope based in the regression equation and in the context of relationship between x and y variables.

For every additional **increase** in **temperature**, the **number of passengers** riding the bus increases by about 4

Use the data from the following table and find the summation for each column, then use the required formula to find regression coefficient, slope and intercept for the regression equation.

	x = Hours Study/week	y = Test Score	x²	y²	x y
1	6	72	36	5184	432
2	9	84	81	7056	756
3	4	56	16	3136	224
4	11	94	121	8836	1034
5	8	79	64	6241	632
	$\Sigma x = 38$	$\Sigma y = 385$	$\Sigma x^2 = 318$	$\Sigma y^2 = 30453$	$\Sigma xy = 3078$

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}} = \frac{(5)(3078) - (38)(385)}{\sqrt{5(318) - (38)^2} \sqrt{5(30453) - (385)^2}} = 0.9896$$

$$\text{Slope} = a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{5(3078) - (38)(385)}{5(318) - (38)^2} = 5.21$$

$$y - \text{itc} = b = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} = \frac{(385)(318) - (38)(3078)}{5(318) - (38)^2} = 37.44$$