*YOUR TA’S NAME*:

*Lecture Worksheet*

*Tuesday 11/3/2020*

**MAIN POINTS OF LECTURE**

1. A scatterplot is a diagram that displays the co-variation of two continuous variables as a set of points on a Cartesian coordinate system.



1. Bivariate regression involves drawing a line through the points on the scatterplot that comes closest to the points on the Y dimension
2. Regression analysis involves estimating an equation that
	1. …describes how, on average, the response variable (Y) is related to the predictor variable (X)
	2. …allows us to make predictions about the value of the response variable (Y) given a specified value of the predictor variable (X)
3. The prediction equation expresses the ith individual’s value of dependent variable Y as a function of predictor variable X
4. The linear regression model  recognizes deviations (or errors, ei) from the prediction equation
5. The values of intercept a and slope bYX that we use to draw the regression line maximize our ability to predict the value of Y (and thus minimizes the prediction errors). Mathematically, we choose the line for which  is smallest. This is the “least squares error sum” criterion and produces ordinary least squares (OLS) estimates of intercept a and slope bYX
6. As with ANOVA, we can ask how much of the variation in Y can be attributed to X and how much is random error
7. The coefficient of determination (R2YX) indicates the proportion of the total variation in Y that is determined by its linear relationship with X



1. The correlation coefficient (rYX) summarize the strength and direction of the linear association between two continuous variables



1. Correlation always ranges from -1 to +1. Correlations between 0 and +1 indicate a positive relationship; if rYX=+1, then there is a perfect positive association. Correlations between -1 and 0 indicate a negative association; if rYX=-1, then there is a perfect negative association. If rYX=0, there is absolutely no association

**QUESTIONS**

1. [From the recorded lecture] Describe the relationship depicted in this scatterplot



 [See the recorded lecture for the answer]

1. [From the recorded lecture]:



 [See the recorded lecture for the answer]

1. [From the synchronous lecture] What was the correlation between (1) students’ perceived chances of being infected with the coronavirus that causes COVID-19 and (b) students’ perceived chances of being hospitalized if they were to be infected with the coronavirus?

 [See the recorded synchronous session for the answer]