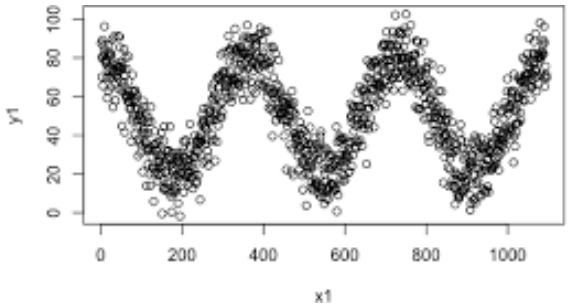
MACHINE LEARNING – WORKSHEET 8 (REGRESSION)

All the questions in this worksheet have one or more than one correct answers. Choose all the correct options to answer your questions.

1. X represents the set of all predictors (X1, X2, X3...) and Y represents a continuous response variable. Then in regression, we try to find a function f such that
   Y=f(X)+Є
   Where Є is random error term
   Which of the following terms are true regarding the Є term?
   A) The mean of the Є is zero
   B) Є can be either or negative
   C) Є is the error which is introduced because there may be some predictor variables apart from the given variable which are effecting the response variable.
   D) None of the above.
   Answer: B)
2. What does this error term Є account for?
   A) It covers up the errors introduced because of the variables which are not present in the given predictors set X but they are influencing the response variable.
   B) It also covers up the measurement errors in the response variable.
   C) It covers up the error introduced because we can never find the exact or true f(X).
   D) None of the above.
   Answer: B) & C)
3. X represents the set of all predictors (X1, X2, X3...) and Y represents a continuous response variable. Then in regression, we try to find a function f such that
   Y=f(X)+Є
   Where Є is random error term
   Which of the following is true regarding the random error Є?
   A) The distribution of random error will be Gaussian.
   B) The random error will be dependent on y.
   C) The random error will be dependent on X.
   D) The random will be independent of X.
4. In which of the following scenarios we can use regression technique?
   A) To predict the sales of a product based on the marketing budget on different media like TV, Newspaper, Radio, etc.
   B) To understand the relationship between the response variable sales and the predictor variables marketing budget on different media like TV, Newspaper, Radio, etc.
   C) To predict the stock price of a company based on the previous stocks prices.
   D) None of the above
   Answer: A),B) & C)
5. X represents the set of all predictors (X1, X2, X3...) and Y represents a continuous response variable. Then in regression, we try to find a function f such that
   Y=f(X)+Є
   Where Є is random error term
   Choose the correct options from following:
   A) The random error can be reduced by Least squares method.
   B) The random error can be reduced if we try to fit a non-linear curve to the data.
   C) The random error is independent of the curve to be used or fit of the curve.
   D) All of the above
   Answer: D
6. Which of the following methods is most commonly used for fitting a curve to data?
   A) Minimum error loss
   B) Least Squares Method
   C) Minimising Euclidean distances
   D) None of the above

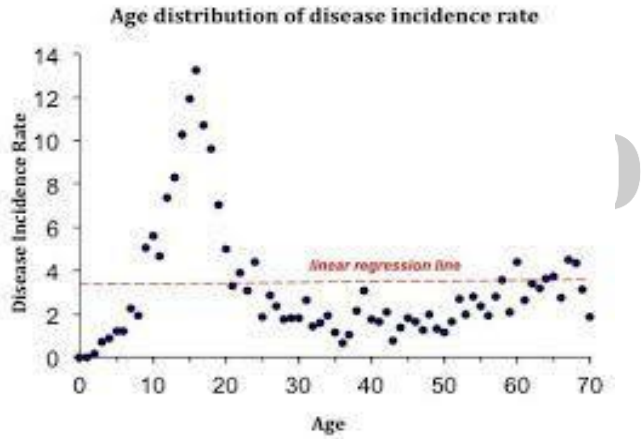
Answer : B & C

7. Consider the following image and answer the following question?



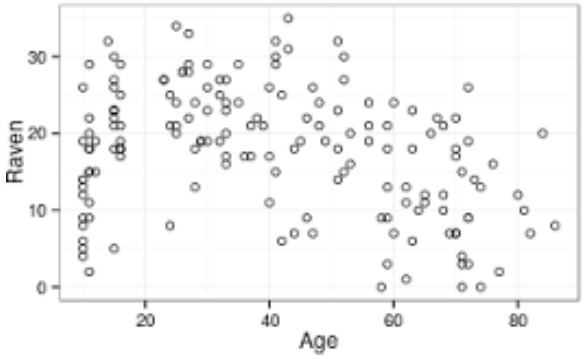
Which of the following functions should be used here for regression?
y1= response variable
x1= predictor variable
A) y1 = B*x1 + C, where B and C are parameters.
B) y1 = B where B is a parameter.
C) y1 = B* Sin(x) + C, where B and C are parameters
D) None of the above
Answer:C)

8. Consider the following image and answer the following question?



Choose the functions is best suited here for regression?
A) Linear function of AGE
B) Polynomial function of AGE with degree>1
C) Cosine function of AGE
D) none of them
Answer:B)

9. Consider the following image and answer the following question?

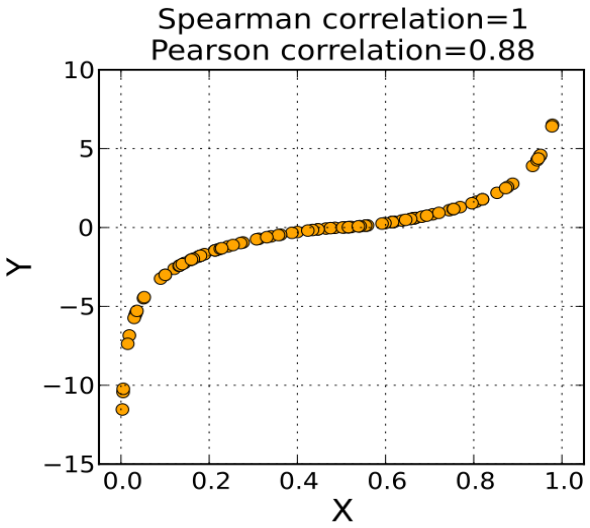


which of the following is true regarding the Pearson correlation coefficient P?

A) P = 1
B) P > 1
C) -1 < P < 0
D) None of the above

Answer:C)

10. Consider the following image and answer the question:



Which of the following is true regarding the data in above picture?

A) Pearson Correlation Coefficient P = 1
B) Pearson Correlation Coefficient P > 1
C) Spearman correlation coefficient = 1
D) None of the above

Answer:C)