



PES University, Bangalore
(Established under Karnataka Act No. 16 of 2013)

SRN

--	--	--	--	--	--	--	--	--	--

UE15CS322
Instructor -
Chidambaran
Iyer

END SEMESTER ASSESSMENT (ESA)
B.TECH. V SEMESTER- Dec 2017

UE15CS322 – DATA ANALYTICS

Time: 3 hrs	Answer All Questions	Max Marks: 100
-------------	----------------------	----------------

General Instructions: All questions are compulsory

I
A)

- (i) Explain in words Principal Component Analysis (2 marks)
- (ii) Suppose that the first and second central moment of a random variable can be classified as either as high or low. Now as a data analytics student, define a random variable for each of the professional below and the professional's desired classification (with respect to the first and second moment) for the random variable.
- a) Shop keeper (2 marks)
 - b) Product Quality Control officer (2 marks)
- (iii) In a throw of two dice, compute and draw the probability density function for the random variable X. X is defined as the sum of the numbers shown on two dice. (4 marks)

B) Using appropriate mathematical notations, define and explain

- (i) Discrete joint probability density function (2 marks)
- (ii) Conditional probability density function (2 marks)
- (iii) Normal distribution (2 marks)
- (iv) Student's t-distribution (2 marks)
- (v) F- distribution (2 marks)

II
A)

- (i) Define using appropriate mathematical notations, an unbiased estimator and a minimum variance estimator? What are these properties called? (3 marks)
- (ii) Write down any five assumptions of the classical linear regression model. (5 marks)
- (iii) Write down the formula for the coefficient of determination (R^2). What does it measure? Ensure you explain all the terms used. (2 marks)

[illegible]

(i) From a sample of 209 firms, the following regression results with the dependent variable being log (salary) were obtained :

where salary = salary of CEO

; RoE = return on equity in percent

; RoS = return on firm's stock.

- iii

- IV

- (i) In the context of recommendation systems, describe cold start. (3 marks)
- (ii) Why is item based collaborative filtering more popular than user-based collaborative filtering. (3 marks)
- (iii) How does the concept of support help us identify a frequent itemset? Explain in detail. (4 marks)

[illegible]

202

- (2+3 marks)

A)

- (4 marks)

B)

- (2 marks)