

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

UE18CS312

OCTOBER 2020: IN SEMESTER ASSESSMENT B Tech FIFTH SEMESTER TEST - 1

UE18CS312 (4 credit subject) - Data Analytics

		Time; 2 Hrs			Answ	er All Q	uestions		Max N	Marks: 60	
1.	a)	An online certifica of CSE. The nur country at the end	nber o	f registi	ations a	ınd numl	ber of suc	cessful ce	rtifications		4 (2+2)
				<u> </u>			<u> </u>	L			
		Number of registrations	44	101	386	4,904	12,106	74,696	1,02,458	12,524	:
		Successful certifications	6	59	174	359	18,036	72,599	96,239	6,980	
		(ii) A potentia certificatio Assuming	how wo I regist ns betw the de	ould we rant wa veen the stailed o	represer nts to ar e fifth and lata for f	nt the data swer the d seventh fifth and	a provided i question: ' n semester	n the table 'Is the inco students s mesters is	?	number of ignificant?"	
	b)	(ii) The organ	ny anor nizers d	naly? S	ubstantia ourse ha	ate your a ve realiz		a has not			3
	c)	The range of scor the marks scored scale map to in the	are re	scaled :							3
2.	a)	bar fro (ii) A restaura	factory nbly lin om eve ant has	produc es (two ry other placed	es ten d for each line. a feedb	lifferent f n flavour) ack card	lavours of A taste te	the choco ster selec able and a	plates and t ts a random	to choose	4 (2+2)
	b)	For the following ratio and discrete/					s as nume	ric/ catego	orical, ordina	al/ interval/	3
		(i) Movie rati (ii) When boo requir (iii) Temperat	oking a ed for t	flight the pass	icket, re: enger (y	es/ no)					
	c)	Twenty engineers following six featu Hearing (v) Sens Briefly outline the	res: (i) ory mo	Intellig tor cod	ence (ii) ordinatio	Conforn on and (v	nance to pi i) Persever	rocedure ance.	(iii) Eyesigh	nt (iv)	3 (2+1)

		groups of twenty p											la dana in i	tho	4 (2+1	
а	a)	Taste testers Ama	n and M	lani ha	ve ra	ted th	e qua	lity of	food	at a res	stauran	it on s	six days iii	uie	+1)	
		week as follows:													.,	
ĺ			 		147	Th	F	Sa	1							
		Day	M		W	<u>Th</u>	1	3	-					1		
		Rating(Aman)	4		3	5 5	2	2	-							
		Rating(Mani)	3	3	2	5_		1	J							
	-	Given: mean and	_tandor	l dovio	tion c	of ratio	uas. IIv	man≒ 3	3. GAn	_{ian} = 1,2	291, µм	_{aní} = 2	.833,			
		Given: mean and $\sigma_{Mani} = 1.067$, cor	Stanuare	, uevia	ant/A	man	Mani)	= 0.72	258		•			1		
	1													1		
		(i) What are	β₀ and β e linear	3 ₁ if we reg re ss	musi sion v	t pred vith th	ict Am e follo	an's n wing i	ating mode	in term 1?	s of Ma	ani's r	ating using	the		
			Ra	ting(A	man)) = β ₀	+ β1 • F	Rating	(Mar	ıi)						
		(ii) What is t	ha aaaffi	cient O	f dete	ermina	ation fo	or this	mod	el?						
		(ii) What is t	ne coem me	asure f	he inf	fluenc	e that	Mani'	s rati	ng of th	e food	on Th	nursday has	s on		
		(III) NOW car	nodel? (S	Suaaes	st the	test o	r statis	stic th	at ca	n be us	ed for t	his.)]		
															3	
+	b)	The correlation I	etween	two va	ariabl	es (#\	views	for a	videc	and a	verage	#vide	eos posted	per	3	
	-,	month) on a vide	eo sharir	ng platf	orm	is fou	nd to I	e po:	sitive	y corre	lated. A	Answe	er the follo	wing		
		questions and (b	riefly) su	bstanti	iate y	our at	rswer:									
								ī	نامامه		fficient	hetw	een #nosti	inas/		
		(i) Is it n ed	essarily	true t	hat t	he Pe	earson	'S COI	relati	on coe	han it is	to ()	for this data	a?		
١		mon	th and #	views o	on a \	video i	would	(i) Is it necessarily true that the Pearson's correlation coefficient between #postings month and #views on a video would to be closer to 1 than it is to 0 for this data?								
		(ii) Can we assume there is no cause-effect relationship between #postings per month									onth					
١	(ii) Can we assume there is no cause-effect relationship between #postings per month and #views on a channel because correlation does not imply causation?										tween	#post	ings per m	onth		
		(ii) Can we and	assume #views o	there on a ch	is no anne	beca beca	se-effe ause c	ct rela	ations	ship bei	tween	#post	ings per m	onth		
		and	#views o	on a ch	anne	beca	ause c	ct rela	ation: tion c	ship be oes no	tween i t imply	#post causa	ings per mation?	<u> </u>	3	
	c)	and For each of the	#views o	on a ch	rplots	l beca	e whet	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	and For each of the	#views o	on a ch	rplots	l beca	e whet	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the fand, if it is not	#views of following, what to	on a ch scatte	rplots	s, state y = on(s)	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to	on a ch scatte	rplots	s, state y = on(s) egress	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the fand, if it is not	#views of following, what to	on a ch scatte	rplots	s, state y = on(s)	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to	on a ch scatte	rplots rmatic ear re	s, state y = on(s) egress	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to	scatte	rplots	s, state y = on(s) egress	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to	on a ch scatte	rplots rmatic ear re	s, state y = on(s) egress	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to	scatte	rplots	s, state y = on(s) egress	ause control e whet β₀+β may b	ct rela	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
	c)	For each of the tand, if it is not amenable for me	#views of following, what to odeling v	scatte	rplots	s, state y = on(s) egress	e whete β ₀ + β may be ion.	ct relation the re	ations tion c e da	ship be oes no a is sui	tween to table for	#post causa or line	ation? ar regression	onun	3	
west to the second seco	c)	and For each of the fand, if it is not amenable for more (i)	#views of following, what to odeling v	scatte ransfor vith line	rplots	I beca s, state y = on(s) egress (ii)	e whele β ₀ + β may be sion.	her th	ation of	to the	timply table for variab	#post causa or line le(s)	ings per mation? ar regression make this	on data		
1		and For each of the tand, if it is not amenable for me (i)	#views of following, what to odeling v	scatte ransfor vith line	rplots	I beca s, state y = on(s) egress (ii)	e whete β ₀ + β may be ion.	her th	e dai	to the	table for variab	#post causa or line le(s)	ar regression make this	on data	4 (2+	
1.	c)	and For each of the sand, if it is not amenable for me (i) Write the linear	#views of following, what to odeling v	scatte ransfor vith line	rplots rmatic ear re	s, state y = on(s) egress (ii) for co	e whete β ₀ + β may be ion.	her the appearance app	estiniable	to the	table for variab	#post causa or line le(s)	ar regression make this ctor in a muent variable	on data	4 (2+	
4+.		and For each of the fame and, if it is not amenable for more (i) Write the linear linear regression	#views of following what to odeling variations algebra in system	scatte ransfor vith line	rplots rmatic ear re	for co	e whete β ₀ + β. may be silven.	ther the sea ap	estiniiable	to the	table for variab	ta vecepend	ar regression make this control in a mulent variable 0.01). If the	on data ultiple es. In here is	4 (2+	
1.		and For each of the fame and, if it is not amenable for more (i) Write the linear linear regression	#views of following what to odeling variations algebra in system	scatte ransfor vith line	rplots rmatic ear re	for co	e whete β ₀ + β. may be silven.	ther the sea ap	estiniiable	to the	table for variab	ta vecepend	ar regression make this control in a mulent variable 0.01). If the	on data ultiple es. In here is	4 (2+	
1.		and For each of the sand, if it is not amenable for me (i) Write the linear	#views of following, what to deling very algebra on syster below, ice a to do the following to do the foll	scatte ransfor vith line ic equa n to pr dentify nis, list	rplots rmatic ear re	for co	e whete β ₀ + β. may be silven.	ther the sea ap	estiniiable	to the	table for variab	ta vecepend	ar regression make this control in a mulent variable 0.01). If the	on data ultiple es. In here is	4 (2+	
1.		Write the linear linear regression the table given insufficient data of regression controls.	#views of following what to deling what to deling what to deling what is algebra an system below, ich to do the cefficient	scatte ransfor vith line ic equa n to pr dentify nis, list	rplots rmatic ear re	for co	e whele β ₀ + β may be ion. mputilipende es that other contents of the cont	her the Appendix Appe	estine ignification	to the nate of susing cant (for essary)	table for variab	ta vecepend	ar regression? ar regression make this ctor in a multiple ent variable 0.01). If the	on data ultiple es. In here is	4 (2+	
1.		For each of the fand, if it is not amenable for more (i) Write the linear linear regression the table given insufficient data of regression control.	#views of following what to deling what to deling what to deling what is algebra in system below, it is deficient to defficient coefficient.	scatte ransfor vith line ic equa m to pr dentify nis, list	rplots rmatic ear re	for co	e whete β ₀ + β may be ion. mputili pende es that other contents of the con	her the Appendix Appe	estiniable	nate of susing cant (foessary	table for variab	ta vecepend on a mine	ar regression? ar regression make this ctor in a make this onumber that the significant controls in the significant control control controls in the significant control control controls in the significant control co	on data ultiple es. In here is	4 (2+	
1.		For each of the fand, if it is not amenable for more (i) Write the linear linear regression the table given insufficient data of regression contacts.	#views of following what to deling what to deling what to deling what is algebra in system below, it is a to do the cefficient contact of the ceffic	scatte ransfor vith line ic equa n to pr dentify nis, list is.	rplots rmatic ear re	for co	e whete β ₀ + β may be sited. The computing pendees that other computing the computing pendees that other computing pendees the computing pendees that other c	her the Appendix Appe	estiniable	nate of s using cant (for essary)	table for variab the Be 5 inder an all to dete	ta vecepend oha =	ar regressions make this control of the signification.	on data ultiple es. In here is	4 (2+	
1.		Write the linear linear regression the table given insufficient data of regression colors.	#views of following what to deling what to deling what to deling what to deling what to deficient to deficient what the deficient which was the deficie	ic equanto pridentify his, list is.	rplots rmatic ear re	for co	e whele β ₀ + β may be ion. mputing pende es that other contact the contact	her the Appendix Appe	estiniable	nate of s using cant (for essary 15.8881	the Be 5 inder an alp	ta vecepend on a remine	ar regression? ar regression ar re	on data ultiple es. In here is	4 (2+	
4.		Write the linear linear regression the table given insufficient data of regression colors. Term Constant X_1 X_2	#views of following what to deling variation and a system below, in the control of the control o	scatte ransfor vith line ic equa n to pr dentify nis, list s. ef 0.166 25	rplots rmatic ear re	for co	e whete β ₀ + β may be ion. The computing pende set that other computing pende set that ot	her the Appendix Appe	estiniable	nate of s using cant (for essary 1.7495 5.5232	the Be 5 inder an alp	ta vecepend on a mine	ar regressions make this make this on the signification of the significa	on data ultiple es. In here is	4 (2+	
1.		Write the linear linear regression the table given insufficient data of regression colors.	#views of following what to deling what to deling what to deling what is algebra in system below, it is a to do the cefficient condition of th	scatte ransfor vith line ic equa n to pr dentify nis, list s. ef 0.166 25	rplots rmatic ear re	for co 4 de eature what c	e whete β ₀ + β may be ion. The computing pende set that other computing pende set that ot	her the Appendix Appe	estiniable	nate of s using cant (for essary 15.8881	the Be 5 inder an all to dete	ta vecepend oha = rmine	ar regression? ar regression ar re	on data ultiple es. In here is	4 (2+	

4.	b)	Rajesh has designed a visible in the night sky ba	ogistic regress sed on the hum	ion classifier idity reported	to predi	ict the day:	likelih	ood of s	itars being	3 (1+2)		
		logit (p) = log(p/(1-p)) = f	₀ + β ₁ *humidit	y , where p is	the prob	ability	stars a	re visible	e at night.			
		Given that β ₀ = 1.8185 an	d $\beta_1 = -0.0665$,	answer the fo	ollowing	questic	ns:					
		(i) What does the va	lue of β ₀ mean	?	1. 200	المالين المالي		n ora vic	sible in the			
		(ii) If humidity on a night sky acc	day = 25, what ording to this m		ability wi	ith whic	ii Star	s ale vis	sible in the			
	c)	In a collection of 1000 sm All the 100 precious stores' by a logistic recolassifier, clearly labeling	nes along with ression model the rows and	100 other r . Write the columns. Wh	ocks ha entries c nat furthe	ive bee of the er step	en clas confus s shou	sified as ion mat Id be ta	s 'precious rix for this			
		the receiver operator cha	racteristics (Ro	C) for this log	istic regi	ression	mode	l?				
		With a schematic sketo	h-i-fl., -l	ribe the key	pharact	tarietice	of th	e level	trend and	1 4 (3+1)		
5.	a)	seasonality components	of an additive ti	me series dat	ta. What	are cy	clic co	nponen	ts and, why	-		
		are they usually not acco	unted for in mo	dels for time :	series da	ata?						
	b)	For the data given below	, use MAPE to	b) For the data given below, use MAPE to compare the forecast accuracy of single exponentia								
	•	smoothing (SES) with alpha = 0.7 with the forecast accuracy of the simple moving average										
		smoothing (SES) with a	pha = 0.7 with	the forecast	accurac	cy of th	e simp	le movi	ng average	9		
		(SMA) with a window size	e = 3 for time	the forecast points t=5,6,7	accurac 7. [You d	can us	e simp e the v	ole movi values o	ng average	e 		
		smoothing (SES) with a (SMA) with a window size available to make the for	e = 3 for time	the forecast points t=5,6,7	accurac 7. [You d	can us	e simp e the v	ole movi values o	ng average	9		
		(SMA) with a window size	e = 3 for time	the forecast points t=5,6,7	accurac 7. [You d	can us	e simp e the v	ole movi values o	ng average	9		
		(SMA) with a window six available to make the for	e = 3 for time ecasts for SMA	the forecast points t=5,6, and for SES	accurac 7. [You o assume	can use the for	e simpe the vecast,	ole movi values o	ng average	9		
		(SMA) with a window six available to make the for	te = 3 for time ecasts for SMA	the forecast points t=5,6, and for SES	accurac 7. [You dassume	can use the for	e simpe the vecast,	ole movi values o	ng average			
	c)	(SMA) with a window six available to make the for T yt Suggest an application for	te = 3 for time ecasts for SMA 1 2 10 11 or each of the for	the forecast points t=5,6, and for SES	accurac 7. [You dassume	the for	e simple the vecast,	ole moving alues of F4=y4.]	ng average f y that are	3		
	c)	(SMA) with a window six available to make the for T yt Suggest an application for (i) Croston's methology	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold	the forecast points t=5,6, and for SES	accurac 7. [You dassume	the for	e simple the vecast,	ole moving alues of F4=y4.]	ng average f y that are			
	c)	(SMA) with a window six available to make the for T yt Suggest an application for	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold	the forecast points t=5,6, and for SES	accurac 7. [You dassume	the for	e simple the vecast,	ole moving alues of F4=y4.]	ng average f y that are			
	c)	(SMA) with a window six available to make the for T yt Suggest an application for (i) Croston's methor (ii) Holt-Winter's methor (iii) ARIMA	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod	the forecast points t=5,6, and for SES 3 4 12 16 ollowing techn	accurace 7. [You can assume 5 17 17 17 17 17 17 17	the for 19 model	e simple the vecast,	ole moving alues on F4=y4.]	ng average f y that are	3		
6	c)	(SMA) with a window six available to make the for T yt Suggest an application for (i) Croston's methor (ii) Holt-Winter's methor (iii) Holt-Winter's methor (iiii) Holt-Winter's methor (iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod	the forecast points t=5,6, and for SES 3 4 12 16 ollowing techn	accurace 7. [You can assume 5 17 17 17 17 17 17 17	the for 19 model	e simple the vecast,	ole moving alues on F4=y4.]	ng average f y that are	3		
6		(SMA) with a window six available to make the for T yt Suggest an application for (i) Croston's methor (ii) Holt-Winter's methor (iii) ARIMA Write the equation correction (iii) ARIMA(0,1,0)	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod sponding to the	the forecast points t=5,6, and for SES 3 4 12 16 ollowing techn	accurace 7. [You can assume 5 17 17 17 17 17 17 17	the for 19 model	e simple the vecast,	ole moving alues on F4=y4.]	ng average f y that are	3		
6	(a)	(SMA) with a window six available to make the for T yt Suggest an application for (i) Croston's methor (ii) Holt-Winter's methor (iii) ARIMA Write the equation correction (i) ARIMA(0,1,0) (ii) ARIMA(1,0,1) Which model is better are	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod sponding to the did thod	the forecast points t=5,6, and for SES 3 4 12 16 bllowing technology two models of	accurace 7. [You can assume 5 17 17 17 17 19 19 19 19	the for 19 model	e simple the vecast, 7 20 I time s	ole moving alues on F4=y4.]	ng average f y that are	3 : 4 (2+2)		
6	(a)	(SMA) with a window six available to make the for T T yt Suggest an application for (i) Croston's method (ii) Holt-Winter's mediii) ARIMA Write the equation correct (i) ARIMA(0,1,0) (ii) ARIMA(1,0,1)	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod sponding to the did thod	the forecast points t=5,6, and for SES 3 4 12 16 ollowing techn	accurace 7. [You can assume 5 17 17 17 17 19 19 19 19	the for 19 model	e simple the vecast, 7 20 I time s	eries da	ng average f y that are	3 : 4 (2+2)		
6	(a)	Suggest an application for (i) Croston's method (ii) Holt-Winter's med (iii) ARIMA Write the equation correction (i) ARIMA(0,1,0) (ii) ARIMA(1,0,1) Which model is better an incomplete the second of the content of t	te = 3 for time ecasts for SMA 1 2 10 11 or each of the fold thod sponding to the days and the days are th	the forecast points t=5,6, and for SES 3 4 12 16 bllowing technology two models of	accurace 7. [You can assume 5 17 17 17 17 19 19 19 19	the for 19 model	e simple the vecast, 7 20 I time s	eries da	ng average f y that are	3 : 4 (2+2)		