

Name:	 UPES UNIVERSITY WITH A PURPOSE
Enrolment No:	

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, May 2019

Course: Business Intelligence Program: BTech – CS (BAO) Course Code: CSIB 387	Semester: VI Time 03 hrs. Max. Marks: 100
--	--

Instructions: All questions are compulsory. Options are given for questions 9 and 11.

SECTION A

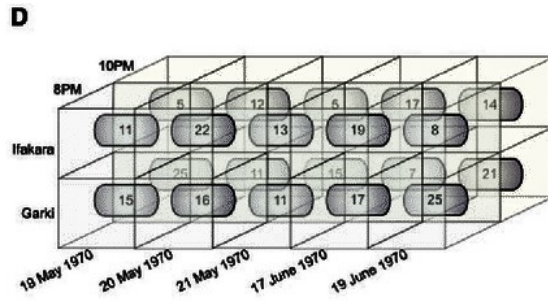
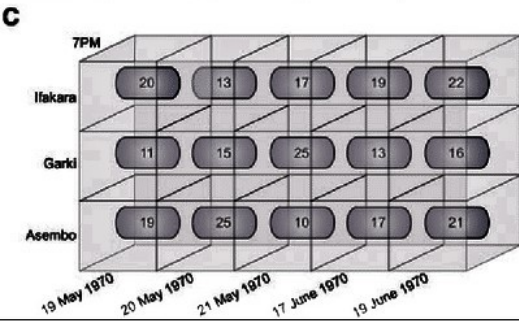
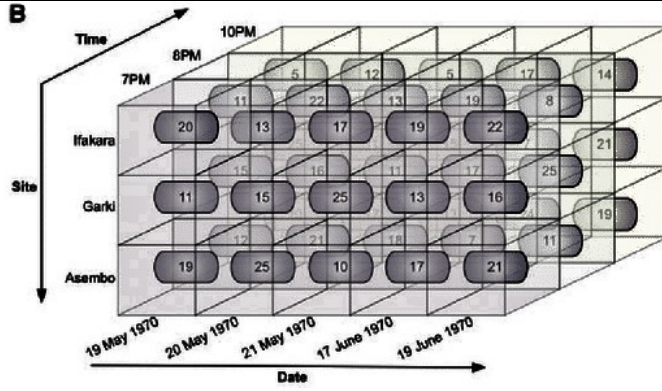
S. No.		Marks	CO
Q 1	Who coined the term BI? Comment on what is the future of BI.	4	CO1
Q 2	Steve Rogers is a Business Analyst and Tony Stark works as a Data Scientist in IBM. Describe how their roles differ. Do you think they will collaborate for any project taken by the company?	4	CO2
Q 3	What is decision support system? What are its three tiers?	4	CO3
Q 4	<p>You got a dataset depicting the popularity of two graphic novels given by a critic which contains three variables.</p> <ol style="list-style-type: none"> 1) Time of survey (in dd-mm-yy format) 2) Rating of ‘Marvel’ (in range between 0 to 10) 3) Rating of ‘DC’ (in range between 0 to 10) <p>The data is collected every day since 1970. You need to graphically represent the data in a chart. What will you use? And why?</p>	4	CO4
Q 5	Differentiate between OLAP and OLTP.	4	CO1

SECTION B

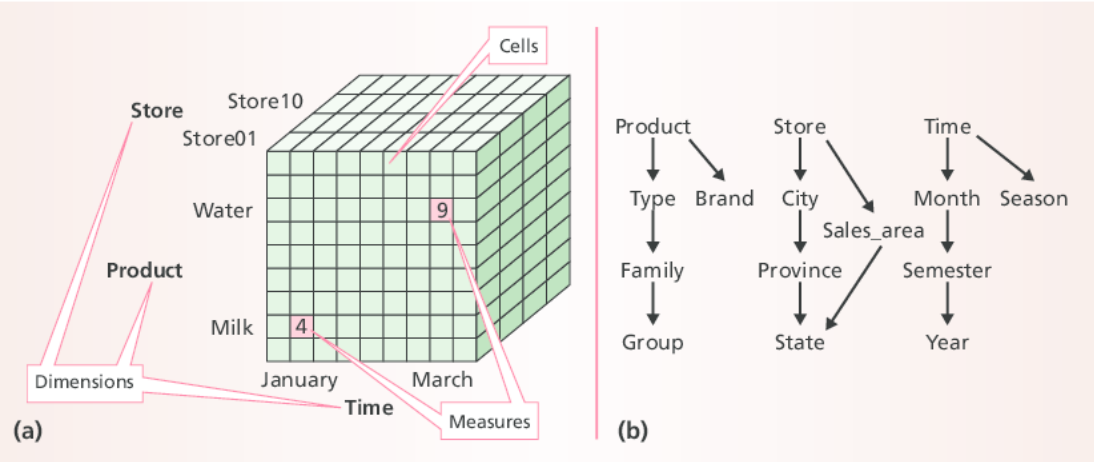
Q 6	What do you mean by analytics? What are its kinds? Explain each of them with proper examples.	10	CO2
Q 7	<p>Consider the figure given below:</p> <p>Here Table: A represents a data set and Table: B its corresponding data cube.</p> <p>What does Table: C and Table: D represent?</p> <p>Explain how are the Tables(C and D) are created in your own words.</p>	10	CO4

A

Date	Time	Site	Abundance
19 May 1970	7PM	Ifakara	20
20 May 1970	7PM	Ifakara	13
21 May 1970	7PM	Ifakara	17
17 June 1970	7PM	Ifakara	19
19 June 1970	7PM	Ifakara	22
19 May 1970	7PM	Garki	11
20 May 1970	7PM	Garki	15
21 May 1970	7PM	Garki	25
17 June 1970	7PM	Garki	13
19 June 1970	7PM	Garki	16
...
19 May 1970	10PM	Asembo	10
20 May 1970	10PM	Asembo	17
21 May 1970	10PM	Asembo	13
17 June 1970	10PM	Asembo	24
19 June 1970	10PM	Asembo	19



Q 8 Suppose you have been given the following data cube(a) from a company and hierarchies on dimensions(b):



Explain how cross tabulation can help to deal with the different hierarchies?

Q 9 Suppose an institute offers only one diploma course on machine learning of duration 1 year. They collected some data for the past few years as given below:

10

CO3

10

CO5

	2016	2017	2018	2019
Students admitted	104	116	120	100
Students graduated	87	92	100	75
Research Grants applied	3	7	6	9
Research Grants received	0	3	5	4
Total Classes	72	65	64	70
Total Cost	54000	60000	64000	60000
Faculty Training Sessions	9	7	2	5

To measure the institute's performance, the management decided on some 'Key Performance Indicators' (KPIs) defined as:

- 1) Graduation Rate: How many students graduating in that year.
- 2) Research Grant: Percentage of research grant received over applied.
- 3) Cost per Student: The amount of money required to educate one student.

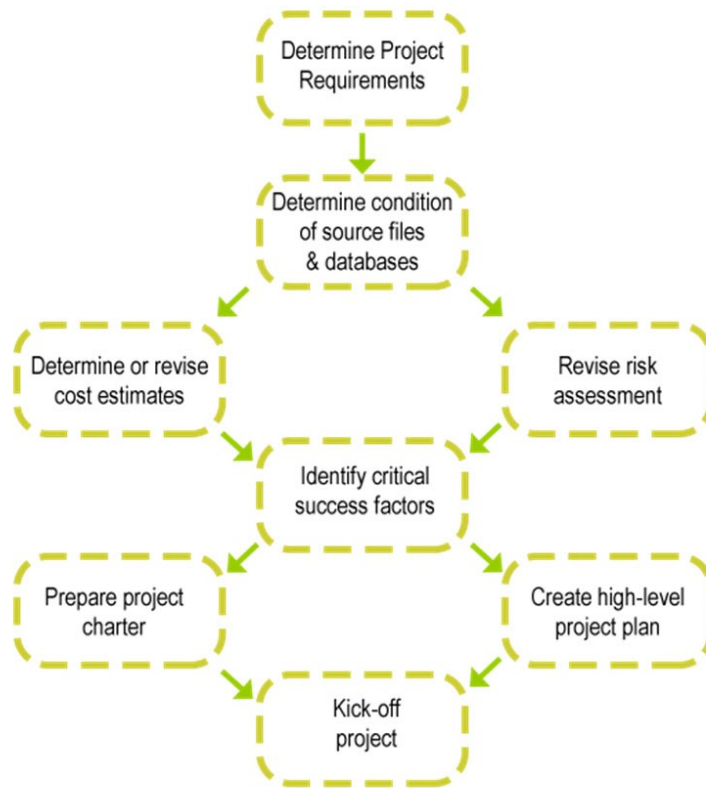
Compute the KPIs for each year. Where do you think these KPIs will be used in the Kaplan-Norton Balanced Scorecard?

OR

Discuss the Gestalt's dashboard design principles. Which graphs should be avoided and which should be used when and how.

SECTION-C

Q 10	<p>The steps are required to plan a BI project are given in the figure below.</p> <p>The company that you are working for is a stocks investment firm. How are you going forward to implement a BI project?</p>	20	CO4
------	---	-----------	------------



Q 11 Dealing with a BI solution architecture, you are asked to create an architecture of UPES. Explain how Centralized and Decentralized architecture will differ. Explain other alternative architectures. How can you expand UPES BI potential?

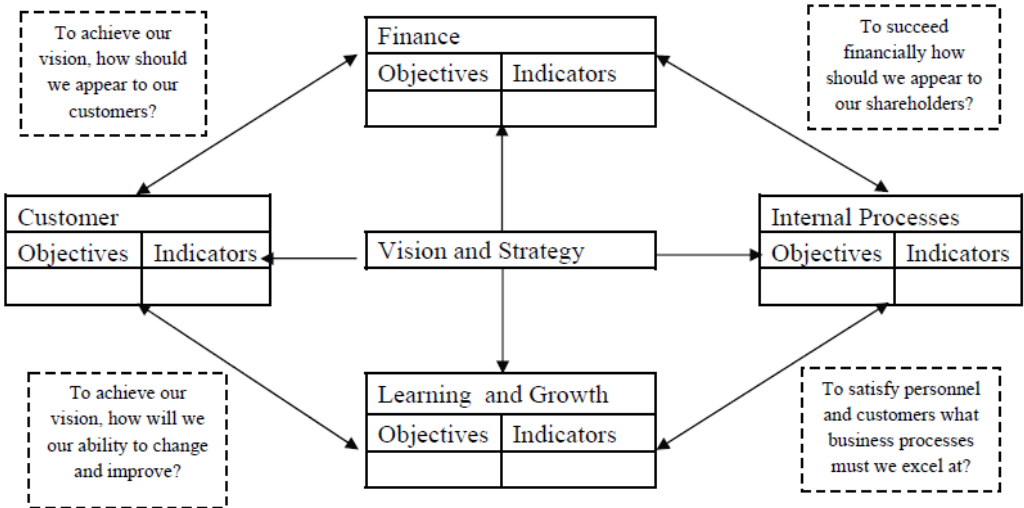
20 CO5

OR


A graphic novel company named ‘Marvel’ looked at its data as shown as below:

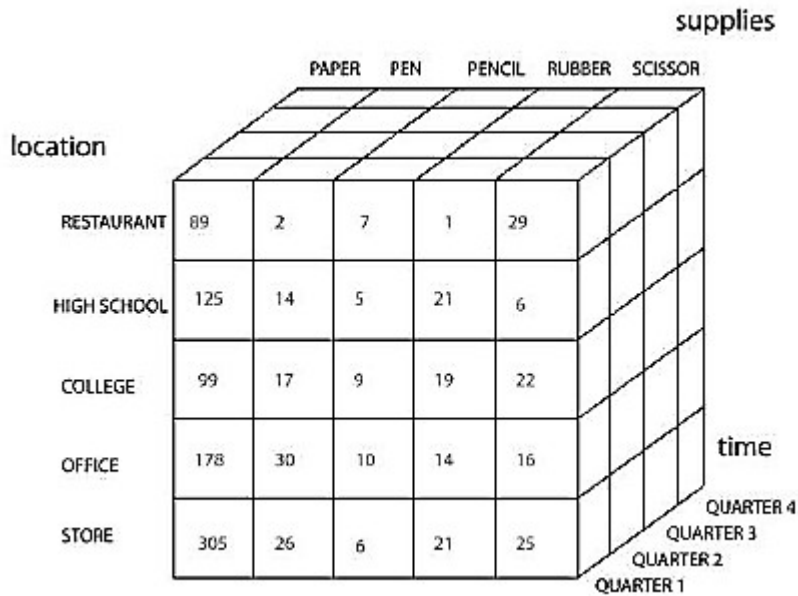
Customer complaints	100
Customer complaints resolved	70
Total Customers till last year	10000
Total Customers till today	98000
Printing Cost last year	8000
Printing Cost this year	7000
New characters introduced last year	12
New characters introduced this year	7
Incentives given to cartoonist last year	10000
Incentives given to cartoonist this year	12000
Average cost of one comic last year	95
Average cost of one comic this year	105

Based on the data, you are now asked to fill up the Objectives and Indicators in each of the four sections given below:



What kind of initiatives could you take to increase your target in each section?
 (Mention creating another column named Target in each section of the balanced scorecard)

Name: Enrolment No:			
UNIVERSITY OF PETROLEUM AND ENERGY STUDIES End Semester Examination, May 2019			
Course: Business Intelligence Program: BTech – CS (BAO) Course Code: CSIB 387		Semester: VI Time 03 hrs. Max. Marks: 100	
Instructions: All questions are compulsory. Options are given for questions 9 and 11.			
SECTION A			
S. No.		Marks	CO
Q 1	Steve Rogers is a Business Analyst and Tony Stark works as a Data Scientist in IBM. Describe how their roles differ. Do you think they will collaborate for any project taken by the company?	4	CO2
Q 2	Differentiate between OLAP and OLTP.	4	CO1
Q 3	<p>You got a dataset depicting the popularity of two graphic novels given by a critic which contains three variables.</p> <ol style="list-style-type: none"> 1) Time of survey (in dd-mm-yy format) 2) Rating of ‘Marvel’ (in range between 0 to 10) 3) Rating of ‘DC’ (in range between 0 to 10) <p>Now the data is collected every day since 1970. You need to graphically represent the data in a chart. What will you use? And why?</p>	4	CO4
Q 4	What is decision support system? What are its three tiers?	4	CO3
Q 5	Consider the following data cube:	4	CO1



What do you understand from it? (dimensions, attributes, measures)

SECTION B

Q 6 “A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management’s decision-making process.”—W. H. Inmon. Explain with your own understanding the definition put forward. **10** **CO2**

Q 7 A company has provided the following data in form of some tables: **10** **CO3**

Table:1

STORE					
Store_id	City	State	District_id	region_id	manager
700121	Dehradun	Uttarakhand	612	788182	Kapil
700122	Roorkee	Uttarakhand	654	788182	Ayush
700123	Dehradun	Uttarakhand	612	788182	Kapil
700124	Guwahati	Assam	891	781271	Manish
700125	Mussoorie	Uttarakhand	653	788182	Ajay

Table:2

PRODUCT						
Product_id	Brand	Colour	Size	Manufacturer_id	Store_id	Shipment_date
80128911	Nike	Black	L	51788182	700121	23-12-2017
80128912	Reebok	Blue	L	51788183	700122	17-06-2018
80128913	Puma	Red	M	51788182	700123	30-11-2019
80128914	Bombay dye	Black	L	51788182	700124	12-04-2018
80128915	Levis	Black	S	51788185	700125	12-04-2018

You are asked to prepare a data cube of Store, Product and Time dimensions. How will you design the fact table if it is to be represented in a star schema? Will there be

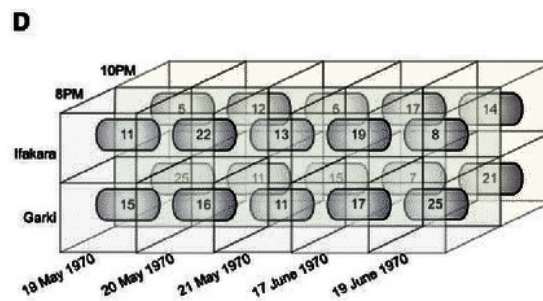
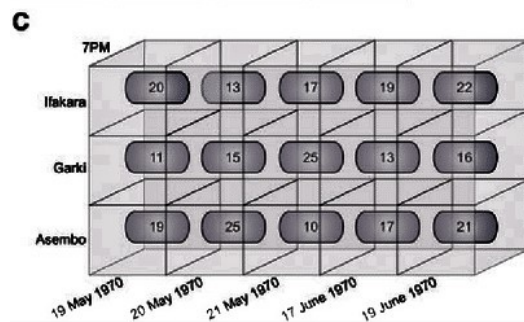
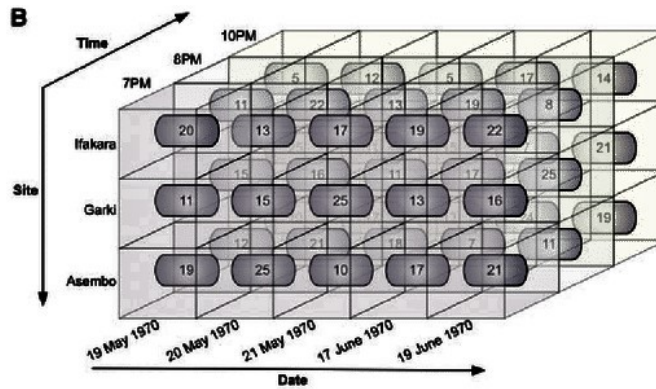
any difference if you represent it using the fact constellation schema?

Q 8

Consider the figure given below:
 Here Table: A represents a data set and Table: B its corresponding data cube.
 What does Table: C and Table: D represent?
 Explain how are the Tables(C and D) are created in your own words.

A

Date	Time	Site	Abundance
19 May 1970	7PM	Ifakara	20
20 May 1970	7PM	Ifakara	13
21 May 1970	7PM	Ifakara	17
17 June 1970	7PM	Ifakara	19
19 June 1970	7PM	Ifakara	22
19 May 1970	7PM	Garki	11
20 May 1970	7PM	Garki	15
21 May 1970	7PM	Garki	25
17 June 1970	7PM	Garki	13
19 June 1970	7PM	Garki	16
...
19 May 1970	10PM	Asembo	10
20 May 1970	10PM	Asembo	17
21 May 1970	10PM	Asembo	13
17 June 1970	10PM	Asembo	24
19 June 1970	10PM	Asembo	19



10

CO4

Q 9

Discuss the Gestalt's visual perception principles applied to Dashboard design.
 Which graphs should be avoided and which should be used when and how.

OR

Suppose an institute offers only one diploma course on machine learning of duration 1 year. They collected some data for the past few years as given below:

	2016	2017	2018	2019
Students admitted	104	116	120	100
Students graduated	87	92	100	75
Research Grants applied	3	7	6	9
Research Grants received	0	3	5	4
Total Classes	72	65	64	70
Total Cost	54000	60000	64000	60000
Faculty Training Sessions	9	7	2	5

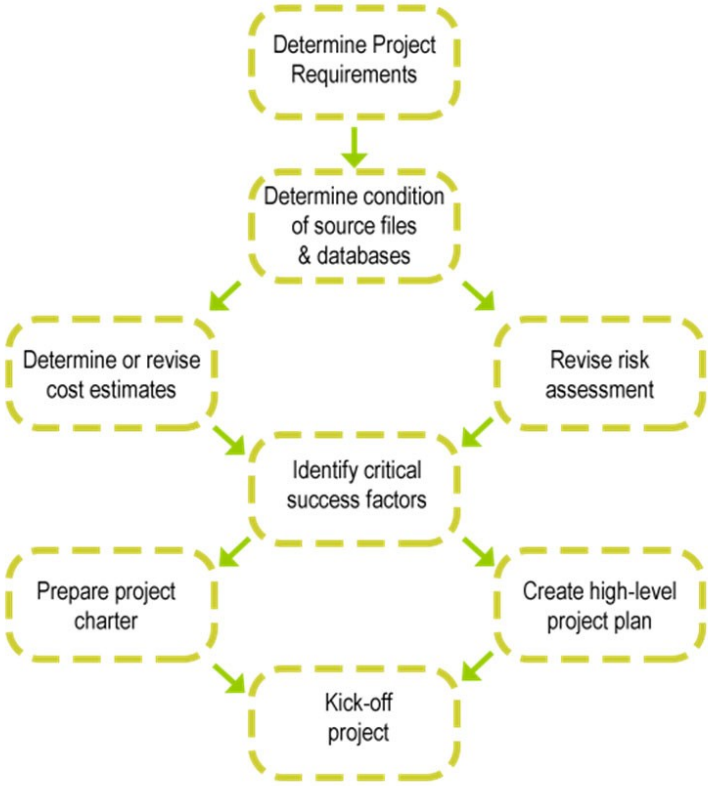
To measure the institute's performance, the management decided on some 'Key Performance Indicators' (KPIs) defined as:

10

CO5

	<p>1) Graduation Rate: How many students graduating in that year. 2) Research Grant: Percentage of research grant received over applied. 3) Cost per Student: The amount of money required to educate one student.</p> <p>Compute the KPIs for each year. Where do you think these KPIs will be used in the Kaplan-Norton Balanced Scorecard?</p>		
--	---	--	--

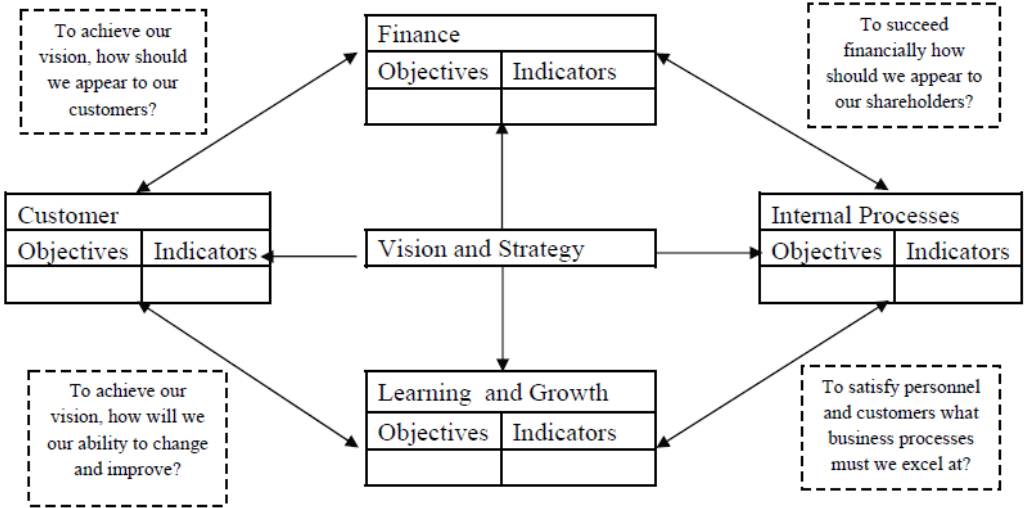
SECTION-C

<p>Q 10</p>	<p>The steps are required to plan a BI project are given in the figure below.</p> <p>The company that you are working for is a News reporting company dealing with a project on 2019 election prediction and survey. How are you going forward to implement a BI project?</p>  <pre> graph TD A[Determine Project Requirements] --> B[Determine condition of source files & databases] B --> C[Determine or revise cost estimates] B --> D[Revise risk assessment] C --> E[Identify critical success factors] D --> E E --> F[Prepare project charter] E --> G[Create high-level project plan] F --> H[Kick-off project] G --> H </pre>	<p>20</p>	<p>CO4</p>
-------------	---	------------------	-------------------

<p>Q 11</p>	<p>An IPL cricket team named ‘Mumbai Indians’ looked at its data as shown as below:</p>	<p>20</p>	<p>CO5</p>
-------------	---	------------------	-------------------

Stakeholders complaints resolved	65
Total viewers till last year	10000
Total viewers till today	98000
Team maintenance Cost last year	80000
Team maintenance Cost this year	70000
New players introduced last year	5
New players introduced this year	3
Incentives given to team last year	10000
Incentives given to team this year	12000
Average cost of one player last year	5700
Average cost of one player this year	6900

Based on the data, you are now asked to fill up the Objectives and Indicators in each of the four sections given below:



What kind of initiatives could you take to increase your target in each section for the coming season? (Mention creating another column named Target in each section of the balanced scorecard)

OR

Dealing with a BI solution architecture, you are asked to create an architecture of UPES. Explain how Centralized and Decentralized architecture will differ. Explain other alternative architectures. How can you expand UPES BI potential?