Name: **Enrolment No:**	UPES UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM & ENERGY STUDIES

Mid Semester Examination (Online) – June, 2021

Program: MBA (Business Analytics) **Semester: II**

Subject/Course: Business Intelligence **Max. Marks: 100**

Course Code: DSBA 7006 **Duration: 3 Hours**

IMPORTANT INSTRUCTIONS

1. ***Use of calculator is allowed.***
2. *Differentiation in marks will be based on to-the-point answers.*
3. *Please note Writing sentences that misguide the examiner from the actual answer will lead to deduction of marks. So write less but accurate answers. Stick to the instructions given in the question paper.*
4. *In case of any confusion, take an assumption and mention the assumption taken.*
5. *In case of further confusion, feel free to contact the faculty in-charge.*

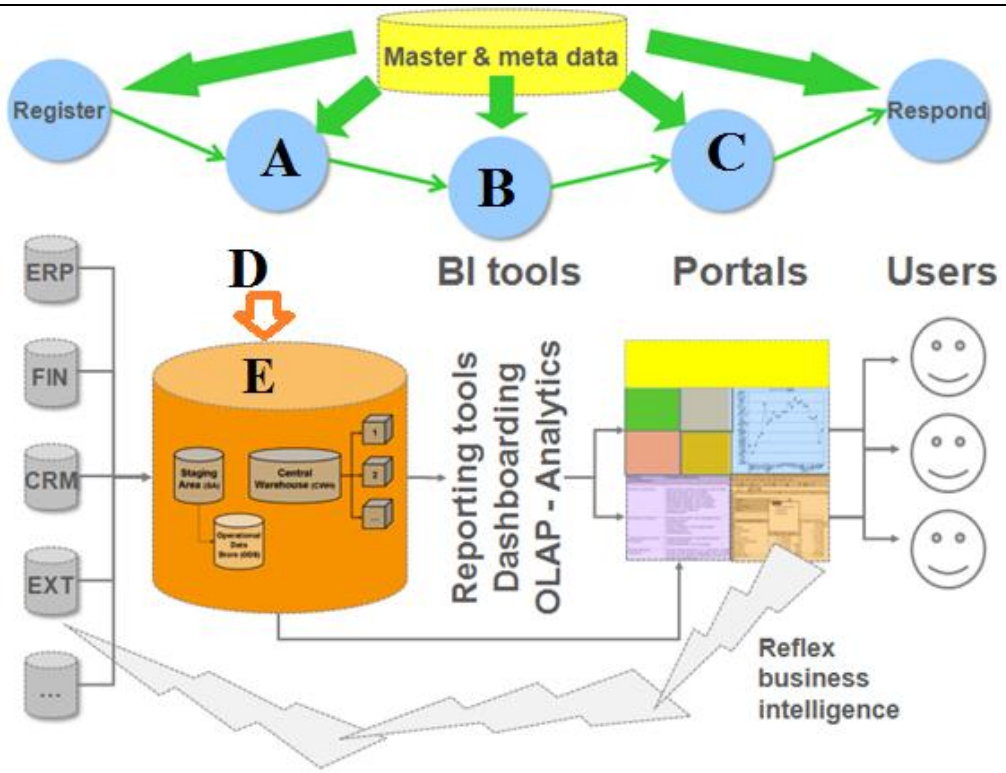
Q.No	SECTION A **1. Each Question will carry 5 Marks** **2. Instruction: Complete the statement / Select the correct answer(s)**	Marks	Cos
1.	Fill in the blanks using the options below: Business Intelligence is a set of methodologies that transform _____ data into meaningful _____. It is generally _____ in nature. The main function of business intelligence is _____ and _____ based on the given dataset. Options: processed, raw, unstructured, information, descriptive, prescriptive, predictive, observation, analysis, interpretation, reporting, algorithmic.	5	CO1

<table>
<tr>
<td>2</td>
<td>Match the following:

<table>
<tr><td>A. SAS</td><td>i) set of database tables to store application data of the Cognos.</td></tr>
<tr><td>B. Content Store</td><td>ii) used for reporting and data analysis and is considered a core component of business intelligence.</td></tr>
<tr><td>C. Content Manager</td><td>iii) Competitor of Cognos</td></tr>
<tr><td>D. Content database</td><td>iv) helps to understand the latest trends, compare data, and assess business performance for multidimensional analysis</td></tr>
<tr><td>E. Cognos Analysis Studio</td><td>v) self-contained database server which is used to host the content store database in demo environments when enterprise DMBS in unavailable</td></tr>
<tr><td></td><td>vi) better understand the product, customer, and organizational needs.</td></tr>
<tr><td></td><td>vii) Manage storage and retrieval of report specifications, configuration data, published packages</td></tr>
</table>
</td>
<td>5</td>
<td>CO3</td>
</tr>
<tr>
<td>3.</td>
<td>True and False: (If false, mention and briefly explain the correct answer with example; If true, explain with example. No marks without explanation)

Generalization requires aggregation whereas specialization requires segregation of data.</td>
<td></td>
<td>CO3</td>
</tr>
<tr>
<td>4.</td>
<td>True and False: (If false, mention and briefly explain the correct answer with example; If true, explain with example. No marks without explanation)

Scientific data processing applies standard relational database and batch processing, whereas Commercial data processing does sorting and summarization using a wide variety of processing tools</td>
<td>5</td>
<td>CO2</td>
</tr>
<tr>
<td>5.</td>
<td>Complete the figure by determining the missing values from A to E:</td>
<td>5</td>
<td>CO1</td>
</tr>
</table>

			
6.	Fill in the blanks from the given options: Data in the real world is dirty because they are _____, noisy, and _____. Noisy data is obtained due to the process of data _____, _____ and _____. Options: wrong, incomplete, misled, unintended, inconsistent, software problems, collection, difference in various sources, functional dependency violation, entry, analysis problem, transmission.	5	CO2
	SECTION B **1. Each question will carry 10 marks** **2. Instruction: Write short / brief notes**		
7.	Mention and explain any five of the eight components of Cognos Software.	10	CO3
8.	Explain the different types of information systems that are used for storage and analysis of data, with one example for each. (2.5 X 4)	10	CO1
9.	Short notes with example: a. Nominal data attribute b. Ordinal data attribute, c. Noisy data and	10	CO1

<table>
<tr><td></td><td>d. Transaction data. (2.5 X 4)</td><td></td><td></td></tr>
<tr><td>10.</td><td>Explain in details using bullet points, how to handle: (i) Missing data (ii) Noisy data
(5 X 2)</td><td>10</td><td>CO2</td></tr>
<tr><td>11.</td><td>Explain the strategies for: (5 X 2)
a. Dimensionality reduction
b. Data transformation</td><td>10</td><td>CO2</td></tr>
<tr><td></td><td>SECTION C
1. Each question will carry 20 marks
2. Instruction: Write long answer (800 words maximum)</td><td></td><td></td></tr>
<tr><td>12.</td><td>Given is a set of attributes and the kind of data they represent
<table><tr><th>Attribute</th><th>Description</th><th>Data type</th></tr><tr><td>Region</td><td>Provides the region East, West, North, South</td><td>Categorical</td></tr><tr><td>Customer Name</td><td>Customer name for the order</td><td>Nominal</td></tr><tr><td>Account Manager</td><td>Account manager in charge of the order</td><td>Nominal</td></tr><tr><td>PO Number</td><td>Order number</td><td>Nominal</td></tr><tr><td>Order Date</td><td>Date of order</td><td>Numeric, Nominal</td></tr><tr><td>Ship Date</td><td>Date of shipment of the order</td><td>Numeric, Nominal</td></tr><tr><td>Internet order</td><td>Whether it is internet order or not – True or False</td><td>Categorical, ordinal</td></tr><tr><td>Product Category</td><td>Product category in the order</td><td>Nominal</td></tr><tr><td>Quantity</td><td>Quantity of a particular product</td><td>Numeric Value</td></tr><tr><td>Revenue</td><td>Revenue obtained from the product category for the order</td><td>Numeric Value</td></tr><tr><td>Year</td><td>Year of the order – 2007 to 2012</td><td>Nominal</td></tr></table>
a) Steps to visualize the percentage of revenue year-wise as a portion of the total revenue earned over the five years. (6)
b) Steps to visualize the total quantity ordered for each product category. Also change the colour of the bars in the bar graph to green. (6)
c) Mention the entire step as to how to set up the conditional formatting and highlight the product categories which have more than 100,000 quantities sold over the years as green but anything less than 10,000 sold is highlighted as red. (8)
Please note: For a and b, What kind of a graph you should choose is also to be determined by you and it contains marks.</td><td>20</td><td>CO4</td></tr>
</table>