

Question 1

A data analyst uses Excel and Tableau to identify patterns and correlations in a dataset to draw conclusions. Which phase of the analytic process is the analyst currently working on?

presenting the results

preparing the data

investigating the data



analyzing the data

Question 2

Match the task and purpose to the appropriate Big Data analytics method.

Categories:

analyzing customer subscription cancellations and correlating them with customer comments and ratings to determine the most common reasons why users cancel subscriptions

A

analyzing customer ordering history to find popular products within last 6 months

B

analyzing historical data of temperature to prepare latest weather forecast

C

reviewing sales data of top 5 competitors to determine the marketing strategy of own product for next quarter

D

Options:



diagnostic analysis



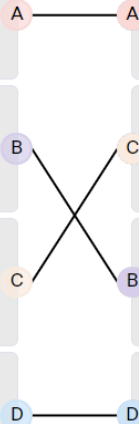
predictive analytics



descriptive analytics



prescriptive analytics



Question 3

Which statement is true concerning data analytics?

Prescriptive analytics uses historical data to provide regular reports on events that have already happened.

Descriptive analytics uses a feedback system to track the outcome of actions taken.



Predictive analytics uses simulation models and forecasting to suggest what could happen.

Descriptive analytics attempts to predict what will happen next.

Question 4

Match the type of data analysis with the question it can answer.

Categories:

Predictive

Prescriptive

Diagnostic

Descriptive

A

B

C

D

Options:



What might happen in the future?



Why did something happen?



What should be done next?



What happened?

Question 5

Which Microsoft Excel formula would correctly multiply a value in cell A1 with a value in cell A19?

=MULTIPLY(A1:A19)

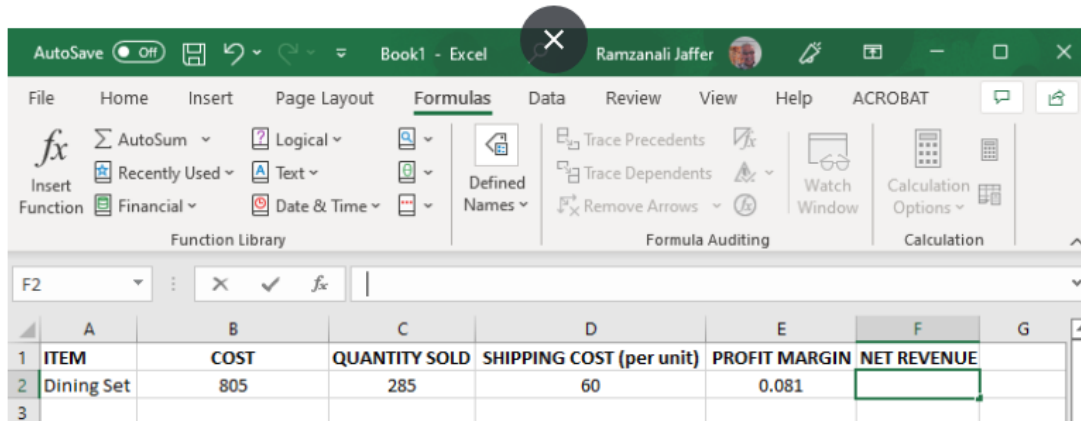


=A1*A19

=x A1:A19

=A1x A19

Question 6



The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Formulas'. The formula bar shows a formula being entered in cell F2. The spreadsheet data is as follows:

	A	B	C	D	E	F	G
1	ITEM	COST	QUANTITY SOLD	SHIPPING COST (per unit)	PROFIT MARGIN	NET REVENUE	
2	Dining Set	805	285	60	0.081		
3							

Refer to the exhibit. A data analyst wants to create a formula in Microsoft Excel that will automatically calculate the net revenue from the sales of dining sets. Which formula inserted in cell F2 will accomplish this?



`=C2*(B2-D2)*E2`

`=B2*C2-D2*E2`

`=B2-D2+C2*E2`

`=(B2*C2)-(D2*C2)*E2`

Question 7

What are three resources provided by the Kaggle web site? (Choose three.)



community competitions

API for major social media networks



publicly available datasets



code

data analysis integration with Excel

data analysis integration with Tableau

Question 8

What are two types of continuous variables? (Choose two.)

ordinal



ratio

discrete

nominal



interval

Question 9

Which two tasks are performed as part of the transform step of the ETL data process? (Choose two.)



standardizing formats such as date, time, and location

communicating results to decision-makers

collecting and recording the data



removing the blank records

transferring the data into a database

Question 10

An analyst in an online order company is researching products that online customers spend the most time browsing for on the website but do not buy. The result of the analysis will enable the company to release quick discount notices on the website to encourage customers to buy those products. What relevant data is required to do the research?



streaming customer-viewing data

customer order history

brands of products

inventory of products

Question 11

Which statement is true about the ELT and ETL process?

In the ELT process, the load step occurs after the transform step.

In the ETL process, the load step occurs before the transform step.

In the ETL process, the extract step occurs after the load step.



In the ELT process, the transform step occurs on the stored data as it is used.

Question 12

Match the ETL process with its description.

Categories:

transform

load

extract

Options:

✓

 is the process of gathering the data to be analyzed

✓

 is the process of transferring data to a database

✓

 is the process of conditioning data into a usable form, such as removing duplicate data, providing missing data, and correcting any errors in the data.

```
graph LR; T[transform] --- O3[is the process of conditioning data...]; L[load] --- O2[is the process of transferring data to a database]; E[extract] --- O1[is the process of gathering the data to be analyzed];
```

Question 13

Which option in Microsoft Excel represents a formula with an absolute reference?

=(B\$=(C\$1\$-F\$1\$)

=(A1-\$B1)

=(B\$1-B\$3)



=(A\$1-\$B\$1))

Question 14

Match the Microsoft Excel function with its description.

Categories:		Options:
LEN	A	B it is used to return the number of blank or non-blank cells in a range ✓
COUNTA	B	A it is used to return the number of characters in a cell ✓
CONCATENATE	C	F it is used to convert text to all capital letters (UPPERCASE) ✓
SUMIF	D	E it is used to return a value or the reference to a value from within a table or range ✓
INDEX	E	D it is used to create a sum of values based on multiple criteria ✓
UPPER	F	C it is used to combine the data from one or more columns into a single column ✓

Question 15

How can data analysts use the Conditional Formatting tool in Excel to aid in data analysis of bike sales records?

<input type="checkbox"/>	to separate the contents of the product description column into separate columns
<input type="checkbox"/>	to list sales order numbers from lowest to highest
<input checked="" type="checkbox"/>	to highlight values in the revenue column that are above average
<input type="checkbox"/>	to show only sales from the United States

Question 16

A data analyst needs to organize sales data for analysis. Which Excel function can order the data by date sold, with the most recent sales listed first?



Sort & Filter

Find & Select

Conditional Formatting

Text to Columns

Question 17

Which action is taken during the data investigation step of the data analysis lifecycle?

Determine if there is enough data to complete the analysis.

Transform the data into a format appropriate for the analysis methods and tools.



Determine if the data is complete and relevant for the analysis.

Identify patterns, relationships, and correlations in the data.

Question 18

A learner is analyzing a large volume of data in a Microsoft Excel spreadsheet and wishes to find duplicate data values. The data has been organized in a table where each row has different but related forms of data in each column. Which Microsoft Excel function can be used to do this?

SUM

ISNA



VLOOKUP

IF

Question 19

What is a statistical analysis result that a descriptive statistical analysis will not provide?



predictions made about other data sets that are not in the population

calculation of the smallest or largest value

calculation of how widely dispersed the data is

visualization of particular data trends occurring

Question 20

What is an outlier in data analytics?

it is a point in inferential data that is numeric

it is a point in descriptive data that is historical

it is a pointer to the sample data set



it is a value or data point that varies significantly from others in the data set

Question 21

Match the type of database to the best description.

Categories:

relational database

spreadsheet

flat file database

A

B

C

Options:



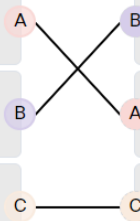
it stores data in a sheet using a tabular format of columns, rows, and cells



it requires a SQL to store, manipulate, and retrieve data



it stores records in a single file with no hierarchical structure



User
user_id
content_id
user_name
user_email
user_dob

Refer to the exhibit. Which option will correctly display the user names and email addresses of users in this table?

SELECT User
FROM user_id, user_name, user_email

SELECT user_id, user_name, user_email
WHERE User



SELECT user_id, user_name, user_email
FROM User

SELECT User
WHERE user_id, user_name, user_email

Question 23

Drag the SQL Data Manipulation Language (DML) statements to the correct description.

Categories:

ORDER BY

WHERE

FROM

SELECT

A

B

C

D

Options:



is a conditional statement used to filter your returned results



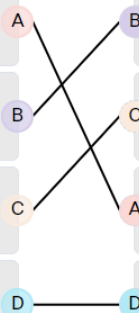
returns the selection from the table name specified



used to structure the resulting query in ascending (ASC) order by default



returns the fields specified



Question 24

What is referenced by the FROM statement in an SQL query?

the column where the data is located

the row where the data is located



the table containing the data

the database to be queried

Question 25

What does a field refer to in database schema terminology?

a unique record within the database

a collection of attributes of an individual item

a collection data in rows and columns



a specific attribute of an individual record

Question 26

What is the purpose of the DML language commands in SQL?

to create, modify or delete objects in the database schema

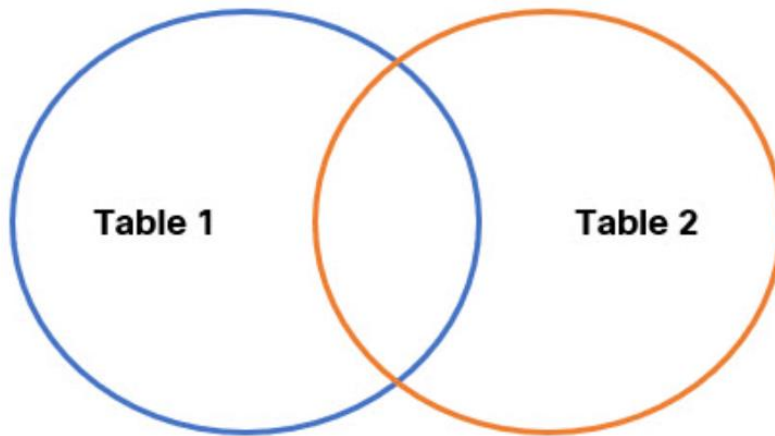
to manage updates to the database storage

to change rights and permissions to data in the database



to select, update, insert or delete records and values from a table

Question 27



Refer to the exhibit. Match the JOIN operation with the proper description.

Categories:

LEFT JOIN
FULL JOIN
RIGHT JOIN
INNER JOIN

Options:

<input checked="" type="checkbox"/>	Returns only records where the selected fields have matching values in Table 1 and Table 2.
<input checked="" type="checkbox"/>	Returns all records from Table 2 and the matched records from Table 1.
<input checked="" type="checkbox"/>	Returns all records from Table 1 and only the matched records from Table 2.
<input checked="" type="checkbox"/>	Returns all records from Table 1 and Table 2.

Movie table:
MovieID – primary key
Title – name of the movie
Date – date the movie was released
Cost – total cost of production
Review table:
ID – Review ID – primary key
Movie Title – name of the movie
Author Name – name of the reviewer
Comment – review text
Date – date review is posted
Score – number of stars awarded

Refer to the exhibit. A data analyst writes a SQL query to extract information from multiple tables in the Movies database. To complete the ON command, what needs to be entered in place of the question mark in this query?

```
SELECT m.Title, m.Date, r.AuthorName, r.Score, r.Comment, r.Date
FROM Review AS r
JOIN Movie AS m
ON m.Title = ?;
```

r.Date;



r.MovieTitle;

r.Score;

r.ID;

r.Comment;

Question 29

What Tableau Public feature is used to consolidate multiple views and visualizations?

screens



dashboards

worksheets

objects

Question 30

Why would you use Tableau in your data analysis?

To display data in a tabular format.



To create visualizations.

To perform mathematical calculations using formulas and built-in functions.

To communicate with and query databases.

Question 31

Which two are benefits of organizing information from multiple datasets into a dashboard? (Choose two)



The information is easily analyzed.

The information is easily secured.



The information is easily visualized.

The information is easily tracked.

Question 32

You are preparing a presentation that needs a visualization showing the relative levels of coffee production in the coffee regions of the world. Which type of visualization in Tableau is well suited for displaying this type of data?

bar chart



bubble map

area graph

heat map

Question 33

Which statement is true in respect of avoiding bias in data analytics?

use a small sample set that supports your hypothesis

don't allow the data to confirm your conclusions

focus on the larger patterns and trends and include outliers in the overall analysis



avoid selecting only data and methods that support your assumptions

Question 34

What assurance does data integrity provide?

Data is recoverable.



Data is unaltered in transit.

Data access is restricted.

Data is available when requested.

Question 35

Which type of bias in data analysis can be caused by the influence of outliers?

confirmation



information

selection

interpretation

Question 36

What are two methods that ensure confidentiality? (Choose two.)

availability

authorization

nonrepudiation



encryption

integrity



authentication

Question 37

Which three features should your data project portfolio contain to ensure that it is considered favorably by employers and recruiters? (Choose three.)

there are links to software, so non-technical reviewers can replicate your analysis



it is organized and easy to navigate



there are links to all your projects on the homepage

it uses a layout and presentation that showcases your web design and markup skills



each project has a clear outcome or conclusion

it contains many multi-colored charts and graphics that show the results of the data analysis

Question 38

Why is the Jupyter Notebook tool useful when developing and testing data analysis software?

markdown text explains the operation of SQL queries and displays the results

when the Python code is compiled, each step is displayed



program execution output is displayed dynamically in real-time

program output is displayed in the standalone Jupyter Notebook application

Question 39

Which data analytic tool can create interactive documents containing executable program code and markdown text?

Kaggle



Jupyter Notebooks

Tableau

Excel

Question 40

What are the three functions of programming languages optimized for data analysis? (Choose three)



data cleaning

data communications across networks

fast response time to user input



statistical analysis

virtualization of processing hardware



data visualization