***Plan, implement, and manage a solution for data analytics (10–15%)***

Plan a data analytics environment

* Identify requirements for a solution, including components, features, performance, and capacity stock-keeping units (SKUs)
* Recommend settings in the Fabric admin portal
* Choose a data gateway type
* Create a custom Power BI report theme

Implement and manage a data analytics environment

* Implement workspace and item-level access controls for Fabric items
* Implement data sharing for workspaces, warehouses, and lakehouses
* Manage sensitivity labels in semantic models and lakehouses
* Configure Fabric-enabled workspace settings
* Manage Fabric capacity

Manage the analytics development lifecycle

* Implement version control for a workspace
* Create and manage a Power BI Desktop project (.pbip)
* Plan and implement deployment solutions
* Perform impact analysis of downstream dependencies from lakehouses, data warehouses, dataflows, and semantic models
* Deploy and manage semantic models by using the XMLA endpoint
* Create and update reusable assets, including Power BI template (.pbit) files, Power BI data source (.pbids) files, and shared semantic models

***Prepare and serve data (40–45%)***

Create objects in a lakehouse or warehouse

* Ingest data by using a data pipeline, dataflow, or notebook
* [Create and ma](https://data-mozart.com/mastering-dp-600-exam-create-and-manage-onelake-shortcuts/)[nage shortcuts](https://data-mozart.com/mastering-dp-600-exam-create-and-manage-onelake-shortcuts/)
* Implement file partitioning for analytics workloads in a lakehouse
* Create views, functions, and stored procedures
* Enrich data by adding new columns or tables

Copy data

* Choose an appropriate method for copying data from a Fabric data source to a lakehouse or warehouse
* Copy data by using a data pipeline, dataflow, or notebook
* Add stored procedures, notebooks, and dataflows to a data pipeline
* Schedule data pipelines
* Schedule dataflows and notebooks

Transform data

* Implement a data cleansing process
* Implement a star schema for a lakehouse or warehouse, including Type 1 and Type 2 slowly changing dimensions
* Implement bridge tables for a lakehouse or a warehouse
* [Denormaliz](https://data-mozart.com/mastering-dp-500-dp-600-exam-optimize-data-model-by-using-denormalization/)[e data](https://data-mozart.com/mastering-dp-500-dp-600-exam-optimize-data-model-by-using-denormalization/) \*\*\* [W](https://youtu.be/bV8RixJgKfE)[atch on YouTube](https://youtu.be/bV8RixJgKfE)
* Aggregate or de-aggregate data
* Merge or join data
* Identify and resolve duplicate data, missing data, or null values
* Convert data types by using SQL or PySpark
* Filter data

Optimize performance

* Identify and resolve data loading performance bottlenecks in dataflows
* Identify and resolve data loading performance bottlenecks in notebooks
* Identify and resolve data loading performance bottlenecks in SQL queries
* Implement performance improvements in dataflows, notebooks, and SQL queries
* Identify and resolve issues with Delta table file sizes

***Implement and manage semantic models (20-25%)***

Design and build semantic models

* [Choose a storage mode, including Direct Lake](https://data-mozart.com/what-do-allen-iverson-and-direct-lake-have-in-common/) \*\*\* [Watch o](https://youtu.be/S-YwZAQ7nIM)[n YouTube](https://youtu.be/S-YwZAQ7nIM)
* Identify use cases for DAX Studio and Tabular Editor 2
* [Implement a star schema for a semantic model](https://data-mozart.com/mastering-dp-600-exam-implement-a-star-schema-for-power-bi-semantic-model/) \*\*\* [Watch on YouTube](https://youtu.be/gMmgI84GIY0)
* Implement relationships, such as bridge tables and many-to-many relationships
* Write calculations that use DAX variables and functions, such as iterators, table filtering, windowing, and information functions
* Implement calculation groups
* Implement dynamic strings
* Implement field parameters
* [Design and build a large](https://data-mozart.com/mastering-dp-500-dp-600-exam-design-and-build-a-large-format-dataset/)[format dataset](https://data-mozart.com/mastering-dp-500-dp-600-exam-design-and-build-a-large-format-dataset/) \*\*\* [W](https://youtu.be/4gjoMA1TaWI)[atch on YouTube](https://youtu.be/4gjoMA1TaWI)
* [Design and build composite](https://data-mozart.com/mastering-dp-500-dp-600-exam-design-and-build-composite-models/)[models](https://data-mozart.com/mastering-dp-500-dp-600-exam-design-and-build-composite-models/)
* [Aggregations in Power BI](https://data-mozart.com/mastering-dp-500-dp-600-exam-aggregations-in-power-bi/) \*\*\* [Watc](https://youtu.be/aP2Jfdkn7BY)[h on YouTube](https://youtu.be/aP2Jfdkn7BY)
* Implement and validate dynamic row-level security
* [Implement and validate object-level security](https://data-mozart.com/mastering-dp-600-exam-implement-and-validate-object-level-security/) \*\*\* [W](https://youtu.be/YDmWPWhPa3g)[atch on YouTube](https://youtu.be/YDmWPWhPa3g)

Optimize enterprise-scale semantic models

* [Implement performance improvements in queries and report visuals](https://data-mozart.com/mastering-dp-500-dp-600-exam-performance-improvements-in-queries-and-visuals/)
* [Improve DAX performance by using DAX Studio](https://data-mozart.com/mastering-dp-500-dp-600-exam-troubleshoot-dax-performance-with-dax-studio/)
* [Optimize a semantic model by using](https://data-mozart.com/mastering-dp-500-dp-600-exam-optimize-data-model-with-tabular-editor/)[Tabular Editor 2](https://data-mozart.com/mastering-dp-500-dp-600-exam-optimize-data-model-with-tabular-editor/)
* [Implement incremental refr](https://data-mozart.com/mastering-dp-500-dp-600-exam-implementing-incremental-refresh/)[esh](https://data-mozart.com/mastering-dp-500-dp-600-exam-implementing-incremental-refresh/) \*\*\* [Watch on YouTu](https://youtu.be/uvZVvNIwzTE)[be](https://youtu.be/uvZVvNIwzTE)

***Explore and analyze data (20-25%)***

Perform exploratory analytics

* Implement descriptive and diagnostic analytics
* Integrate prescriptive and predictive analytics into a visual or report
* Profile data

Query data by using SQL

* Query a lakehouse in Fabric by using SQL queries or the visual query editor
* Query a warehouse in Fabric by using SQL queries or the visual query editor
* Connect to and query datasets by using the XMLA endpoint

https://learn.microsoft.com/en-us/credentials/certifications/exams/dp-600/

Questions -  
  
  
Q1. You have a Microsoft Power BI report that contains two metrics starting balance and ending balance. Additional fields are used for customer segment.  
  
You need to recommend a visual to display the value breakdown between the starting and ending balance of each customer segment. Which visual should you recommend?  
  
A. ribbon chart  
B. stacked bar chart  
C. Treemap  
D. waterfall chart  
  
  
  
Q2. You plan to analyze lakehouse data by using a Fabric notebook and PySpark. You load the data to a DataFrame by running the following code. df = spark.sql(“SELECT \* FROM Product”)  
You need to display top 100 rows from the dataframe. Which PySpark command should you run?  
  
A. describe(df.limit(100))  
B. df.describe(100)  
C. df.printSchema(100)  
D. display(df.limit(100))  
  
  
  
Q3. You are developing a Microsoft Power BI semantic model. Two tables in the data model are not connected in a physical relationship. You need to establish a virtual relationship between the tables. Which DAX function should you use?  
  
A. CROSSFILTER()  
B. PATH()  
C. TRETAS()  
D. USERELATIONSHIP()  
  
  
  
Q4. You are developing a complex semantic model that contains more than 20 date columns. You need to conform the date format for all the date columns as quickly as possible. What should you use?  
  
A. ALM toolkit  
B. DAX Studio  
C. Tabular Editor  
D. Vertipaq Analzer  
  
  
Q5. You have Fabric Workspace that contains a complex sematic model for a Power BI report. You need to optimize the semantic model for analytical queries and use denormalization to reduce the model complexity and the number of joins between tables. Which tables should you denormalize?  
  
A. dimension tables on the same level of granularity  
B. fact tables on same level of granularity  
C. role-playing dimension tables  
D. Snowflake dimension tables  
  
  
  
Q6. You have a Fabric tenant that contains a lakehouse. You plan to use Fabric notebook and PySpark to read sales data and save the data as a Delta tables named Sales. The table must be partitioned by Sales Year and Quarter. You load the sales data to a dataframe named df that contains a Year column and a Quarter column. Which command should you run next?  
  
A. df.write.mode(“overwrite”).format(“delta”).partitionBy(“Year,Quarter”).save(“Files/Sales”)  
B. df.write.mode(“overwrite”).format(“delta”).partitionBy(“Year,Quarter”).save(“Table/Sales”)  
C. df.write.mode(“overwrite”).format(“parquet”).partitionBy(“Year,Quarter”).save(“Files/Sales”)  
D. df.write.mode(“overwrite”).format(“parquet”).partitionBy(“Year,Quarter”).save(“Tables/Sales”)  
  
  
  
  
Q7. You are managing a set of dataflow Gen2 queries that are currently ingesting tables into Fabic lakehouse. You need to ensure that the tables are optimize for direct lake connections that will be used by connected semantic models. What should you do?  
  
A. Apply incremental refresh policy to the sematic model refreshes.  
B. Run the VACUUM command.  
C. Use OPTIMZE to apply V-order.  
D. Use the shortcuts to the lakehouse tables from the semantic models.

Q8. Your Fabric-enbaled workspace has several items, including a lakehouse, dataflows, notebooks, and Power BI reports. If you add a user to the workspace role “Contributor”, what will the user be able to do?  
  
A. Create and edit items for not manage workspace.  
B. Manage all items with full admin rights  
C. View all items, but not create or edit items.   
  
  
  
Q. What is the difference between a 1. delta table created using shortcut and 2. a managed table? Can SQL endpoint read tables created using both the options?  
  
  
You need to identify the rows of a DataFrame named df\_customers in which any of the columns (axis 1 of the DataFrame) are NULL.

Q. You have a Fabric warehouse.

You are writing a T-SQL statement to retrieve data from a table named Sales to display the highest sales amount for specific customers.

SELECT CustomerKey

, SalesAmount

, \*\*<target1>\*\* OVER(ORDER BY SalesAmount DESC) AS Ranking

FROM dbo.Sales

WHERE CustomerKey IN (1, 2, 3)

You need to ensure that after ties for SalesAmount, the next Sales amount increments the Ranking value by one.

The following is an example of the expected result.

CustomerKey|SalesAmount|Ranking

1|100|1

2|100|1

1|80|2

Which function should you use for <target1> in the T-SQL statement?

Select only one answer.

DENSE\_RANK()

NTILE()

RANK()

ROW\_NUMBER()

Q. You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a data pipeline named Pipeline1 that runs in the US-West Azure region. Workspace1 also contains a semantic model named SemanticModel1 and a warehouse named Warehouse1.

You need to ensure that Pipeline1 runs at midnight (12:00 AM), and that the schedule is set to the UTC-0 time zone.

How should you configure the schedule for Pipeline1?

Select only one answer.

Add a data pipeline notebook activity to convert the US West time zone to UTC-0.

For Pipeline1, set the scheduler time zone to UTC-0.

For SemanticModel1, set the schedule time zone to UTC-0.

For Warehouse1, set the scheduler time zone to UTC-0.

Next   
  
  
  
Q. You have a Fabric workspace named Workspace1.

You plan to create a data pipeline to ingest data into Workspace1.

You need to ensure that the pipeline activity supports parameterization.

Which two activities support parameterization in the data pipeline UI? Each correct answer presents part of the solution.

Select all answers that apply.

Dataflow Gen2

KQL activity

notebooks

SQL stored procedures

user-defined functions   
  
  
  
Q. You have a Fabric tenant that contains a lakehouse named Lakehouse1.

A notebook named Notebook1 is used to ingest and transform data from an external data source before loading the data to Lakehouse1.

You need to ensure that the process meets the following requirements:

* Runs daily at 7:00 AM.
* Attempts to rerun the process two more times if a source file is unavailable.

The solution must minimize development effort.

What should you do?

Select only one answer.

Create a pipeline. Add Notebook1 to the pipeline. Schedule the pipeline to run daily.

Create a pipeline. Add Notebook1 to the pipeline three times. Connect the activities so that the second and third notebook activities run if the first activity fails. Schedule the pipeline to run daily.

Create a pipeline. Add Notebook1 to the pipeline three times. Connect the activities so that the second notebook activity runs if the first activity fails and the third activity runs if the second activity fails. Schedule the pipeline to run daily.

From the Notebook schedule settings, configure daily runs and retry settings.

Q. You have a Fabric tenant that contains a lakehouse named Lakehouse1.

A SELECT query from a managed Delta table in Lakehouse1 takes longer than expected to complete. The table receives new records daily and must keep change history for seven days.

You notice that the table contains 1,000 Parquet files that are each 1 MB.

You need to improve query performance and reduce storage costs.

What should you do from Lakehouse explorer?

Select only one answer.

Manually delete any files that have a creation date that is older than seven days.

Select **Maintenance** and run the OPTIMIZE command.

Select **Maintenance** and run the OPTIMIZE command as well as the VACUUM command with a retention policy of seven days.

Select **Maintenance** and run the VACUUM command with a retention policy of seven days.