[Testlet 1 Case study Contoso, Ltd. - Format Texte 6](#_Toc199876388)

[Testlet 2 Case study Litware, Inc. - Format Texte 9](#_Toc199876389)

[QUESTIONS 14](#_Toc199876390)

[Question 1 14](#_Toc199876391)

[REPONSE 14](#_Toc199876392)

[Question 2 14](#_Toc199876393)

[REPONSE 15](#_Toc199876394)

[Question 3 16](#_Toc199876395)

[Question 4 16](#_Toc199876396)

[Question 5 17](#_Toc199876397)

[Question 6 18](#_Toc199876398)

[Question 7 18](#_Toc199876399)

[Question 8 Multiple 18](#_Toc199876400)

[Question 9 Multiple 19](#_Toc199876401)

[Question 10 Multiple QMA10 20](#_Toc199876402)

[Question 11 Multiple 20](#_Toc199876403)

[Question 12 Multiple 21](#_Toc199876404)

[Question 13 21](#_Toc199876405)

[Question 14 22](#_Toc199876406)

[Question 15 22](#_Toc199876407)

[Question 16 23](#_Toc199876408)

[Question 17 Multiple 23](#_Toc199876409)

[Question 18 24](#_Toc199876410)

[Question 19 24](#_Toc199876411)

[Question 20 25](#_Toc199876412)

[Question 21 27](#_Toc199876413)

[Question 22 Multiple QMA 27](#_Toc199876414)

[Question 23 28](#_Toc199876415)

[Question 24 28](#_Toc199876416)

[Question 25 29](#_Toc199876417)

[Question 26 29](#_Toc199876418)

[Question 27 31](#_Toc199876419)

[Question 28 31](#_Toc199876420)

[Question 29 32](#_Toc199876421)

[Question 30 32](#_Toc199876422)

[Question 31 33](#_Toc199876423)

[Question 32 33](#_Toc199876424)

[Question 33 33](#_Toc199876425)

[Question 34 Multiple QM34 34](#_Toc199876426)

[Question 35 34](#_Toc199876427)

[Question 36 Multiple QM34 35](#_Toc199876428)

[Question 37 Multiple QM34 35](#_Toc199876429)

[Question 38 35](#_Toc199876430)

[Question 39 36](#_Toc199876431)

[Question 40 37](#_Toc199876432)

[Question 41 Multiple QM34 39](#_Toc199876433)

[Question 42 39](#_Toc199876434)

[Question 44 41](#_Toc199876435)

[Question 45 41](#_Toc199876436)

[Question 46 43](#_Toc199876437)

[Question 47 43](#_Toc199876438)

[Question 48 44](#_Toc199876439)

[Question 49 44](#_Toc199876440)

[Question 50 45](#_Toc199876441)

[Question 51 46](#_Toc199876442)

[Question 52 46](#_Toc199876443)

[Question 53 46](#_Toc199876444)

[Question 54 47](#_Toc199876445)

[Question 55 48](#_Toc199876446)

[Question 56 49](#_Toc199876447)

[Question 57 50](#_Toc199876448)

[Question 58 50](#_Toc199876449)

[Question 59 51](#_Toc199876450)

[Question 60 51](#_Toc199876451)

[Question 61 52](#_Toc199876452)

[Question 62 52](#_Toc199876453)

[Question 63 53](#_Toc199876454)

[Question 64 53](#_Toc199876455)

[Question 65 54](#_Toc199876456)

[Question 66 56](#_Toc199876457)

[Question 67 57](#_Toc199876458)

[Question 68 58](#_Toc199876459)

[Question 69 58](#_Toc199876460)

[Question 70 58](#_Toc199876461)

[Question 71 60](#_Toc199876462)

[Question 72 61](#_Toc199876463)

[Question 73 61](#_Toc199876464)

[Question 74 62](#_Toc199876465)

[Question 75 62](#_Toc199876466)

[Question 76 63](#_Toc199876467)

[Question 77 63](#_Toc199876468)

[Question 78 64](#_Toc199876469)

[Question 79 66](#_Toc199876470)

[Question 80 66](#_Toc199876471)

[Question 81 67](#_Toc199876472)

[Question 82 67](#_Toc199876473)

[Question 83 68](#_Toc199876474)

[Question 84 69](#_Toc199876475)

[Question 85 69](#_Toc199876476)

[Question 86 72](#_Toc199876477)

[Question 87 73](#_Toc199876478)

[Question 88 74](#_Toc199876479)

[Question 89 76](#_Toc199876480)

[Question 90 77](#_Toc199876481)

[Question 91 78](#_Toc199876482)

[Question 92 80](#_Toc199876483)

[Question 93 81](#_Toc199876484)

[Question 94 81](#_Toc199876485)

[Question 95 82](#_Toc199876486)

[Question 96 83](#_Toc199876487)

[Question 97 83](#_Toc199876488)

[Question 98 83](#_Toc199876489)

[Question 99 84](#_Toc199876490)

[Question 100 84](#_Toc199876491)

[Question 101 87](#_Toc199876492)

[Question 102 87](#_Toc199876493)

[Question 103 88](#_Toc199876494)

[REPONSE 88](#_Toc199876495)

[Question 104 88](#_Toc199876496)

[REPONSE 89](#_Toc199876497)

[Question 105 89](#_Toc199876498)

[REPONSE 89](#_Toc199876499)

[Question 106 90](#_Toc199876500)

[REPONSE 91](#_Toc199876501)

[Question 107 93](#_Toc199876502)

[REPONSE 94](#_Toc199876503)

[Question 108 94](#_Toc199876504)

[REPONSE 94](#_Toc199876505)

[Question 109 95](#_Toc199876506)

[REPONSE 97](#_Toc199876507)

[Question 110 97](#_Toc199876508)

[REPONSE 98](#_Toc199876509)

[Question 111 100](#_Toc199876510)

[REPONSE 100](#_Toc199876511)

[Question 112 101](#_Toc199876512)

[REPONSE 101](#_Toc199876513)

[Question 113 101](#_Toc199876514)

[REPONSE 102](#_Toc199876515)

[Question 114 102](#_Toc199876516)

[Which syntax should you use in a notebook to access the Research division data for Productline1? 103](#_Toc199876517)

[A.spark.read.format(“delta”).load(“Tables/productline1/ResearchProduct”) 103](#_Toc199876518)

[B.spark.sql(“SELECT \* FROM Lakehouse1.ResearchProduct ”) 103](#_Toc199876519)

[C.external\_table(‘Tables/ResearchProduct) 103](#_Toc199876520)

[D.external\_table(ResearchProduct) 103](#_Toc199876521)

[REPONSE 103](#_Toc199876522)

[Question 115 103](#_Toc199876523)

[REPONSE 104](#_Toc199876524)

[Page 24 105](#_Toc199876525)

[Question 116 105](#_Toc199876526)

[REPONSE 105](#_Toc199876527)

[Question 117 106](#_Toc199876528)

[REPONSE 106](#_Toc199876529)

[Question 118 107](#_Toc199876530)

[REPONSE 107](#_Toc199876531)

[Question 119 107](#_Toc199876532)

[REPONSE 108](#_Toc199876533)

[Question 120 108](#_Toc199876534)

[REPONSE 109](#_Toc199876535)

[Question 121 110](#_Toc199876536)

[REPONSE 110](#_Toc199876537)

[ANNEXES 111](#_Toc199876538)

[1. 📂 Accéder aux données dans Lakehouse, Warehouse, Eventhouse 111](#_Toc199876539)

[2. 📚 Définitions et cas d’usage 111](#_Toc199876540)

[3. ⚖️ Tableau comparatif des modèles de données dans Fabric 111](#_Toc199876541)

[4. 🧩 Résumé rapide – Quel outil pour quelle tâche ? 112](#_Toc199876542)

[BASE DIAGRAMME 112](#_Toc199876543)

[Query Folding indicator 114](#_Toc199876544)

|  |
| --- |
| Testlet 1 Case study Contoso, Ltd. - Format Texte  **Case study Contoso, Ltd**  **Overview  Contoso, Ltd. is a US-based health supplements company. Contoso has two divisions named Sales and Research. The Sales division contains two departments named Online Sales and Retail Sales. The Research division assigns internally developed product lines to individual teams of researchers and analysts.**  **Existing Environment:  Identity Environment:  Contoso has a Microsoft Entra tenant named contoso.com. The tenant contains two groups named ResearchReviewersGroup1 and ResearchReviewersGroup2.**  **Data Environment:  Contoso has the following data environment: The Sales division uses a Microsoft Power BI Premium capacity.  - The semantic model of the Online Sales department includes a fact table named Orders that uses Import made. In the system of origin, the OrderID value represents the sequence in which orders are created. - The Research department uses an on-premises, third-party data warehousing product. - Fabric is enabled for contoso.com. - An Azure Data Lake Storage Gen2 storage account named storage1 contains Research division data for a product line named Productline1. The data is in the delta format. - A Data Lake Storage Gen2 storage account named storage2 contains Research division data for a product line named Productline2. The data is in the CSV format.**  **Requirements:  Planned Changes:  Contoso plans to make the following changes:  - Enable support for Fabric in the Power BI Premium capacity used by the Sales division. - Make all the data for the Sales division and the Research division available in Fabric. - For the Research division, create two Fabric workspaces named Productline1ws and Productline2ws. - In Productline1ws, create a lakehouse named Lakehouse1. - In Lakehouse1, create a shortcut to storage1 named ResearchProduct.**  **Data Analytics Requirements:  Contoso identifies the following data analytics requirements:  - All the workspaces for the Sales division and the Research division must support all Fabric experiences. - The Research division workspaces must use a dedicated, on-demand capacity that has per-minute billing. - The Research division workspaces must be grouped together logically to support OneLake data hub filtering based on the department name. - For the Research division workspaces, the members of ResearchReviewersGroup1 must be able to read lakehouse and warehouse data and shortcuts by using SQL endpoints. - For the Research division workspaces, the members of ResearchReviewersGroup2 must be able to read lakehouse data by using Lakehouse explorer. - All the semantic models and reports for the Research division must use version control that supports branching.**  **Data Preparation Requirements:  Contoso identifies the following data preparation requirements:  - The Research division data for Productline1 must be retrieved from Lakehouse1 by using Fabric notebooks. - All the Research division data in the lakehouses must be presented as managed tables in Lakehouse explorer.**  **Semantic Model Requirements:  Contoso identifies the following requirements for implementing and managing semantic models:  - The number of rows added to the Orders table during refreshes must be minimized. - The semantic models in the Research division workspaces must use Direct Lake mode.**  **General Requirements:  Contoso identifies the following high-level requirements that must be considered for all solutions:  - Follow the principle of least privilege when applicable. - Minimize implementation and maintenance effort when possible.** |

|  |
| --- |
| **🔷 Architecture actuelle et changements prévus**   * **Sales Division** utilise déjà **Power BI Premium**. * **Fabric est activé** sur le tenant Entra contoso.com. * Les données Research sont dans deux **Data Lake Storage Gen2** :   + storage1 (delta format → Productline1)   + storage2 (CSV format → Productline2) * Deux workspaces Fabric prévus : Productline1ws et Productline2ws.   **🔷 Données & Stockage**   * Dans Productline1ws :   + Création d’un **Lakehouse1**   + Création d’un **shortcut** vers storage1 nommé ResearchProduct. * **Données Research** à présenter en **managed tables** dans Lakehouse explorer. * Accès aux données via **SQL endpoints** (Group1) ou **Lakehouse Explorer** (Group2).   **🔷 Capacités Fabric**   * **Sales** doit activer Fabric sur sa capacité Power BI Premium. * **Research** doit utiliser une **capacité Fabric dédiée, à facturation à la minute** (≠ Power BI Premium classique). * Les workspaces Research doivent permettre un **filtrage logique par département** dans **OneLake Data Hub**.   **🔷 Sécurité et accès**   * **Group1** : Lecture des données Lakehouse et Warehouse **via SQL endpoints**. * **Group2** : Lecture des données Lakehouse **via Lakehouse Explorer**. * Suivre le **principe du moindre privilège**.   **🔷 Modèles sémantiques**   * Sales :   + Orders en Import → limiter les lignes ajoutées à chaque refresh (→ **incrémental** conseillé). * Research :   + Tous les modèles doivent être en **Direct Lake**.   + Gestion via **version control avec support du branching** (→ Git integration).   **🔷 Préparation des données**   * Les données de **Productline1** doivent être traitées via **Fabric notebooks**. * Toutes les données dans les lakehouses doivent être exposées comme **managed tables**. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Testlet 2 Case study Litware, Inc. - Format Texte **study Litware, Inc. Overview  Litware, Inc. is a manufacturing company that has offices throughout North America. The analytics team at Litware contains data engineers, analytics engineers, data analysts, and data scientists.**  **Existing Environment:  Fabric Environment:  Litware has been using a Microsoft Power BI tenant for three years. Litware has NOT enabled any Fabric capacities and features.**  **Available Data:  Litware has data that must be analyzed as shown in the following table.**  **Image 1: Table of Data Sources**   |  |  |  | | --- | --- | --- | | **Description** | **Original source** | **Total size** | | Customer data | Customer relationship management (CRM) system | 50 MB | | Product data | Customer relationship management (CRM) system | 200 MB | | Customer satisfaction surveys | SurveyMonkey | 500 GB |   **The Product data contains a single table and the following columns.**  **Image 2: Table of Product Data Columns**   |  |  | | --- | --- | | **Name** | **Data type** | | ProductID | Integer | | ProductName | String | | ProductCategory | String | | ListPrice | Decimal |   **The customer satisfaction data contains the following tables:  - Survey - Question - Response**  **For each survey submitted, the following occurs:  - One row is added to the Survey table. - One row is added to the Response table for each question in the survey.**  **The Question table contains the text of each survey question. The third question in each survey response is an overall satisfaction score. Customers can submit a survey after each purchase.**  **User Problems:  The analytics team has large volumes of data, some of which is semi-structured. The team wants to use Fabric to create a new data store.**  **Product data is often classified into three pricing groups: high, medium, and low. This logic is implemented in several databases and semantic models, but the logic does NOT always match across implementations.**  **Requirements:  Planned Changes:  Litware plans to enable Fabric features in the existing tenant. The analytics team will create a new data store as a proof of concept (PoC). The remaining Litware users will only get access to the Fabric features once the PoC is complete. The PoC will be completed by using a Fabric trial capacity.**  **The following three workspaces will be created:  - AnalyticsPOC: Will contain the data store, semantic models, reports pipelines, dataflow, and notebooks used to populate the data store - DataEngPOC: Will contain all the pipelines, dataflows, and notebooks used to populate OneLake - DataSciPOC: Will contain all the notebooks and reports created by the data scientists**  **The following will be created in the AnalyticsPOC workspace:  - A data store (type to be decided) - A custom semantic model - A default semantic model - Interactive reports**  **The data engineers will create data pipelines to load data to OneLake either hourly or daily depending on the data source. The analytics engineers will create processes to ingest, transform, and load the data to the data store in the AnalyticsPOC workspace daily. Whenever possible, the data engineers will use low-code tools for data ingestion. The choice of which data cleansing and transformation tools to use will be at the data engineers’ discretion.**  **All the semantic models and reports in the Analytics POC workspace will use the data store as the sole data source.**  **Technical Requirements:  The data store must support the following:  - Read access by using T-SQL or Python - Semi-structured and unstructured data - Row-level security (RLS) for users executing T-SQL queries**  **Files loaded by the data engineers to OneLake will be stored in the Parquet format and will meet Delta Lake specifications.**  **Data will be loaded without transformation in one area of the AnalyticsPOC data store. The data will then be cleansed, merged, and transformed into a dimensional model.**  **The data load process must ensure that the raw and cleansed data is updated completely before populating the dimensional model.**  **The dimensional model must contain a date dimension. There is no existing data source for the date dimension. The Litware fiscal year matches the calendar year. The date dimension must always contain dates from 2010 through the end of the current year.**  **The product pricing group logic must be maintained by the analytics engineers in a single location. The pricing group data must be made available in the data store for T-SQL queries and in the default semantic model.  The following logic must be used:  - List prices that are less than or equal to 50 are in the low pricing group. - List prices that are greater than 50 and less than or equal to 1,000 are in the medium pricing group. - List prices that are greater than 1,000 are in the high pricing group.**  **Security Requirements:  Only Fabric administrators and the analytics team must be able to see the Fabric items created as part of the PoC.**  **Litware identifies the following security requirements for the Fabric items in the AnalyticsPOC workspace:  - Fabric administrators will be the workspace administrators. - The data engineers must be able to read from and write to the data store. No access must be granted to datasets or reports. - The analytics engineers must be able to read from, write to, and create schemas in the data store. They also must be able to create and share semantic models with the data analysts and view and modify all reports in the workspace. - The data scientists must be able to read from the data store, but not write to it. They will access the data by using a Spark notebook - The data analysts must have read access to only the dimensional model objects in the data store. They also must have access to create Power BI reports by using the semantic models created by the analytics engineers. - The date dimension must be available to all users of the data store. - The principle of least privilege must be followed.**  **Both the default and custom semantic models must include only tables or views from the dimensional model in the data store.  Litware already has the following Microsoft Entra security groups:  - FabricAdmins: Fabric administrators - AnalyticsTeam: All the members of the analytics team - DataAnalysts: The data analysts on the analytics team - DataScientists: The data scientists on the analytics team - DataEngineers: The data engineers on the analytics team - AnalyticsEngineers: The analytics engineers on the analytics team**  **Report Requirements:  The data analysts must create a customer satisfaction report that meets the following requirements:  - Enables a user to select a product to filter customer survey responses to only those who have purchased that product. - Displays the average overall satisfaction score of all the surveys submitted during the last 12 months up to a selected date. - Shows data as soon as the data is updated in the data store. - Ensures that the report and the semantic model only contain data from the current and previous year. - Ensures that the report respects any table-level security specified in the source data store. - Minimizes the execution time of report queries.** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **🔷 Vue d'ensemble**   * **Litware, Inc.** : Entreprise manufacturière nord-américaine. * Équipe analytics composée de **data engineers, analytics engineers, data analysts et data scientists**. * Utilise **Power BI** depuis 3 ans, mais **Fabric n’était pas encore activé**.   **🔷 Données disponibles**   | **Type de données** | **Source** | **Taille** | | --- | --- | --- | | Customer data | CRM | 50 MB | | Product data | CRM | 200 MB | | Customer satisfaction surveys | SurveyMonkey | 500 GB |  * **Product data** : une seule table avec 4 colonnes (ProductID, ProductName, ProductCategory, ListPrice). * **Satisfaction surveys** : 3 tables (Survey, Question, Response) — la **3e question** est une **note de satisfaction globale**. * **Fréquence des soumissions** : une enquête par achat.   **🔷 Problèmes actuels**   * Volumes importants de données, certains semi-structurés. * **Logique de classification des prix (high, medium, low)** dupliquée de façon incohérente dans plusieurs systèmes.   **🔷 Changements prévus**   * Activation des **fonctionnalités Fabric** sur le tenant Power BI existant. * Réalisation d’un **PoC** (Proof of Concept) dans une **capacité Fabric d’essai**. * Accès aux fonctionnalités Fabric limité à l’équipe analytics pendant le PoC. * Création de **3 workspaces Fabric** :   + AnalyticsPOC : data store, modèles sémantiques, rapports, pipelines, notebooks.   + DataEngPOC : pipelines, dataflows, notebooks pour chargement vers OneLake.   + DataSciPOC : notebooks et rapports des data scientists.   **🔷 Exigences techniques – Data Store**  Le **data store** du workspace AnalyticsPOC doit :   * Accepter **T-SQL** et **Python** en lecture. * Supporter **données semi-structurées et non structurées**. * Implémenter la **RLS** pour les requêtes T-SQL. * Les fichiers doivent être au format **Parquet**, **conformes à Delta Lake**. * Charger les données brutes sans transformation, puis les transformer vers un **modèle dimensionnel**. * Le processus de chargement doit **finaliser le brut et le cleansed** avant d'alimenter le modèle dimensionnel. * Le **modèle dimensionnel** doit inclure une **dimension date** :   + Pas de source externe pour la date.   + Couvrir **de 2010 à la fin de l’année courante**. * La logique des groupes de prix (low, medium, high) doit être **centralisée** :   + ≤ 50 → low   + 50 & ≤ 1 000 → medium   + 1 000 → high   + Disponible via T-SQL et dans le **modèle sémantique par défaut**.   **🔷 Modèles Sémantiques**   * Création dans AnalyticsPOC :   + Un **modèle sémantique personnalisé**.   + Un **modèle sémantique par défaut**. * Les deux modèles ne doivent utiliser **que les vues ou tables du modèle dimensionnel** du data store. * Les rapports et modèles doivent :   + Réfléter **uniquement les données des années en cours et précédente**.   + **Afficher les données dès qu’elles sont mises à jour** dans le data store.   + **Respecter la sécurité (RLS/table-level)** du data store.   + Être **performants (temps de requête minimisé)**.   **🔷 Sécurité & Accès**   | **Groupe Entra Azure** | **Rôle & Accès (dans AnalyticsPOC)** | | --- | --- | | FabricAdmins | Admins du workspace | | DataEngineers | Lecture/écriture sur data store, **pas d'accès aux rapports ou modèles** | | AnalyticsEngineers | Lecture/écriture + création de schémas dans le data store, création/partage de modèles, accès complet aux rapports | | DataScientists | Lecture **uniquement** sur data store via **notebooks Spark** | | DataAnalysts | Lecture **du modèle dimensionnel uniquement** ; création de rapports à partir des modèles | | AnalyticsTeam | Membres des groupes ci-dessus |  * Le **principe du moindre privilège** doit être suivi. * La **dimension date** doit être accessible à tous les utilisateurs du data store.   **🔷 Rapport Satisfaction Client (créé par les Data Analysts)**  Ce rapport doit :   * **Filtrer** les réponses selon un produit sélectionné. * **Afficher la moyenne** de satisfaction globale sur les **12 derniers mois jusqu’à une date sélectionnée**. * **Montrer les données immédiatement après mise à jour**. * **Limiter les données** aux années **N et N-1**. * **Respecter la sécurité** table-level du data store. * **Optimiser les performances d’exécution** des requêtes. |

# QUESTIONS

## Question 1

Case study Contoso

You need to ensure that Contoso can use version control to meet the data analytics requirements and the general requirements.

What should you do?

1. Store at the semantic models and reports in Data Lake Gen2 storage.
2. Modify the settings of the Research workspaces to use a GitHub repository.
3. Modify the settings of the Research division workspaces to use an Azure Repos repository.
4. Store all the semantic models and reports in Microsoft OneDrive.

### REPONSE

|  |
| --- |
| **C.Modify the settings of the Research division workspaces to use an Azure Repos repository.**  Même si GitHub est en bonne voie, **la bonne réponse reste encore "Azure Repos"** car :   * C’est **la seule solution officiellement "généralement disponible" (GA)** pour tous les artefacts. * **La préversion de GitHub** peut ne pas être encore activée ou stable dans certains environnements clients. * Les **exigences de fiabilité** et de **support officiel** dans les certifications Microsoft privilégient Azure Repos. |
| ✔️ **Dans l’examen,** **Toujours penser "Azure Repos"** dès qu’on parle d’intégrer des artefacts Power BI / Fabric dans un système de version avec branches Git. |

## Question 2

You have a Fabric warehouse that contains a table named Staging.Sales. Staging.Sales contains the following columns.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data type** | **Nullable** |
| ProductID | Integer | No |
| ProductName | Varchar(30) | No |
| SalesDate | Datetime2(6) | No |
| WholesalePrice | Decimal(18,2) | Yes |
| Amount | Decimal(18,2) | Yes |

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName and has a summarized Amount that is higher than 10,000.

|  |
| --- |
| A:  SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount  FROM Staging.Sales  WHERE DATEPART(YEAR,SaleDate) = '2023'  GROUP BY ProductID, ProductName  HAVING SUM(Amount) > 10000  B:  SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount  FROM Staging.Sales  GROUP BY ProductID, ProductName  HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000  C:  SELECT ProductIO, ProductName, SUM(Amount) AS TotalAmount  FROM Staging.Sales  WHERE DATEPART(YEAR,SaleDate) = ‘2023’ AND SUM(Amount) > 10000  D:  SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount  FROM Staging.Sales  WHERE DATEPART(YEAR,SaleDate) = ‘2023'  GROUP BY ProductID, ProductNane  HAVING TotalAmount > 10000 |

Which query should you use?

A, B, C or D

### REPONSE

|  |
| --- |
| **A**. SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount  FROM Staging.Sales  WHERE DATEPART(YEAR,SaleDate) = '2023'  GROUP BY ProductID, ProductName  HAVING SUM(Amount) > 10000  **Pourquoi ?**   * On filtre bien les ventes de l’année 2023 avec WHERE DATEPART(YEAR, SaleDate) = '2023'. * On regroupe les résultats par ProductID et ProductName. * On garde seulement les groupes dont la somme de Amount dépasse 10 000 grâce à HAVING SUM(Amount) > 10000.   **Et les autres ?**   * **B** : essaie de filtrer l’année dans HAVING — ça ne marche pas, car on ne peut pas utiliser une colonne non agrégée ici. * **C** : utilise SUM(Amount) dans WHERE — interdit en SQL. * **D** : essaie d’utiliser l’alias TotalAmount dans HAVING — ce n’est pas autorisé. |

## Question 3

You have a Fabric tenant that contains JSON files in OneLake. The files have one billion items.

You plan to perform time series analysis of the items.

You need to transform the data, visualize the data to find insights, perform anomaly detection, and share the insights with other business users. The solution must meet the following requirements:

• Use parallel processing.

• Minimize the duplication of data.

• Minimize how long it takes to load the data.

What should you use to transform and visualize the data?

A.the PySpark library in a Fabric notebook

B.the pandas library in a Fabric notebook

C.a Microsoft Power BI report that uses core visuals

## Question 4

You have a Fabric tenant that contains customer churn data stored as Parquet files in OneLake. The data contains details about customer demographics and product usage.

You create a Fabric notebook to read the data into a Spark DataFrame. You then create column charts in the notebook that show the distribution of retained customers as compared to lost customers based on geography, the number of products purchased, age, and customer tenure.

Which type of analytics are you performing?

A.diagnostic

B.descriptive

C.prescriptive

D.predictive

## Question 5

HOTSPOT -

You have a Fabric tenant that contains a semantic model. The model contains data about retail stores.

You need to write a DAX query that will be executed by using the XMLA endpoint. The query must return the total amount of sales from the same period last year.

How should you complete the DAX expression? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

|  |
| --- |
| EVALUATE  VAR \_Lysales = \_\_\_1 ?\_\_\_ ( [Total Sales], SAMEPERTODLASTYEAR ( ‘Orders'[Order Date] ))  Choices   1. CALCULATE 2. CALCULATETABLE 3. FILTER 4. SUMMARIZE 5. SUMMARIZECOLUMNS 6. RETURN   \_\_\_ 2? \_\_\_\_   1. [Total Sales] 2. \_LYSales 3. (\_LYSales } 4. VAR |

Page 2

## Question 6

You have a Fabric workspace named Workspace1 that contains a dataflow named Dataflow1. Dataflow1 returns 500 rows of data.

You need to identify the min and max values for each column in the query results.

Which three Data view options should you select? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A.Show column value distribution

B.Enable column profile

C.Show column profile in details pane

D.Show column quality details

E.Enable details pane

## Question 7

You have a Fabric tenant that contains a Microsoft Power BI report.

You are exploring a new semantic model.

You need to display the following column statistics:

• Count

• Average

• Null count

• Distinct count

• Standard deviation

Which Power Query function should you run?

A.Table.schema

B.Table.view

C.Table.FuzzyGroup

D.Table.Profile

## Question 8 Multiple

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a Delta table named Customer.

When you query Customer, you discover that the query is slow to execute. You suspect that maintenance was NOT performed on the table.

You need to identify whether maintenance tasks were performed on Customer.

Solution: You run the following Spark SQL statement:

DESCRIBE DETAIL customer -

Does this meet the goal?

A.Yes

B.No

## Question 9 Multiple

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a new semantic model in OneLake.

You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression:

df.explain().show()

Does this meet the goal?

A.Yes

B.No

## Question 10 Multiple QMA10

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.

|  |
| --- |
| EVALUATE  FILTER (  VALUES ( Customer[Customer Name] ),  CALCULATE ( COUNTROWS ( ‘Order Item’ ) ) > 0  )  ORDER BY Customer[Customer Name] |

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

ISEMPTY ( RELATEDTABLE ( 'Order Item' ) )

Does this meet the goal?

A.Yes

B.No

## Question 11 Multiple

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

NOT ISEMPTY ( CALCULATETABLE ( 'Order Item ' ) )

Does this meet the goal?

A.Yes

B.No

## Question 12 Multiple

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

CALCULATE ( COUNTROWS ( 'Order Item' ) ) >= 0

Does this meet the goal?

A.Yes

B.No

## Question 13

HOTSPOT -

You have a data warehouse that contains a table named Stage.Customers. Stage.Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer.

You need to write a T-SQL query that will return the customer ID, name. postal code, and the last updated time of the most recent row for each customer ID.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

|  |
| --- |
| Answer Area  WITH CUSTOMERBASE AS (  SELECT [CustomerID]  , [CustomerName]  , [PostalCode]  , [LastUpdated]  , x = \_\_\_ 1? \_\_\_ OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)   1. LAST Value 2. NTITLE() 3. ROW\_NUMBER()   FROM [LakehousePOC] . [dbo] . [CustomerChanges]  )  SELECT CustomerID, CustomerName, PostalCode, LastUpdated  FROM CUSTOMERBASE  \_\_\_2? \_\_\_\_   1. Having Max(LastUpdated) = 1 2. WHERE LastUpdated = Max(LastUpdated) 3. WHERE X = 1 |

## Question 14

Case study Contoso

You need to migrate the Research division data for Productline2. The solution must meet the data preparation requirements.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Answer Area**

|  |
| --- |
| df=spark.read.format('csv').options(header=”true”,inferSchema=”true”).load([“fsi@storage1.dfs.core.windows.net/files/productline2”])  df.write.mode("overwrite").format(  **1 - csv**  **2 - delta**  **3 - parquet**  ).save(  **1 - productline2**  **2 - Tables/productline2**  **3 - Tables/research/productline2**  ) |

## Question 15

Which syntax should you use in a notebook to access the Research division data for Productline1?

A.spark.read.format(“delta”).load(“Files/ResearchProduct”)

B.spark.sql(“SELECT \* FROM Lakehouse1.ResearchProduct ”)

C.spark.sql(“SELECT \* FROM Lakehouse1.Tables.ResearchProduct ”)

D.external\_table(ResearchProduct)

Page 4

## Question 16

HOTSPOT

-

You have a Fabric workspace that uses the default Spark starter pool and runtime version 1.2.

You plan to read a CSV file named Sales\_raw.csv in a lakehouse, select columns, and save the data as a Delta table to the managed area of the lakehouse. Sales\_raw.csv contains 12 columns.

You have the following code.

|  |
| --- |
| (spark  .red  .format (“csv”)  .option(“heade”, ‘true’)  .load ("Files/sales\_raw.csv”)  .select (‘SalesOrderNumber’, ‘OrderDate’, ‘CustomerName’, ‘unitPrice’)  .withcolumn("Year”, year (“OrderDate”) )  .write  .partitionBy (‘Year’)  .SaveAsTable (“sales”)  ) |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Answer Area

|  |  |
| --- | --- |
| Statement | Answer Y or N |
| The Spark engine will read only the ‘SalesOrderNumber,‘OrderDate’,’CustomerName’,"UnitPrice' columns from Sales\_raw.csv. | Y/N |
| The Year column replaces the OrderDate column in the table. | Y/N |
| Adding inferSchema=’true’ to the options will increase the execution time of the query | Y/N |

## Question 17 Multiple

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a new semantic model in OneLake.

You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression:

df.describe().show()

Does this meet the goal?

A.Yes

B.No

## Question 18

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A.Switch

B.KQL

C.Append variable

D.Lookup

## Question 19

HOTSPOT

You have a Fabric tenant that contains a semantic model named model1. The two largest columns in model1 are shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Cardinality** | **Compressed size** | **Description** |
| TransactionKey | 2,500,350,000 | 160 GB | Contains an incrementing surrogate key for a fact table named SalesTransaction |
| SaleDateTime | 6,850,273 | 120 GB | Contains the datetime (to the nearest second) of when a sale occurred |

You need to optimize model1. The solution must meet the following requirements:

• Reduce the model size.

• Increase refresh performance when using Import mode.

• Ensure that the datetime value for each sales transaction is available in the model.

What should you do on each column? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

TransactionKey: \_\_\_\_1? \_\_\_\_

1. Change the data type of the column.
2. Remove the column.
3. Split the column.
4. Truncate the data.

SaleDateTime : \_\_\_\_2?\_\_\_\_

1. Change the data type of the column.
2. Remove the column.
3. Split the column.
4. Truncate the data.Question

## Question 20

DRAG DROP

-

You have a Fabric tenant that contains a data warehouse named DW1. DW1 contains a table named DimCustomer. DimCustomer contains the fields shown in the following table.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **ID** | **CustomerKey** | **CustomerAltKey** | **Title** | **FirstName** | **LastName** | **AddressLine1** | **City** |
| 1 | 29546 | christopher1@contoso.com | Mr. | Christopher | Beck | 93-2501,BlackfriarsRoad | London |
| 2 | 29587 | john12@contoso.com | Mr. | John | Brooks | 99-3ForestWorks | London |
| 3 | 29612 | richard2@contoso.com | Mr. | Richard | Byham | KnavesBeechesInd | High Wycombe |

You need to identify duplicate email addresses in DimCustomer. The solution must return a maximum of 1,000 records.

Which four T-SQL statements should you run in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

1. **SELECT TOP(1000) CustomerAltKey, COUNT(\*)**
2. **GROUP BY CustomerAltKey**
3. **FROM DimCustomer**
4. **SELECT CustomerAltKey, COUNT(\*)**
5. **LIMIT 1000**
6. **HAVING COUNT(\*) > 1**
7. **WHERE COUNT(\*) > 1**

Answer AREA :

\_\_\_1? \_\_\_\_

\_\_\_2? \_\_\_\_

\_\_\_3? \_\_\_\_

\_\_\_4? \_\_\_\_

Page 5

## Question 21

HOTSPOT

-

You have a Fabric tenant that contains a warehouse named WH1.

You run the following T-SQL query against WH1.

|  |
| --- |
| SELECT e.[WWI Employee ID],  e.Employee,  e.[Preferred Name],  gdr.[WWI Employee ID] AS [Direct Report ID],  gdr.Employee AS [Direct Report]  FROM Dimension.Employee AS e  OUTER APPLY Dimension.GetDirectReports(e.[Employee Key]) AS gdr; |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statement** | **Answer Y or N ?** |
| **1. Dimension.GetDirectReports is a scalar T-SQL function.** | Y/N |
| **2. The Dimension.GetDirectReports function will run only once when the query runs.** | Y/N |
| **3. The output rows will include at least one row for each row in the Dimension.Employee table.** | Y/N |

## Question 22 Multiple QMA

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.

|  |
| --- |
| EVALUATE  FILTER (  VALUES ( Customer[Customer Name] ),  CALCULATE ( COUNTROWS ( ‘Order Item’ ) ) > 0  )  ORDER BY Customer[Customer Name] |

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

NOT ( CALCULATE ( COUNTROWS ( 'Order Item' ) ) < 0)

Does this meet the goal?

A.Yes

B.No

## Question 23

Case Study CONTOSO

Which syntax should you use in a notebook to access the Research division data for Productline1?

A.spark.read.format(“delta”).load(“Tables/ResearchProduct”)

B.spark.read.format(“delta”).load(“Files/ResearchProduct”)

C.external\_table(‘Tables/ResearchProduct)

D.external\_table(ResearchProduct)

## Question 24

HOTSPOT -

You have a Fabric tenant.

You plan to create a Fabric notebook that will use Spark DataFrames to generate Microsoft Power BI visuals.

You run the following code.

|  |
| --- |
| from powerbiclient import QuickVisualize, get\_dataset\_config, Report  PBI\_visualize = QuickVisualize(get\_dataset\_config(df))  PBI\_visualize |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statement** | **Answer Y or N ?** |
| **1. The code embeds an existing Power BI Report** | Y/N |
| **2. The code creates a Power BI Report** | Y/N |
| **3. The code displays a summary of the data frame** | Y/N |

## Question 25

You have a Microsoft Power BI Premium Per User (PPU) workspace that contains a semantic model.

You have an Azure App Service app named App1 that modifies row-level security (RLS) for the model by using the XMLA endpoint.

App1 requires users to sign in by using their Microsoft Entra credentials to access the XMLA endpoint.

You need to configure App1 to use a service account to access the model.

What should you do first?

A.Add a managed identity to the workspace.

B.Modify the XMLA Endpoint setting.

C.Upgrade the workspace to Premium capacity.

D.Add a managed identity to App1.

## Question 26

HOTSPOT

You have a Fabric tenant that contains a warehouse named WH1.

You have source data in a CSV file that has the following fields:

• SalesTransactionID

• SaleDate

• CustomerCode

• CustomerName

• CustomerAddress

• ProductCode

• ProductName

• Quantity

• UnitPrice

You plan to implement a star schema for the tables in WH1. The dimension tables in WH1 will implement Type 2 slowly changing dimension (SCD) logic.

You need to design the tables that will be used for sales transaction analysis and load the source data.

Which type of target table should you specify for the CustomerName, CustomerCode, and SaleDate fields? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer AREA

CustomerCode : \_\_\_\_1? \_\_\_\_\_\_

1. Dimension
2. Fact
3. Factless fact
4. Junk dimension

CustomerName : \_\_\_\_2? \_\_\_\_\_\_

1. Dimension
2. Fact
3. Factless fact
4. Junk dimension

SalesDate : \_\_\_\_3? \_\_\_\_\_\_

1. Dimension
2. Fact
3. Factless fact
4. Junk dimension

## Question 27

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A.Get metadata

B.Copy data

C.Lookup

D.Append variable

## Question 28

HOTSPOT -

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows.

You run the following T-SQL statement.

CREATE TABLE test.FactSales AS CLONE OF dbo.FactSales;

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statement** | **Answer Y or N ?** |
| **1. A replica dbo.FactSales is created in the test schema by copying the metadata only** | Y/N |
| **2. Additional schema changes to dbo.FactSales will also apply to test.FactSales** | Y/N |
| **3. Additional data changes to dbo.FactSales will also apply to test.FactSales** | Y/N |

## Question 29

You have an Amazon Web Services (AWS) subscription that contains an Amazon Simple Storage Service (Amazon S3) bucket named bucket1.

You have a Fabric tenant that contains a lakehouse named LH1.

In LH1, you plan to create a OneLake shortcut to bucket1.

You need to configure authentication for the connection.

Which two values should you provide? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A.the shared access signature (SAS) token

B.the secret access key

C.the access ID

D.the access key ID

E.the certificate thumbprint

## Question 30

You have a Fabric tenant that contains a data warehouse.

You need to load rows into a large Type 2 slowly changing dimension (SCD). The solution must minimize resource usage.

Which T-SQL statement should you use?

A.UPDATE AND INSERT

B.MERGE

C.TRUNCATE TABLE and INSERT

D.CREATE TABLE AS SELECT

Page 7

## Question 31

Case study CONTOSO

Which syntax should you use in a notebook to access the Research division data for Productline1?

A.spark.sql("SELECT \* FROM Lakehouse1.Tables.ResearchProduct")

B.spark.read.format("delta").load("Tables/productline1/ResearchProduct")

C.external\_table(ResearchProduct)

D.spark.read.format("delta").load("Tables/ResearchProduct")

## Question 32

You have a Fabric tenant that contains a lakehouse named LH1.

You create new tables in LH1.

You need to ensure that the tables are added automatically to the default semantic model.

What should you do?

A.Disable Query Caching for the default semantic model.

B.From the settings pane of LH1, enable Sync the default Power BI semantic model.

C.Enable Refresh for the default semantic model.

D.From the Endorsement and discovery settings of LH1, select Make discoverable.

## Question 33

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A.Append variable

B.Lookup

C.Copy data

D.KQL

## Question 34 Multiple QM34

Your network contains an on-premises Active Directory Domain Services (AD DS) domain named contoso.com that syncs with a Microsoft Entra tenant by using Microsoft Entra Connect.

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the USERNAME() function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You update the measure to use the USEROBJECTID() function.

Does this meet the goal?

A.Yes

B.No

## Question 35

You are the administrator of a Fabric workspace that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:

Table1: A Delta table created by using a shortcut

Table2: An external table created by using Spark

Table3: A managed table -

You plan to connect to Lakehouse1 by using its SQL endpoint.

What will you be able to do after connecting to Lakehouse1?

A.Read Table3.

B.Update the data Table3.

C.Read Table2.

D.Update the data in Table1.

Page 8

## Question 36 Multiple QM34

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the USERNAME() function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You update the measure to use the USERPRINCIPALNAME() function.

Does this meet the goal?

A.Yes

B.No

## Question 37 Multiple QM34

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the USERNAME() function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You add user objects to the list of synced objects in Microsoft Entra Connect.

Does this meet the goal?

A.Yes

B.No

## Question 38

HOTSPOT

You have a Fabric tenant that contains a PySpark notebook named Notebook1.

You define sas\_token as a variable in the first cell of Notebook1 and store a shared access signature (SAS) token in the variable.

In the second cell, you run the following code.

|  |
| --- |
| customers\_path = 'wasbs://contacts@contoso.blob.core.windows.net/customers?{sas\_token}'  customers = spark.read.parquet(customers\_path)  customers.write \  .format("delta") \  .mode("overwrite") \  .saveAsTable("Customers") |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statement** | **Answer Y or N ?** |
| **1. Customers is a panda Dataframe** | Y/N |
| **2. If a delta table named customers does NOT exist, an error will be generated** | Y/N |
| **3. The source data is located in the customer folder in a container named contatcs** | Y/N |

## Question 39

You have a Fabric tenant that contains a warehouse named DW1 and a lakehouse named LH1. DW1 contains a table named Sales.Product. LH1 contains a table named Sales.Orders.

You plan to schedule an automated process that will create a new point-in-time (PIT) table named Sales.ProductOrder in DW1. Sales.ProductOrder will be built by using the results of a query that will join Sales.Product and Sales.Orders.

You need to ensure that the types of columns in Sales.ProductOrder match the column types in the source tables. The solution must minimize the number of operations required to create the new table.

Which operation should you use?

A.INSERT INTO

B.CREATE TABLE AS SELECT (CTAS)

C.CREATE TABLE AS CLONE OF

D.CREATE MATERIALIZED VIEW AS SELECT

## Question 40

HOTSPOT

You have a Fabric tenant that contains the semantic model shown in the following exhibit.

Representation visuelle de “Semantic model / Setting »

|  |  |  |
| --- | --- | --- |
| **Menu Properties** | **Menu Data** | |
| **General** | **Data Tab** | **Model Tab** |
|  |  | ( i ) View and organize all of the items in your semantic model. Learn how. |
| **Name :** |  | **Semantic model : (detail)** |
| Semantic model |  | . Calculation groups (0) |
| **Description :** |  | > Cultures (1) |
| Enter a description |  | > Measures (1) |
| **Server :** |  | > Perspectives (0) |
| powerbi://api.powerbi.com/v1.0/myorg/Dev%20Workspace%2001 |  | > Relationships (2) |
| **Database ID :** |  | . Roles (0) |
| a36b1c0d-e543-49ff-852d-ee4bd860118f |  | > Tables (5) |
| **Compatibility Level :** |  |  |
| 1604 |  |  |
| **Cultures :** |  |  |
| en-US |  |  |
| **Discourage implicit measures YES/NO :** |  |  |
| No |  |  |
| **Direct Lake Behavior list :** |  |  |
| Direct Lake Only |  |  |

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Answer Area

1. Querying SQL views from the warehouse of the model will cause a fallback to [answer choice \_\_\_ 1? \_\_\_] mode.
   1. Automatic
   2. Direct Lake
   3. DirectQuery
2. Row and column security is [answer choice \_\_\_2?\_\_\_\_].
   1. configured to use object-level security (OLS)
   2. configured to use row-level security (RLS)
   3. undefined

Page 9

## Question 41 Multiple QM34

Note: This section contains one or more sets of questions with the same scenario and problem. Each question presents a unique solution to the problem. You must determine whether the solution meets the stated goals. More than one solution in the set might solve the problem. It is also possible that none of the solutions in the set solve the problem.

After you answer a question in this section, you will NOT be able to return. As a result, these questions do not appear on the Review Screen.

Your network contains an on-premises Active Directory Domain Services (AD DS) domain named contoso.com that syncs with a Microsoft Entra tenant by using Microsoft Entra Connect.

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the USERNAME() function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You create a role in the model.

Does this meet the goal?

A.Yes

B.No

## Question 42

You have a Fabric tenant that contains two workspaces named Workspace1 and Workspace2 and a user named User1.

You need to ensure that User1 can perform the following tasks:

• Create a new domain.

• Create two subdomains named subdomain1 and subdomain2.

• Assign Workspace1 to subdomain1.

• Assign Workspace2 to subdomain2.

The solution must follow the principle of least privilege.

Which role should you assign to User1?

A.domain admin

B.domain contributor

C.Fabric admin

D.workspace Admin

Question 43  
You have a Fabric tenant that contains three users named User1, User2, and User3. The tenant contains a security group named Group1. User1 and User3 are members of Group1.  
  
The tenant contains the workspaces shown in the following table.

| **Name** | **Admin** |
| --- | --- |
| **Workspace1** | **User1** |
| **Workspace2** | **User2** |

The tenant contains the domains shown in the following table.

| **Name** | **Admin** |
| --- | --- |
| **Domain1** | **User1** |
| **Domain2** | **User2** |

User1 creates a new workspace named Workspace3.  
  
You assign Domain1 as the default domain of Group1.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| Statements | answer |
| User2 is assigned the Contributor role for Workspace3. | Y/N |
| User3 is assigned the Viewer role for Workspace3. | Y/N |
| User3 is assigned the Contributor role for Workspace1. | Y/N |

## Question 44

You have a Fabric warehouse named Warehouse1 that contains a table named Table1. Table1 contains customer data.

You need to implement row-level security (RLS) for Table1. The solution must ensure that users can see only their respective data.

Which two objects should you create? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A.DATABASE ROLE

B.STORED PROCEDURE

C.CONSTRAINT

D.FUNCTION

E.SECURITY POLICY

## Question 45

You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1. Lakehouse1 contains a table named Table1. Table1 contains the following data.

|  |  |
| --- | --- |
| **Column name** | **Description** |
| TransactionID | Contains a unique ID for each transaction |
| Date | Contains the date of each transaction |
| ProductID | Contains a unique ID for each product |
| ProductColor | Contains a descriptive attribute that describes the color of each product |
| ProductName | Contains a unique name for each product |
| SalesAmount | Contains the sales amount of each transaction |

You need to perform the following actions:

• Load the data from Table1 into a star schema.

• Create a product dimension table named DimProduct and a fact table named FactSales.

Which three columns should you include in DimProduct?

A.ProductColor, ProductID, and ProductName.

B.ProductName, SalesAmount, and TransactionlD.

C.Date, ProductID, and TransactionlD.

D.ProductID, ProductName, and SalesAmount

Page 10

## Question 46

DRAG DROP

-

You have a Fabric workspace that contains a Dataflow Gen2 query. The query returns the following data.

|  |  |  |
| --- | --- | --- |
| **Column name** | **Data type** | **Description** |
| CustomerID | Whole number | Contains unique customer IDs |
| CustomerName | Text | Contains customer names |
| VersionDate | Date/Time | Contains record timestamps |

You need to filter the results to ensure that only the latest version of each customer’s record is retained. The solution must ensure that no new columns are loaded to the semantic model.

Which four actions should you perform in sequence in Power Query Editor? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

|  |
| --- |
| **Actions** |
| 1.Filter the query where the version date value equals the max version date value. |
| 2.Remove the max version date column. |
| 3.Expand the All Rows detail column. |
| 4.Group by CustomerID and calculate the max version date per customer ID. |
| 5.Remove duplicates based on CustomerID. |
| 6.Group by CustomerID, use the All Rows operation, and calculate the max version date per customer ID. |

ANSWER ordered :

|  |
| --- |
| 1 |
| 2 |
| 3 |
| 4 |

## Question 47

You have a Fabric tenant that contains a warehouse.

You use a dataflow to load a new dataset from OneLake to the warehouse.

You need to add a PowerQuery step to identify the maximum values for the numeric columns.

Which function should you include in the step?

A.Table.MaxN

B.Table.Max

C.Table.Range

D.Table.Profile

## Question 48

HOTSPOT

You have a Microsoft Power Bl project that contains a file named definition.pbir. definition.pbir contains the following JSON.

|  |
| --- |
| {  "version": "1.0",  "datasetReference": {  "byPath": {  "path": "./Sales.Dataset"  },  "byConnection": null  }  } |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statements** | **Answer Area Y or N** |
| definition.pbir is in the PBIR-Legacy format. | Y/N |
| The semantic model referenced by definition.pbir is located in the Power BI service. | Y/N |
| When the related report is opened, Power BI Desktop will open the semantic model in full edit mode. | Y/N |

## Question 49

You have a Fabric tenant that contains a workspace named Workspace1.

You plan to deploy a semantic model named Model1 by using the XMLA endpoint.

You need to optimize the deployment of Model1. The solution must minimize how long it takes to deploy Model1.

What should you do in Workspace1?

A.Select Small semantic model storage format.

B.Select Users can edit data models in the Power BI service.

C.Set Enable Cache for Shortcuts to On.

D.Select Large semantic model storage format.

## Question 50

You have a Microsoft Power BI project that contains a semantic model.

You plan to use Azure DevOps for version control.

You need to modify the .gitignore file to prevent the data values from the data sources from being pushed to the repository.

Which file should you reference?

A.unappliedChanges.json

B.cache.abf

C.localSettings.json

D.model.bim

Page 11

## Question 51

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 uses Pro license mode and contains a semantic model named Model1.

You need to ensure that Model1 supports XMLA connections.

Which setting should modify?

A.Users can edit data models in the Power BI service

B.Enforce strict access control for all data connection types

C.Enable Cache for Shortcuts

D.License mode

## Question 52

You have a Fabric tenant that contains a complex semantic model. The model is based on a star schema and contains many tables, including a fact table named Sales.

You need to visualize a diagram of the model. The diagram must contain only the Sales table and related tables.

What should you use from Microsoft Power BI Desktop?

A.data categories

B.Data view

C.Model view

D.DAX query view

## Question 53

HOTSPOT

-

You have a Fabric tenant that contains a lakehouse named LH1.

You need to deploy a new semantic model. The solution must meet the following requirements:

• Support complex calculated columns that include aggregate functions, calculated tables, and Multidimensional Expressions (MDX) user hierarchies.

• Minimize page rendering times.

How should you configure the model? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Mode : \_\_\_\_1? \_\_\_\_

1. Direct Lake
2. Direct Query
3. Import

Query Caching : \_\_\_\_2?\_\_\_\_\_

1. Capacity default
2. Off
3. On

## Question 54

HOTSPOT

-

You have the following KQL query.

|  |
| --- |
| Sales  | where Status == "Cancelled"  | where OrderDate >= ago(30d)  | summarize TotalSales = sum(SalesAmount) by ProductCategory  | where TotalSales > 0 |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statements** | **Answer Area Y or N** |
| The query excludes sales that have a Status of Cancelled. | Y/N |
| The query calculates the total sales of each product category for the last 30 days. | Y/N |
| The query includes product categories that have had zero sales during the last 30 days. | Y/N |

## Question 55

You have a Fabric warehouse that contains a table named SalesOrderDetail, SalesOrderDetail contains three columns named OrderQty, ProductID and SalesOrderlD. SalesOrderDetail contains one row per combination of SalesOrderlD and ProductID.

You need to calculate the proportion of the total quantity of each sales order represented by each product within the sales order.

|  |
| --- |
| 1. SELECT SalesOrderID, ProductID, OrderQty   CAST(1. \* OrderQty / SUM(OrderQty) OVER(PARTITION BY ProductID) \* 100  AS DECIMAL(5,2)) AS PercentbyProductID  FROM Sales.SalesOrderDetail;   1. SELECT SalesOrderID, ProductID, OrderQty,   CAST(1. \* OrderQty / SUM(OrderQty) OVER(ORDER BY SalesOrderID) \* 100  AS DECIMAL(5,2)) AS PercentbyProductID  FROM Sales.SalesOrderDetail;   1. SELECT SalesOrderID, ProductID, OrderQty,   CAST(1. \* OrderQty / SUM(OrderQty) OVER(ORDER BY ProductID) \* 100  AS DECIMAL(5,2)) AS PercentbyProductID  FROM Sales.SalesOrderDetail;   1. SELECT SalesOrderID, ProductID, OrderQty,   CAST(1. \* OrderQty / SUM(OrderQty) OVER(PARTITION BY SalesOrderID) \* 100  AS DECIMAL(5,2)) AS PercentbyProductID  FROM Sales.SalesOrderDetail; |

Which T-SQL statement should you run?

A, B, C or D

Page 12

## Question 56

HOTSPOT

-

You have a Fabric lakehouse named Lakehouse1 that contains the following data.

|  |  |  |
| --- | --- | --- |
| **Table name** | **Column name** | **Data type** |
| dbo.publicholidays | countryOrRegion | Varchar(8000) |
| dbo.publicholidays | holidayName | Varchar(8000) |
| dbo.publicholidays | Date | Date |
| dbo.sales | OrderDate | Date |
| dbo.sales | Quantity | Float |
| dbo.sales | UnitPrice | Float |

You need build a T-SQL statement that will return the total sales amount by OrderDate only for the days that are holidays in Australia. The total sales amount must sum the quantity multiplied by the price on each row in the dbo.sales table.

How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

SELECT s.OrderDate,

\_\_\_1?\_\_\_\_ AS TotalSalesAmt

Choice:

1. Sum(s.Quantity \* s.UnitPrice)
2. Sum(s.Quantity) \* s.UnitPrice
3. Sum(s.quantity) \* s.UnitPrice
4. Sum(s.Quantity) \* Sum(s.UnitPrice)

FROM [Lakehouse1].[dbo].[sales] AS s

\_\_\_2?\_\_\_\_ JOIN [Lakehouse1].[dbo].[publicholidays] AS ph ON s.OrderDate = ph.date

Choice:

1. Cross
2. Full outer
3. Inner
4. Right outer

WHERE ph.countryOrRegion = 'Australia'

GROUP BY s.OrderDate

## Question 57

You have a Fabric tenant that contains a machine learning model registered in a Fabric workspace.

You need to use the model to generate predictions by using the PREDICT function in a Fabric notebook.

Which two languages can you use to perform model scoring? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

A.T-SQL

B.DAX

C.Spark SQL

D.PySpark

## Question 58

HOTSPOT

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a warehouse named DW1. DW1 contains two tables named Employees and Sales. All users have read access to Dw1.

You need to implement access controls to meet the following requirements:

• For the Sales table, ensure that the users can see only the sales data from their respective region.

• For the Employees table, restrict access to all Personally Identifiable Information (PII).

• Maintain access to unrestricted data for all the users.

What should you use for each table? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Employees:**

1. Column-level security
2. Item permissions
3. Row-level security (RLS)
4. Workspace permissions

**Sales:**

1. Column-level security
2. Item permissions
3. Row-level security (RLS)
4. Workspace permissions

## Question 59

You need to create a Microsoft Power BI file that will be used to create multiple reports. The solution must meet the following requirements:

• The file must include predefined data source connections.

• The file must include the report structure and formatting.

• The file must NOT contain any data.

Which file format should you use?

A.PBIT

B.PBIDS

C.PBIX

D.PBIP

## Question 60

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a single semantic model that has two Microsoft Power BI reports.

You have a Microsoft 365 subscription that contains a data loss prevention (DLP) policy named DLP1.

You need to apply DLP1 to the items in Workspace1.

What should you do?

A.Create a workspace identity.

B.Apply a certified endorsement to the semantic model.

C.Apply sensitivity labels to the semantic model and reports.

D.Apply a master data endorsement to the semantic model.

PAGE 13

## Question 61

You have a Fabric tenant that contains a workspace named Workspace1 and a user named User1. User1 is assigned the Contributor role for Workspace1.

You plan to configure Workspace1 to use an Azure DevOps repository for version control.

You need to ensure that User1 can commit items to the repository.

Which two settings should you enable for User1? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A.Users can sync workspace items with GitHub repositories

B.Users can create and use Data workflows

C.Users can create Fabric items

D.Users can synchronize workspace items with their Git repositories

## Question 62

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a data pipeline named Pipeline1 and a lakehouse named Lakehouse1.

You perform the following actions:

• Create a workspace named Workspace2.

• Create a deployment pipeline named DeployPipeline1 that will deploy items from Workspace1 to Workspace2.

• Add a folder named Folder1 to Workspace1.

• Move Lakehouse1 to Folder1.

• Run DeployPipeline1.

Which structure will Workspace2 have when DeployPipeline1 is complete?

A.\Folder1\Pipeline1 \Folder1\Lakehouse1

B.\Pipeline1 \Lakehouse1

C.\Pipeline1 \Folder1\Lakehouse1

D.\Folder1\Lakehouse1

## Question 63

Your company has a finance department.

You have a Fabric tenant, an Azure Storage account named storage1, and a Microsoft Entra group named Group1. Group1 contains the users in the finance department.

You need to create a new workspace named Workspace1 in the tenant. The solution must meet the following requirements:

• Ensure that the finance department users can create and edit items in Workspace1.

• Ensure that Workspace1 can securely access storage1 to read and write data.

• Ensure that you are the only admin of Workspace1.

• Minimize administrative effort.

You create Workspace1.

Which two actions should you perform next? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A.Assign the Contributor role to Group1.

B.Create a workspace identity.

C.Assign the Admin role to yourself.

D.Assign the Contributor role to each finance department user.

## Question 64

You have a Fabric tenant that contains the workspaces shown in the following table.

|  |  |
| --- | --- |
| **Name** | **Contents** |
| Workspace\_DEV | Lakehouse1 |
| Notebook1 |
| Pipeline1 |
| SemanticModel1 |
| Workspace\_TEST | Lakehouse2 |
| Notebook2 |
| SemanticModel1 |

You have a deployment pipeline named Pipeline1 that deploys items from Workspace\_DEV to Workspace\_TEST. In Pipeline1, all items that have matching names are paired.

You deploy the contents of Workspace\_DEV to Workspace\_TEST by using Pipeline1.

What will the contents of Workspace\_TEST be once the deployment is complete?

A.Lakehouse1 - Lakehouse2 - Notebook1 - Notebook2 - Pipeline1 - SemanticModel1

B.Lakehouse1 - Notebook1 - Pipeline1 - SemanticModel1

C.Lakehouse2 - Notebook2 - SemanticModel1

D.Lakehouse2 - Notebook2 - Pipeline1 - SemanticModel1

## Question 65

HOTSPOT

-

You have a Fabric tenant that contains a workspace named Workspace\_DEV. Workspace\_DEV contains the semantic models shown in the following table.

|  |  |
| --- | --- |
| **Name** | **Scheduled refresh policy** |
| Model1 | Configured |
| Model2 | Not Configured |

Workspace\_DEV contains the dataflows shown in the following table.

|  |  |  |
| --- | --- | --- |
| **Name** | **Type** | **Scheduled refresh policy** |
| DF1 | Dataflow Gen1 | Configured |
| DF2 | Dataflow Gen2 | Not Configured |

You create a new workspace named Workspace\_TEST.

You create a deployment pipeline named Pipeline1 to move items from Workspace\_DEV to Workspace\_TEST.

You run Pipeline1.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statements** | **Answer Area Y or N** |
| DF1 will be deployed to Workspace\_TEST. | Y/N |
| Data from Model1 will be deployed to Workspace\_TEST. | Y/N |
| The scheduled refresh policy for Model1 will be deployed to Workspace\_TEST. | Y/N |

Page 14

## Question 66

HOTSPOT

You have a Fabric warehouse that contains two tables named DimDate and Trips.

DimDate contains the following fields.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Description** |
| DateID | Date | Unique identifier for the date |
| IsHoliday | Bit | Used to denote that the date is a statutory holiday |

Trips contains the following fields.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Description** |
| tripID | Integer | Trip identifier |
| tripDistance | Integer | Miles travelled |
| dateID | Date | Date of trip |

You need to compare the average miles per trip for statutory holidays versus non-statutory holidays.

How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

SELECT d.IsHoliday, \_\_\_1?\_\_\_\_ as MilesPerTrip

Choice :

1. **(count(t.tripID) / sum(t.tripDistance))**
2. **(sum(t.tripDistance) / count(t.tripID))**
3. **Sum(t.tripDistance) / t.tripID)**

FROM dbo.Trips t

INNER JOIN dbo.DimDate d ON t.dateID = d.DateID

\_\_\_2?\_\_\_\_ d.IsHoliday

Choice :

1. **group by**
2. **order by**
3. **where**

## Question 67

HOTSPOT

-

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named LH1 and a warehouse named DW1. LH1 contains a table named signindata that is in the dbo schema.

You need to create a stored procedure in DW1 that deduplicates the data in the signindata table.

How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

CREATE PROCEDURE dbo.usp\_GetPerson

AS \_\_\_1? \_\_\_\_

Choice :

1. **BEGIN**
2. **BEGIN DISTRIBUTED TRANSACTION**
3. **SET**

BEGIN DISTRIBUTED TRANSACTION

SET

SELECT \_\_\_ 2ç\_\_\_ PersonID, FirstName, LastName

Choice :

1. **DISTINCT**
2. **GROUP BY**
3. **TOP (100) PERCENT WITH TIES**

FROM LH1.dbo.signindata

GROUP BY PersonID, FirstName, LastName

TOP (100) PERCENT WITH TIES

END

## Question 68

You are analyzing the data in a Fabric notebook.

You have a Spark DataFrame assigned to a variable named df.

You need to use the Chart view in the notebook to explore the data manually.

Which function should you run to make the data available in the Chart view?

A.displayHTML

B.show

C.write

D.display

## Question 69

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A.Get metadata

B.Switch

C.Lookup

D.Append variable

## Question 70

HOTSPOT

-

You have the following T-SQL statement.

|  |
| --- |
| SELECT  ProductCategory,  Region,  SUM(CASE WHEN RefundStatus <> 'Refunded' THEN SalesAmount ELSE 0 END) AS TotalRevenue  FROM  Sales  WHERE  YEAR(TransactionDate) = YEAR(GETDATE())  GROUP BY  ProductCategory, Region; |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statements** | **Answer Area Y or N** |
| The statement returns Region values when a Sales item has a RefundStatus of Refunded. | Y/N |
| The statement only returns TransactionDate values that occurred during the current year. | Y/N |
| The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded. | Y/N |

Page 15

## Question 71

HOTSPOT

-

You have a Fabric warehouse that contains the following data.

|  |  |  |
| --- | --- | --- |
| **Table name** | **Primary key** | **Foreign key** |
| Customer | CustomerID | SalesRegion.SalesRegionID |
| CustomerAddress | CustomerID, AddressID | Customer.CustomerID, Address.AddressID |
| Address | AddressID | None |
| SalesRegion | SalesRegionID | None |

The data has the following characteristics:

• Each customer is assigned a unique CustomerID value.

• Each customer is associated to a single SalesRegion value.

• Each customer is associated to a single CustomerAddress value.

• The Customer table contains 5 million rows.

• All foreign key values are non-null.

You need to create a view to denormalize the data into a customer dimension that contains one row per distinct CustomerID value. The solution must minimize query processing time and resources.

How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

|  |
| --- |
| Create View DimCustomer  As  Select  C.CustomerID,  C.CustomerName,  C.CompanyName,  SR.SalesRegion,  A.City,  A.StateProvince,  A.CountryRegion,  A.PostalCode  from SalesLT.Customer C  \_\_\_1?\_\_\_\_ SalesLT.SalesRegion SR on C.SalesRegionID = SR.SalesRegionID  Choice:   1. Full outer join 2. Inner join 3. Left outer join   inner join [SalesLT].[CustomerAddress] CA on CA.customerid = C.customerid  inner join [SalesLT].[Adress] A on \_\_\_2? \_\_\_\_  Choice:   1. A.AdressID = CA.AdressID 2. C.AdressID = CA.AdressID 3. C.customerid = A.customerid   where CA.AddressType = 'Main Office' |

## Question 72

You have a query in Microsoft Power BI Desktop that contains two columns named Order\_Date and Shipping\_Date.

You need to create a column that will calculate the number of days between Order\_Date and Shipping\_Date for each row.

Which Power Query function should you use?

A.DateTime.LocalNow

B.Duration.Days

C.Duration.From

D.Date.AddDays

## Question 73

You have a Fabric tenant that contains a semantic model.

You need to modify object-level security (OLS) for the model.

What should you use?

A.the Fabric service

B.Microsoft Power BI Desktop

C.ALM Toolkit

D.Tabular Editor

## Question 74

You have a Microsoft Power BI semantic model that contains a measure named TotalSalesAmount. TotalSalesAmount returns a sales revenue amount that is translated into a selected currency.

You need to ensure that the value returned by TotalSalesAmount is formatted to use the correct currency symbol.

What should you include in the solution?

A.a field parameter

B.a linguistic schema

C.a dynamic format string

D.the WINDOW DAX function

## Question 75

You plan to use Fabric to store data.

You need to create a data store that supports the following:

• Writing data by using T-SQL

• Multi-table transactions

• Dynamic data masking

Which type of data store should you create?

A.KQL database

B.lakehouse

C.warehouse

D.semantic model

Page 16

## Question 76

DRAG DROP

-

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 uses the Pro license mode and contains a semantic model named Model1.

You have an Azure DevOps organization.

You need to enable version control for Workspace1. The solution must ensure that Model1 is added to the repository.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

|  |
| --- |
| **Actions** |
| 1.Connect Workspace1 to a Git provider |
| 2.Sync Workspace1 with the repository |
| 3.Configure branch policies in Azure DevOps. |
| 4.Ceate a deployment pipeline |
| 5.Assign Workspace1 to a fabric capacity |

Answer area

|  |
| --- |
|  |
|  |
|  |

## Question 77

HOTSPOT

-

You have a Fabric tenant that contains a workspace named Enterprise. Enterprise contains a semantic model named Model1. Model1 contains a date parameter named Date1 that was created in Power Query.

You build a deployment pipeline named Enterprise Data that includes two stages named Development and Test. You assign the Enterprise workspace to the Development stage.

You need to perform the following actions:

• Create a workspace named Enterprise [Test] and assign the workspace to the Test stage.

• Configure a rule that will modify the value of Date1 when changes are deployed to the Test stage.

Which two settings should you use? To answer, select the appropriate settings in the answer area.

NOTE: Each correct answer is worth one point.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Section** | **Développement (Development)** | **Test** | | --- | --- | --- | | **Nom du pipeline** | **Enterprise Data** |  | | **Environnement** | Enterprise |  | | **Composants liés** | - Warehouses (Preview) : 0  - Environments (Preview) : 0  - Dashboards (Preview) : 0  - Lakehouses (Preview) : 0 | Aucun composant encore déployé | | **Actions** | - [✏️] Modifier le stage  - **Publish app**  - **Deploy** | - [✏️] Modifier le stage  - [🔗] Assign a workspace | | **Déploiement** | Bouton vert **Deploy** | Zone : « Deploy to this stage or Assign a workspace »  ⬇ Sélecteur déroulant : « Select »  🔗 Lien : « Why can't I see all my workspaces? »  Texte grisé : « Assign a workspace » | | **Navigation globale** | Deployment history – Manage Access – Settings |  | |

## Question 78

HOTSPOT

-

You have two Microsoft Power BI queries named Employee and Retired Roles.

You need to merge the Employee query with the Retired Roles query. The solution must ensure that duplicate rows in each query are removed.

Which column and Join Kind should you use in Power Query Editor? To answer, select the appropriate options in the answer area.

NOTE: Each correct answer is worth one point.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Interface de fusion de tables (Merge)**  **Titre principal**  **Answer Area** **Merge**  *Select a table and matching columns to create a merged table.*  **🧾 Table 1 : Employee**   | **ID** | **Division** | **Title Detail** | **Role** | **Department** | **Department Number** | **Division Number** | | --- | --- | --- | --- | --- | --- | --- | | 198 | Services | Marketing | Vice President | Marketing | 22 | 14 | | 299 | Services | Business Development | Vice President | Sales | 17 | 14 | | 169 | Services | Human Resources | Vice President | Human Resources | 13 | 14 | | 146 | Services | HSE & Risk Management | Vice President | Safety | 16 | 14 | | … | … | … | … | … | … | … |   **🧾 Table 2 : Retired Roles**   | **Division** | **Role** | | --- | --- | | Services | Technical Analyst | | Services | Technical Support Analyst | | Services | Senior Engineer | | Services | Chief Technical Officer | | Services | Applied Technology Specialist |   **🔄 Zone de jointure (Join Kind)**   | **Option sélectionnée par défaut** | | --- | | **Left Outer (all from first, matching from second)** |   🔽 Autres options disponibles :   | **Type de jointure** | **Description** | | --- | --- | | Left Outer | All from first, matching from second | | Right Outer | All from second, matching from first | | Full Outer | All rows from both | | Inner | Only matching rows | | Left Anti | Rows only in first | | Right Anti | Rows only in second |   **✅ Boutons d'action**   * **OK** * **Cancel** |

## Question 79

You have a Fabric tenant that contains a Microsoft Power BI report named Report1. Report1 includes a Python visual.

Data displayed by the visual is grouped automatically and duplicate rows are NOT displayed.

You need all rows to appear in the visual.

What should you do?

A.Reference the columns in the Python code by index.

B.Modify the Sort Column By property for all columns.

C.Add a unique field to each row.

D.Modify the Summarize By property for all columns.

## Question 80

You have a Fabric tenant that contains a workspace named Workspace1 and a user named User1. Workspace1 contains a warehouse named DW1.

You share DW1 with User1 and assign User1 the default permissions for DW1.

What can User1 do?

A.Build reports by using the default dataset.

B.Read data from the tables in DW1.

C.Connect to DW1 via the Azure SQL Analytics endpoint.

D.Read the underlying Parquet files from OneLake.

Page 17

## Question 81

DRAG DROP

-

You have a Fabric workspace named Workspace1.

You have three groups named Group1, Group2, and Group3.

You need to assign a workspace role to each group. The solution must follow the principle of least privilege and meet the following requirements:

• Group1 must be able to write data to Workspace1, but be unable to add members to Workspace1.

• Group2 must be able to configure and maintain the settings of Workspace1.

• Group3 must be able to write data and add members to Workspace1, but be unable to delete Workspace1.

Which workspace role should you assign to each group? To answer, drag the appropriate roles to the correct groups. Each role may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

|  |
| --- |
| **Roles** |
| 1.Admin |
| 2.Contributor |
| 3.Member |
| 4.Viewer |

Answer area

|  |  |
| --- | --- |
| Group1: |  |
| Group2: |  |
| Group3: |  |

## Question 82

You have a Fabric tenant named Tenant1 that contains a lakehouse named Lakehouse1.

You need to add data to Lakehouse1 from a CSV file in an Azure Storage account outside of Fabric. The solution must minimize development effort.

What should you use to add the data?

A.copy job

B.shortcut

C.pipeline

D.Dataflow Gen2

## Question 83

HOTSPOT

-

You have a Fabric warehouse that contains a table named Table1. Table1 contains the following data.

|  |  |
| --- | --- |
| **Column name** | **Data type** |
| item\_id | Integer |
| item\_name | varchar(50) |
| item\_description | varchar(50) |
| purchase\_date | Datetime |

You need to create a T-SQL statement that meets the following requirements:

• Outputs the item name of each item and returns a null value if the item name is longer than 20 characters

• Outputs the PurchaseDate value in the format of МММ dd, yy

How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer as Itemid

\_\_\_1?\_\_\_\_ as ItemName

Choice :

1. .convert(varchar(20), item\_name)
2. .convert(varchar(max), item\_name)
3. .try.cast(item\_name as varchar(20))

, item.description as ItemDecription

\_\_\_ 2? \_\_\_\_ as PurchaseDate

Choice :

1. .convert(varchar, purchase\_date, 7)
2. .convert(varchar, purchase\_date, 109)
3. .convert(varchar, purchase\_date, 112)

FROM

Table1

WHERE

Item\_type ) @Itemtype\_parameter

## Question 84

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A.Append variable

B.Script

C.Stored procedure

D.Get metadata

## Question 85

HOTSPOT

You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 contains the following data.

|  |  |  |
| --- | --- | --- |
| **Table name** | **Column name** | **Data type** |
| Employee | EmployeeID | Int |
| Employee | EmployeeName | Varchar(128) |
| Employee | EmployeePosition | Varchar(64) |
| Contract | EmployeeID | Int |
| Contract | ContractType | Varchar(64) |
| Contract | StartDate | Datetime2 |
| Contract | EndDate | Datetime2 |

You need to create a T-SQL statement that will denormalize the tables and include the ContractType and StartDate attributes in the results. The solution must meet the following requirements:

• Include attributes from matching rows in the Contract table.

• Ensure that all the rows from the Employee table are preserved.

• Return the total number of employees per contract type for all the contract types that have more than two employees.

How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

WITH result AS (

SELECT

e.EmployeeID,

e.EmployeeName,

e.EmployeePosition,

c.ContractType

FROM Employee AS e

\_\_\_1? \_\_\_\_ Contract AS c ON c.EmployeeID = e.EmployeeID

Choice 1 :

1. CROSS JOIN
2. INNER JOIN
3. LEFT OUTER JOIN
4. RIGHT OUTER JOIN

)

SELECT

COUNT(DISTINCT EmployeeID) AS TotalEmployees, ContractType

FROM result

GROUP BY ContractType

\_\_\_2? \_\_\_\_ COUNT(DISTINCT EmployeeID) > 2

Choice 2 :

1. CONTAINS
2. HAVING
3. LIMIT
4. WHERE

Page 18

## Question 86

HOTSPOT

-

You have a KQL database that contains a table named Readings.

You need to query Readings and return the results shown in the following table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| City | Area | MeterReading | Datetime | PrevMeterReading | PrevDatetime |
| Copenhagen | Area1 | 2700 | 2022-04-20 10:00:00 | None | None |
| Copenhagen | Area2 | 2720 | 2022-04-20 11:00:00 | 2700 | 2022-04-20 10:00:00 |
| Copenhagen | Area1 | 2750 | 2022-04-20 12:00:00 | 2720 | 2022-04-20 11:00:00 |
| Copenhagen | Area2 | 2780 | 2022-04-20 13:00:00 | 2750 | 2022-04-20 12:00:00 |

How should you complete the query? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

|  |
| --- |
| Readings  | filter City == "Copenhagen"  | sort by Datetime  |\_ Choice 1? \_ PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)  Choice 1:   1. evaluate 2. extend 3. lookup 4. project 5. summarize   | \_\_ Choice 2? \_ City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime  Choice 2:   1. evaluate 2. extend 3. lookup 4. project 5. summarize |

## Question 87

HOTSPOT

-

You have a Fabric workspace that contains a warehouse named DW1. DW1 contains the following tables and columns.

|  |  |
| --- | --- |
| **Table name** | **Column name** |
| SalesOrderDetail | ProductID |
| SalesOrderDetail | ModifiedDate |
| SalesOrderDetail | OrderQty |
| Product | ProductID |
| Product | Name |

You need to summarize order quantities by year and product. The solution must include the yearly sum of order quantities for all the products in each row.

How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

|  |
| --- |
| SELECT  \_ Choice 1? \_ (SO.ModifiedDate AS DATE) AS OrderDate,  Choice 1:   1. CAST 2. CONVERT 3. YEAR   P.Name AS ProductName,  SUM(SO.OrderQty) AS OrderQty  FROM [dbo].[SalesOrderDetail] SO  INNER JOIN [dbo].[Product] P  ON P.ProductID = SO.ProductID  GROUP BY  \_ Choice 2? \_  Choice 2 :   1. CUBE(YEAR(SO.ModifiedDate), P.Name) 2. GROUPING SETS ((YEAR(SO.ModifiedDate), P.Name), (YEAR(SO.ModifiedDate))) 3. ROLLUP(YEAR(SO.ModifiedDate), P.Name) 4. YEAR(SO.ModifiedDate), P.Name   ORDER BY OrderDate |

REPONSE

|  |
| --- |
| L'objectif est de résumer les quantités de commandes par année et par produit, en incluant une ligne pour chaque année et produit, ainsi qu'une ligne pour chaque année avec le total de tous les produits confondus.  **Choix 1 : Sélection de la colonne d'année dans le SELECT**   * **YEAR(SO.ModifiedDate)** : C'est le bon choix car il extrait uniquement l'année, ce qui est nécessaire pour regrouper les données par année.   **Choix 2 : Comment regrouper pour obtenir les totaux par produit et les totaux par année**   * **GROUPING SETS ((YEAR(SO.ModifiedDate), P.Name), (YEAR(SO.ModifiedDate)))** : C'est le meilleur choix car il permet d'obtenir exactement ce qui est demandé :   + Une ligne pour chaque combinaison année et produit.   + Une ligne pour chaque année avec le total de tous les produits.   **Réponse finale**  Voici la requête T-SQL qui répond exactement aux exigences :  Copier  SELECT  YEAR(SO.ModifiedDate) AS OrderDate,  P.Name AS ProductName,  SUM(SO.OrderQty) AS OrderQty  FROM [dbo].[SalesOrderDetail] SO  INNER JOIN [dbo].[Product] P  ON P.ProductID = SO.ProductID  GROUP BY  GROUPING SETS ((YEAR(SO.ModifiedDate), P.Name), (YEAR(SO.ModifiedDate)))  ORDER BY OrderDate |

## Question 88

DRAG DROP -

You have a Fabric tenant that contains a semantic model. The model contains data about retail stores.

You need to write a DAX query that will be executed by using the XMLA endpoint. The query must return a table of stores that have opened since December 1, 2023.

How should you complete the DAX expression? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

VALUES :

1. DEFINE
2. EVALUATE
3. FILTER
4. SUMMARIZE
5. TABLE

ANSWER AREA

|  |
| --- |
| \_\_ VALUE1 \_\_  VAR SalesSince = DATE (2023, 12, 01)  \_\_ VALUE2 \_\_  FILTER (  \_\_ VALUE3 \_\_ (Store, Store[Name], Store[OpenDate]),  Store[OpenDate] >= SaleSince  ) |

REPONSE

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Analyse ligne par ligne :**  **✅ Ligne 1 :** DEFINE  VAR SalesSince = DATE (2023, 12, 01)  On déclare une variable → **DEFINE** est utilisé pour déclarer des mesures ou des variables dans une requête DAX. 👉 **VALUE1 = DEFINE**  **✅ Ligne 2 :** EVALUATE  C’est le mot-clé **obligatoire** pour exécuter la partie qui retourne un résultat tabulaire dans une requête DAX. 👉 **VALUE2 = EVALUATE**  **✅ Ligne 3 :** FILTER (  SUMMARIZE(Store, Store[Name], Store[OpenDate]),  Store[OpenDate] >= SalesSince  )   * **FILTER()** est utilisé pour **restreindre les lignes** selon une condition logique. * **SUMMARIZE()** construit un tableau contenant les colonnes Store[Name] et Store[OpenDate] à partir de la table Store.   👉 **VALUE3 = SUMMARIZE**  **✅ Réponse complète :**   | **Position** | **Réponse** | | --- | --- | | **VALUE1** | DEFINE | | **VALUE2** | EVALUATE | | **VALUE3** | SUMMARIZE | | **✅ Requête finale en DAX :**  **DEFINE**  **VAR SalesSince = DATE(2023, 12, 01)**  **EVALUATE**  **FILTER(**  **SUMMARIZE(Store, Store[Name], Store[OpenDate]),**  **Store[OpenDate] >= SalesSince**  **)** |  | |

## Question 89

You have a Fabric workspace named Workspace1.

Workspace1 contains multiple semantic models, including a model named Model1. Model1 is updated by using an XMLA endpoint.

You need to increase the speed of the write operations of the XMLA endpoint.

What should you do?

A.Delete any unused semantic models from Workspace1.

B.Select Large semantic model storage format for Workspace1.

C.Configure Model 1 to use the Direct Lake storage format.

D.Delete any unused columns from Model1.

REPONSE

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Option** | **Pertinence pour améliorer la vitesse des écritures XMLA** | | --- | --- | | A | Aucune | | B ✅ | **Directement liée à l’objectif** (même si modèle "petit") | | C | Incompatible avec XMLA write | | D | Utile pour optimiser l’espace, **pas pour les écritures XMLA** |   **PLUS** :   **Le format "Large semantic model" n’est pas réservé aux très grands modèles.**   * Il **active des optimisations mémoire et des capacités d’écriture XMLA avancées**, quelle que soit la taille actuelle. * Il permet d’**écrire plus rapidement dans le modèle**, surtout via des partitions (ce qui est typique des scénarios XMLA write).    **Microsoft recommande ce format pour toute écriture XMLA à haut débit** :   * Même pour des modèles de taille moyenne * Car le **mode standard** a des limites sur la mémoire, la parallélisation, et la gestion des partitions.    **Supprimer des colonnes inutilisées** :   * Cela **réduit la taille**, ce qui est bien… * Mais **n’accélère pas intrinsèquement les opérations d’écriture** via XMLA. * Et **aucune indication** dans la question ne dit que les colonnes inutiles sont responsables du ralentissement. |

## Question 90

You have a Fabric workspace named Workspace1 that is assigned to a newly created Fabric capacity named Capacity1.

You create a semantic model named Model1 and deploy Model1 to Workspace1.

You need to publish changes to Model1 directly from Tabular Editor.

What should you do?

A.For Workspace1, enable Git integration.

B.For Model1, enable external sharing.

C.For Workspace1, create a managed private endpoint.

D.For Capacity1, set XMLA Endpoint to Read Write.

REPONSE

|  |
| --- |
| **✅ Réponse correcte :**  **D. For Capacity1, set XMLA Endpoint to Read Write**  **🔍 Explication détaillée :**  Tu veux **publier des modifications dans un modèle sémantique (Model1) depuis Tabular Editor**.  👉 Pour cela, **Tabular Editor utilise l’XMLA endpoint** pour se connecter et pousser les changements directement dans le modèle dans Fabric (comme c’était le cas avec Power BI Premium). |

Page 19

## Question 91

HOTSPOT

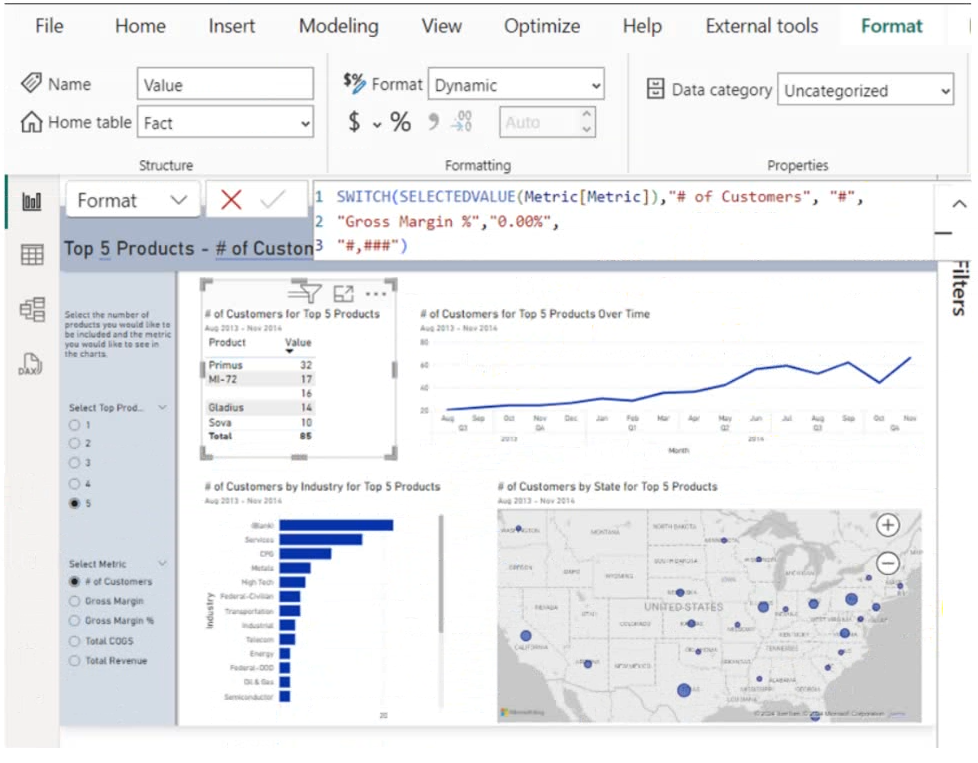
-

You are creating a report and a semantic model in Microsoft Power BI Desktop.

The Value measure has the expression shown in the following exhibit.

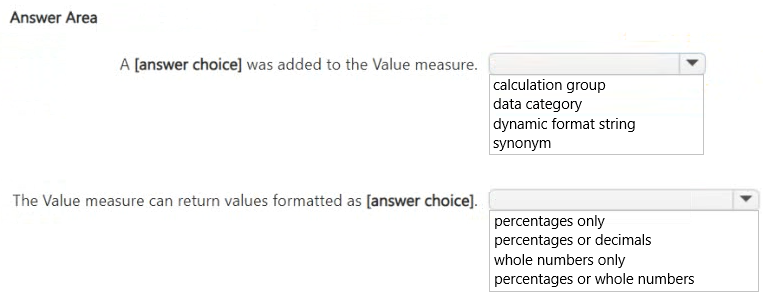
Code DAX :

|  |
| --- |
| SWITCH(  SELECTEDVALUE(Metric[Metric]),  "# of Customers", "#",  "Gross Margin %", "0.00%",  “#,"#,###"  ) |



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.



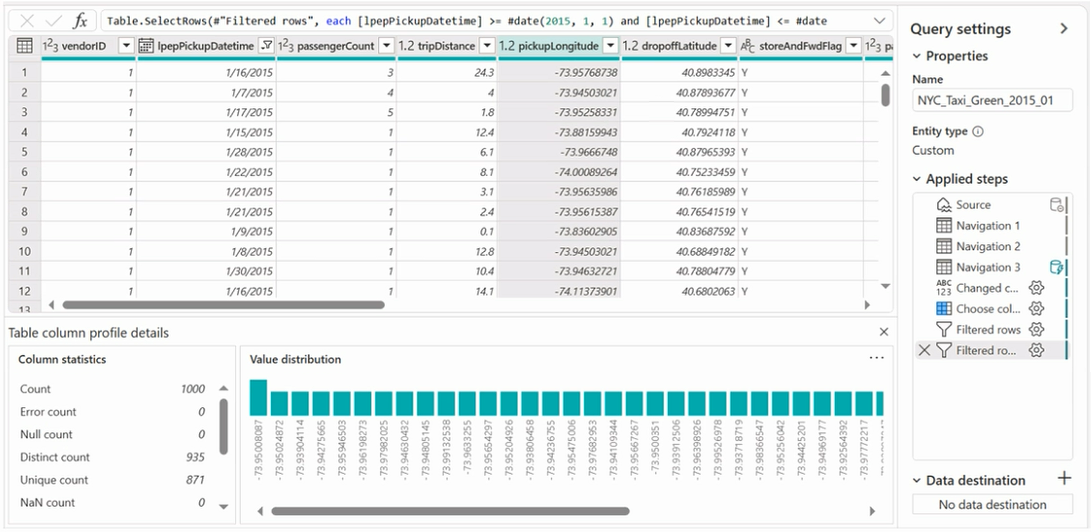
REPONSE

|  |
| --- |
| 1. **Premier choix : "A [answer choice] was added to the Value measure."**    * **L'expression utilise SELECTEDVALUE(Metric[Metric]) pour déterminer le format à appliquer. Cela signifie que la mesure "Value" utilise une logique conditionnelle basée sur la valeur sélectionnée dans la colonne Metric[Metric].**    * **Choix correct : dynamic format string** 2. **Deuxième choix : "The Value measure can return values formatted as [answer choice]."**    * **L'expression retourne trois types de formats :**      + **"#" pour le nombre de clients ("# of Customers").**      + **"0.00%" pour le pourcentage de marge brute ("Gross Margin %").**      + **"#,###" pour un format de nombre entier avec séparateurs de milliers.**    * **Cela signifie que la mesure "Value" peut retourner des valeurs formatées en tant que pourcentages, nombres entiers, ou nombres avec séparateurs de milliers.**    * **Choix correct : percentages or whole numbers**  * **SELECTEDVALUE(Metric[Metric]) :** Cette fonction retourne la valeur actuellement sélectionnée dans la colonne Metric[Metric]. * **SWITCH** : Cette fonction évalue la valeur sélectionnée et retourne un format spécifique en fonction de cette valeur. Par exemple :   + Si la valeur sélectionnée est **"# of Customers"**, elle retourne **"#"** comme format.   + Si la valeur sélectionnée est **"Gross Margin %"**, elle retourne **"0.00%"** comme format.   + Par défaut, si aucune des valeurs spécifiées n'est sélectionnée, elle retourne **"#,###"** comme format. |

## Question 92

You have a Fabric workspace named Workspace1 that contains a dataflow named Dataflow1. Dataflow1 has a query that returns 2,000 rows.

You view the query in Power Query as shown in the following exhibit.



**vendorID | lpepPickupDatetime ... | pickupLongitude**

**1 | 1/6/2023 | 73.99341**

**1 | 23/4/2023 | 73.99322**

**1 | 12/6/2024 | 73.99311**

**....**

|  |  |
| --- | --- |
| **Table Column Profile Details** |  |
| **Column Statistics** |  |
|  |  |
| **Count** | **1000 rows.** |
| **Error count** | **0 errors.** |
| **Null count** | **0 null values.** |
| **Distinct count** | **935 distinct values.** |
| **Unique count** | **871 unique values.** |
| **NaN count** | **0 NaN values.** |

Value Distribution: (exhibit )

A histogram showing the distribution of values for the pickupLongitude column. The values range approximately from -73.99 to -73.94

| ██

| ██

| ██ ██ ██ ██

| ██ ██ ██ ██

|\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

-73.99... -73.99…-73.99… ....

What can you identify about the pickupLongitude column?

A.The column has duplicate values.

B.All the table rows are profiled.

C.The column has missing values.

D.There are 935 values that occur only once.

## Question 93

HOTSPOT –

Case Study CONTOSO

You need to recommend a solution to group the Research division workspaces.

What should you include in the recommendation? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Grouping method :

1. Capacity
2. Domain
3. Tenant

Tool :

1. OneLake data hub
2. The Fabric Admin portal
3. The Microsoft Entra admin center

## Question 94

You have a Fabric tenant named Tenant1 that contains a workspace named WS1. WS1 uses a capacity named C1 and contains a dataset named DS1.

You need to ensure read-write access to DS1 is available by using XMLA endpoint.

What should be modified first?

A.the DS1 settings

B.the WS1 settings

C.the C1 settings

D.the Tenant1 settings

## Question 95

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 is assigned to a Fabric capacity.

You need to recommend a solution to provide users with the ability to create and publish custom Direct Lake semantic models by using external tools. The solution must follow the principle of least privilege.

Which three actions in the Fabric Admin portal should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A.From the Tenant settings, set Allow XMLA Endpoints and Analyze in Excel with on-premises datasets to Enabled.

B.From the Tenant settings, set Allow Azure Active Directory guest users to access Microsoft Fabric to Enabled.

C.From the Tenant settings, select Users can edit data model in the Power BI service.

D.From the Capacity settings, set XMLA Endpoint to Read Write.

E.From the Tenant settings, set Users can create Fabric items to Enabled.

F.From the Tenant settings, enable Publish to Web.

Page 20

## Question 96

You are creating a semantic model in Microsoft Power BI Desktop.

You plan to make bulk changes to the model by using the Tabular Model Definition Language (TMDL) extension for Microsoft Visual Studio Code.

You need to save the semantic model to a file.

Which file format should you use?

A.PBIP

B.PBIX

C.PBIT

D.PBIDS

## Question 97

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains three schemas named schemaA, schemaB, and schemaC.  
  
You need to ensure that a user named User1 can truncate tables in schemaA only.  
  
How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

**ANWSER AREA**

**GRANT \_\_\_1 ?\_\_\_\_\_\_\_ ON \_\_\_\_2 ?\_\_\_\_\_\_\_ TO User1 ;**

| GRANT | Choice 1 | ON | Choice 2 | TO User1; |
| --- | --- | --- | --- | --- |
|  | 1- ALTER |  | 1 - DATABASE::schemaA |  |
|  | 2 - CONNECT |  | 2 - OBJECT::schemaA |  |
|  | 3 - EXECUTE |  | 3 - SCHEMA::schemaA |  |

## Question 98

You plan to deploy Microsoft Power BI items by using Fabric deployment pipelines. You have a deployment pipeline that contains three stages named Development, Test, and Production. A workspace is assigned to each stage.

You need to provide Power BI developers with access to the pipeline. The solution must meet the following requirements:

Ensure that the developers can deploy items to the workspaces for Development and Test.

Prevent the developers from deploying items to the workspace for Production.

Follow the principle of least privilege.

Which three levels of access should you assign to the developers? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A.Build permission to the production semantic models

B.Admin access to the deployment pipeline

C.Viewer access to the Development and Test workspaces

D.Viewer access to the Production workspace

E.Contributor access to the Development and Test workspaces

F.Contributor access to the Production workspace

## Question 99

You have a Fabric workspace that contains a DirectQuery semantic model. The model queries a data source that has 500 million rows.

You have a Microsoft Power Bi report named Report1 that uses the model. Report1 contains visuals on multiple pages.

You need to reduce the query execution time for the visuals on all the pages.

What are two features that you can use? Each correct answer presents a complete solution,

NOTE: Each correct answer is worth one point.

A.user-defined aggregations

B.automatic aggregation

C.query caching

D.OneLake integration

## Question 100

You have a Fabric tenant that contains 30 CSV files in OneLake. The files are updated daily.

You create a Microsoft Power BI semantic model named Model1 that uses the CSV files as a data source. You configure incremental refresh for Model1 and publish the model to a Premium capacity in the Fabric tenant.

When you initiate a refresh of Model1, the refresh fails after running out of resources.

What is a possible cause of the failure?

A.Query folding is occurring.

B.Only refresh complete days is selected.

C.XMLA Endpoint is set to Read Only.

D.Query folding is NOT occurring.

E.The delta type of the column used to partition the data has changed.

REPONSE

|  |
| --- |
| **D. Query folding is NOT occurring.**  **Résumé de la situation :**   * Tu as **30 fichiers CSV** dans **OneLake**, mis à jour quotidiennement. * Tu as un **semantic model Power BI** avec **actualisation incrémentielle**. * Le **rafraîchissement échoue** par manque de ressources.   **💥 Ce que fait l’actualisation incrémentielle :**  Elle repose sur **la capacité de la source de données à appliquer des filtres côté source** (sur la colonne de date de partition), pour ne charger **que les nouvelles données** (ex. : la journée d’hier).  Mais pour que cela fonctionne, **Power BI doit pouvoir “pousser” ces filtres vers la source**, ce qu’on appelle le :  **Query folding**  **🛑 Or dans ton cas :**   * Les **CSV dans OneLake** **ne supportent pas** le query folding comme le ferait une base SQL. * Du coup, **Power BI télécharge chaque fichier en entier**, chaque fois. * Donc même si tu demandes une actualisation partielle (incrémentielle), **en pratique il recharge tout**, ce qui **explose la mémoire et échoue**.   **✅ Explication :**  **🔹 Contexte :**   * Tu utilises des **fichiers CSV** comme source. * Tu as configuré **l’actualisation incrémentielle** dans Power BI. * La **modèle échoue à l’actualisation** à cause d’un manque de ressources.   **🔹 Ce que ça implique :**  L’actualisation incrémentielle repose sur la **capacité du moteur Power BI** à **pousser les filtres de dates vers la source de données** — c’est ce qu’on appelle le **query folding**.  📌 **Query folding** = capacité à traduire une requête (ex. : "donne-moi les lignes de la dernière semaine") en une requête **filtrée en amont** (dans la source).  Mais avec **des fichiers CSV dans OneLake**, **Power BI ne peut pas appliquer de query folding** car :   * Les fichiers CSV ne sont **pas des sources relationnelles**, * Ils ne permettent pas de **filtrage serveur**, * Donc **Power BI doit charger toute la donnée**, **puis filtrer localement** → ce qui **consomme énormément de mémoire**.   **❌ Mauvaises réponses, expliquées :**   * **A. Query folding is occurring** → Faux : si c’était le cas, ce serait bénéfique, pas une cause d’échec. * **B. Only refresh complete days is selected** → Ne cause pas d’échec de rafraîchissement ; c’est un filtre logique. * **C. XMLA Endpoint is set to Read Only** → Cela bloque les écritures/modifications via XMLA, **mais pas le rafraîchissement classique via Power BI Service**. * **E. The delta type of the column used to partition the data has changed** → Cela pourrait poser problème **avec des sources Delta Lake**, mais tu travailles avec des fichiers **CSV**, donc ce n’est **pas applicable ici**.   **✅ Conclusion :**  **Réponse correcte : D. Query folding is NOT occurring** → Car **l’absence de query folding** force Power BI à **charger l’intégralité des fichiers CSV chaque jour**, rendant l’actualisation trop lourde. |

Page 20

[Microsoft - DP-600 - Page 21 | Examprepper](https://www.examprepper.co/exam/71/21)

## Question 101

You have a Fabric tenant that uses a Microsoft Power BI Premium capacity.

You need to enable scale-out for a semantic model.

What should you do first?

A.At the semantic model level, set Large dataset storage format to Off.

B.At the tenant level, set Create and use Metrics to Enabled.

C.At the semantic model level, set Large dataset storage format to On.

D.At the tenant level, set Data Activator to Enabled.

REPONSE

|  |
| --- |
| C.At the semantic model level, set Large dataset storage format to On.  To enable scale-out for a semantic model in Power BI Premium, the Large dataset storage format must first be turned On. |

## Question 102

You have a Fabric tenant that contains a warehouse. The warehouse uses row-level security (RLS).

You create a Direct Lake semantic model that uses the Delta tables and RLS of the warehouse.

When users interact with a report built from the model, which mode will be used by the DAX queries?

A.DirectQuery

B.Dual

C.Direct Lake

D.Import

REPONSE

|  |
| --- |
| **A. DirectQuery**.  **🔁 Récapitulatif:**   * On utilise un **semantic model en Direct Lake**. * Ce modèle **pointe vers un entrepôt Fabric** (**warehouse**). * Cet entrepôt applique une **Row-Level Security (RLS)**.   ➡️ Dans ce cas précis, **le mode Direct Lake ne peut pas appliquer la RLS dynamiquement**. ➡️ **Power BI bascule automatiquement le traitement des requêtes DAX en mode DirectQuery** pour faire respecter la sécurité. |

## Question 103

You have a Fabric tenant that contains a complex semantic model. The model is based on a star schema and contains many tables, including a fact table named Sales.

You need to create a diagram of the model. The diagram must contain only the Sales table and related tables.

What should you use from Microsoft Power BI Desktop?

A.data categories

B.Data view

C.Model view

D.DAX query view

### REPONSE

|  |
| --- |
| C.Model view |

## Question 104

HOTSPOT –

Case Study CONTOSO

You need to refresh the Orders table of the Online Sales department. The solution must meet the semantic model requirements.

What should you include in the solution?

A.an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the maximum value of the OrderID column in the destination lakehouse

B.an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the minimum value of the OrderID column in the destination lakehouse

C.an Azure Data Factory pipeline that executes a dataflow to retrieve the minimum value of the OrderID column in the destination lakehouse

D.an Azure Data Factory pipeline that executes a dataflow to retrieve the maximum value of the OrderID column in the destination lakehouse

### REPONSE

|  |
| --- |
| **C. an Azure Data Factory pipeline that executes a dataflow to retrieve the maximum value of the OrderID column in the destination lakehouse**  ✅ **Justification** : Pour minimiser le nombre de lignes ajoutées lors de l’actualisation, il faut charger uniquement les nouvelles commandes. Un **dataflow** permet de lire la valeur **maximale de OrderID** dans le lakehouse, ce qui permet de filtrer les nouvelles lignes à importer. C’est une approche efficace, maintenable et bien adaptée à Fabric. |

## Question 105

You have a Fabric tenant that contains a semantic model. The model uses Direct Lake mode.

You suspect that some DAX queries load unnecessary columns into memory.

You need to identify the frequently used columns that are loaded into memory.

What are two ways to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

A.Use the Analyze in Excel feature.

B.Use the Vertipaq Analyzer tool.

C.Query the $System.DISCOVER\_STORAGE\_TABLE\_COLUMN\_SEGMENTS dynamic management view (DMV).

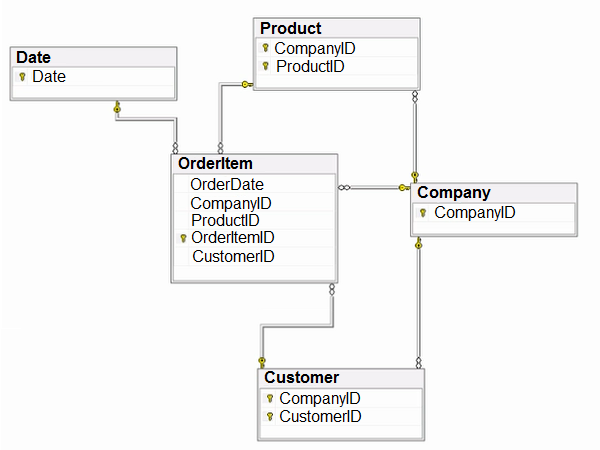
D.Query the DISCOVER\_MEMORYGRANT dynamic management view (DMV).

### REPONSE

|  |
| --- |
| ✅ **B. Use the VertiPaq Analyzer tool** ✅ **C. Query the $System.DISCOVER\_STORAGE\_TABLE\_COLUMN\_SEGMENTS dynamic management view (DMV)**  **✅ Explanation:**  You are using **Direct Lake mode** in Microsoft Fabric, and you want to **identify which columns are unnecessarily being loaded into memory** (i.e., which ones are contributing to memory usage during DAX queries).  **✅ B. VertiPaq Analyzer**   * A powerful tool (typically used via DAX Studio) that allows you to **analyze the internal structure** of your semantic model. * It helps you understand **column sizes, usage, and compression**, and detect **inefficient column usage**. * Very effective for identifying **unused or bloated columns**.   **✅ C. $System.DISCOVER\_STORAGE\_TABLE\_COLUMN\_SEGMENTS DMV**   * This DMV shows **how column segments are loaded into memory**. * You can see which **columns are accessed**, how frequently, and which segments are **materialized** in memory. * Essential for diagnosing memory usage patterns and optimizing queries. |
| **🔧 Comment utiliser un DMV ?**  Tu peux les interroger :   * Avec **DAX Studio** (outil externe pour Power BI) * Ou avec l’onglet **DAX Query View** dans Power BI Desktop (dans certaines versions) * Ou via des outils comme **SQL Server Management Studio (SSMS)** pour des modèles Analysis Services |

## Question 106

You have the source data model shown in the following exhibit.



Voici une représentation sous forme de tableau des tables, de leurs colonnes, et des relations entre elles :

You have the source data model shown in the following exhibit.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table Source** | **Colonnes** | **Colonne Source** | **Type de Clé** | **Table Cible** | **Colonne Cible** | **Type de Clé** |
| **OrderItem (Fact Table)** | OrderDate, CompanyID, ProductID, OrderItemID( PK) , Customer ID | OrderDate | Clé Étrangère | **Date** | Date | Clé Primaire |
| **OrderItem (Fact Table)** | OrderDate, CompanyID, ProductID, OrderItemID( PK) , Customer ID | ProductID | Clé Étrangère | **Product** | ProductID | Clé Primaire |
| **OrderItem (Fact Table)** | OrderDate, CompanyID, ProductID, OrderItemID( PK) , Customer ID | CompanyID | Clé Étrangère | **Company** | CompanyID | Clé Primaire |
| **OrderItem (Fact Table)** | OrderDate, CompanyID, ProductID, OrderItemID( PK) , Customer ID | CustomerID | Clé Étrangère | **Customer** | CustomerID | Clé Primaire |
| ***Product*** | *CompanyId (PK), ProductId (PK)* | *CompanyID* | *Clé Primaire* | ***Company*** | *CompanyID* | *Clé Primaire* |
| ***Customer*** | *CompanyId (PK), CustomerID (PK)* | *CompanyID* | *Clé Primaire* | ***Company*** | *CompanyID* | *Clé Primaire* |

The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.

You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer.

What should you include in the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Answer Area**

The relationship between OrderItem and Product must be based on: \_\_Choice 1 ? \_\_\_

1. The ProductID column
2. Both the CompanyID and the ProductID columns
3. A new key that combines the CompanyID and ProductID columns

The Company entity must be: \_\_Choice 2 ? \_\_\_

1. Omitted
2. Denormalized into the Product entity only
3. Denormalized into the Customer and Product entities

### REPONSE

|  |  |
| --- | --- |
| **Choix 1 : La relation entre OrderItem et Product doit être basée sur…**  ✅ **Both the CompanyID and the ProductID columns**  **Justification :**   * La table **Product** possède une **clé primaire composée** de **CompanyID et ProductID**. * **ProductID seul n'est pas unique**, donc la jointure correcte entre **OrderItem** et **Product** doit utiliser **les deux colonnes** pour garantir l'intégrité des données. * **Dans Fabric, une seule relation active est autorisée**, donc une relation devra être **active**, et l’autre devra être **inactive**, gérée via DAX (voir ci-dessous).   **Choix 2 : L’entité Company doit être…**  ✅ **Denormalized into the Customer and Product entities**  **Justification :**   * La table **Company** est référencée par **Product** et **Customer**, mais **elle n'apporte pas directement de valeur analytique** dans les requêtes. * **Dénormaliser** signifie **intégrer** les attributs de **Company** directement dans **Product et Customer**, évitant ainsi une jointure supplémentaire. * Cela **simplifie les requêtes**, **améliore les performances** et **permet une analyse plus fluide** dans Fabric ou Power BI. * Ainsi, les tables **Customer** et **Product** incluront les attributs nécessaires de **Company**, permettant une analyse sans dépendance explicite à cette table.   **🚨 Contraintes dans Microsoft Fabric**  ✔ **Une seule relation active est permise** → Comme **Product** a une clé composite (**CompanyID et ProductID**), on doit **choisir** laquelle sera active. ✔ **L’autre relation devra être inactive et activée avec DAX** → **USERELATIONSHIP** permet d’activer la relation inactive dans des mesures spécifiques.  **Exemple de mesure DAX pour activer la relation inactive (CompanyID)**  Si la relation active est sur **ProductID**, alors pour utiliser **CompanyID**, on crée une mesure comme ceci :   |  | | --- | | Total Sales with CompanyID =  CALCULATE(  SUM(OrderItem[SalesAmount]),  USERELATIONSHIP(OrderItem[CompanyID], Product[CompanyID])  ) | |
| ***💡 Alternative : Clé substitut unique Si tu veux éviter les relations multiples, une autre option est de créer une nouvelle colonne clé qui combine CompanyID et ProductID :***   |  | | --- | | ProductKey = CONCATENATE(Product[CompanyID], "-" , Product[ProductID]) |   **🚫 Pourquoi ce n’est pas la solution privilégiée ?**  **1️⃣ Ajout d’une colonne inutile**   * **La table Product a déjà une clé composite (CompanyID, ProductID) définie comme clé primaire.** * **Créer une clé artificielle ne change pas la relation logique, mais ajoute une redondance inutile dans le modèle.** * **Les performances ne sont pas améliorées par l’ajout d’une clé combinée.**   **2️⃣ Fabric est optimisé pour les relations sur plusieurs colonnes**   * **Microsoft Fabric supporte les relations sur plusieurs colonnes, donc il n’est pas nécessaire d’unifier les clés en une seule.** * **La gestion des relations via USERELATIONSHIP en DAX permet déjà d’activer la relation inactive au besoin.**   **3️⃣ Risque d’impact sur l’intégrité des données**   * **En concaténant CompanyID et ProductID dans une colonne unique, on perd la séparation logique des entités.** * **Cela pourrait poser des problèmes en cas de filtres ou de modifications du modèle à long terme.**   **✅ Pourquoi garder la relation sur les deux colonnes est meilleur ?**   * **Respecte la clé composite existante sans ajout de structure artificielle.** * **Permet une relation active et une inactive gérée dynamiquement via DAX (USERELATIONSHIP).** * **Garde une meilleure clarté du modèle dimensionnel, facilitant l’analyse.** |
| **💡 Conclusion : La meilleure approche est d’utiliser CompanyID et ProductID comme relation principale, avec une relation inactive gérée en DAX, plutôt que de créer une clé artificielle qui n’apporte aucune simplification réelle et peut compliquer les filtres dans les analyses.** |

## Question 107

You have a Fabric tenant that contains a semantic model named Model1. Model1 uses Import mode. Model1 contains a table named Orders. Orders has 100 million rows and the following fields.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data type** | **Description** |
| OrderID | Integer | Column imported from the source |
| OrderDateTime | Date/Time | Column imported from the source |
| Quantity | Integer | Column imported from the source |
| Price | Decimal | Column imported from the source |
| TotalSalesAmount | Decimal | Calculated column that Multiplies Quantity and Price |
| TotalQuantity | Integer | Measure |

You need to reduce the memory used by Model1 and the time it takes to refresh the model.

Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A.Split OrderDateTime into separate date and time columns.

B.Replace TotalQuantity with a calculated column.

C.Convert Quantity into the Text data type.

D.Replace TotalSalesAmount with a measure.

### REPONSE

|  |
| --- |
| **✅ Explication:**  **✔️ A. Split OrderDateTime into separate date and time columns**   * Raison : Les **colonnes DateTime** ont une **cardinalité élevée** (chaque valeur est souvent unique à la seconde près), ce qui augmente fortement la **taille mémoire** dans Power BI/Fabric. * En **séparant en deux colonnes** (Date, Time), on réduit la cardinalité de chaque colonne => **meilleure compression** => **réduction mémoire et rafraîchissement plus rapide**.   **✔️ D. Replace TotalSalesAmount with a measure**   * Raison : TotalSalesAmount est une **colonne calculée** (multiplication de deux colonnes importées). * Elle est **stockée**, donc prend de la **mémoire** inutilement. * La **remplacer par une mesure** (calculée à la demande) **réduit la taille du modèle** et **accélère le rafraîchissement**. |

## Question 108

You have a Fabric tenant that contains a semantic model.

You need to prevent report creators from populating visuals by using implicit measures.

What are two tools that you can use to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

A.Microsoft Power BI Desktop

B.Tabular Editor

C.Microsoft SQL Server Management Studio (SSMS)

D.DAX Studio

### REPONSE

|  |  |
| --- | --- |
| Pour **empêcher l'utilisation de mesures implicites** (comme les **agrégations automatiques** sur les colonnes numériques), vous devez **désactiver les agrégations par défaut** sur les colonnes du modèle sémantique.  **✔️ A. Power BI Desktop**   * Vous pouvez aller dans **Model view**, sélectionner une colonne, et définir « Hide in report view » * Vous pouvez aller dans **Report** **view**, sélectionner une colonne, et définir **"Aucune agrégation par défaut"** (Default summarization = None). * Cela empêche Power BI de créer automatiquement des mesures implicites à partir de cette colonne.   **✔️ B. Tabular Editor**   * Tabular Editor permet d’interdire l’utilisation des mesures implicites en modifiant les propriétés du modèle sémantique. * Permet de modifier directement les métadonnées du modèle, et de définir :  |  | | --- | | Csharp :  column.SummarizeBy = AggregateFunction.None; |   Très utile pour les modèles complexes ou automatisés. |

## Question 109

You have a Fabric tenant that contains two lakehouses.  
  
You are building a dataflow that will combine data from the lakehouses.  
The applied steps from one of the queries in the dataflow is shown in the following exhibit.

|  |  |
| --- | --- |
| **Image representant Query settings, step Filtered rows sélectionnée et actions possibles**  **Query settings**  **>Properties**  **Name:**  Customers1  **Entity type**  Custom  **>Applied steps ( \* Query Folding Indicators )**  - Source ( FY )  - Navigation 1  - Capitalized each word (FN)  - Appended query (U)  - Changed column type  - Added custom  - **Filtered rows** : ( selected) -> Possible Actions :   |  | | --- | | Edit settings  Rename  Delete  Delete until end  Insert step after  Move before  *Move after ( inactive )*  Extract previous...  *View data source query ( inactive )*  *View query plan ( inactive )*  Properties | |
|  |

**Legende :**

FY : The folding indicator tells you that the query up to this step is evaluated by the data source.

FN : The not-folding indicator tells you that some part of the query up to this step is evaluated outside the data source. You can compare it with the last folding indicator, if there is one, to see if you can rearrange your query to be more performant.

U : Unknown indicators represent an absence of a query plan, either due to an error or attempting to run the query plan evaluation on something other than a table (such as a record, list, or primitive).

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

**Answer Area**

[Answer choice] of the transformation steps in the query will fold.

1 - All

2 - None

3 - Some

The Added custom step will be performed in [answer choice].

1 - each lakehouse's query engine

2 - the Microsoft Power Query engine

3 - the source lakehouse query engine

### REPONSE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| | **Statement** | **Answer Choice** | | --- | --- | | [\_\_\_] of the transformation steps in the query will fold. | **Some** | | The Added custom step will be performed in [\_\_\_]. | **the Microsoft Power Query engine**   * Since the "Added custom" step comes **after** the last folding step and has **no folding indicator**, it will be **evaluated outside the data source**. * That means it will be executed by **Power Query**, not by the lakehouse's query engine. | |

## Question 110

You have a Fabric tenant that contains a lakehouse named Lakehouse’. Lakehouse1 contains a table named Tablet.

You are creating a new data pipeline.

You plan to copy external data to Table’. The schema of the external data changes regularly.

You need the copy operation to meet the following requirements:

Replace Table1 with the schema of the external data.

Replace all the data in Table1 with the rows in the external data.

You add a Copy data activity to the pipeline.

What should you do for the Copy data activity?

A.From the Source tab, add additional columns.

B.From the Destination tab, set Table action to Overwrite.

C.From the Settings tab, select Enable staging.

D.From the Source tab, select Enable partition discovery.

E.From the Source tab, select Recursively.

### REPONSE

**( Attention aux erreurs de nommage dans la consigne sur le nom de la table dans le Lakehouse : ( Table / Tablet / Table’ ) )**

|  |
| --- |
| **B. From the Destination tab, set Table action to Overwrite.**  Explanation:  Overwrite: This option completely replaces the destination table (Table1) with the source data, including the schema. This means that the structure of the table (schema) and the data will be updated to match exactly with that of the external data. |

Page 23

[Microsoft - DP-600 - Page 23 | Examprepper](https://www.examprepper.co/exam/71/23)

## Question 111

You have a Fabric tenant that contains a lakehouse.

You plan to query sales data files by using the SQL endpoint. The files will be in an Amazon Simple Storage Service (Amazon S3) storage bucket.

You need to recommend which file format to use and where to create a shortcut.

Which two actions should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A.Create a shortcut in the Files section.

B.Use the Parquet format

C.Use the CSV format.

D.Create a shortcut in the Tables section.

E.Use the delta format.

### REPONSE

|  |
| --- |
| **B. Use the Parquet format**  bien que l'interrogation SQL soit possible sur tous ces formats, Parquet est souvent préféré pour les analyses de données en raison de son efficacité et de ses performances, tandis que Delta est utilisé pour des scénarios nécessitant des fonctionnalités avancées comme le versionnage et les transactions. Les fichiers CSV, bien que simples, ne sont pas optimisés pour les requêtes analytiques complexes.  **D. Create a shortcut in the Tables section**  **Explication** : Créer un raccourci dans la section Tables vous permet de traiter les fichiers comme des tables SQL, ce qui est idéal pour effectuer des requêtes SQL directement sur les données. |

## Question 112

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a subfolder named Subfolder1 that contains CSV files.

You need to convert the CSV files into the delta format that has V-Order optimization enabled.

What should you do from Lakehouse explorer?

A.Use the Load to Tables feature.

B.Create a new shortcut in the Files section.

C.Create a new shortcut in the Tables section.

D.Use the Optimize feature.

### REPONSE

|  |
| --- |
| * + **The concept of 'sub-folder' is a contextual detail that does not affect the decision regarding the conversion method.**   **A. Use the Load to Tables feature.**  **Explanation:**  In Microsoft Fabric's Lakehouse Explorer, to convert CSV files into the **Delta format** with **V-Order optimization**, you should:   1. **Use the "Load to Tables" feature** from the Files section of the Lakehouse. 2. This feature allows you to load CSV files into managed Delta tables. 3. Once the data is loaded into Delta tables, **V-Order optimization** can be applied automatically or manually using the **Optimize** feature later if needed.   **Why not the others?**   * **B. Create a new shortcut in the Files section**: This only links to external data; it doesn't convert formats. * **C. Create a new shortcut in the Tables section**: Similar to B, it links to existing tables but doesn't perform conversion. * **D. Use the Optimize feature**: This is used **after** the data is already in Delta format, not for converting CSV to Delta. |

## Question 113

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains an unpartitioned table named Table1.

You plan to copy data to Table1 and partition the table based on a date column in the source data.

You create a Copy activity to copy the data to Table1.

You need to specify the partition column in the Destination settings of the Copy activity.

What should you do first?

A.From the Destination tab, set Mode to Append.

B.From the Destination tab, select the partition column.

C.From the Source tab, select Enable partition discovery.

D.From the Destination tabs, set Mode to Overwrite.

### REPONSE

|  |
| --- |
| **D. From the Destination tab, set Mode to Overwrite.**  **Explanation:**  To specify a **partition column** in the **Destination settings** of a **Copy activity** in Microsoft Fabric (or Azure Data Factory/Synapse Pipelines), you must first:   * **Set the write mode to "Overwrite"**. This is because partitioning is only supported when overwriting the destination table. The "Append" mode does not allow specifying partition columns. * **~~A. Set Mode to Append~~**~~: This does not support partitioning.~~   + **~~B. Select the partition column: You can’t do this until the mode is set to Overwrite.~~** * **~~C. Enable partition discovery in Source~~**~~: This is for reading partitioned data, not writing it.~~ |

## Question 114

Case Study CONTOSO

## Which syntax should you use in a notebook to access the Research division data for Productline1?

## A.spark.read.format(“delta”).load(“Tables/productline1/ResearchProduct”)

## B.spark.sql(“SELECT \* FROM Lakehouse1.ResearchProduct ”)

## C.external\_table(‘Tables/ResearchProduct)

## D.external\_table(ResearchProduct)

### REPONSE

|  |
| --- |
| **A. spark.read.format("delta").load("Tables/ResearchProduct")**   1. **spark.read.format("delta")**   Cela indique que tu veux lire des données au **format Delta Lake**.  C’est important car Fabric reconnaît automatiquement les fichiers Delta dans les **shortcuts** comme des tables exploitables.   1. **.load("Tables/ResearchProduct")**   Le chemin "Tables/ResearchProduct" fait référence à un **raccourci (shortcut)** qui a été ajouté dans la **section "Tables"** du Lakehouse.  Même si c’est un raccourci vers un stockage externe (comme Azure Data Lake Gen2), **Fabric le traite comme une table** si les données sont au format Delta ou Parquet.  Ce chemin est **interprété comme un alias logique**, pas comme un chemin de fichier brut. |

## Question 115

You have source data in a folder on a local computer.

You need to create a solution that will use Fabric to populate a data store. The solution must meet the following requirements:

Support the use of dataflows to load and append data to the data store.

Ensure that Delta tables are V-Order optimized and compacted automatically.

Which type of data store should you use?

A.a lakehouse

B.an Azure SQL database

C.a warehouse

D.a KQL database

### REPONSE

* + Remarques :

1. **"Data store"** = un système de stockage de données, comme un **lakehouse**, un **warehouse**, une **base de données SQL**, etc.
2. **V-Order** est une **technique d’optimisation du stockage** utilisée dans Microsoft Fabric (et Delta Lake dans certains cas) pour améliorer les performances de lecture des tables Delta.

|  |
| --- |
| **A. Un lakehouse**  **🎯 Indices décisifs dans la question :**   1. **« Use of dataflows to load and append data »** → Les **dataflows Gen2** dans Microsoft Fabric **chargent les données dans un lakehouse**, pas dans un warehouse ou une base SQL. 2. **« Delta tables »** → Le **format Delta** est **nativement pris en charge uniquement dans les lakehouses** dans Fabric. 3. **« V-Order optimized and compacted automatically »** → L’**optimisation V-Order** et la **compaction automatique** sont des **fonctionnalités spécifiques aux lakehouses** dans Fabric.   **🧠 Astuce pour s’en souvenir :**  Quand on a dans une question :   * **Dataflows** * **Delta format** * **V-Order** * **Compaction automatique**   👉 **Pense immédiatement à "lakehouse"** dans le contexte de Microsoft Fabric |

## Page 24

## Question 116

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code.

|  |
| --- |
| **df.write.partitionBy(“year”, “month”, “day”).mode(“overwrite”).parquet(“Files/SalesOrder”)** |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

|  |  |
| --- | --- |
| **Statements** | **Answer : Y/N** |
| The results will form a hierarchy of folders for each partition key |  |
| The resulting file partitions can be read in parallel across multiple nodes |  |
| The resulting file partitionscill use file compression |  |

### REPONSE

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Statement** | **Answer** | **Explication** | | --- | --- | --- | | **The results will form a hierarchy of folders for each partition key** | **Yes** | Le partitionnement par "year", "month", "day" crée une structure de dossiers imbriqués comme year=2025/month=06/day=03/. | | **The resulting file partitions can be read in parallel across multiple nodes** | **Yes** | Chaque partition est stockée dans un dossier distinct, ce qui permet une lecture parallèle efficace dans un environnement distribué comme Spark. | | **The resulting file partitions will use file compression** | **No** | Par défaut, parquet() n’active **pas automatiquement la compression**. Il faut spécifier un codec (ex. .option("compression", "snappy")) pour cela. | |

## Question 117

You have a Fabric workspace named Workspace1 that contains a data flow named Dataflow1 contains a query that returns the data shown in the following exhibit (in powerquery)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **VendorID** | **2015-01-01** | **2015-01-02** | **2015-01-03** | **2015-01-04** | **2015-01-05** |
| 1 | 16 | 15 | 0 | 9 | 12 |
| 2 | 20 | 15 | 7 | 17 | 20 |

You need to transform the data columns into attribute-value pairs, where columns become rows.

You select the VendorID column.

Which transformation should you select from the context menu of the VendorID column?

A.Group by

B.Unpivot columns

C.Unpivot other columns

D.Split column

E.Remove other columns

### REPONSE

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **C. Unpivot other columns**   |  |  |  | | --- | --- | --- | | **VendorID** | **Attribute** | **Value** | | 1 | 2015-01-01 | 16 | | 1 | 2015-01-02 | 15 | | 1 | 2015-01-03 | 0 | | ... | ... | ... | |

## Question 118

You have a Fabric tenant that contains a data pipeline.

You need to ensure that the pipeline runs every four hours on Mondays and Fridays.

To what should you set Repeat for the schedule?

A.Daily

B.By the minute

C.Weekly

D.Hourly

### REPONSE

|  |
| --- |
| **C. Weekly 🡪 pour choisir les jours !**  **✅ Explication :**  On veut que le pipeline s’exécute **toutes les 4 heures**, **uniquement les lundis et vendredis**.  Voici comment cela se traduit dans les options de planification :   * **"Repeat"** doit être défini sur **Weekly**, car tu veux **choisir des jours spécifiques** (lundi et vendredi). * Ensuite, on pourra configurer :   + Les **jours de la semaine** (Monday, Friday).   + Une **fréquence répétée toutes les 4 heures** dans ces journées. |

## Question 119

You have a Fabric tenant that contains a warehouse.

Several times a day, the performance of all warehouse queries degrades. You suspect that Fabric is throttling the compute used by the warehouse.

What should you use to identify whether throttling is occurring?

A.the Capacity settings

B.the Monitoring hub

C.dynamic management views (DMVs)

D.the Microsoft Fabric Capacity Metrics app

### REPONSE

|  |
| --- |
| **D. the Microsoft Fabric Capacity Metrics app**  Cet outil te permet de :   * Visualiser l’utilisation de la capacité (CPU, mémoire, etc.). * Identifier les périodes de throttling (ralentissement dû à une surcharge). * Analyser les tendances de consommation par workspace, par type de ressource, etc   Lol .. |

## Question 120

You have a Fabric workspace that uses the default Spark starter pool and runtime version 1.2.

You plan to read a CSV file named Sales\_raw.csv in a lakehouse, select columns, and save the data as a Delta table to the managed area of the lakehouse. Sales\_raw.csv contains 12 columns.

You have the following code.

|  |
| --- |
| from pyspark.sql.functions import year  ( spark  .red => erreur de frappe , il s’agit de .read  .format("csv")  .option("heade", "true") => erreur de frappe , il s’agit de "header"  .load("Files/sales\_raw.csv")  .select('SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice')  .withColumn("Year", year("OrderDate"))  .write  .partitionBy('Year')  .saveAsTable("sales")  ) |

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |
| --- | --- |
| **Statements** | **Answer Y/N** |
| The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales\_raw.csv. |  |
| Removing the partition will reduce the execution time of the query. |  |
| Adding inferSchema='true' to the options will increase the execution time of the query. |  |

### REPONSE

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Statement** | **Answer** | | The Spark engine will read only the selected columns   * Spark **lit d’abord toutes les colonnes** du fichier CSV lors du .load(). * Le .select(...) intervient **après** le chargement. * Donc, **les 12 colonnes sont lues**, même si seulement 4 sont sélectionnées ensuite. | No | | Removing the partition will reduce execution time   * Dans certains ca… « Peut » * La partition par "Year" implique :   + Un **shuffle** des données.   + Une **écriture dans plusieurs dossiers**. * Si le volume de données est faible ou modéré, **ne pas partitionner** peut **réduire le temps d’exécution**. * Donc, **oui**, supprimer la partition peut améliorer les performances dans certains cas. | Yes | | Adding inferSchema='true' increases execution time   * inferSchema=true force Spark à **scanner les données** pour **déduire les types de colonnes**. * Cela **ajoute un coût de traitement initial**, surtout pour les gros fichiers CSV. * Donc, **oui**, cela augmente le temps d’exécution. | Yes | |

Page 25

[Microsoft - DP-600 - Page 25 | Examprepper](https://www.examprepper.co/exam/71/25)

## Question 121

You have a Fabric tenant that contains a warehouse.

A user discovers that a report that usually takes two minutes to render has been running for 45 minutes and has still not rendered.

You need to identify what is preventing the report query from completing.

Which dynamic management view (DMV) should you use?

A.sys.dm\_exec\_requests

B.sys.dm\_exec\_sessions

C.sys.dm\_exec\_connections

D.sys.dm\_pdw\_exec\_requests

### REPONSE

|  |
| --- |
| **D.sys.dm\_pdw\_exec\_requests**  Cette DMV permet de :   * Suivre **l’état d’exécution des requêtes**. * Identifier les requêtes **en cours, bloquées ou longues**. * Voir des détails comme le **temps d’exécution**, le **statut**, et les **requêtes SQL** elles-mêmes. |

------------------------- FIN --------------------------------------------------------------------------------

# ANNEXES

## 📂 Accéder aux données dans Lakehouse, Warehouse, Eventhouse

| **Artefact** | **Description** | **Accès** |
| --- | --- | --- |
| **Lakehouse** | Stockage de fichiers (Parquet/Delta), orienté big data | Notebook, SQL Endpoint, Power BI |
| **Warehouse** | Entrepôt SQL relationnel, optimisé pour BI | T-SQL, Power BI, Pipelines |
| **Eventhouse** | Stockage d’événements en temps réel, basé sur Kusto | KQL, Dashboards temps réel |

## 📚 Définitions et cas d’usage

| **Type** | **Description courte** | **Cas d’usage principal** |
| --- | --- | --- |
| **Lakehouse** | Fusion d’un data lake et d’un entrepôt | Big data, machine learning, stockage brut |
| **Warehouse** | Entrepôt de données relationnel performant | Rapports Power BI, modèles sémantiques |
| **Eventhouse** | Plateforme d’analyse en temps réel | Streaming, alertes, surveillance temps réel |

## ⚖️ Tableau comparatif des modèles de données dans Fabric

| **Critère** | **Lakehouse** | **Warehouse** | **Eventhouse** |
| --- | --- | --- | --- |
| **Format** | Fichiers Delta / Parquet | Tables SQL relationnelles | Streaming / événements (KQL) |
| **Latence** | Moyenne (batch / streaming léger) | Faible à moyenne | Très faible (millisecondes) |
| **Volume** | Très grand (pétaoctets) | Moyen à grand | Moyen à très grand |
| **Fréquence** | Batch, micro-batch, parfois streaming | Batch / semi-temps réel | Temps réel continu |
| **Utilisateurs** | Data engineers, data scientists | Analysts, décideurs | Analysts temps réel, IoT teams |
| **Outils** | Notebooks, PySpark, SQL Endpoint | T-SQL, Power BI | KQL, dashboards temps réel |
| **Performance BI** | Moyenne (via modèles Power BI) | Excellente | Bonne sur données fraîche |

## 🧩 Résumé rapide – Quel outil pour quelle tâche ?

| **Tâche** | **Outil recommandé** |
| --- | --- |
| Nettoyage ETL simple | Dataflow Gen2 |
| Ingestion complexe & planifiée | Pipeline |
| Stockage de données historiques massives | Lakehouse |
| BI rapide sur données structurées | Warehouse |
| Analyse de données temps réel | Eventhouse |

## BASE DIAGRAMME

**🔧 Structure conceptuelle du diagramme Fabric**

**🧱 Conteneurs principaux (espaces logiques dans Fabric)**

* **OneLake** (lac de données unifié)
  + Contient : *Lakehouse, Warehouse, Eventhouse*
* **Workspaces Fabric**
  + Contiennent : artefacts, notebooks, pipelines, Power BI, etc.

**🧩 Artefacts (composants principaux)**

**1. Lakehouse**

* **Type** : Stockage fichiers structurés/semi-structurés (Delta/Parquet)
* **Relations** :
  + **Push** depuis Pipelines, Notebooks, Dataflows
  + **Pull** par Notebooks, SQL Endpoint, Power BI, ML
* **Outils liés** : PySpark, Notebooks, Pipelines, Power BI
* **Cas d’usage** : Stockage brut, big data, ML, ingestion massive
* **Conteneur** : OneLake

**2. Warehouse**

* **Type** : Base relationnelle optimisée BI
* **Relations** :
  + **Push** depuis Pipelines, Dataflows
  + **Pull** par Power BI, T-SQL
* **Outils liés** : Power BI, SQL Query Editor, Pipelines
* **Cas d’usage** : Modèle sémantique, reporting rapide
* **Conteneur** : OneLake

**3. Eventhouse**

* **Type** : Ingestion d’événements en temps réel
* **Relations** :
  + **Push** par Event Streams (IoT, Kafka, etc.)
  + **Pull** par Dashboards temps réel (Power BI)
* **Outils liés** : KQL, Dashboards Power BI en streaming
* **Cas d’usage** : Monitoring, alerting, IoT
* **Conteneur** : OneLake

**🔁 Outils / Services transverses**

* **Power BI** : Consomme Warehouse, Lakehouse (via modèle), Eventhouse
* **Notebooks** : Interagit avec Lakehouse (PySpark, SQL), pipelines
* **Pipelines** : Orchestration et push vers Lakehouse, Warehouse
* **Dataflows Gen2** : Préparation ETL simple vers Warehouse
* **Event Streams** : Flux d’entrée vers Eventhouse

**🔀 Flux (sens push/pull)**

| **Source** | **Cible** | **Type** | **Exemple** |
| --- | --- | --- | --- |
| Pipeline | Lakehouse | Push | Ingestion planifiée |
| Pipeline | Warehouse | Push | Chargement structuré |
| Event Stream | Eventhouse | Push | Flux IoT ou Kafka |
| Lakehouse | Notebook | Pull | Analyse data scientist |
| Lakehouse | Power BI | Pull | Modèle via SQL Endpoint |
| Warehouse | Power BI | Pull | DirectQuery / import |
| Eventhouse | Dashboard temps réel | Pull | KQL live queries |

## Query Folding indicator

Le query Folding c’est :

Hey base de données ! Tu peux faire ces étapes toi-même, directement chez toi, avant de m’envoyer les résultats ?”

The Folding query is:

"Hey database! Can you do these steps yourself, right at home, before sending me the results?"

| **Indicateur** | **Icon** | **Description** |
| --- | --- | --- |
| **Pliage (PO)** |  | L’indicateur de repli vous indique que la requête jusqu’à cette étape est évaluée par la source de données. |
| **Pas de (PN) repli** |  | L’indicateur de non-repli vous indique qu’une partie de la requête jusqu’à cette étape est évaluée en dehors de la source de données. Vous pouvez le comparer avec le dernier indicateur de repli, s’il en existe un, pour voir si vous pouvez réorganiser votre requête pour qu’elle soit plus performante. |
| **Peut se (P?) replier** |  | Les indicateurs de repli peuvent être rares. Ils signifient qu’une requête « peut » se replié. Ils indiquent que le repli ou non est déterminé au moment de l’exécution, lors de l’extraction des résultats de la requête et que le plan de requête est dynamique. Ces indicateurs ne s’affichent en général qu’avec des connexions ODBC ou OData. |
| **Opaque (O)** |  | Les indicateurs opaques vous indiquent que le plan de requête résultant n’est pas concluant pour une raison quelconque. Il indique généralement qu’il existe une table « constante » vraie, ou que cette transformation ou connecteur n’est pas prise en charge par les indicateurs et l’outil de plan de requête. |
| **Inconnu (I)** |  | Les indicateurs inconnus représentent une absence d’un plan de requête, soit en raison d’une erreur ou d’une tentative d’exécution de l’évaluation du plan de requête sur une autre table (par exemple, un enregistrement, une liste ou une primitive). |

| **Indicator** | **Icon** | **Description** |
| --- | --- | --- |
| **Folding (FY)** |  | The folding indicator tells you that the query up to this step is evaluated by the data source. |
| **Not folding (FN)** |  | The not-folding indicator tells you that some part of the query up to this step is evaluated outside the data source. You can compare it with the last folding indicator, if there is one, to see if you can rearrange your query to be more performant. |
| **Might fold (F ?)** |  | Might fold indicators are uncommon. They mean that a query "might" fold. They indicate either that folding or not folding is determined at runtime, when pulling results from the query, and that the query plan is dynamic. These indicators likely only appear with ODBC or OData connections. |
| **Opaque (O)** |  | Opaque indicators tell you that the resulting query plan is inconclusive for some reason. It generally indicates that there's a true "constant" table, or that that transform or connector isn't supported by the indicators and query plan tool. |
| **Unknown (U)** |  | Unknown indicators represent an absence of a query plan, either due to an error or attempting to run the query plan evaluation on something other than a table (such as a record, list, or primitive). |