2 DISCUSSIONS2DISCUSSIONS

Disable fullscreen

QUESTION 12 OF 65

**Which of the following statements is true about the version: 2 at the start of a model or source yml file?**

Choose only ONE best answer.

**A**

It is required for all yml files

**B**

It is only for model files

**C**

It is only for source files

**D**

It is not required for all yml files

**This answer is incorrect. The correct answer is 'A'**Answer: A  
  
Explanation:  
  
Model And Source Yml Files Always Start With `Version: 2` Because It Allowed The Structure To Be More Extensible.  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Getting-Started-Dbt-Core

QUESTION 13 OF 65

**A company wants to use dbt to track changes in their sales data. They have only specific columns they want to track for snapshot, some of the columns don't seem important right now. What is the recommended approach to capture this data ?**

Choose only ONE best answer.

**A**

Only include columns that are immediately useful

**B**

Use a "select \*" statement to capture all columns

**C**

Only include columns that are specified in the dbt\_project.yml file

**D**

Use the "source" function to select columns

**This answer is incorrect. The correct answer is 'B'**Answer: B) Use a "select \*" statement to capture all columns.  
  
This is the recommended approach when using snapshots in dbt. By including all columns, even if they don't seem useful at the moment, you can ensure that the snapshot is complete and provides a solid foundation for further analysis and processing.  
  
Reference: https://docs.getdbt.com/docs/build/snapshots

QUESTION 20 OF 65

**When you run dbt models, what will dbt do if the target schema specified does not exist?**

Choose only ONE best answer.

**A**

It will create the schema for you.

**B**

It will prompt you to create the schema manually.

**C**

It will stop the model execution and throw an error.

**D**

It will use the default schema specified in the config file.

**This answer is incorrect. The correct answer is 'A'**Correct Answer: A. It Will Create The Schema For You.  
  
Explanation:  
When Running Dbt Models, Dbt Will Check For The Target Schema Specified Exists,If The Schema Does Not Exist, Dbt Will Create It For You.  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Getting-Started-Dbt-Core

*QUESTION 21 OF 65*

**A data engineer is working on an incremental model in dbt and has specified multiple columns to identify each row uniquely. However, the incremental model run is failing with a null value error. What could be the reason for this error?**

*Choose only* ONE best answer.

**A**

The data engineer has specified the columns as a list instead of a string expression.

**B**

The data engineer has specified the columns as a string expression instead of a list.

**C**

The columns specified in the list contain null values.

**D**

The incremental model materialization is not appropriate for the database.

**This answer is incorrect. The correct answer is 'C'**Answer: C) The columns specified in the list contain null values.  
  
Explanation: When specifying multiple columns to identify each row uniquely, it's important to make sure that each column in the list doesn't contain any null values. If any of the columns contain null values, the incremental model run may fail with a null value error.  
  
Reference: https://docs.getdbt.com/docs/build/incremental-models

QUESTION 26 OF 65

**A marketing company wants to track changes to their customer demographics data over time. They want to be able to identify records that have been updated or deleted, and to keep a historical record of all changes. Which snapshot strategy should they use in their dbt configuration?**

Choose only ONE best answer.

**A**

Timestamp strategy

**B**

Check strategy

**C**

Both timestamp and check strategies

**D**

None of the Above

**This answer is incorrect. The correct answer is 'C'**Answer: C) Both timestamp and check strategies.  
  
The timestamp strategy is designed to detect changes in a record by comparing its timestamp to the previous snapshot. This makes it a good choice for tracking changes over time. The check strategy, on the other hand, is better suited for detecting deleted records and updates. By using both strategies, the marketing company can ensure that they have a comprehensive view of all changes to their customer demographics data over time.  
  
Reference: https://docs.getdbt.com/docs/build/snapshots

**Mani is a data scientist working on a dbt project. She wants to define a variable that can be accessed by all the packages in the project. Which type of variable should she define?**

Choose only ONE best answer.

**A**

Global variable declaration in the dbt\_project.yml file.

**B**

Package-scoped variable declaration in the dbt\_project.yml file.

**C**

Variable's default argument.

**D**

Variable defined with --vars command line argument.

**This answer is incorrect. The correct answer is 'A'**

Answer: A

Explanation: In dbt, variables defined in the global variable declaration in the dbt\_project.yml file are globally scoped and can be accessed by all the packages in the project. Therefore, define the variable as a global variable in the dbt\_project.yml file.

The order of precedence for variable declaration is as follows (highest priority first):

1. The variables defined on the command line with --vars.
2. The package-scoped variable declaration in the dbt\_project.yml file
3. The global variable declaration in the dbt\_project.yml file.
4. The variable's default argument (if one is provided).

Reference: <https://docs.getdbt.com/docs/build/project-variables>

QUESTION 29 OF 65

**What are the arguments passed to the model() function when running a dbt project with dbt run --select python\_model?**

Choose only ONE best answer.

**A**

The model name and the session object

**B**

The dbt Core class and the session object

**C**

The dbt object and the data platform's connection object

**D**

The dbt object and the session object

**This answer is correct.**

Answer: D

Explanation: When running a dbt project with dbt run --select python\_model, dbt will prepare and pass in both arguments (dbt and session) to the model() function.

* **dbt**: A class compiled by dbt Core, unique to each model, enables you to run your Python code in the context of your dbt project and DAG.
* **session**: A class representing your data platform’s connection to the Python backend. The session is needed to read in tables as DataFrames, and to write DataFrames back to tables. In PySpark, by convention, the SparkSession is named spark, and available globally. For consistency across platforms, we always pass it into the model function as an explicit argument called session.

Reference: <https://docs.getdbt.com/docs/build/python-models>

### QUESTION 34 OF 65

**What does the "unused model configurations" warning message in dbt indicate?**

Choose only ONE best answer.

**A**

The configurations have been nested under the project name

**B**

The configurations are trying to be applied to a directory that doesn't exist

**C**

The configurations are missing certain parameters

**D**

The configurations are not being properly applied to the models

**This answer is correct.**Answer: B. The Configurations Are Trying To Be Applied To A Directory That Doesn'T Exist  
  
Explanation:  
When Youy Get An "Unused Model Configurations" Error Message In Dbt This Means, You Might Have Forgotten To Nest Your Configurations Under Your Project Name, Or You Might Be Trying To Apply Configurations To A Directory That Doesn'T Exist.  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Getting-Started-Dbt-Core

### QUESTION 35 OF 65

**Which of the following is an example of a Jinja statement used for control flow?**

Choose only ONE best answer.

**A**

{{ my\_variable }}

**B**

{% for item in my\_list %} ... {% endfor %}

**C**

{% my\_macro() %}

**D**

{{ my\_variable.upper() }}

**This answer is correct.**Answer: B) {% for item in my\_list %} ... {% endfor %}  
  
Explanation: Jinja statements are denoted by curly braces with a percent sign inside {% %}. They are used for control flow, like setting up for loops and if statements, and for defining macros. The example given in option B is a for loop statement, which iterates over the items in a list and executes the code inside the loop for each item.  
  
Options A, C, and D are not examples of Jinja statements used for control flow, but rather for outputting variables, calling macros, and manipulating strings, respectively.  
  
Reference: https://docs.getdbt.com/docs/build/jinja-macros

QUESTION 37 OF 65

**You and your team are using dbt to manage your data models. You need to switch between different locations of the profiles.yml file for each run, but you do not want to type the --profiles-dir option every time you run a dbt command. This is because your team works with different data sources, and the profiles are stored in different locations for each data source. What should you do to avoid having to type the --profiles-dir option every time?**

Choose only ONE best answer.

**A**

Use the --profiles option when executing a dbt command

**B**

Use the DBT\_PROFILES\_DIR environment variable to change the default location of the profiles.yml file

**C**

Update the dbt config file to point to a custom location for the profiles.yml file

**D**

Create a shell alias for the dbt command to include the --profiles-dir option

**This answer is incorrect. The correct answer is 'B'**

Answer: B. Use the DBT\_PROFILES\_DIR environment variable to change the default location of the profiles.yml file.

Explanation: Option B is the best way to change the location of the profiles.yml file without having to type the --profiles-dir option every time you run a dbt command. It allows you to change the default.

There are multiple ways to direct dbt to a different location for your profiles.yml file:

1. Use the --profiles-dir option when executing a dbt command: dbt run --profiles-dir path/to/directory. If using this method, the --profiles-dir option needs to be provided every time you run a dbt command.
2. Use the DBT\_PROFILES\_DIR environment variable to change the default location: specifying this environment variable overrides the directory that dbt looks for your profiles.yml file in. You can specify this by running: export DBT\_PROFILES\_DIR=path/to/directory.

Reference: <https://docs.getdbt.com/docs/getting-started/connection-profiles>

**How can using dbt sources help with version control of data sources?**

Choose only ONE best answer.

**A**

By using a schema .yml file in the dbt project's /models subfolder

**B**

By using a version control system for database tables

**C**

By using a version control system for SQL views

**D**

By using a version control system for stored procedures

**This answer is incorrect. The correct answer is 'A'**Answer: A  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Learning-More/Refactoring-Legacy-Sql

**In DBT, which of the following options can be used in combination with the dbt test command to run a test on a single model?**

Choose ALL answers that apply.

* **A**

--select

* **B**

-s

* **C**

--model

* **D**

-m

**This answer is incorrect. The correct answer is 'A' 'B'**Answer: A and B (--Select And -S).  
  
In Dbt We Can Choose To Test One Model At A Time, To Run A Test A Single Model We Use Methods Very Similar To Running A Single Model: Use The --Select Flag (Or -S Flag), Followed By The Name Of The Model: 1. Dbt Test --Select Model\_Name 2. Dbt Test -S Model\_Name.  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Getting-Started/Building-Your-First-Project/Test-And-Document-Your-Project

### QUESTION 43 OF 65

**You have been working with dbt and BigQuery to manage your data pipelines. Recently, you added a nested column to a table in BigQuery and set the on\_schema\_change parameter to "sync\_all\_columns" in your dbt configuration. However, when running the dbt command, you noticed that the new nested column is not added to the table. What could be the cause of this issue?**

Choose only ONE best answer.

**A**

The on\_schema\_change feature in dbt does not track changes to nested columns.

**B**

You will need to rerun the model using the --full-refresh flag.

**C**

You will need to drop the relation from your data platform and rerun.

**D**

There is an issue with the data being updated.

**This answer is correct.**

Correct answer: A) The on\_schema\_change feature in dbt does not track changes to nested columns.

Explanation: The on\_schema\_change feature in dbt only tracks changes to the top-level columns, and not changes to nested columns. This means that if you add, remove, or modify a nested column in BigQuery, it won't be detected as a schema change even if the on\_schema\_change is set. Therefore, the most likely cause of this issue is that the on\_schema\_change feature in dbt does not track changes to nested columns.

Reference: <https://docs.getdbt.com/docs/build/incremental-models>

### QUESTION 44 OF 65

**Jenny is a data engineer working on a dbt project that involves several SQL queries which do not fit into the mold of dbt models. She wants to use dbt functionality like {{ ref(...) }} to select from models in an environment-agnostic way. Which directory in the dbt project should she store the SQL files to enable this feature?**

Choose only ONE best answer.

**A**

models/

**B**

analyses/

**C**

macros/

**D**

tests/

**This answer is correct.**Answer: B  
  
Explanation: The analysis functionality of dbt allows data teams to version control and collaborate on data transformations in a dbt project, even if the SQL statement doesn't fit into the mold of a dbt model. To use this feature, any .sql files found in the analyses/ directory of a dbt project will be compiled, but not executed, and dbt functionality like {{ ref(...) }} can be used to select from models in an environment-agnostic way.  
  
Reference: https://docs.getdbt.com/docs/build/analyses

**Jenny is a data analyst working on a dbt project that requires a package to address a specific problem area. What is the recommended method to specify the package in dbt?**

Choose only ONE best answer.

**A**

Use a private repository to store the package and specify it in the dbt\_project.yml file.

**B**

Use the latest patch version from a specific minor release when specifying the package in the dbt Hub.

**C**

Use the latest major version when specifying the package in the dbt Hub.

**D**

Use the prerelease version when specifying the package in the dbt Hub.

**This answer is incorrect. The correct answer is 'B'**Answer: B  
  
Explanation: To specify a package in dbt, the recommended method is to use the dbt Hub, which is a registry for dbt packages that uses semantic versioning. When using dbt Hub, you should specify a version for the package and pin it to the latest patch version from a specific minor release. This method is useful for handling duplicate dependencies that dbt may encounter.   
  
Reference: https://docs.getdbt.com/docs/build/packages

### QUESTION 48 OF 65

**How can you ensure that dbt will be able to run with the user credentials you use in your target?**

Choose only ONE best answer.

**A**

By ensuring the user has appropriate privileges

**B**

By assigning a single set of database credentials for all dbt users

**C**

By assigning a separate set of database credentials for each dbt user

**D**

By having different privileges for different data warehouses

**This answer is correct.**Correct Answer: A  
  
Explanation:  
dbt Recommends That Each Dbt User Has Their Own Set Of Database Credentials, Including A Separate User For Production Runs Of Dbt – This Helps Debug Rogue Queries, Simplifies Ownerships Of Schemas, And Improves Security. To Ensure The User Credentials You Use In Your Target Allow Dbt To Run, You Will Need To Ensure The User Has Appropriate Privileges. While The Exact Privileges Needed Varies Between Data Warehouses, At A Minimum Your User Must Be Able To: Read Source Data ,Create Schemas¹  and Read System Tables  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Connection-Profiless

QUESTION 49 OF 65

**You and your team are working on a project that involves using dbt to manage your data pipelines. One of your top priorities is to ensure that sensitive information, such as database credentials, are not compromised. In order to achieve this, you need to take the necessary precautions to prevent unauthorized access to this information. What should you and your team avoid when working with dbt projects to ensure that sensitive information is not compromised?**

Choose ALL answers that apply.

* **A**

Storing profiles.yml within the dbt project

* **B**

Storing profiles.yml in a different version control system

* **C**

Checking credentials in the dbt profiles.yml into your documentation tool.

* **D**

Storing profiles outside of the dbt project

**This answer is incorrect. The correct answer is 'A'  'B' 'C'**

Answer: A, B, C. Avoid Checking credentials in the dbt profiles into version control.

Explanation: The dbt profiles are stored separately to dbt projects to avoid checking credentials into version control. Database credentials are extremely sensitive information and should never be checked into version control.

Reference: <https://docs.getdbt.com/docs/getting-started/dbt-core/>

### QUESTION 53 OF 65

**What is the purpose of running "dbt compile" from the command line?**

Choose only ONE best answer.

**A**

To run a dbt model and update the target database

**B**

To generate documentation for your dbt project

**C**

To view the compiled SQL generated by your dbt models

**D**

To run unit tests for your dbt models

**This answer is correct.**Answer: C) To view the compiled SQL generated by your dbt models  
  
Explanation: Running "dbt compile" from the command line is used to compile the SQL generated by your dbt models. This can be useful for verifying that your Jinja code is generating the correct SQL code and identifying any errors in your code. The compiled SQL files can be found in the "target/compiled/{project name}/" directory. It's recommended to use a split-screen in your code editor to view both files at once.  
  
  
Reference: https://docs.getdbt.com/docs/build/jinja-macros

**You are reviewing a dbt project that has a source defined in the schema.yml file. The source is in a database that is different from the default target database being used in the profiles.yml file. How can you reference thissource in a downstream model?**

Choose only ONE best answer.

**A**

Use the 'schema' and 'identifier' properties of the source schema.yml to specify the database

**B**

Use the 'database' property of the source schema.yml to specify the database

**C**

Use the 'name' property of the source schema.yml to specify the database

**D**

Use the 'table' property of the source schema.yml to specify the database

**This answer is incorrect. The correct answer is 'B'**

Answer: B

Explanation: When you define a source in the schema.yml file, you can use the database property to define the database for the source.

Option B is the correct way to reference a source with a different database in a downstream model. Option A, C, and D are incorrect because they do not address the need to use the database property.

Reference: <https://docs.getdbt.com/docs/build/sources>

**You and your team are working on a project using dbt to manage your data pipelines. One of the requirements of the project is to capture the current state of a data source at a specific point in time. To achieve this, you need to define a snapshot strategy in dbt. What is the purpose of defining a snapshot strategy in dbt?**

Choose only ONE best answer.

**A**

To specify which sources should be archived.

**B**

To define the models that should be built using dbt.

**C**

To provide a consistent view of the data even when the original data changes.

**D**

To specify how dbt should capture the current state of a data source at a specific point in time.

**This answer is correct.**

Answer: D

Explanation: The purpose of defining a snapshot strategy in dbt is to specify how dbt should capture the current state of a data source at a specific point in time. The snapshot strategy is defined in the snapshot config and can be set to check, time, or timestamp.

Option A is incorrect because defining which sources should be archived is done in the archive config. Option B is incorrect because defining models is done in the model config.

Option C is incorrect because providing a consistent view of the data is the benefit of snapshotting, not the purpose of defining a snapshot strategy.

Reference: <https://docs.getdbt.com/docs/build/snapshots>

QUESTION 56 OF 65

**In DBT, which of the following methods can be used to run a test on a single model?**

Choose ALL answers that apply.

* **A**

dbt test --select model\_name

* **B**

dbt test -s model\_name

* **C**

dbt test model\_name

* **D**

dbt run model\_name

**This answer is incorrect. The correct answer is 'A' 'B'**Answer: A, B (dbt Test --Select Model\_Name And dbt Test -S Model\_Name.  
  
In Dbt We Can Choose To Test One Model At A Time, To Run A Test A Single Model We Use Methods Very Similar To Running A Single Model: Use The --Select Flag (Or -S Flag), Followed By The Name Of The Model: 1. Dbt Test --Select Model\_Name 2. Dbt Test -S Model\_Name.  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Getting-Started/Building-Your-First-Project/Test-And-Document-Your-Project

**You are a data person working on a dbt project. You have a .sql file in the tests directory of your project that asserts that the total payment amount is positive. How does dbt determines the name of the test?**

Choose only ONE best answer.

**A**

dbt uses an internal hash key.

**B**

The name of the test is determined by the name of the .sql file

**C**

dbt uses a sequenctial name based on your dbt\_project.yml config.

**D**

dbt uses a sequenctial name based on your profiles.yml config.

**This answer is correct.**

Answer: B. The name of the test is determined by the name of the .sql file.

Explanation: Singular tests in dbt are defined in .sql files located in the tests directory specified in the test-paths configuration. The name of the test is determined by the name of the .sql file. Therefore, if the file is named assert\_total\_payment\_amount\_is\_positive.sql, the test is also called assert\_total\_payment\_amount\_is\_positive.

Reference: <https://docs.getdbt.com/docs/build/tests>

### QUESTION 61 OF 65

**Which of the following features does dbt provide to assist with handling mutable records in a data source when trying to reconstruct historic values?**

Choose only ONE best answer.

**A**

Snapshotting raw data for a point in time

**B**

Modifying records to match their historic values

**C**

Modifying records to match their historic values

**D**

Ignoring mutable records

**E**

Automatically updating records to their current values

**This answer is correct.**Answer: A. Snapshotting Raw Data For A Point In Time.  
  
Often, Records In A Data Source Are Mutable, In That They Change Over Time. This Can Be Difficult To Handle In Analytics If You Want To Reconstruct Historic Values. Dbt Provides A Mechanism To Snapshot Raw Data For A Point In Time, Through Use Of Snapshots.  
  
 Reference: Https://Docs.Getdbt.Com/Docs/Introduction

**A data engineer is building a dbt project and wants to ensure that the snapshots are complete and accurate. What is the recommended approach for including columns in the snapshot?**

Choose only ONE best answer.

**A**

Only include columns that are immediately useful

**B**

Use a "select \*" statement to capture all columns

**C**

Only include columns that are specified in the dbt\_project.yml file

**D**

Use the "source" function to select columns

**This answer is correct.**Answer: B) Use a "select \*" statement to capture all columns.  
  
This is the recommended approach when working with snapshots in dbt. By including as many columns as possible, you can ensure that the snapshot is complete and accurate, even if some columns don't seem useful at the moment. This approach provides a solid foundation for further analysis and processing, and ensures that the snapshot can be used effectively in the future.  
  
Reference: https://docs.getdbt.com/docs/build/snapshots

### QUESTION 64 OF 65

**What is the advantage of using Jinja in dbt SQL models over python models?**

Choose only ONE best answer.

**A**

Jinja allows for better performance of SQL models.

**B**

Jinja provides access to the project's context, which is not available to Python models.

**C**

Jinja allows for easier integration with external systems.

**D**

Jinja allows for easier implementation of constraints and validations on data.

**This answer is correct.**Answer: B  
  
Explanation:  
In dbt, project context refers to the configuration settings and project-level information that define the behavior and characteristics of a dbt project. The project context is generally defined in the dbt\_project.yml file, which is located in the root directory of the project.  
  
The advantage of using Jinja in dbt SQL models is that Jinja provides access to the project's context, which is not available to Python models. This allows for more flexibility and customization in the SQL code, as well as the ability to integrate external systems.  
  
  
Option A is incorrect because Jinja does not necessarily improve the performance of SQL models, and the performance depends on the specific use case and the complexity of the query.  
  
Option C is incorrect because Jinja does not have any direct impact on the integration with external systems.  
  
Option D is incorrect because constraints and validations on data can be implemented using SQL in dbt, using the test configuration block in your model file.  
  
Reference: https://docs.getdbt.com/docs/build/python-models

### QUESTION 65 OF 65

**What is the first step in migrating your code 1:1 into dbt?**

Choose only ONE best answer.

**A**

Initiating a "dbt compile"

**B**

Auditing the code

**C**

Moving the SQL code into a .sql file under the /models directory

**D**

Executing the query and materializing it as a table

**This answer is correct.**Answer: C  
  
Reference: Https://Docs.Getdbt.Com/Docs/Get-Started/Learning-More/Refactoring-Legacy-Sql