dbt Architect Certification Exam Study Guide

dbt Labs

How to use this study guide

This is the official study guide for the **dbt Architect Certification Exam** from the team at dbt Labs. While the guide suggests a sequence of courses and reading material, we recommend using it to supplement (rather than substitute) real-world use and experience with dbt.

The **exam overview** will provide high-level details on the exam format. We recommend being mindful of the number of questions and time constraints.

The **topic outline** will provide a clear list of the topics that are assessed on the exam. dbt Subject Matter Experts (SMEs) used this topic outline to write and review all of the exam items that you will find on the exam.

The **sample exam questions** provide an example of the format to expect for the exam. The types of questions you can expect on the exam include:

- Multiple-choice
- Fill-in-the-blank
- Matching
- Hotspot
- Build list
- Discrete Option Multiple Choice (DOMC)

The **learning path** will walk you through a suggested series of courses, readings, and documentation. We will also provide guidance on the types of experiences to build while using dbt on a real-world project.

Finally, the **additional resources** section will provide additional places to continue your learning.

Exam Overview

The **dbt Architect Certification Exam** assesses your ability to design secure, scalable dbt implementations, with a focus on environment orchestration, role-based access control, integrations with other tools, and collaborative development workflows aligned with best practices.

We recommend that you have SQL proficiency and at least six months of experience administrating an Enterprise dbt account. This foundational knowledge will help you navigate the exam confidently.

Logistics

- **Duration:** 2 Hours
- **Format & Registration: online proctored**
- **Length:** 65 questions
- **Passing Score**: 65% or higher. You will know your score immediately after completion of the exam.
- **Price:** $200*
- **Language:** Only English at this time
- **Certification Expiration:** The certification expires 2 years after the date awarded.
- **Supported browsers: Caveon Web browsers**

*The price applies only to the online proctored exam; there is no additional fee for Coalesce premium attendees. Discounts are available for dbt Labs SI Partners.

Scoring

The exam is scored on a point basis, with 1 point for each correct answer, and 0 points for incorrect answers. All questions are weighted equally.

An undisclosed number of unscored questions will be included on each exam. These are unmarked and indistinguishable from scored questions. Answers to these questions will be used for research purposes only, and will not count towards the score.

Retakes & Cancellations

If you do not pass the exam, you may schedule a retake. You will need to pay a registration fee for each retake. You can reschedule or cancel without penalty on Talview up to 24 hours before your scheduled exam. We will not issue refunds for no-shows.

Accommodations

Please contact **Talview** with any accommodation requests.

Topic Outline

The **dbt Architect Certification Exam** has been designed to assess the following topics and sub-topics.

Topic 1: Configuring dbt data warehouse connections

- Understanding how to connect the warehouse
- Configuring IP whitelist
- Creating and testing a connection for the project
- Authenticating through OAuth to access the data in dbt
- Adding Client ID and Secret for OAuth

Topic 2: Configuring dbt git connections

- Connecting the git repo to dbt
- Setting up integrations with git providers

Topic 3: Creating and maintaining dbt environments

- Understanding access control to different environments
- Determining when to use a service account
- Rotating key pair authentication via the API
- Understanding environment variables
- Creating a new dbt deployment environment
- Setting a default schema/dataset for environments
- Understanding custom branches and how to configure them to environments
- Configuring dbt to allow deferral to other environments

Topic 4: Creating and maintaining job definitions

- Set up a CI job with deferral
- Understanding steps within a dbt job
- Scheduling a job to run on schedule
- Implementing run commands in the correct order
- Creating a new dbt job
- Configuring optional settings such as environment variable overrides, threads, deferral, target name, and dbt version override
- Generating documentation on a job that populates the dbt Catalog page
- Configuring jobs to be triggered after other dbt jobs (job chaining)
- Configuring Advanced CI
- Configuring self-deferral
- Understanding when to use which type of job deferral

Topic 5: Configuring dbt security and licenses

- Creating service tokens for API access
- Assigning permission sets
- Creating license mappings
- Adding and removing users
- Adding SSO application for dbt Enterprise
- Creating and assigning RBAC

Topic 6: Setting up monitoring and alerting for jobs

- Setting up email notifications
- Using Webhooks for event-driven integrations with other systems

Topic 7: Setting up a dbt mesh and leveraging cross-project references

- Setting up additional dbt projects
- Understanding how environment types relate to cross-project references
- Utilizing model governance

Topic 8: Configuring and using dbt Catalog

- Using dbt Catalog to understand the current lineage, troubleshoot issues and optimise cost and performance
- Using dbt Catalog to find public models and cross-project references

Sample exam questions

Sample Exam Question 1

Why should you avoid using SSO-based connections for staging and production deployment environments in dbt?

A. **They rely on authentication sessions that can expire or require re-authentication, leading to unexpected job failures in automated deployments.**

B. **They cannot connect to cloud data warehouses due to incompatible authentication protocols.**

C. **They restrict the ability to manage environment variables required for production deployments.**

D. **They can only be used by users with the Account Admin role.**

Explanation: dbt supports authenticating via OAuth in Development environments only, where users can authorize their Development credentials using Single Sign-On. Once a user has authorized dbt with their data platform, via their identity provider, the user will receive (behind the scenes) an Access Token which can then be used to open a data platform connection and execute queries in the dbt Studio.

The lifetime of the token is dictated by the data platform integration. When a user's token expires, the user will need to re-authorize with the data platform to continue development in dbt. For this reason, it is best practice to utilize a service account for deployment credentials.

B is incorrect because it is technically possible to connect to your warehouse this way, it is just not advised, so dbt does not allow it.

C is incorrect because the method of authentication does not restrict the usage of environment variables.

D is incorrect because SSO OAuth can be used by any user with a Developer license, regardless of permissions granted.

Link to docs

Sample Exam Question 2

You have configured a snapshot's `target_schema` to be `prod_snapshot`.

You want to limit who can write to this schema.

How can you allow developers to read from the `prod_schema` in dev but allow read and write privileges for the production job?

A. **Use a `post-hook` to grant read and write privileges in the snapshot config.**

B. **Use a service account with read and write privileges in the deployment environment settings.**

C. **Use an `on-run-end` to grant read and write privileges in the snapshot config.**

D. **Assign the developer the Job Admin role in dbt.**

Explanation: This question boils down to the fact that different users or service accounts should have different permissions to operate on different objects in the database. Developers should have read access to `prod_schema`, but not write access to `prod_snapshot`. The production deployment credentials should have read and write to `prod_schema` and `prod_snapshot`. This is handled succinctly by utilizing different credentials for each environment, and granting the appropriate permissions within the data platform to each user / service account.

For answers A and C, utilizing a `post-hook` or `on-run-end` hook won't help, because the production deployment credentials need write access before it can create or alter the table in the `prod_snapshot` schema. You could grant privileges afterwards to certain groups, but that doesn't solve the ask at hand.

D is incorrect because assigning the dbt Job Admin role will allow the developer to create and edit jobs, runs, environment variables, and data warehouse configs, but that does not affect their personal access to objects in the data database.

Link to docs

Sample Exam Question 3

Inspect this environment variable configuration:

Key	Project Default	Production	Development
`DBT_ENV_VALUE`	`not_set`	`prod`	`dev`

Within a job there is an override to this variable with the value `finance`.

Which value will be used when running the dbt job if this job configuration is removed?

A. **the dbt project default `not_set`.**

B. **the default argument supplied to the `env_var` Jinja function in code.**

C. **the value assigned to the environment specified in the job description.**

D. **the job-level environment variable is required and cannot be removed.**

Explanation: The Project Default column is recommended when you want to supply a catch-all default or add a project-wide token or secret. Values set at the environmental level take priority over the project-level default value, and as seen in the above example, dbt will interpret the variable value differently in your Production vs. Development environment.

A job inherits the values set at the environment level, but a job-level override will take precedence for the environment in which it runs. If you remove the job-level override, it will default back to the environment variable, in this case, the value `prod`.

A is incorrect because Project Default is lower in the order of precedence than the environment itself.

B is incorrect because the Jinja function default argument is lower in precedence than both Project Default and the environment.

D is incorrect because the job-level environment variable is not required and can be removed.

Link to docs

Sample Exam Question 4

Congratulations! Your company has expanded to Europe.

As part of this expansion, IT plans to set up another data warehouse in the EU while leveraging the same Enterprise dbt account across regions.

What should you do to achieve this?

A. **purchase another account for the EU region.**

B. **configure a connection to the EU — no further action is required.**

C. **upgrade your account to Business Critical and establish a PrivateLink connection.**

D. **configure dbt Mesh Cross Platform to connect to the EU warehouse.**

Explanation: In a multi-Region Setup, it would be prudent to set up two separate PrivateLink endpoints in the regions in which you are established. You will then configure multiple projects to use the respective private endpoints. This helps you maintain data sovereignty (EU data stays in EU) and isolation (no data traverses public networks).

A is incorrect because you will not need to purchase another account.

B is incorrect, because there is further design and action required to make this work.

C is correct because you will need PrivateLink to meet data residency/compliance goals (e.g., GDPR) and this is only possible with Business Critical Enterprise plan or higher.

D is incorrect because you do not want to set up Cross Platform mesh, as you will want data in the EU to stay segregated from the other region's data.

Link to docs

Learning Path:

This is **not the only way** to prepare for the exam, but just one recommended path for an advanced dbt user to prepare for the exam. Each checkpoint provides a logical set of courses to complete, readings and documentation to digest, and real-world experience to seek out. We recommend this order, but things can be reorganized based on your learning preferences.

Checkpoint 0 - Prerequisites

dbt is a tool that brings together several different technical skills in one place. We recommend starting this path after you've obtained the badge from the **dbt Fundamentals course**.

Additionally, being aware of the **Analytics Development Lifecycle (ADLC)** provides an excellent implementation framework for architecting your dbt Architecture.

Checkpoint 1 - Setting up dbt Connections

Resources:

- Courses:
  - dbt and Databricks for Admins
  - dbt and BigQuery for Admins
  - dbt and Snowflake for Admins
  - dbt and Redshift for Admins
- Readings:
  - dbt Platform Architecture
  - Maximum override: Configuring unique connections in dbt Platform
  - Version control basics
  - Getting Started with git Branching Strategies and dbt
- Documentation:
  - About data platform connections
  - Supported data platforms
  - Access, Regions & IP addresses
  - Set up Snowflake OAuth
  - Set up Databricks OAuth
  - Set up BigQuery OAuth
  - Set up external OAuth

    - Connect to GitHub
    - Connect to GitLab
    - Connect to Azure DevOps
    - Connect with Git clone

- Experience
    - Understanding how to connect the warehouse including configuring IP whitelist
    - Authenticating through OAuth to access the data in dbt Studio
    - Connecting the git repo to dbt

Checkpoint 2 - Configuring and Managing Projects

Resources:

- Courses:
    - Advanced Deployment

- Readings:
    - dbt environments
    - More time coding, less time waiting: Mastering defer in dbt
    - To defer or to clone, that is the question

- Documentation:
    - About user access
    - About environment-level permissions
    - Environment Variables
    - Custom schemas
    - Custom target names
    - Customize dbt models database, schema and alias
    - Using threads
    - Deployment environments
    - How do I use the 'Custom Branch' settings in a dbt Environment?
    - Jobs in dbt Platform
    - Customizing CI/CD with custom pipelines
    - Get started with Continuous Integration tests
    - Advanced CI
    - Using defer in dbt Platform

- Experience
  - Understanding and configuring access control across different environments and projects
  - Determining when to use a service account and rotating key pair authentication via the API
  - Creating new dbt deployment environment with customisations using environment variables, custom branches, custom schemas/datasets and deferrals
  - Setting up different dbt jobs (scheduled, continuous integration, merge or via APIs)
  - Configuring Advanced CI

Checkpoint 3 - Security & Monitoring

Resources:

- Courses:
  - Webhooks
  - Role-based access control: dbt and Okta
  - SSO dbt and Entra ID
  - SSO for dbt Platform and Google Workspace
  - SSO for dbt Platform and Okta
- Readings:
  - Single sign-on (SSO) Overview
- Documentation:
  - Service account tokens
  - Users and licenses
  - Self-service Team account permissions
  - Enterprise permissions
  - About user access
  - Migrating to Auth0 for SSO
  - Set up SSO with SAML 2.0
  - Set up SSO with Okta
  - Set up SSO with Google Workspace
  - Set up SSO with Microsoft Entra ID
  - Set up SCIM
  - Job notifications
  - Webhooks for your jobs

- Experience
  - Creating Service tokens for API access
  - Assigning permission sets, creating license mappings and managing users
  - Adding SSO application for dbt enterprise, including creating and assigning role-based access control (RBAC)
  - Setting up email notifications
  - Using Webhooks for event-driven integrations with other systems

Checkpoint 4 - Governance and Discovery

Resources:

- Courses:
  - dbt Mesh
  - dbt Catalog

- Readings:
  - How Hybrid Mesh unlocks dbt collaboration at scale
  - Column-Level Lineage, Model Performance, and Recommendations: ship trusted data products with dbt Catalog
  - Implementing your mesh plan

- Documentation:
  - Intro to dbt Mesh
  - Project dependencies
  - Model access
  - Discover data with dbt Catalog

- Experience
  - Setting up additional dbt projects with cross-project references
  - Using dbt Catalog to understand the current lineage, troubleshoot issues and optimise cost and performance
  - Using dbt Catalog to find public models and cross-project references

Additional Resources:

- dbt Slack
  - #dbt-certification
  - #learn-on-demand
  - #advice-dbt-for-power-users
  - #dbt-deployment-and-orchestration
- If you are a dbt Labs partner or enterprise client, contact your partner manager or account team for additional benefits.
- Want to use email? Contact certification@dbtlabs.com