Certification Exam

☆

You have passed the quiz!

Your score:
54 of 75 Correct (72%)

Elapsed time:
59 minutes

75 of 75 questions answered

Hide Answers

Question 1: ✔ Correct answer
What is the purpose of the PythonOperator in Airflow?

☐ ~~To deploy Python packages to remote servers~~

☐ ~~To schedule Python scripts to run on a regular basis~~

☐ ~~To manage Python virtual environments for Airflow tasks~~

☑ **To execute Python code as a task in a DAG**

Question 2: ✖ Incorrect answer
Examine the DAG below:

```
 from airflow import DAG
 from airflow.operators.python_operator import PythonOperator
 from datetime import datetime

 with DAG(
     'basic_python_dag',
     schedule_interval='@daily',
     catchup=False
 ):

     python_task = PythonOperator(
         python_callable=lambda: print('Hello Learner!')
     )
```

Which of the following issues would cause this DAG to produce an error? (select all that apply)

☑ **The PythonOperator is missing a task_id parameter**

☑ ~~The value assigned to the python_callable parameter has an error~~

☐ ~~The tags parameter is missing~~

☑ **The start_date parameter is missing**

Question 3: ✔ Correct answer
What is the default timezone in Airflow?

☐ ~~The user's local timezone~~

☐ ~~EST~~

☑ **UTC**

☐ ~~AEST~~

Question 4: ✔ Correct answer
What are the core architectural components of Airflow? (select all that apply)

☑ **A web server**

☑ **A scheduler**

☑ **A metadata database**

☑ **An executor**

☐ ~~DAGs~~

☐ ~~A NoSQL database~~

Question 5: ✔ Correct answer
Once a DAG has a unique ID, the next step is to define when it will start being scheduled. Let's consider the scenario where the goal is to start scheduling the DAG as of 2023/01/01 . Which of the following parameters would accomplish this?

☐ ~~schedule_interval=datetime(2023, 1, 1)~~

☐ ~~start_date='2023/01/01'~~

☐ ~~start_date=datetime(1 , 1, 2023)~~

☑ **start_date=datetime(2023, 1 , 1)**

Question 6: ✘ Incorrect answer
What mode is considered best practice to use when a sensor's `poke_interval` parameter value is set to a very short duration (< 5 minutes)?

☑ ~~reschedule~~

☐ **poke**

Question 7: ✔ Correct answer
A member of a data team has a SQL script that aggregates data about their company's previous day's performance. They want to schedule the SQL script to run every day at midnight. Is this scenario a use case for Airflow?

☐ ~~No~~

☑ **Yes**

Question 8: ✔ Correct answer
Examine the DAG Below:

```
 from airflow import DAG
 from airflow.operators.bash import BashOperator
 from datetime import datetime, timedelta

 dag = DAG(
     'basic_bash_dag',
     schedule_interval='@daily',
     start_date=datetime(2023, 1, 1),
     catchup=False,
 )

 task_1 = BashOperator(
     task_id='task_1',
     bash_command='ls',
     dag=dag,
 )

 task_2 = BashOperator(
     task_id='task_2',
     bash_command='pwd',
     dag=dag,
 )

 task_1 >> task_2
```

Assuming the current date is **2023/01/04** , how many DAG runs will occur right after the DAG is triggered?

☐ 0

☐ 4

☑ **1**

☐ 2

Question 9: ✖ Incorrect answer

Assume a data team member has added **task_b** to the DAG below:

```
 from airflow import DAG
 from airflow.operators.bash_operator import BashOperator
 from datetime import datetime

 dag = DAG(
     'data_dag',
     start_date=datetime(2023, 1, 1),
     catchup=False,
 )

 task_a = BashOperator(
     task_id='task_a',
     bash_command='echo "Executing task A"',
     dag=dag,
 )

 #This task was added
 task_b = BashOperator(
     task_id='task_b',
     bash_command='echo "Executing task B"',
     dag=dag,
 )

 task_a >> task_b
```

The team member doesn't see the new task in the DAG on the Airflow UI. What could be the cause?

☐ **Airflow has a default delay of up to 30 seconds to detect a new task in a DAG.**

☐ ~~The schedule_interval parameter is missing~~

☐ ~~The task dependency setup is incorrect.~~

☑ ~~Airflow has a default delay of up to 5 minutes to detect a new task in a DAG.~~

Question 10: ✔ Correct answer
Examine the Airflow UI below:

	DAG ⇅	Owner ⇅	Runs ⓘ	Schedule	Last Run ⓘ	Next Run ⇅ ⓘ	Recent Tasks ⓘ
⬤	check_dag data-engineering-team	airflow	◯◯③◯◯	@daily ⓘ	2023-08-08, 16:08:58 ⓘ	2023-08-08, 00:00:00 ⓘ	◯◯◯◯◯③◯◯◯◯◯◯◯◯◯

Is it possible to determine which tasks have failed across all DAG runs of a single DAG using the "Recent Tasks" column?

☑ **No**

☐ ~~Yes~~

Question 11: ✔ Correct answer
Which of the following are valid values for scheduling a DAG in Airflow? (select all that apply)

☑ **Timedelta**

☑ **Cron**

☐ ~~Integers~~

☐ ~~Calendar~~

Question 12: ✔ Correct answer

A member of a data team writes a Python script that downloads data from an API and stores it in a database on a daily basis. Is this scenario a use case for Airflow?

☑ **Yes**

☐ ~~No~~

Question 13: ✔ Correct answer

Which of the following Airflow CLI commands is necessary to run to create the metadata database?

☐ ~~airflow db check~~

☑ **airflow db init**

☐ ~~airflow db start~~

☐ ~~airflow db reset~~

Question 14: ✔ Correct answer

What is the role of the executor in Airflow ?

☐ ~~It executes tasks.~~

☑ **It defines how tasks are executed and on which system.**

☐ ~~It deletes unwanted tasks.~~

☐ ~~It schedules tasks.~~

Question 15: ✔ Correct answer

Assume a DAG needs to process data one month before the value of the DAG's `start_date` parameter. Which of the following Airflow CLI commands is the best way to accomplish this?

☐ ~~airflow dags trigger~~

☑ **airflow dags backfill**

☐ ~~It is not possible in Airflow~~

☐ ~~airflow dags run~~

Question 16: ✔ Correct answer

What method allows an Airflow user to get data from an XCom?

☐ ~~By executing the method xcom_add~~

☑ **By executing the method xcom_pull**

Question 17: ✖ Incorrect answer

In a bank setting, a DAG is designed to process awaiting transfers. To accomplish this, the DAG should be triggered every 4 hours, starting from the date **datetime(2023, 1, 1)** . What value should be assigned to the **schedule_interval** DAG parameter to achieve this scheduling frequency?

☑ schedule_interval='0 4 * * *'

☐ schedule_interval='/4 * * * *'

☐ **schedule_interval=timedelta(hours=4)**

☐ schedule_Interval='4'

Question 18: ✔ Correct answer

What language is Airflow primarily written in?

☐ C++

☐ Java

☑ **Python**

☐ Ruby

Question 19: ✔ Correct answer

Examine the DAG below:

```
from airflow import DAG
from airflow.operators.bash_operator import BashOperator
from datetime import datetime

with DAG(
    'data_dag',
    schedule='@daily'
    catchup=False
):

    task_a = BashOperator(
        task_id='task_a',
        bash_command='echo "Executing task A"'
    )

    task_b = BashOperator(
        task_id='task_b',
        bash_command='echo "Executing task B"'
    )

  task_a >> task_b
```

Would this DAG be able to run properly?

☐ ~~Yes~~

☑ **No**

Question 20: ✔ Correct answer
After several weeks of testing, a data team has decided to use Airflow in their tech stack. Before getting started, they need to communicate with their DevOps team to set up an Airflow instance on Linux Ubuntu OS. However, the DevOps team needs to know which version of Python is currently running on the environment that the data team has been testing.

Which of the following Airflow CLI commands would quickly allow the team to know what version of Python the environment is running?

☐ ~~airflow standalone~~

☐ ~~airflow python version~~

☐ ~~airflow version~~

☑ **airflow info**

Question 21: ✔ Correct answer
Assume a scenario where a company is running a website, and a specific DAG should only be triggered when a customer hits a button on a form. This means that the DAG should not be scheduled. Which of the following values for the `schedule_interval` parameter of the DAG accomplishes this?

☑ **schedule_interval=None**

☐ ~~schedule_interval='@once'~~

☐ ~~schedule_interval='@never'~~

☐ ~~schedule_interval=''~~

Question 22: ✔ Correct answer
In which of the following places do Airflow users set up the dependencies between tasks in a DAG?

☐ ~~In a separate configuration file~~

☐ ~~In the Airflow UI~~

☐ ~~In a Python file in the 'includes' directory~~

☑ **In the DAG file itself**

Question 23: ✔ Correct answer
An Airflow variable was created using an environment variable and is not visible on the Airflow UI. Can it still be accessed/used in DAGs?

☑ **Yes**

☐ ~~No~~

Question 24: ✘ Incorrect answer
What types of data can be stored in an Airflow variable? (select all that apply)

☐ **Integers**

☑ **JSON**

☐ ~~Python Dataclasses~~

☐ **Strings**

Question 25: ✘ Incorrect answer
Which of the following is a valid way to create a data pipeline in Airflow? (select all that apply)

☐ **Using the DAG class and defining the dag parameter in an operator**

☑ ~~Creating a Python function called DAG~~

☑ **Using with DAG()**

☑ **Using the @dag decorator**

Question 26: ✔ Correct answer
Which of the following DAG `schedule_interval` parameter Cron values would cause a DAG to run every day at midnight? (select all that apply)

☑ **@daily**

☐ ~~0 0 * * 0~~

☑ **0 0 * * ***

☐ ~~0 * * * *~~

Question 27: ✔ Correct answer
A member of a data team is tasked with checking whether multiple data pipelines have failed every morning. If they have, then they have to rerun them manually. They are looking for a data tool that offers automated notifications when something goes wrong with the data pipelines. Is this scenario a use case for Airflow?

☐ ~~No~~

☑ **Yes**

Question 28: ✔ Correct answer
If a DAG has a `schedule_interval` parameter with the value of `@daily` and a `catchup` parameter with the value of `True`, how far back in time will Airflow try to catch up on missed DAG runs?

☑ **All days from the set `start_date` parameter value of the DAG**

☐ ~~One week~~

☐ ~~One day~~

☐ ~~One month~~

Question 29: ✔ Correct answer

What task dependency relationship results in the following DAG?
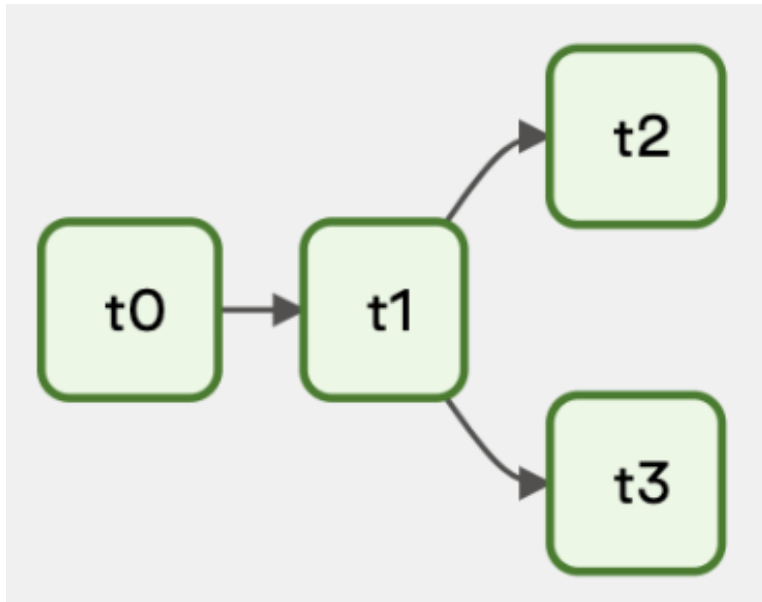
☐ ~~[t0, t1] >> [t2, t3]~~

☐ ~~[t0, t1] >> t2 >> t3~~

☐ ~~t0 >> t1 >> t2 >> t3~~

☑ **t0 >> t1 >> [t2, t3]**

Question 30: ✔ Correct answer

Assume a new DAG was just added to an Airflow instance. What is the default amount of time required for the new DAG to appear on the Airflow UI?

☐ ~~1 second~~

☐ ~~1 minute~~

☑ **5 minutes**

☐ ~~30 seconds~~

Question 31: ✘ Incorrect answer

Take a look at the code below from a `.env` file that creates a Snowflake connection:

```
AIRFLOW_CONN_SNOWFLAKE_CONN='snowflake://LOGIN:PASSWORD@/?account=xy12345&region=eu-centra
```

What is the connection Id?

☐ ~~CONN_SNOWFLAKE_CONN~~

☑ ~~AIRFLOW_CONN_SNOWFLAKE_CONN~~

☐ ~~SNOWFLAKE~~

☐ **SNOWFLAKE_CONN**

Question 32: ✔ Correct answer

What is the purpose of setting dependencies between tasks in Airflow?

☐ To increase the parallelism of task execution.

☐ To improve the performance of the Airflow scheduler.

☑ **To ensure that tasks are executed in the correct order.**

☐ To reduce the likelihood of task failures.

Question 33: ✔ Correct answer
Which of the following is a valid way to create a connection in Airflow? (select all that apply)

☑ **Airflow CLI**

☑ **Environment Variables**

☑ **Airflow UI**

☐ Inside of a XCom

Question 34: ✔ Correct answer
What is an Airflow provider?

☐ Operators that are part of core Airflow, like the PythonOperator and the BashOperator.

☐ A configuration in Airflow that allows users to control the number of tasks running at a time.

☐ A Docker image that installs Airflow.

☑ **A separate Python package that extends Airflow's functionality, in most cases by integrating with external systems or services.**

Question 35: ✘ Incorrect answer
What method allows an Airflow user to add data to XComs?

☐ By executing the method xcom_add

☑ By executing the method xcom_pull

☐ **By executing the method xcom_push**

☐ By providing the data to the xcom parameter of the DAG object

Question 36: ✘ Incorrect answer
Examine the DAG below:

```
 from airflow import DAG
 from airflow.operators.bash_operator import BashOperator
 from datetime import datetime

 # Define the DAG instance
 dag = DAG(
     'basic_bash_dag',
     start_date=datetime(2023, 1, 1),
     schedule_interval='@daily',
     catchup=False,
 )

 # Define the BashOperator tasks
 task_1 = BashOperator(
     task_id='task_1',
     bash_command='echo "Executing task 1"',
     dag=dag,
 )

 task_2 = BashOperator(
     task_id='task_2',
     bash_command='echo "Executing task 2"',
     dag=dag,
 )

 # Set task dependencies
 task_1 >> task_2
```

Which of the following ways allows the code to avoid defining a DAG object in all operators (`dag=dag`)? (select all that apply)

- ☐ **By defining tasks within a function decorated by @dag**
- ☐ ~~By creating a default_argument dictionary~~
- ☐ ~~By giving all tasks the same task_id~~
- ☑ **By instantiating the DAG object with the context manager with**

Question 37: ✔ Correct answer
What is the purpose of the `default_args` parameter in a DAG?

- ☐ ~~It specifies the schedule interval of a DAG.~~
- ☐ ~~It specifies the operator a DAG will use.~~
- ☐ ~~It specifies a dictionary of default parameters for at max one task in a DAG.~~
- ☑ **It specifies a dictionary of default parameters for all tasks in a DAG.**

Question 38: ✔ Correct answer

Assume there is a DAG with four tasks labeled t0, t1, t2, and t3. Is writing the DAG's task dependency as `t0 >> t1 >> t2 >> t3` equivalent to writing `t3 << t2 << t1 << t0`?

☐ ~~No~~

☑ **Yes**

Question 39: ✔ Correct answer

Assume an Airflow instance has a variable named `my_json` which stores a JSON value. Which of the following would fetch the variable in a DAG as a JSON object?

☐ ~~Variable.get(my_json)~~

☐ ~~Variable.retrieve('my_json', deserialize_json=True)~~

☐ ~~Variable.get('my_json')~~

☑ **Variable.get('my_json', deserialize_json=True)**

Question 40: ✔ Correct answer

What task dependency relationship results in the following DAG? (select all that apply)
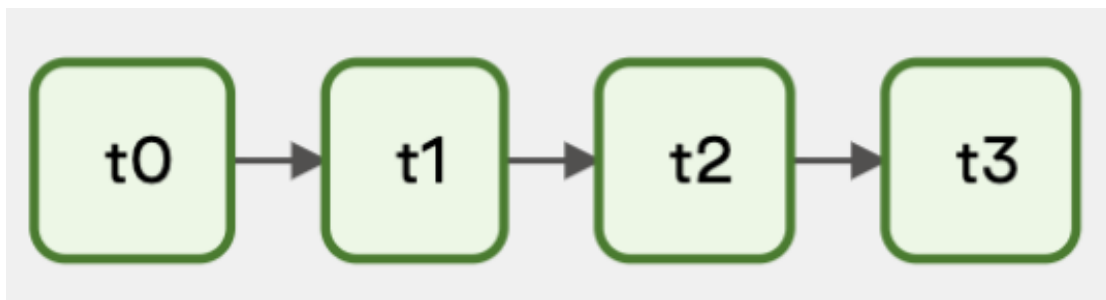
☐ ~~t0 >> [t1, t2, t3]~~

☐ ~~t0 >> [t1, t2] >> t3~~

☑ **t3 << t2 << t1 << t0**

☑ **t0 >> t1 >> t2 >> t3**

Question 41: ✘ Incorrect answer

What is a limitation of using XComs? (select all that apply)

☐ ~~XComs are the only way to share data between tasks.~~

☑ **The amount of data that can be given to an XCom is limited.**

☐ **There must be a defined DAG with at least one associated task to create an XCom.**

☑ ~~XComs can only be used with a maximum of two tasks in a DAG.~~

Question 42: ✔ Correct answer

Assume a data team member is building a connection in Airflow. They notice the provider they are trying to use (e.g., Presto) is missing from the dropdown in the UI (shown below). What should they do?

Conn Id *	presto

Conn Type *	Postgres ▲
	🔍
	~~Google Dataprep~~
	HTTP
	Hashicorp Vault
	IMAP
	Kubernetes Cluster Connection
	Mesos Framework ID
	MySQL
	Postgres

Description

Host

Schema

Login

Password

Port

Extra

Save 💾   ←

☐ ~~Restart Airflow~~

☑ **Find and install the provider package**

☐ ~~Find a different provider because if the provider is missing in the dropdown, then it is impossible to use it in Airflow~~

☐ ~~Reinstall Airflow~~

Question 43: ✔ Correct answer
A CEO asks their company's data team to create a dashboard that analyzes user behavior in the company's product. Which of the following best describes how Airflow can help create that dashboard?

☐ ~~Airflow can be used as the product's primary database for user information which powers the dashboard.~~

☐ ~~Airflow can be embedded directly into a product to analyze user click events and present the information as a dashboard.~~

☐ ~~Airflow can be used to build the analytics dashboard directly in the Airflow UI.~~

☑ **Airflow can be used to schedule a data pipeline that ingests user data into a data warehouse that powers the dashboard.**

Question 44: ✔ Correct answer
What is the typical journey of a task?

☐ ~~Queued → Scheduled → Running → Success/Failure~~

☑ **No status → Scheduled → Queued → Running → Success/Failure**

☐ ~~Scheduled → Queued → Running → Success/Failure~~

☐ ~~No status → Queued → Scheduled → Running → Success/Failure~~

Question 45: ✖ Incorrect answer
Assume there is a DAG with four tasks labeled t0, t1, t2, and t3. Would writing the DAG's task dependency as `[t0, t1] >> [t2, t3]` work?

☑ ~~Yes~~

☐ **No**

Question 46: ✔ Correct answer
What is the purpose of the `catchup` parameter in a DAG?

☐ ~~It determines the schedule interval for the DAG.~~

☐ ~~It determines whether the DAG should be scheduled to run in the future.~~

☐ ~~It determines the number of retries for failed tasks.~~

☑ **It determines whether the DAG should be run retroactively for missed intervals.**

Question 47: ✖ Incorrect answer
Examine the Airflow UI below:

What does the "Last Run" column represent?

☐ ~~The interval of time between each DAG run.~~

☐ **The start date/time of the latest DAG run.**

☑ ~~The end date/time of the most recent DAG run.~~

☐ ~~The date when the DAG is scheduled to run next.~~

Question 48: ✔ Correct answer
Which of the following best describes what a Sensor Operator does?

☐ ~~It provides various utility functions to support workflow management~~

☐ ~~It evaluates conditions and determines the path a workflow should take~~

☑ **It waits for a certain condition to be met before proceeding with a workflow**

Question 49: ✖ Incorrect answer

What Airflow CLI command runs a task without checking for dependencies or recording its state in the metadata database?

☑ airflow tasks start

☐ **airflow tasks test**

☐ airflow standalone

☐ airflow dags test

Question 50: ✔ Correct answer

A DAG was designed to fetch data from files coming from different data sources. The DAG has to be triggered every day at 7:00 AM. What value should be assigned to the `schedule_interval` DAG parameter to achieve this scheduling frequency?

☐ schedule_interval=timedelta(hours=6)

☑ **schedule_interval='0 7 * * *'**

☐ schedule_interval='7 * * * *'

☐ schedule_interval=timedelta(hours=7)

Question 51: ✖ Incorrect answer

What task dependency relationship results in the following DAG?
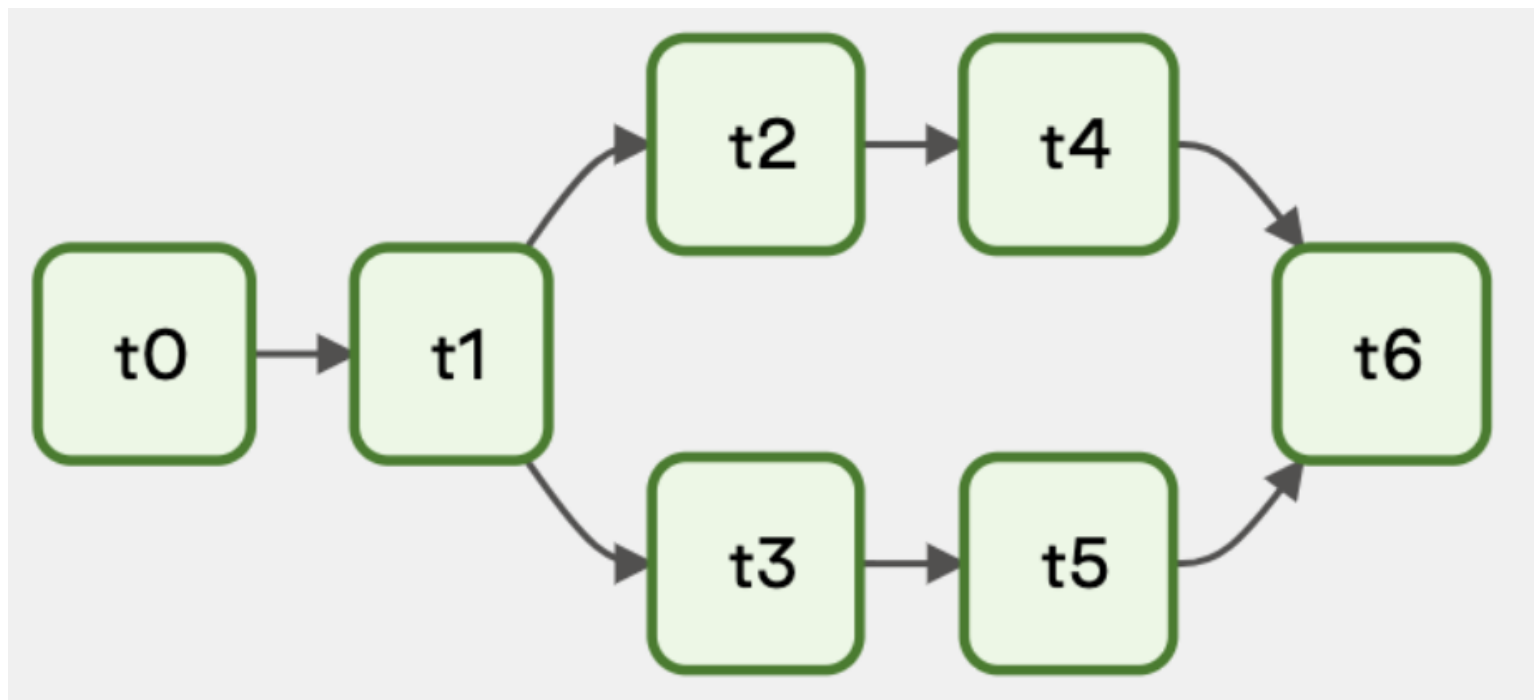
☐ t0 >> t1 >> [t2, t3] >> [t4, t5] >> t6

☐ [t0, t1] >> t2 >> t3 >> t4 >> [t5, t6]

☑ chain(t0, t1, [t2, t4], [t3, t5], t6)

☐ **chain(t0, t1, [t2, t3], [t4, t5], t6)**

Question 52: ✔ Correct answer
What is the purpose of an Airflow variable? (select all that apply)

☐ ~~To store large amounts of data~~

☐ ~~To store a XCom~~

☑ **To reuse a value across multiple DAGs**

☑ **To avoid hardcoding a value**

Question 53: ✔ Correct answer
Examine the Airflow error below:

> ❗ **DAG Import Errors (1)**   ⌃
>
> Broken DAG: [/usr/local/airflow/dags/test.py] Traceback (most recent call last):   ⌃
>   File "/usr/local/lib/python3.9/site-packages/airflow/models/baseoperator.py", line 1039, in dag
>   dag.add_task(self)
>   File "/usr/local/lib/python3.9/site-packages/airflow/models/dag.py", line 2328, in add_task
>   raise AirflowException("DAG is missing the start_date parameter")
> airflow.exceptions.AirflowException: DAG is missing the start_date parameter

Which of the following is likely to resolve this error?

☐ ~~Importing the correct DAG object~~

☐ ~~Changing the format of the start_date parameter~~

☑ **Adding a start_date parameter to the DAG**

☐ ~~There is not enough information from the error to solve the error~~

Question 54: ✔ Correct answer
Examine the Airflow UI component below:

Recent Tasks ⓘ

What does this component represent?

- ☐ ~~The status of all future DAG runs.~~
- ☐ ~~Only failed, success, or skipped task statuses from all active or most recent DAG runs.~~
- ☑ **All of the statuses of tasks from all active or most recent DAG runs.**
- ☐ ~~The status of all previous DAG runs.~~

Question 55: ✖ Incorrect answer
Assume a member of a data team working with Airflow added a new DAG file into the `/dags` folder but reports that it doesn't show up on the Airflow UI. Which of the following could be causing the issue? (select all that apply)

- ☑ **The code in the DAG does not contain any words, including "airflow" or "dag".**
- ☑ **Airflow has a default delay of 5 minutes before Airflow acknowledges the new DAG file.**
- ☑ ~~The Airflow instance must be restarted for the DAG to show up on the UI.~~
- ☑ **The .airflowignore file has the name of the file in it.**

Question 56: ✖ Incorrect answer
What is the default value for the `start_date` parameter of a DAG?

- ☐ **The parameter has to be explicitly set and has no default value.**
- ☐ ~~The current date and time of a users location (e.g., New York).~~
- ☐ ~~The current UTC date and time.~~

☑ ~~Midnight on January 1, 1900.~~

Question 57: ✔ Correct answer
Which of the following is the correct way to define an Airflow variable using an environment variable with the name `customer_folder` and the value 'my_folder'?

☐ ~~CUSTOMER_FOLDER='my_folder'~~
☐ ~~AIRFLOW_CUSTOMER_FOLDER='my_folder'~~
☐ ~~AIRFLOW_VARIABLE_CUSTOMER_FOLDER='my_folder'~~
☑ **AIRFLOW_VAR_CUSTOMER_FOLDER='my_folder'**

Question 58: ✖ Incorrect answer
Assume an Airflow variable with the name `airtable_api_key` was created via the Airflow UI. Will the variable's value be hidden on the Airflow UI?

☐ **Yes**
☑ ~~No~~

Question 59: ✔ Correct answer
Take a look at the code below from a `.env` file that creates a PostgreSQL connection:

```
AIRFLOW_CONN_POSTGRES='my-conn-type://login:password@host:port/schema?param1=val1&param2=v
```

Is this the correct way to create a connection using an environment variable?

☑ **Yes**
☐ ~~No~~

Question 60: ✔ Correct answer
In a typical Airflow setup, what folder is parsed by the scheduler for new DAG files?

☐ ~~The /airflow folder~~
☑ **The /dags folder**
☐ ~~The /plugins folder~~
☐ ~~The /include folder~~

Question 61: ✔ Correct answer
Assume there is a DAG with defined task dependencies running in an Airflow instance. What would be the most helpful Airflow UI view for an overview of the dependencies?

☐ ~~Grid view~~
☐ ~~DAGs view~~
☐ ~~Gant view~~

☑ **Graph view**

Question 62: ✔ Correct answer
What mode is considered best practice to use when a sensor's `poke_interval` parameter value is set to greater than 5 minutes?

☐ ~~poke~~

☑ **reschedule**

Question 63: ✔ Correct answer
Assume a DAG experienced a bug, and a data team needed to pause the DAG to fix it. The bug took 3 days to fix. The DAG is scheduled to be triggered every day and has a DAG `start_date` parameter value of `2023/01/05 00:00`. The current date after the fix is `2023/01/08 10:00`. If the DAG is unpaused; how many running DAG runs will occur if the DAG `catchup` parameter value is set to `True`?

☐ ~~2~~

☑ **3**

☐ ~~4~~

☐ ~~1~~

Question 64: ✖ Incorrect answer
Examine the `SqlSensor` below:

```
waiting_for_table = SQLSensor(
    task_id='waiting_for_table',
    conn_id='postgres',
    sql='sql/CHECK_TABLE.sql',
    poke_interval=30
)
```

The goal of the sensor is to expect data to arrive in a SQL table every 30 minutes. Is the sensor configured correctly to accomplish this goal?

☑ ~~Yes~~

☐ **No**

Question 65: ✖ Incorrect answer
Examine the DAG below:

```
 from airflow import DAG
 from airflow.operators.python_operator import PythonOperator
 from datetime import datetime

 with DAG(
     'five_task_dag',
     start_date=datetime(2023, 1, 1)
     catchup=False,
 ):

     task_a = PythonOperator(
         task_id='task_a',
         python_callable=lambda: print('Executing Task A'),
     )

     task_b = PythonOperator(
         task_id='task_b',
         python_callable=lambda: print('Executing Task B'),
     )

     task_c = PythonOperator(
         task_id='task_c',
         python_callable=lambda: print('Executing Task C'),
     )

     task_d = PythonOperator(
         task_id='task_d',
         python_callable=lambda: print('Executing Task D'),
     )

     task_e = PythonOperator(
         task_id='task_e',
         python_callable=lambda: print('Executing Task E'),
     )

     [task_a, task_b] >> task_e
     [task_c, task_d] >> task_e
```

How many tasks will run at the same time if this DAG is unpaused?

☑ ~~2~~

☐ ~~3~~

☐ ~~5~~

☐ **4**

Question 66: ✔ Correct answer
Examine the DAG below:

```
 from airflow.decorators import dag, task
 from airflow.utils.dates import datetime


 @dag(start_date=datetime(2023, 1, 1), catchup=False)
 def my_dag():

     @task
     def my_task():
         print('My task executed')
```

This DAG doesn't show up on the Airflow UI. Why?

☑ **my_dag isn't called**

☐ ~~The start_date parameter is incorrectly formatted~~

☐ ~~The catchup parameter is set to False~~

☐ ~~The schedule_interval parameter is missing~~

Question 67: ✔ Correct answer
What is the default timeout value of a sensor?

☐ ~~1 minute~~

☐ ~~1 day~~

☐ ~~7 hours~~

☑ **7 days**

Question 68: ✔ Correct answer
Assume an Airflow user lives in New York, which is in the timezone UTC-5, and defines the `start_date` parameter of a DAG with a value of `datetime(2023, 1, 1)` . Will the DAG be triggered at 00:00 on January 1st, 2023, in New York?

☐ ~~Yes~~

☑ **No**

Question 69: ✔ Correct answer
What does it mean when a DAG's task has the status of queued?

☑ **The task has been assigned to an Executor and is awaiting a worker.**

☐ ~~The task is executed immediately.~~

☐ ~~The task is waiting to be scheduled.~~

☐ ~~The task is running in a worker.~~

Question 70: ✖ Incorrect answer

Which parameters are optional when creating a new DAG object without using the `@dag` decorator? (select all that apply)

☑ **schedule_interval**

☐ ~~dag_id~~

☐ ~~start_date~~

☐ **default_args**

Question 71: ✔ Correct answer

Assume there is a DAG with five tasks labeled t0, t1, t2, t3, t4, and t5. Would writing the DAG's task dependency as `[t0, t1] >> t2 >> [t3, t4] >> t5` work?

☑ **Yes**

☐ ~~No~~

Question 72: ✔ Correct answer

Which of the following statements are true about a DAG in Airflow? (select all that apply)

☑ **A DAG's nodes define its tasks, and the directed edges define the dependencies between those tasks.**

☐ ~~A DAG is a programming language used for data analysis.~~

☐ ~~It is a good practice to have loops in a DAG.~~

☑ **DAG stands for Directed Acyclic Graph.**

Question 73: ✔ Correct answer

Which feature in Airflow allows tasks to communicate with each other and share data?

☐ ~~Connections~~

☐ ~~Hooks~~

☑ **XComs**

☐ ~~Variables~~

Question 74: ✔ Correct answer

Which of the following best describes what a Transfer Operator does?

☑ **It moves or transfers data between different sources, destinations, or storage systems**

☐ ~~It evaluates conditions and determines the path the workflow should take~~

☐ ~~It intakes and transforms data into a different format or structure~~

☐ ~~It sends notifications or alerts during the execution of a workflow~~

Question 75: ✖ Incorrect answer

Examine the Airflow UI below:

DAGs

All 6	Active 6	Paused 0		× extract ⊗		Search DAGs		Auto-refresh	⟳	

There is no DAG with the tag "processing" visible. Why?

☐ ~~There is a 5-minute default delay to see a new DAG~~

☐ ~~Airflow is experiencing an error~~

☑ ~~There is a 30-second default delay to see a new DAG~~

☐ **There is a filter applied for tags**

Was this page helpful?

😠 😟 😐 🙂 😍

Want feedback like this? Try Hotjar