coursera

Search in course [Search]

English ⌄   🔔   S ⌄

ETL and Data Pipelines wi... > Module 5 > Peer Review: Project Sub...

‹ Previous   Next ›

Peer-graded Assignment: Peer Review: Project Submission and Peer Review

ℹ️ It looks like this is your first peer-graded assignment. Learn more   ✕

Instructions   **My submission**   **Discussions**

Build ETL Data Pipelines with BashOperator using Apache Airflow

Submitted on August 6, 2024   Shareable Link

PROMPT

Task 1.1: Submit screenshot for the **Define the DAG arguments** task. (2 pts)

Upload the **dag_args.jpg** (or .png) file.

Define the DAG arguments

```
#defining DAG arguments
# You can override them on a per-task basis during operator initialization
default_args = {
    'owner': 'Shrey',
    'start_date': days_ago(0),
    'email': ['randonguy@gmail.com'],
    'email_on_failure': True,
    'retries': 1,
    'retry_delay': timedelta(minutes=5),
}
```

RUBRIC

Did the learner complete creating the DAG and schedule it on Apache Airflow?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

TIP: If the screenshot appears small and is hard to read try zooming in by pressing "Ctrl" and "+" keys together (Mac: "Command" and "+"), or Right-click on the image and "View Image" (Firefox) or "Open Image in new Tab" (Chrome).

○ 0 points
   Incorrect. The learner did not submit the screenshot or there is no entry for the DAG arguments visible in the screenshot.

○ 1 point
   Partially Correct. The screenshot shows at least 3 of the below arguments defined with proper values:

   • owner
   • start_date
   • email
   • email_on_failure
   • email_on_retry
   • retries
   • retry_delay

◉ **2 points**   SP
   Correct. The screenshot shows all the 7 arguments below defined with proper values.

   • owner
   • start_date
   • email

- email_on_failure
- email_on_retry
- retries
- retry_delay

PROMPT

Task 1.2: Submit screenshot for the **Define the DAG** task. (2 pts)

Upload the **dag_definition.jpg** (or .png) file.

Define the DAG

```
# define the DAG
dag = DAG(
    'ETL_toll_data',
    default_args=default_args,
    description='Apache Airflow Final Assignment',
    schedule_interval=timedelta(days=1),
)
```

RUBRIC

Did the learner define the DAG correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit the screenshot or there is no entry for the DAG parameters in the screenshot.

○ 1 point
Partially Correct. The screenshot shows at least 2 of the parameters below defined with proper values:

- DAG id
- Schedule
- default_args
- description

◉ **2 points**   SP
Correct. The screenshot shows all the 4 parameters below defined with proper values:

- DAG id
- Schedule
- default_args
- description

PROMPT

Task 2.1: Submit screenshot for the **Create a task to unzip data** task. (2 pts)

Upload the **unzip_data.jpg** (or .png) file.

Create a task to unzip data

```
unzip_data = BashOperator(
    task_id='unzip_data',
    bash_command='tar -xzf /home/project/airflow/dags/finalassignment/tolldata.tgz',
    dag=dag,
)
```

RUBRIC

Did the learner create the **unzip** task correctly using **tar** command?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit the screenshot or there is no entry for the **unzip** task visible in the screenshot.

○ 1 point
Partially Correct. The screenshot shows the **unzip** task has been created with BashOperator but does not use **tar** command.

◉ **2 points**   SP
Correct. The screenshot shows the **unzip** task has been created with BashOperator and uses **tar** command

to unzip.

Task 2.2: Submit screenshot for the **Create a task to extract data from csv file** task. (2 pts)

Upload the **extract_data_from_csv.jpg** (or .png) file.

Create a task to extract data from csv file

```
extract_data_from_csv = BashOperator(
    task_id='extract_data_from_csv',
    bash_command='cut -d"," -f1-4 /home/project/airflow/dags/finalassignment/vehicle-data.csv > \
    |/home/project/airflow/dags/finalassignment/csv_data.csv',
    dag=dag,
```

Did the learner create the **extract data from csv task** correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

- ○ 0 points
  Incorrect. The learner did not submit the screenshot or there is no entry for the extract task visible in the screenshot.

- ○ 1 point
  Partially Correct. The screenshot shows the "**extract data from csv**" task has been created with BashOperator but does not use the **cut** command.

- ◉ **2 points** SP
  Correct. The screenshot shows the "**extract data from csv**" task has been created with BashOperator and uses the **cut** command.

Task 2.3: Submit screenshot for the **Create the extract data from tsv file** task. (2 pts)

Upload the **extract_data_from_tsv.jpg** (or .png) file.

Create the extract data from tsv file

```
extract_data_from_tsv = BashOperator(
    task_id='extract_data_from_tsv',
    bash_command='cut -d"," -f5-7 /home/project/airflow/dags/finalassignment/vehicle-data.csv > \
    |/home/project/airflow/dags/finalassignment/tsv_data.csv',
    dag=dag,
)
```

Did the learner create the **extract data from tsv file task** correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

- ○ 0 points
  Incorrect. The learner did not submit the screenshot or there is no entry for the **extract data** task visible in the screenshot.

- ○ 1 point
  Partially Correct. The screenshot shows the "**extract data**" task has been created with BashOperator but does not use the **cut** command.

- ◉ **2 points** SP
  Correct. The screenshot shows the "**consolidate data**" task has been created with BashOperator and uses the **cut** command.

Task 2.4: Submit screenshot for the **Create a task to extract data from fixed width file** task. (2 pts)

Upload the **extract_data_from_fixed_width.jpg** (or .png) file.

Did the learner create the **extract data from fixed width file** task correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

Create a task to extract data from fixed width file

```
extract_data_from_fixed_width = BashOperator(
    task_id='extract_data_from_fixed_width',
    bash_command='cut -c50-67 /home/project/airflow/dags/finalassignment/payment-data.txt > \ |tr " " "," > \
    /home/project/airflow/dags/finalassignment/fixed_width_data.csv',
    dag=dag,
```

○ 0 points
Incorrect. The learner did not submit the screenshot or there is no entry for the extract task visible in the screenshot.

○ 1 point
Partially Correct. The screenshot shows the **"extract_data_from_fixed_width"** task has been created with BashOperator but does not use the **cut** command.

◉ 2 points   SP
Correct. The screenshot shows the **"extract_data_from_fixed_width"** task has been created with BashOperator and uses the **cut** command.

PROMPT

Task 2.5: Submit screenshot for the **Create a task to consolidate data** extracted from previous tasks. (2 pts)

Upload the **consolidate_data.jpg** (or .png) file.

Create a task to consolidate data

```
consolidate_data = BashOperator(
    task_id='consolidate_data',
    bash_command='paste /home/project/airflow/dags/finalassignment/extracted_data.csv > \ paste /home/project/airflow/dags/finalassignment/tsv_data.csv \
    > paste /home/project/airflow/dags/finalassignment/fixed_width_data.csv > \ paste /home/project/airflow/dags/finalassignment/extracted_data.csv >
    dag=dag,
```

RUBRIC

Did the learner create the **consolidate data** task correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit the screenshot or there is no entry for the transform task visible in the screenshot.

○ 1 point
Partially Correct. The screenshot shows the **"consolidate"** task has been created with BashOperator but does not use `**paste**` command.

◉ 2 points   SP
Correct. The screenshot shows the **"consolidate"** task has been created with BashOperator and uses `**paste**` command.

PROMPT

Task 2.6: Submit screenshot for the **Transform the data** task. (2 pts)

Upload the **transform.jpg** (or .png) file.

Transform the data

```
transform_data = BashOperator(
    task_id='transform_data',
    bash_command='tr "[a-z]" "[A-Z]" < /home/project/airflow/dags/finalassignment/extracted_data.csv \
    > /home/project/airflow/dags/finalassignment/staging/transformed_data.csv',
    dag=dag,
```

RUBRIC

Did the learner create the **transform data** task correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit the screenshot or there is no entry for the transform task visible in the screenshot.

○ 1 point
Partially Correct. The screenshot shows the **"transform"** task has been created with BashOperator but does

created with BashOperator but does
not use `tr` command.

○ **2 points**
Correct. The screenshot shows the
"**transform**" task has been created
with BashOperator and uses `tr`
command.

SP

PROMPT

Task 2.7: Submit screenshot for the **Define the task pipeline** task. (1 pt)

Upload the **task_pipeline.jpg** (or .png) file.

Define the task pipeline

```
#Task pipeline
unzip_data >> extract_data_from_csv >> extract_data_from_tsv >> extract_data_from_fixed_width >>consolidate_data >> transform_data
```

RUBRIC

Did the learner **define the task pipeline** correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit
the screenshot or the screenshot
shows less than 6 tasks or tasks in the
wrong order.

○ **1 point**
Correct. The screenshot shows all 6
tasks in the correct order as follows if
the:

SP

- unzip_data
- extract_data_from_csv
- extract_data_from_tsv
- extract_data_from_fixed_width
- consolidate_data
- transform_data

PROMPT

Task 3.1: Submit screenshot for the **Submit the DAG** task. (1 pt)

Upload the **submit_dag.jpg** (or .png) file.

Submit the DAG

```
theia@theiadocker-shreyans2012:/home/project$ export AIRFLOW_HOME=/home/project/airflow
 cp ETL_toll_data.py $AIRFLOW_HOME/dags
theia@theiadocker-shreyans2012:/home/project$ 
```

RUBRIC

Did the learner **submit** the DAG correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria below.

○ 0 points
Incorrect. The learner did not submit
the screenshot showing that the DAG
has been submitted through CLI or
Web UI.

○ **1 point**
Correct. The learner has submitted
the screenshot showing that the DAG
has been submitted through CLI or
Web UI.

SP

PROMPT

Task 3.2: Submit screenshot for the **Unpause and trigger the DAG** task. (3 pts)

Upload the **unpause_trigger_dag.jpg** (or .png) file.

RUBRIC

Did the learner unpause and trigger the DAG correctly?

Review the screenshot uploaded by the learner and grade this question based on the criteria

Unpause and trigger the DAG

below.

○ 0 points
Incorrect. The learner did not submit
the screenshot showing that the DAG
has been unpaused and triggered CLI
or Web UI.

○ 1 point
Partially Correct. The learner has
submitted the screenshot showing
that the DAG has been unpaused but
not triggered through CLI or Web UI.

◉ **3 points**   SP
Correct. The learner has submitted
the screenshot showing that the DAG
has been unpaused and triggered
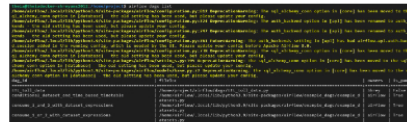through CLI or Web UI.

PROMPT

Task 3.3: Submit screenshot for the **List the DAG
tasks** task. (2 pts)

Upload the **dag_tasks.jpg** (or .png) file.

List the DAG tasks

RUBRIC

Did the learner get the **list of tasks** in the DAG
correctly?

Review the screenshot uploaded by the learner
and grade this question based on the criteria
below.

○ 0 points
Incorrect. The learner did not submit
the screenshot showing that the list
of tasks in the DAG through CLI or
Web UI.

○ 1 point
Partially Correct. The learner has
submitted the screenshot showing
that the list of tasks in the DAG
through CLI or Web UI but **not all the
6 tasks**.

◉ **2 points**   SP
Correct. The learner has submitted
the screenshot showing that the list
of tasks in the DAG through CLI or
Web UI with **all the 6 tasks**.

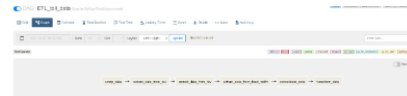**PROMPT**

Task 3.4: Submit screenshot for the **Monitor the
DAG** task. (2 pts)

Upload the **dag_runs.jpg** (or .png) file.

Monitor the DAG

RUBRIC

Did the learner create the **monitor DAG** task
correctly?

Review the screenshot uploaded by the learner
and grade this question based on the criteria
below.

○ 0 points
Incorrect. The learner did not submit
the screenshot or the tasks are
shown with status other than
"Success".

○ 1 point
Partially Correct. The screenshot
shows at least one of the tasks in the
DAG run as **"Success"**.

2 points

○ **2 points**
Correct. The screenshot shows one or more DAG runs with all the tasks as **"Success"**.

SP

Start new attempt

Comments
Comments left for the learner are visible only to that learner and the person who left the comment.

SS

Share your thoughts...

👍 **Like**   👎 **Dislike**   🏳 **Report an issue**