

```

# import the libraries

from datetime import timedelta
# The DAG object; we'll need this to instantiate a DAG
from airflow.models import DAG
# Operators; you need this to write tasks!
from airflow.operators.bash_operator import BashOperator
# This makes scheduling easy
from airflow.utils.dates import days_ago

#defining DAG arguments

# You can override them on a per-task basis during operator
initialization
default_args = {
    'owner': 'your_name_here',
    'start_date': days_ago(0),
    'email': ['your_email_here'],
    'retries': 1,
    'retry_delay': timedelta(minutes=5),
}

# defining the DAG

# define the DAG
dag = DAG(
    'my-first-dag',
    default_args=default_args,
    description='My first DAG',
    schedule_interval=timedelta(days=1),
)

# define the tasks

# define the first task
download = BashOperator(
    task_id='download',
    bash_command='curl "https://cf-courses-data.s3.us.cloud-object-
storage.appdomain.cloud/IBM-DB0250EN-
SkillsNetwork/labs/Apache%20Airflow/Build%20a%20DAG%20using%20Airflow/web-
server-access-log.txt" -o web-server-access-log.txt',
    dag=dag,
)

# define the second task
extract = BashOperator(
    task_id='extract',
    bash_command='cut -d":" -f1,4 /home/project/web-server-access-log.txt
> /home/project/airflow/dags/extracted-data.txt',
    dag=dag,
)

# define the third task
transform = BashOperator(

```

```

        task_id='transform',
        bash_command='tr "[a-z]" "[A-Z]" <
/home/project/airflow/dags/extracted-data.txt >
/home/project/airflow/dags/capitalized.txt',
        dag=dag,
    )

# define the fourth task
load = BashOperator(
    task_id='load',
    bash_command='zip capitalized.zip capitalized.txt
dag=dag,
)
```

```

# task pipeline
download >> extract >> transform >> load
```

```

#####
#Instructions to run
###
# 1. Connect to Apache Airflow --> open Apache webserver
# 2. Create a python file and add code
# 3. Open terminal and add the following commands:
### a. to set the AIRFLOW_HOME.
### export AIRFLOW_HOME=/home/project/airflow
### echo $AIRFLOW_HOME
### b. to submit the DAG that was created
### export AIRFLOW_HOME=/home/project/airflow
### cp my_first_dag.py $AIRFLOW_HOME/dags
### d. list out all dags
### airflow dags list
### e. monitor by going to the airflow
```