

Cloudera CDP-3001

CDP Data Developer - Certification Exam

For More Information – Visit link below:

<https://www.examsempire.com/>

Product Version

1. Up to Date products, reliable and verified.
2. Questions and Answers in PDF Format.



<https://examsempire.com/>

Visit us at: <https://www.examsempire.com/cdp-3001>

Latest Version: 6.0

Question: 1

Which is the right fit for ELT (D+Pipeline)?

- A. In ELT Data is copied from System-A to system-B with state change.
- B. In EVT Data is copied from System-A to System-B without state change.
- C. ELT Data is copied from System-A to System-B with where both System-A and System-B can be data Generator.
- D. In ELT Date is copied from System-A to System-B with where can be data Generator.

Answer: B

Question: 2

Please Map the following

Data Curation	Sequencing and Managing of the Data.
Data Optimization	Formatting, Correcting and transforming of the data.
Data Orchestration	Structuring the data as per need.

Answer:

Data Curation	Sequencing and Managing of the Data.
Data Optimization	Formatting, Correcting and transforming of the data.
Data Orchestration	Structuring the data as per need.

Question: 3

Which of the following solution on CDP platform can be used to create a complex data pipeline?

- A. CDP COW
- B. CDP CDE
- C. CDP ODB
- D. CDP ML

- E. CDP DataFlow
- F. CDP Replication Manager
- G. CDP Workload XM

Answer: B

Question: 4

Which of the following component is used in Cloudera CDP, Data Engineering to orchestrate multi-stage

- A. Apache spark
- B. Apache Atlas
- C. Apache Navigator
- D. Apache Airflow
- E. Apache Hive

Answer: D

Question: 5

Using Cloudera Data Engineering, you can create complex data pipelines and that is based on which of the following platform?

- A. Apache Airflow
- B. Apache Spark
- C. Apache Nig
- D. Apache Sqoop

Answer: B

Question: 6

If you are building a cost-effective Data Pipeline on CDP Cloudera Data Engineering platform, then which of the following statement is correct?

- A. Data pipelines that are bursty in nature can leverage the private cloud CDE service.
- B. Data pipelines that are bursty in nature can leverage the public cloud CDE service.
- C. Longer running persistent Data Pipeline loads can run on-prem.
- D. Longer running persistent Data Pipeline loads can run public Cloud CDE Service.

Answer: BC

Question: 7

You have just build a new job for CDP Cloudera Data Engineering and that is written using Python 3 and saved in to HadoopExam.py file. Which of the following is a correct command using CDE CLI to run this job ad-hoc?

- A. `cde spark submit --conf "spark.pyspark.python=python3" HadoopExam.py`
- B. `cde spark run --conf "spark.pyspark.python=python3" HadoopExam.py`
- C. `cde spark submit --conf "spark.pyspark.python=python3" --id HadoopExam.py`
- D. `cde spark run --conf "spark.pyspark.python=python3" --id HadoopExam.py`

Answer: A

Question: 8

You have created a resource called "HadoopExam_ETL" which has one Spark Job defined in HadoopExam.py file. You want this job to run using Python 3, Every hour between 01-Jan-2022 to 31-Dec-2022. Which of the following is correct command for the same?

- A.
`cde job create --name "HadoopExam_ETL" \
--type spark \
--conf "spark.pyspark.python=python3" \
--application-file, "HadoopExam.py"\
--cron-expression "0*/1 * * *"\
schedule -enabled "true" \
--schedule-start "2022-01-01 " \
schedule-end "2022-12-31" \
--mount-l-resource "HadoopExam"`
- B.
`cde job create --name "HadoopExam" \
--type spark \
--conf "spark.pyspark.python=python3" \
--application-file "HadoopExam.py" \
--cron-expression "0*/1 * * *"\
--schedule-enabled '*false' \
--schedule-start "2022-01-01" \
--schedule-end "2022-12-31" \
--mount-l-resource "HadoopExam E T L"`
- C.
`cde job create --name "HadoopExam" \
--type spark \
--conf " spark. pyspark.python— python 3" \`

```
—application-file "HadoopExam.py"\
—cron-expression "0 * /1 ***"\
—schedule-enabled "true" \
--schedule-start "2022-01-01 " \
—schedule-end "2022-12-31" \
—mount-1 -resource "HadoopExam_ELT"
D.
cde job create --name "HadoopExam" \
—type job \
--conf "spark.pyspark.python=python3" \
—cron-expression "0 * /1 ***"\
—schedule-enabled "true" \
„schedule-start "2022-01-01 " \
-schedule-end "2022-12-31" \
—mount-1 -resource "HadoopExam_ELT"
```

- A. A
- B. B
- C. C
- D. D

Answer: C

Question: 9

Which of the following command line option can help you to find the output of your CDP Cloudera Data Engineering job using CLI?

- A. cde job list --filter 'name[like]%HadoopExam%'
- B. cde run list --filter 'job[like]%HadoopExam%'
- C. cde run logs --type "driver/stdout" --id #
- D. cde run monitor --type "driver/stdout" --id #
- E. cde run outputs --type "driver/stdout" --id #

Answer: C

Question: 10

What is the resource file contains in case CDP Data Engineering and correct statement about this?

- A. In CDP DE, resources are automatically mounted and available to all Spark Executors.
- B. As part of resource, you can upload Jars file.
- C. As part of resource, you can upload Configurations file.
- D. As part of resource, you can upload Python egg file.

Answer: ABCD

Thank You for Trying Our Product
Special 16 USD Discount Coupon: NSZUBG3X

Email: support@examsempire.com

**Check our Customer Testimonials and ratings
available on every product page.**

Visit our website.

<https://examsempire.com/>