EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 42 DISCUSSIO..

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 42
Topic #: 1
[All Professional Machine Learning Engineer Questions]

You work for an advertising company and want to understand the effectiveness of your company's latest advertising campaign. You have streamed 500 MB of campaign data into BigQuery. You want to query the table, and then manipulate the results of that query with a pandas dataframe in an AI Platform notebook.
What should you do?

A. Use AI Platform Notebooks' BigQuery cell magic to query the data, and ingest the results as a pandas dataframe.

B. Export your table as a CSV file from BigQuery to Google Drive, and use the Google Drive API to ingest the file into your notebook instance.

C. Download your table from BigQuery as a local CSV file, and upload it to your AI Platform notebook instance. Use pandas.read_csv to ingest he file as a pandas dataframe.

D. From a bash cell in your AI Platform notebook, use the bq extract command to export the table as a CSV file to Cloud Storage, and then use gsutil cp to copy the data into the notebook. Use pandas.read_csv to ingest the file as a pandas dataframe.

Show Suggested Answer

by zosoabi at *June 10, 2021, 3:31 p.m.*

Comments

Type your comment...

Submit

zosoabi Highly Voted 3 years, 4 months ago

A: no "CSV" found in provided link https://cloud.google.com/bigquery/docs/bigquery-storage-python-pandas

upvoted 26 times

Sum_Sum Highly Voted 11 months, 3 weeks ago

Selected Answer: A

A is the google recommended answer. And what you should use
C is what the intern does ...

upvoted 5 times

sharth 10 months ago

Dude, I laughed so hard

upvoted 2 times

PhilipKoku Most Recent 5 months ago

Selected Answer: A

A) Magic command

upvoted 1 times

M25 1 year, 6 months ago

Selected Answer: A

Went with A

upvoted 2 times

SergioRubiano 1 year, 7 months ago

Selected Answer: A

A, Using the command %%bigquery df

upvoted 1 times

Dunnoth 1 year, 8 months ago

Why not D? using BQ notebook magic would be ok for a single time use. but usually a DS would reload the data multiple time, and every time you need to stream 500mb data to the notebook instance from BQ. Isn't it cheaper to store the data as a csv in a bucket?

upvoted 2 times

John_Pongthorn 1 year, 9 months ago

Selected Answer: A

```
%%bigquery df
SELECT name, SUM(number) as count
FROM `bigquery-public-data.usa_names.usa_1910_current`
GROUP BY name
ORDER BY count DESC
LIMIT 3

print(df.head())
```

upvoted 4 times

hiromi 1 year, 11 months ago

Selected Answer: A

A
https://cloud.google.com/bigquery/docs/visualize-jupyter

upvoted 2 times

Sachin2360 2 years, 4 months ago

Answer : A . Refer to this link for details: https://cloud.google.com/bigquery/docs/bigquery-storage-python-pandas
First 2 points talks about querying the data.
Download query results to a pandas DataFrame by using the BigQuery Storage API from the IPython magics for BigQuery in a Jupyter notebook.
Download query results to a pandas DataFrame by using the BigQuery client library for Python.
Download BigQuery table data to a pandas DataFrame by using the BigQuery client library for Python.
Download BigQuery table data to a pandas DataFrame by using the BigQuery Storage API client library for Python.

upvoted 2 times

Mohamed_Mossad 2 years, 4 months ago

Selected Answer: A

https://googleapis.dev/python/bigquery/latest/magics.html#ipython-magics-for-bigquery

upvoted 2 times

NickNtaken 2 years, 6 months ago

Selected Answer: A

this is the simplest and most straightforward way read BQ data into Pandas dataframe.

upvoted 3 times

mmona19 2 years, 6 months ago

Selected Answer: C

both A and C is technically correct. C has more manual step and A has less. The question does not ask which requires least effort. so C is clear answer

upvoted 1 times

wish0035 1 year, 10 months ago

"A and C are valid, but C is more difficult than A. they don't ask to be easier so I will go with the more difficult". WHAAAT? Google best practices are always: easier > harder. Even they encourage you to skip ML if you don't need ML.

upvoted 2 times

SlipperySlope 2 years, 8 months ago

Selected Answer: C

C is the correct answer due to the size of the data. It wouldn't be possible to download it all into an in memory data frame.

upvoted 1 times

u_phoria 2 years, 4 months ago

500mb of data into a pandas dataframe generally isn't a problem, far from it.

upvoted 2 times

ggorzki 2 years, 9 months ago

Selected Answer: A

IPython magics for BigQuery
https://cloud.google.com/bigquery/docs/bigquery-storage-python-pandas

upvoted 1 times

NamitSehgal 2 years, 10 months ago

I agree with A

upvoted 1 times

Y2Data 3 years, 1 month ago

Just load it

https://googleapis.dev/python/bigquery/latest/magics.html

upvoted 2 times

Social Media

Facebook , Twitter
YouTube , Reddit
Pinterest

We are the biggest and most updated IT certification exam material website.
Using our own resources, we strive to strengthen the IT professionals community for free.