

In [1]:

```
!pip install --upgrade google-cloud
!pip install --upgrade google-cloud-bigquery
!pip install --upgrade google-cloud-storage
```

```
Requirement already satisfied: google-cloud in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (0.34.0)
Requirement already satisfied: google-cloud-bigquery in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(3.4.1)
Requirement already satisfied: requests<3.0.0dev,>=2.21.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packa
ges (from google-cloud-bigquery) (2.27.1)
Requirement already satisfied: google-cloud-core<3.0.0dev,>=1.4.1 in /Users/suraa/opt/anaconda3/lib/python3.9/si
te-packages (from google-cloud-bigquery) (2.3.2)
Requirement already satisfied: google-resumable-media<3.0.0dev,>=0.6.0 in /Users/suraa/opt/anaconda3/lib/python3.
9/site-packages (from google-cloud-bigquery) (2.4.0)
Requirement already satisfied: python-dateutil<3.0.0dev,>=2.7.2 in /Users/suraa/opt/anaconda3/lib/python3.9/site-p
ackages (from google-cloud-bigquery) (2.8.2)
Requirement already satisfied: protobuf!=3.20.0,!>3.20.1,!>4.21.0,!>4.21.1,!>4.21.2,!>4.21.3,!>4.21.4,!>4.21.5,<
5.0.0dev,>=3.19.5 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from google-cloud-bigquery) (4.21.1
2)
Requirement already satisfied: packaging<22.0.0dev,>=14.3 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packa
ges (from google-cloud-bigquery) (21.3)
Requirement already satisfied: google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5 in /Users/
suraa/opt/anaconda3/lib/python3.9/site-packages (from google-cloud-bigquery) (2.11.0)
Requirement already satisfied: grpcio<2.0.0dev,>=1.47.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from google-cloud-bigquery) (1.51.1)
Requirement already satisfied: proto-plus<2.0.0dev,>=1.15.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-pac
kages (from google-cloud-bigquery) (1.22.1)
Requirement already satisfied: google-auth<3.0.0dev,>=2.14.1 in /Users/suraa/opt/anaconda3/lib/python3.9/site-pack
ages (from google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-bigquery) (2.15.
0)
Requirement already satisfied: googleapis-common-protos<2.0.0dev,>=1.56.2 in /Users/suraa/opt/anaconda3/lib/python
3.9/site-packages (from google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-big
query) (1.57.0)
Requirement already satisfied: grpcio-status<2.0.0dev,>=1.33.2 in /Users/suraa/opt/anaconda3/lib/python3.9/site-pa
ckages (from google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-bigquery) (1.5
1.1)
Requirement already satisfied: rsa<5,>=3.1.4 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from goo
gle-auth<3.0.0dev,>=2.14.1->google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-b
igquery) (4.7.2)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from google-auth<3.0.0dev,>=2.14.1->google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->googl
e-cloud-bigquery) (0.2.8)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from google-auth<3.0.0dev,>=2.14.1->google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->googl
e-cloud-bigquery) (4.2.2)
Requirement already satisfied: six>=1.9.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from google
-auth<3.0.0dev,>=2.14.1->google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-bigq
uery) (1.16.0)
Requirement already satisfied: google-crc32c<2.0.0dev,>=1.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packa
ges (from google-resumable-media<3.0.0dev,>=0.6.0->google-cloud-bigquery) (1.1.2)
Requirement already satisfied: cffi>=1.0.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from googl
e-crc32c<2.0.0dev,>=1.0->google-resumable-media<3.0.0dev,>=0.6.0->google-cloud-bigquery) (1.15.1)
Requirement already satisfied: pyparsing in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from cffi>=
1.0.0->google-crc32c<2.0.0dev,>=1.0->google-resumable-media<3.0.0dev,>=0.6.0->google-cloud-bigquery) (2.21)
Requirement already satisfied: pyrsing!=3.0.5,>=2.0.2 in /Users/suraa/opt/anaconda3/lib/python3.9/site-package
s (from packaging<22.0.0dev,>=14.3->google-cloud-bigquery) (3.0.4)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (f
rom pyasn1-modules>=0.2.1->google-auth<3.0.0dev,>=2.14.1->google-api-core[grpc]!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.
0.0dev,>=1.31.5->google-cloud-bigquery) (0.4.8)
Requirement already satisfied: idna<4,>=2.5 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from requ
ests<3.0.0dev,>=2.21.0->google-cloud-bigquery) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from requests<3.0.0dev,>=2.21.0->google-cloud-bigquery) (1.26.13)
Requirement already satisfied: charset-normalizer~>2.0.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packag
es (from requests<3.0.0dev,>=2.21.0->google-cloud-bigquery) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (fro
m requests<3.0.0dev,>=2.21.0->google-cloud-bigquery) (2022.9.24)
Requirement already satisfied: google-cloud-storage in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(2.7.0)
Requirement already satisfied: requests<3.0.0dev,>=2.18.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packa
ges (from google-cloud-storage) (2.27.1)
Requirement already satisfied: google-api-core!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5 in /Users/suraa
/opt/anaconda3/lib/python3.9/site-packages (from google-cloud-storage) (2.11.0)
Requirement already satisfied: google-auth<3.0.0dev,>=1.25.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-pack
ages (from google-cloud-storage) (2.15.0)
Requirement already satisfied: google-resumable-media>=2.3.2 in /Users/suraa/opt/anaconda3/lib/python3.9/site-pa
ckages (from google-cloud-storage) (2.4.0)
Requirement already satisfied: google-cloud-core<3.0.0dev,>=2.3.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site
-packages (from google-cloud-storage) (2.3.2)
Requirement already satisfied: googleapis-common-protos<2.0.0dev,>=1.56.2 in /Users/suraa/opt/anaconda3/lib/python
3.9/site-packages (from google-api-core!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-storage)
(1.57.0)
Requirement already satisfied: protobuf!=3.20.0,!>3.20.1,!>4.21.0,!>4.21.1,!>4.21.2,!>4.21.3,!>4.21.4,!>4.21.5,<
5.0.0dev,>=3.19.5 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from google-api-core!=2.0.*,!=2.1.
*,!=2.2.*,!=2.3.0,<3.0.0dev,>=1.31.5->google-cloud-storage) (4.21.12)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from google-auth<3.0.0dev,>=1.25.0->google-cloud-storage) (0.2.8)
Requirement already satisfied: rsa<5,>=3.1.4 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from goo
gle-auth<3.0.0dev,>=1.25.0->google-cloud-storage) (4.7.2)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages
(from google-auth<3.0.0dev,>=1.25.0->google-cloud-storage) (4.2.2)
Requirement already satisfied: six>=1.9.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from google
-auth<3.0.0dev,>=1.25.0->google-cloud-storage) (1.16.0)
Requirement already satisfied: google-crc32c<2.0.0dev,>=1.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packa
ges (from google-resumable-media>=2.3.2->google-cloud-storage) (1.1.2)
Requirement already satisfied: cffi>=1.0.0 in /Users/suraa/opt/anaconda3/lib/python3.9/site-packages (from googl
e-crc32c<2.0.0dev,>=1.0->google-resumable-media>=2.3.2->google-cloud-storage) (1.15.1)
```

Requirement already satisfied: pycparser in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from cffi>=1.0.0->google-crc32c<2.0dev,>=1.0->google-resumable-media>=2.3.2->google-cloud-storage) (2.21)

Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from pyasn1-modules>=0.2.1->google-auth<3.0dev,>=1.25.0->google-cloud-storage) (0.4.8)

Requirement already satisfied: idna<4,>=2.5 in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0.0dev,>=2.18.0->google-cloud-storage) (3.4)

Requirement already satisfied: charset-normalizer==2.0.0 in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0.0dev,>=2.18.0->google-cloud-storage) (2.0.4)

Requirement already satisfied: certifi>=2017.4.17 in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0.0dev,>=2.18.0->google-cloud-storage) (2022.9.24)

Requirement already satisfied: urllib3<1.27,>=1.21.1 in /Users/suraaaj/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0.0dev,>=2.18.0->google-cloud-storage) (1.26.13)

In [2]:

```
import os
os.environ["GOOGLE_APPLICATION_CREDENTIALS"]="BigQuery_Python_for_class.json"
```

In [3]:

```
from google.cloud import bigquery

client = bigquery.Client()
```

In [4]:

```
suraaaj = """ select * from `bigquery-public-data.covid19_open_data.covid19_open_data` limit 10
"""
results = client.query(suraaaj).to_dataframe()

results.head()
```

Out[4]:

	location_key	date	place_id	wikidata_id	datacommons_id	country_code	country_name	iso_3166_1_alpha_2	iso_3166_1_alpha_3
0	AR	2020-04-03	ChIJZ8b99fXKvJURqA_wKpl3Lz0	Q414	country/ARG	AR	Argentina	AR	ARG
1	AR	2020-07-03	ChIJZ8b99fXKvJURqA_wKpl3Lz0	Q414	country/ARG	AR	Argentina	AR	ARG
2	AR	2020-02-19	ChIJZ8b99fXKvJURqA_wKpl3Lz0	Q414	country/ARG	AR	Argentina	AR	ARG
3	AR	2020-04-07	ChIJZ8b99fXKvJURqA_wKpl3Lz0	Q414	country/ARG	AR	Argentina	AR	ARG
4	AR	2020-01-16	ChIJZ8b99fXKvJURqA_wKpl3Lz0	Q414	country/ARG	AR	Argentina	AR	ARG

5 rows × 701 columns

In [5]:

```
ds=client.dataset("hacker_news", project='bigquery-public-data')
```

In [6]:

```
# Construct a BigQuery client object.
client = bigquery.Client()

dataset_id = 'bigquery-public-data.hacker_news'

tables = client.list_tables(dataset_id) # Make an API request.

print("Tables contained in '{}':".format(dataset_id))
for table in tables:
    print("{}.{}.{}".format(table.project, table.dataset_id, table.table_id))
```

```
Tables contained in 'bigquery-public-data.hacker_news':
bigquery-public-data.hacker_news.comments
bigquery-public-data.hacker_news.full
bigquery-public-data.hacker_news.full_201510
bigquery-public-data.hacker_news.stories
```

In [7]:

```
client = bigquery.Client()
def runner(sql):
    results = client.query(sql).to_dataframe()
    return results
```

In [8]:

```
#Give me the user retention using comments table
```

In [9]:

```
comments="select * from bigquery-public-data.hacker_news.comments limit 10"

runner(comments)
```

Out[9]:

	id	by	author	time	time_ts	text	parent	deleted	dead	ranking
0	388422	None	None	1228601919	2008-12-06 22:18:39+00:00	None	388384	True	<NA>	0
1	1712701	None	None	1285079381	2010-09-21 14:29:41+00:00	None	1710966	True	<NA>	0
2	9977901	None	None	1438288897	2015-07-30 20:41:37+00:00	None	9975137	True	<NA>	0
3	9347728	None	None	1428589592	2015-04-09 14:26:32+00:00	None	9347727	True	<NA>	0
4	9673798	None	None	1433663787	2015-06-07 07:56:27+00:00	None	9673780	True	<NA>	0
5	147820	None	None	1206634249	2008-03-27 16:10:49+00:00	None	147579	True	<NA>	0
6	296931	None	None	1220726144	2008-09-06 18:35:44+00:00	None	296865	True	<NA>	0
7	3084074	None	None	1317997112	2011-10-07 14:18:32+00:00	None	3083928	True	<NA>	0
8	2004459	None	None	1292340836	2010-12-14 15:33:56+00:00	None	2004398	True	<NA>	0
9	3069513	None	None	1317705687	2011-10-04 05:21:27+00:00	None	3069501	True	<NA>	0

In [10]:

```
# FACEBOOK - MEDIUM
```

In [11]:

```
retention="""
    WITH BASE AS
    (
        SELECT author as id, EXTRACT (MONTH FROM time_ts) as month_
        from bigquery-public-data.hacker_news.comments
        group by 1,2
    ),
    base_2 as (
        select a.id as id , a.month_ as current_month, b.month_ as next_month
        from base a
        left join base b
        on a.id=b.id
        and
        a.month_ = b.month_ -1
    ),
    base_3 as (
        select current_month,
        count( case when next_month is not null then id end) as retained,
        count(id) as total
        from base_2
        group by 1
    )
    select *, round((retained/ total *100), 2) as per_ret from base_3
    order by current_month asc"""
```

In [12]:

```
runner (retention)
```

Out[12]:

	current_month	retained	total	per_ret
0	1	36895	61243	60.24
1	2	38507	62065	62.04
2	3	40350	67684	59.62
3	4	39949	67798	58.92
4	5	39366	65943	59.70
5	6	39795	65554	60.71
6	7	40336	68311	59.05
7	8	39739	67707	58.69
8	9	39210	66755	58.74
9	10	36464	65415	55.74
10	11	35055	58966	59.45
11	12	0	58517	0.00

In []:

In []: