Deep Learning Walkthrough - 02

Code in github.com/google-aai/sc17

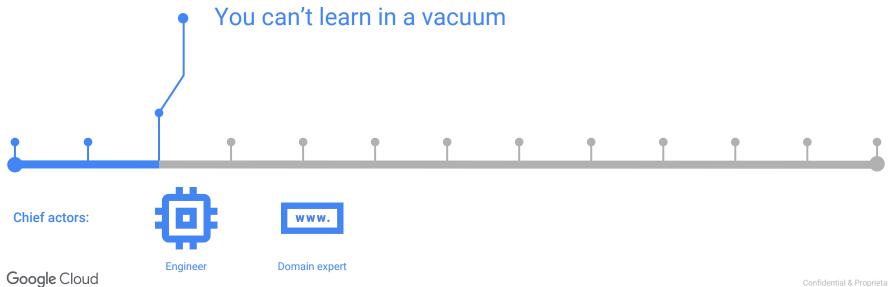
Cassie Kozyrkov

Chief Decision Scientist, Google Cloud GitHub: kozyrkov; Twitter: @quaesita



Google Cloud

Step 2 | Get your data





```
Secure https://ssh.cloud.google.com/projects/super-188716/zones/us-east1-c/instances/super-188716-compute-instance?authus...
                                                                                           sh run_step_2a_query.sh
             @super-188716-compute-instance:~/sc17/cats$ sh run step 2a query.sh
(env)
Waiting on bgjob r2b58c9818095dc06 0000016075203e0a 1 ... (145) Current Status. DoNE
                            original url
 https://c7.staticflickr.com/9/8894/18277258689 5acf8b0a8a o.jpg
                                                                               8.526988257706591E-6
 https://c2.staticflickr.com/3/2803/4020824567 0b7e36bfcd o.jpg
                                                                              1.3883631782896089E-5
 https://c7.staticflickr.com/4/3318/3188958741 235b507552 o.jpg
                                                                               2.702642935203947E-5
 https://farm2.staticflickr.com/6171/6235175628 7f2b84a4fc o.jpg
                                                                               4.646176122080625E-5
 https://c5.staticflickr.com/8/7153/6740118299 laacb29588 o.jpg
                                                                              6.0176315697990646E-5
 https://farm7.staticflickr.com/5346/17204131514 007a763656 o.jpg
                                                                               6.279622240379368E-5
 https://c7.staticflickr.com/4/3751/9788632466 aa5204ba9f o.jpg
                                                                               6.490346933474278E-5
 https://farml.staticflickr.com/2593/4148830705 b01d9c83e8 o.jpg
                                                                                6.60085155750334E-5
 https://farml.staticflickr.com/4045/4218663050_bd64e12b8a_o.jpg
                                                                               8.965109192482501E-5
 https://cl.staticflickr.com/4/3200/2350820461 alf528269d o.jpg
                                                                               9.404920838229636E-5
 https://c2.staticflickr.com/4/3224/2932224208 e799727148 o.jpg
                                                                              1.0307303246426258E-4
 https://c8.staticflickr.com/8/7105/7364627244 5ef122e36a o.jpg
                                                                              1.1777178603544769E-4
```

1.5044281923115577E-4

1.5933507039811445E-4

1.8154501185296823E-4

1.9661932271120008E-4

2.5992094483573425E-4

2.6391619407029054E-4

2.7360821774990995E-4

2.8884169378939895E-4

2.9849741811061037E-4

2.974360479515159E-4

2.993901183343528E-4

3.080053757211765E-4

2.68379707281996E-4

1.673010517302398E-4

1.880243997225417E-4

2.3681323217839E-4

https://farm6.staticflickr.com/8109/8657170264 72a8146898 o.jpg

https://farml.staticflickr.com/4034/4613672388 402a6de86c o.jpg

https://farm3.staticflickr.com/7726/17339965475 e681255eb0 o.jpg

https://farm4.staticflickr.com/195/499321718 b13b311e45 o.jpg

https://cl.staticflickr.com/1/67/203231915 df34a12f3d o.jpg

https://farm3.staticflickr.com/35/124377903 8d2463be2d o.jpg

https://c4.staticflickr.com/4/3012/2947965305 d617988305 o.jpg

https://cl.staticflickr.com/4/3644/3559941164 de8840e814 o.jpg

https://c2.staticflickr.com/3/2746/4479280283 0937d4fle7 o.jpg

https://farm6.staticflickr.com/6160/6145126622 dbc68ec602 o.jpg

https://farm6.staticflickr.com/3856/14394042358 04e8040e77 o.jpg

https://farm5.staticflickr.com/8610/16181928586 bc12b486f4 o.jpg

https://farm4.staticflickr.com/5612/15559754015 981c5b64b0 o.jpg

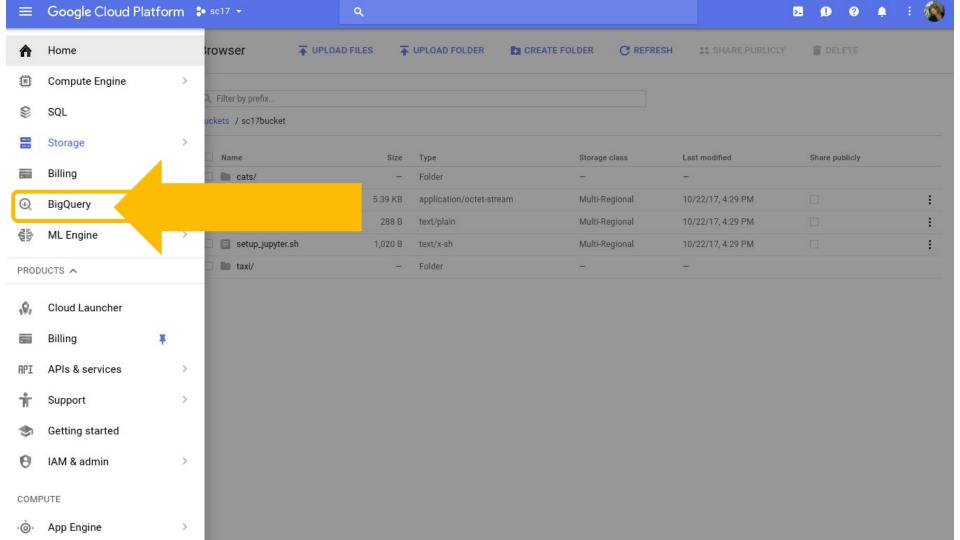
https://c2.staticflickr.com/6/5084/5236877415 316d4398ed o.jpg

https://c3.staticflickr.com/1/153/379515012 9a99bb1584 o.jpg

https://c2.staticflickr.com/9/8264/8666029623 caa21476ea o.jpg

https://c7.staticflickr.com/1/171/470614612 79aae80715 o.jpg

Populate a BigQuery table [Behind the scenes]



COMPOSE QUERY

Query History

Job History

Filter by ID or label

sc17

No datasets found in this project.

Please create a dataset or select a new project from the menu above.

Welcome to BigQuery!

Query is a web service that lets you do interactive analysis of massive datasets—up to plable and easy to use, BigQuery lets developers and businesses tap into powerful

bad some of your own data into a table using the menu on the

To get star. /ing options:

- Read our Bigo. start guide
- Run a query against ou...sample data by clicking "Compose Query"

Create new dataset

Switch to project

tions in BigQuery

50 %

Refresh

?

UI, see the BigQuery Web UI guide

nyc-tlc

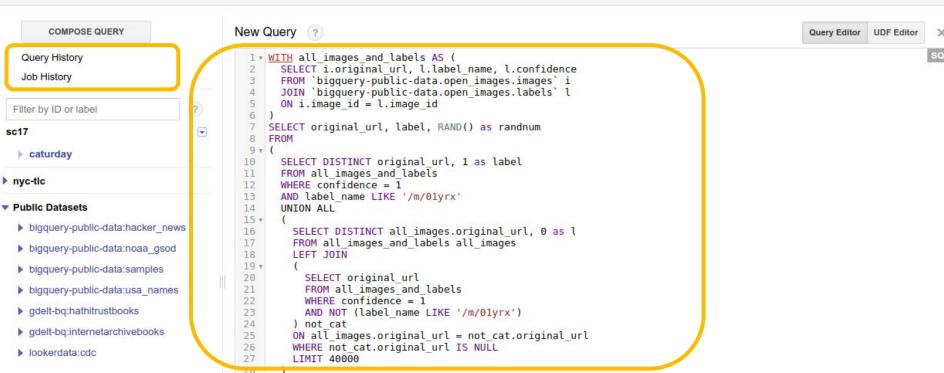
Public Datasets

- bigquery-public-data:hacker_news
- bigquery-public-data:noaa gsod
- ▶ bigquery-public-data:samples
- bigguery-public-data:usa names
- bigquery-public-data:usa_name
- gdelt-bq:internetarchivebooks

gdelt-bg:hathitrustbooks

loekerdata:cds





Google Cloud

29

RUN QUERY

ORDER BY randnum

Save Query

Error: Encountered " "WITH" "WITH "" at line 1, column 1. Was expecting: <EOF> [Try using standard SQL (https://cloud.google.com/bigguery/docs/reference/standard-sql/enabling-standard-sql)]

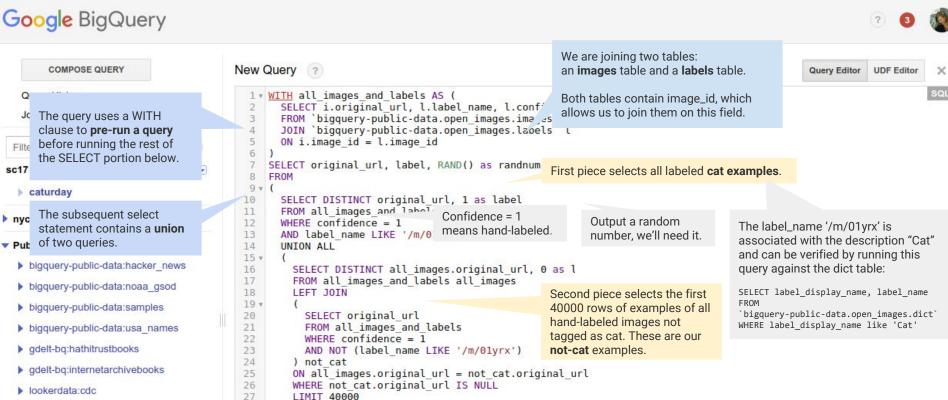
Format Query

Show Options

Save View







ORDER BY randnum

Error: Encountered " "WITH" "WITH "" at line 1, column 1. Was expecting: <EOF> [Try using standard SQL

(https://cloud.google.com/bigguery/docs/reference/standard-sql/enabling-standard-sql)]

Google Cloud



28 29



Save Query





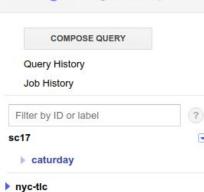












- ▼ Public Datasets
- bigquery-public-data:hacker_news
- bigquery-public-data:noaa_gsod
- bigguery-public-data:samples
- bigquery-public-data:usa_names

Query Priority

SQL Dialect

RUN QUERY

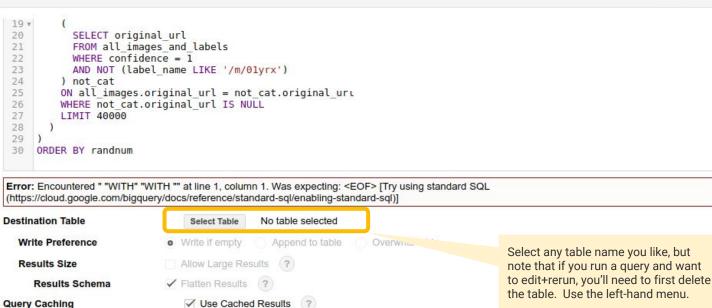
UDF Source URIs

Maximum Billing Tier

Maximum Bytes Billed

Save Query

- gdelt-bq:hathitrustbooks
- ▶ gdelt-bq:internetarchivebooks
- lookerdata:cdc



Interactive

Project Default

Project Default

✓ Use Legacy SQL

Save View

Edit

Batch

Format Query

?

?

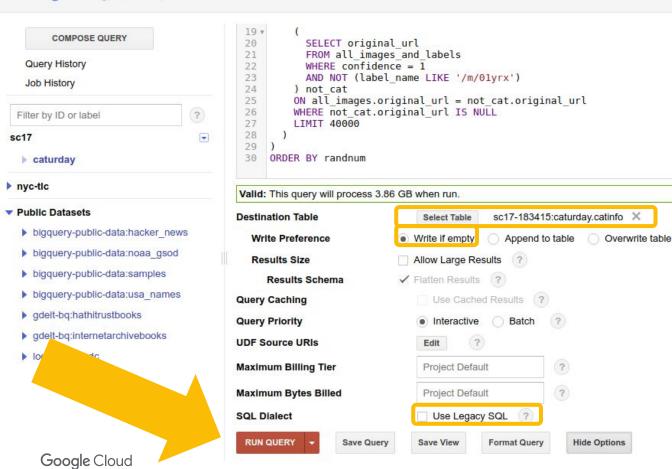
Hide Options



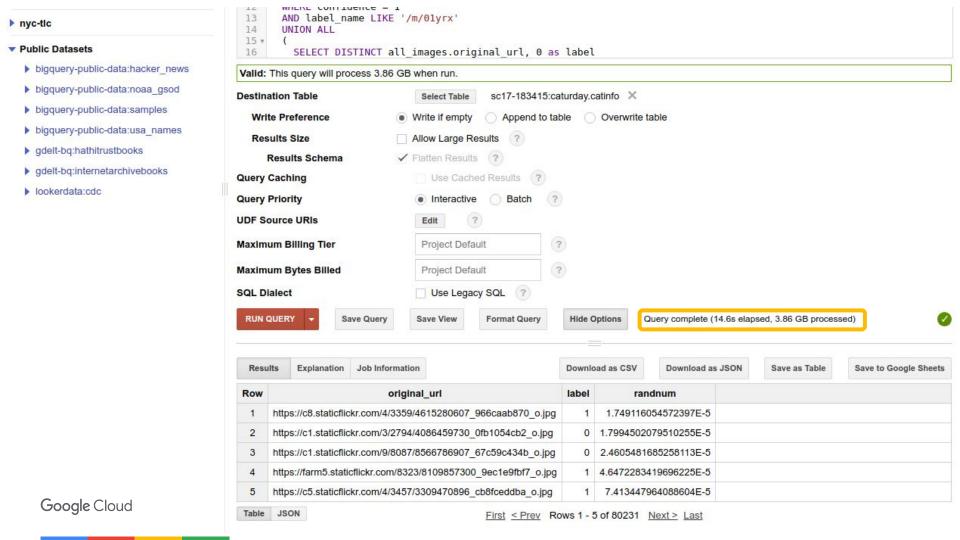


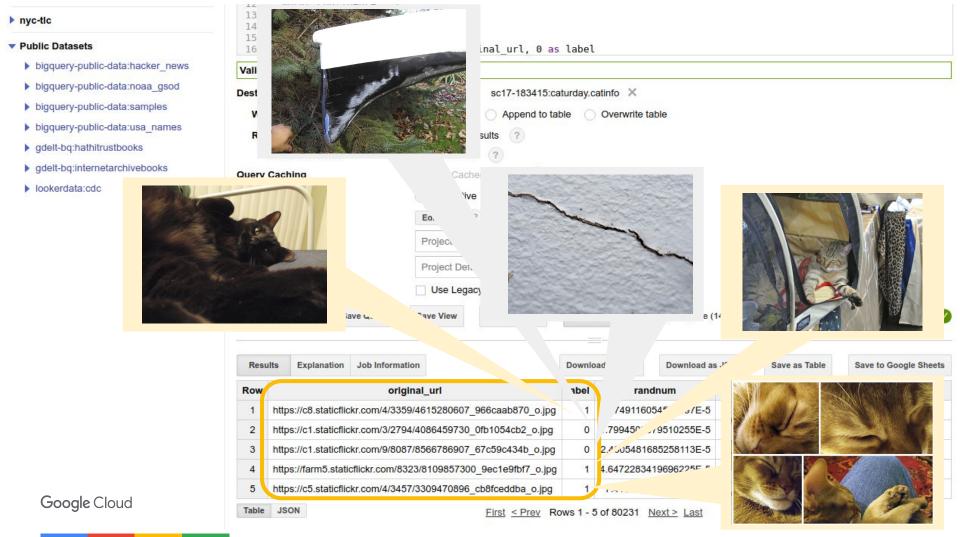


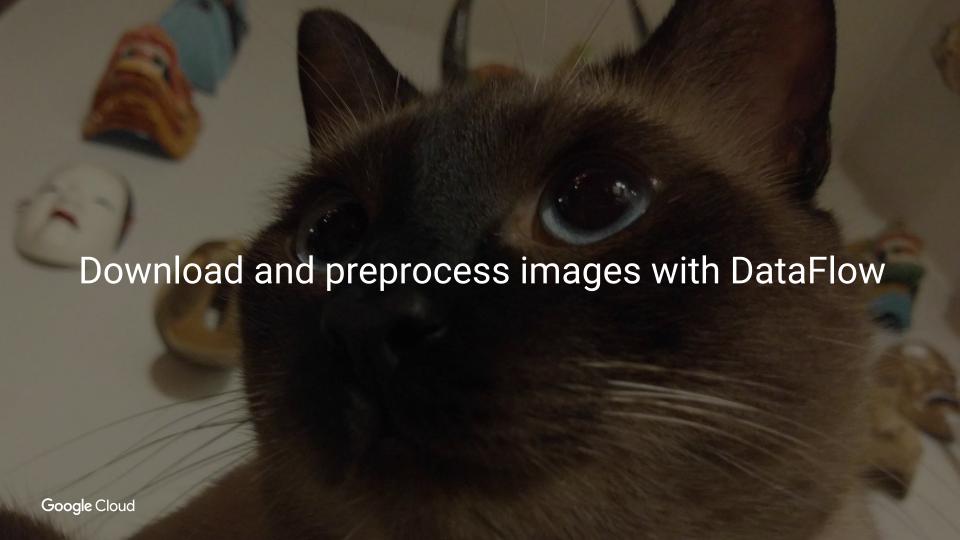












Secure https://ssh.cloud.google.com/projects/super-188716/zones/us-east1-c/instances/super-188716-compute-instance?authus...

(env) user | super-188716-compute-instance:~/sc17/cats\$ sh run_step_2b_get_images.sh \$PROJECT dataset catinfo \$BUCKET catimages

No handlers could be found for logger "oauth2client.contrib.multistore_file"

sh run_step_2b_get_images.sh \$PROJECT dataset catinfo \$BUCKET catimages

into filter bad or absent images images into fetch images from urls INFO:root:2017-12-20T18:30:34.097Z: JOB MESSAGE DETAILED: (1f4f7cec0d1c2dc6): Fusing consumer write images to stora

ge into resize and pad images

INFO:root:2017-12-20T18:30:34.121Z: JOB MESSAGE DEBUG: (1f4f7cec0d1c2044): Workflow config is missing a default res

Secure https://ssh.cloud.google.com/projects/super-188716/zones/us-east1-c/instances/super-188716-compute-instance?authus...

INFO:root:2017-12-20T18:30:34.074Z: JOB MESSAGE DETAILED: (1f4f7cec0d1c2b48): Fusing consumer filter bad or absent

INFO:root:2017-12-20T18:30:34.043Z: JOB MESSAGE DETAILED: (1f4f7cec0d1c28ca): Fusing consumer resize and pad images

INFO:root:2017-12-20T18:30:34.149Z: JOB MESSAGE DEBUG: (1f4f7cec0d1c22c2): Adding StepResource setup and teardown t

INFO:root:2017-12-20T18:30:34.458Z: JOB MESSAGE BASIC: (54dc4c9c134f9124): Executing operation read rows from cat i

INFO:root:2017-12-20T18:30:34.539Z: JOB MESSAGE BASIC: (348e57263115b6ad): Starting 50 workers in us-central1-f... INFO:root:2017-12-20T18:30:34.573Z: JOB MESSAGE BASIC: (903c4cdd1507102b): BigQuery query issued as job: "dataflow job 257583136778705608". You can check its status with the bg tool: "bg show -j --project id=super-188716 dataflow"

INFO:root:2017-12-20T18:30:42.350Z: JOB MESSAGE DETAILED: (6199526dfc07429f): Autoscaling: Raised the number of wor

INFO:root:2017-12-20T18:30:58.331Z: JOB MESSAGE DETAILED: (6199526dfc0744c4): Autoscaling: Raised the number of wor

INFO:root:2017-12-20T18:30:58.350Z: JOB MESSAGE DETAILED: (6199526dfc07470a): Resized worker pool to 46, though goa

INFO:root:2017-12-20T18:31:03.652Z: JOB MESSAGE DETAILED: (6199526dfc074022): Autoscaling: Raised the number of wor

INFO:root:2017-12-20T18:31:03.680Z: JOB MESSAGE DETAILED: (6199526dfc074268): Resized worker pool to 48, though goa

INFO:root:2017-12-20T18:31:08.996Z: JOB MESSAGE DETAILED: (6199526dfc0744a3): Autoscaling: Raised the number of wor

INFO:root:2017-12-20T18:30:34.182Z: JOB MESSAGE DEBUG: (1f4f7cec0d1c2540): Adding workflow start and stop steps.

nfo table+fetch images from urls+filter bad or absent images+resize and pad images+write images to storage INFO:root:2017-12-20T18:30:34.503Z: JOB MESSAGE DEBUG: (348e57263115b477): Starting worker pool setup.

INFO:root:2017-12-20T18:30:34.224Z: JOB MESSAGE DEBUG: (1f4f7cec0d1c27be): Assigning stage ids. INFO:root:2017-12-20T18:30:34.392Z: JOB MESSAGE DEBUG: (54dc4c9c134f910d): Executing wait step start3

INFO:root:Job 2017-12-20 10 30 29-8607060276961813628 is in state JOB STATE RUNNING

kers to 0 based on the rate of progress in the currently running step(s).

kers to 46 based on the rate of progress in the currently running step(s).

kers to 48 based on the rate of progress in the currently running step(s).

kers to 50 based on the rate of progress in the currently running step(s).

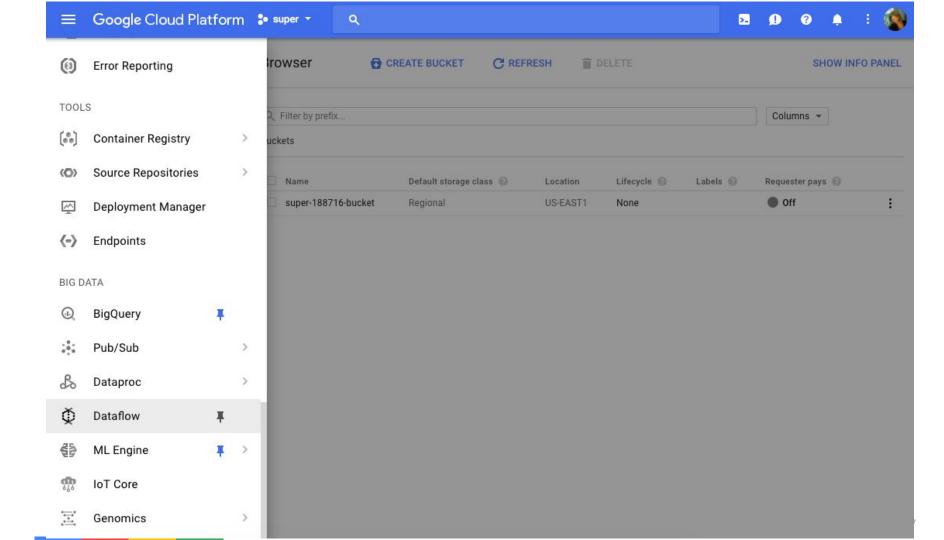
ource spec.

o workflow graph.

job 257583136778705608".

1 was 50. This could be a quota issue.

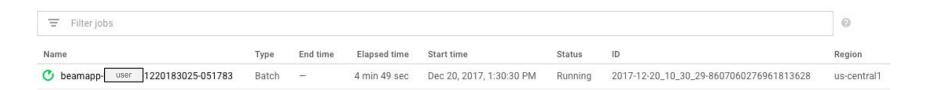
1 was 50. This could be a quota issue.





Jobs

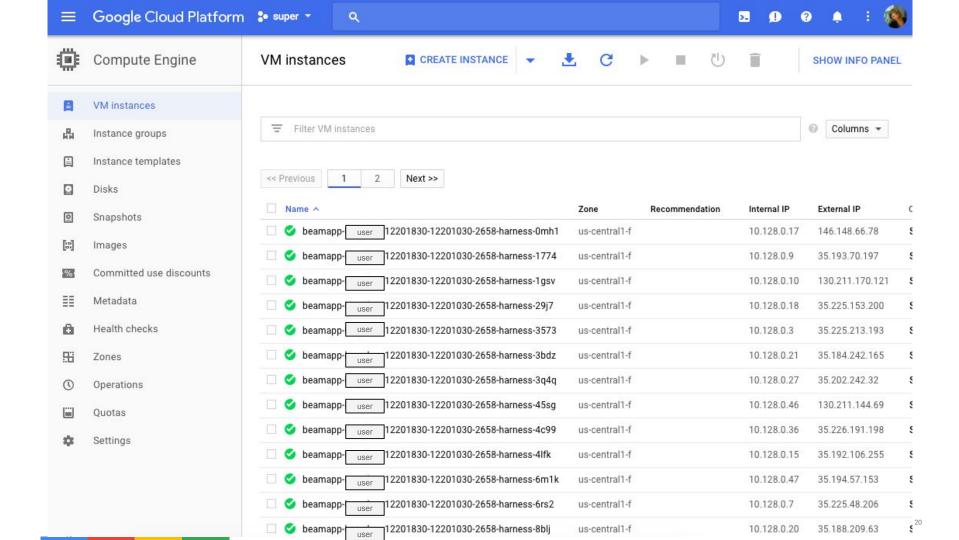
+ CREATE JOB FROM TEMPLATE

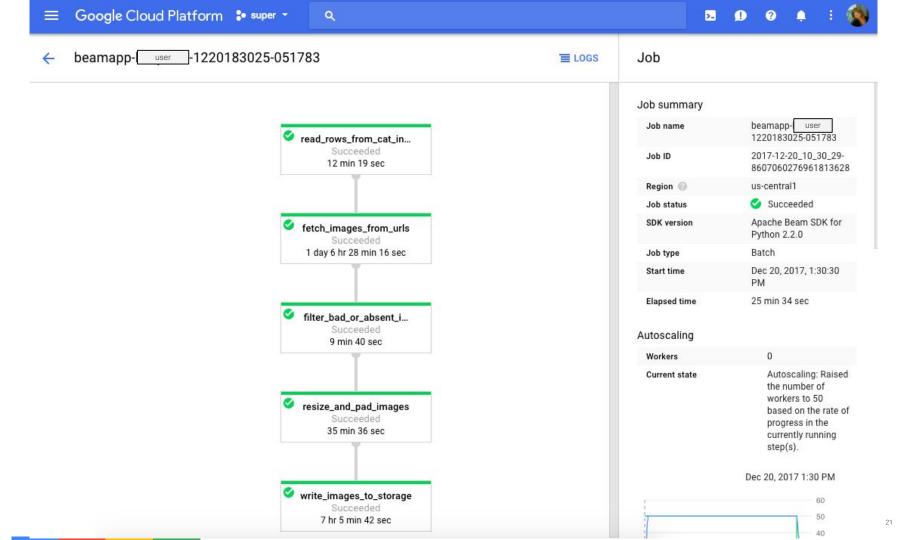


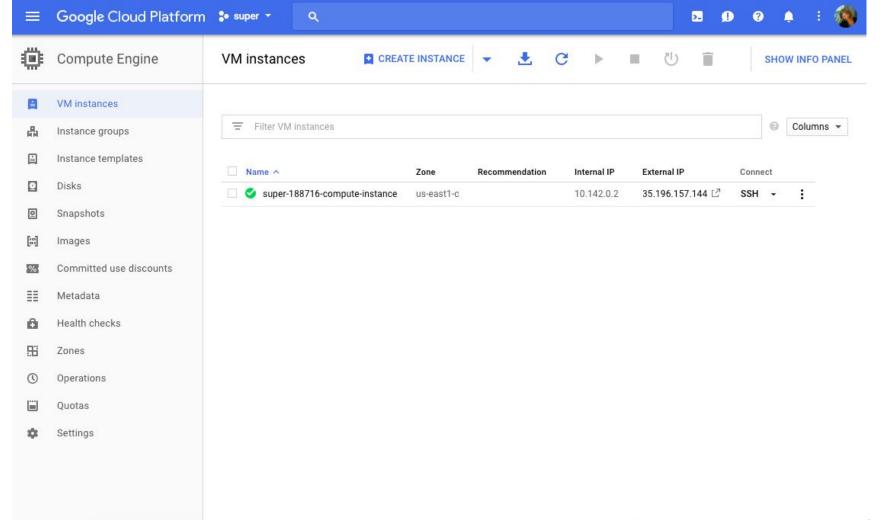
Contidential & Proprietary

18

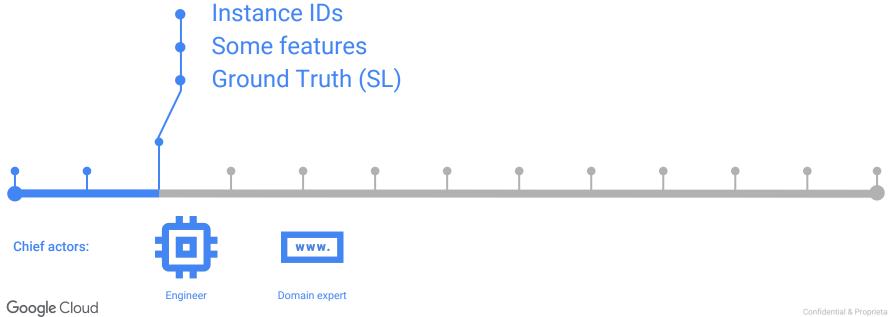








Step 2 is finished | You now have code for gathering



Step 2 is finished | You now have code for gathering

