

MODUL PRAKTIKUM

KODE IS429 – BIG DATA ANALYTICS
PROGRAM SARJANA S1 SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA

Gedung B Lantai 5, Kampus UMN
Jl. Scientia Boulevard, Gading Serpong, Tangerang-15811 Indonesia
Telp: +62-21.5422.0808 (ext. 1803), email: ict.lab@umn.ac.id, web: umn.ac.id



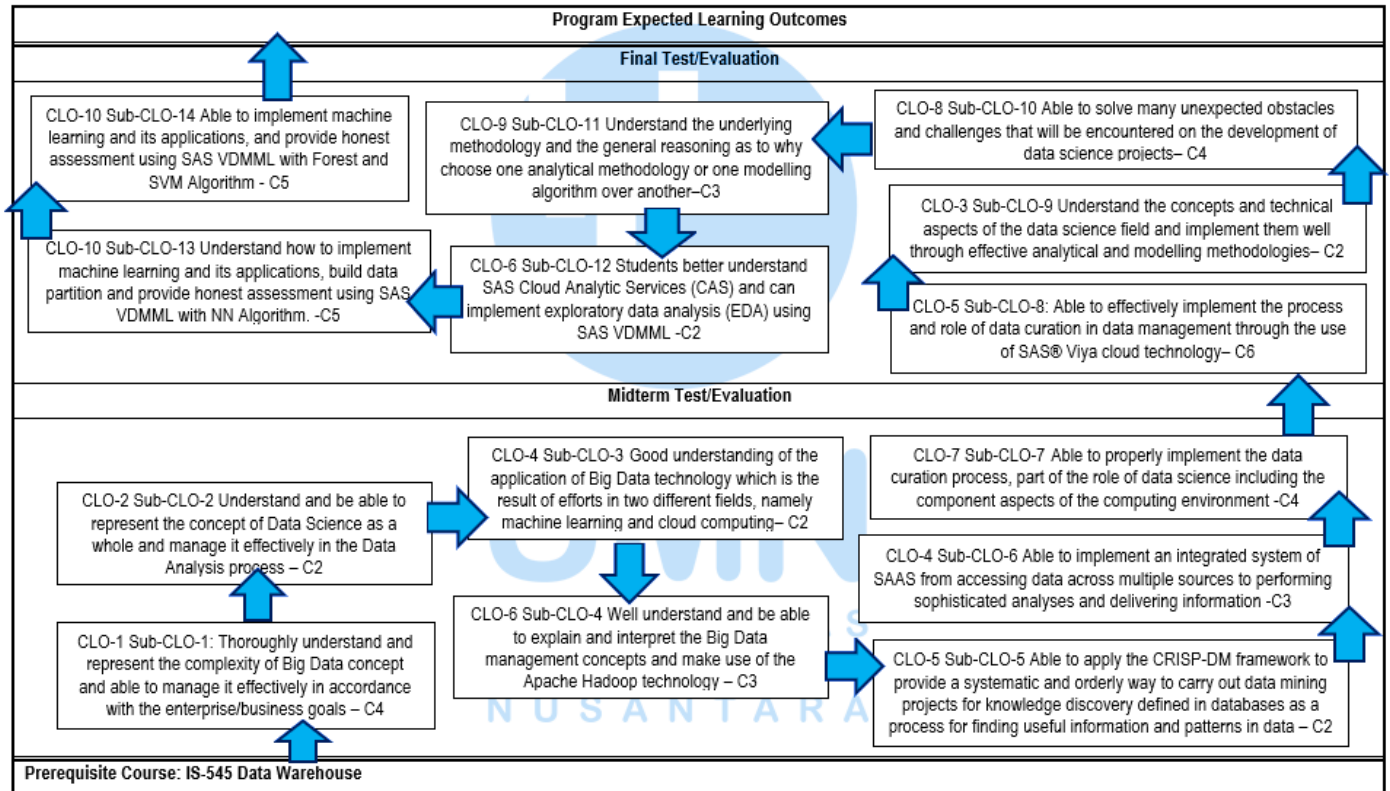
TABLE OF CONTENTS

TABLE OF CONTENTS.....	2
PRACTICUM LEARNING ACHIEVEMENTS	3
MODUL 1 GOOGLE BIGQUERY.....	4
MODUL 2 GOOGLE BIGQUERY ANALYSIS.....	9
MODUL 3 GOOGLE DATA STUDIO	18
MODUL 4 SAS® VIYA PROGRAMMING INTERFACE: SAS® STUDIO.....	38
MODUL 5 SAS® STUDIO: SAS BASIC PROGRAMMING	55
MODUL 6 SAS® STUDIO INTERFACE: EXPLORING DATA	69
MODUL 7 SAS® DATA PREPARATION APPLICATION	84
MODUL 8 SELF-SERVICE DATA PREPARATION IN SAS® VIYA®	94
MODUL 9 SELF-SERVICE DATA PREPARATION IN SAS® VIYA®	110
MODUL 10 SAS® Viya® STEP-BY-STEP DATA SCIENCE PROJECT	PART 1126
MODUL 11 SAS® Viya® STEP-BY-STEP DATA SCIENCE PROJECT	PART 2139
MODUL 12 SAS® VDMML Exploratory Data Analysis (EDA)	154
MODUL 13 SAS® VDMML: DATA PARTITION & NEURAL NETWORK MODELS.....	168
MODUL 14 SAS® VDMML: BUILD & COMPARE SEVERAL TYPES OF MODELS	183

PRACTICUM LEARNING ACHIEVEMENTS

Learning Outcome Stages Chart

Learning Outcome Stages Chart



MODUL 1

GOOGLE BIGQUERY

THEME DESCRIPTION

Mahasiswa memahami dan mampu menggunakan *multicloud serverless data warehouse* Google BigQuery untuk pengelolaan big data.

WEEKLY LEARNING OUTCOMES (SUB-LESSONS)

CLO-1-Sub-CLO-1 : Thoroughly understand and represent the complexity of Big Data concept and able to manage it effectively in accordance with the enterprise/business goals, C-4.

Through the processing steps consisting of:

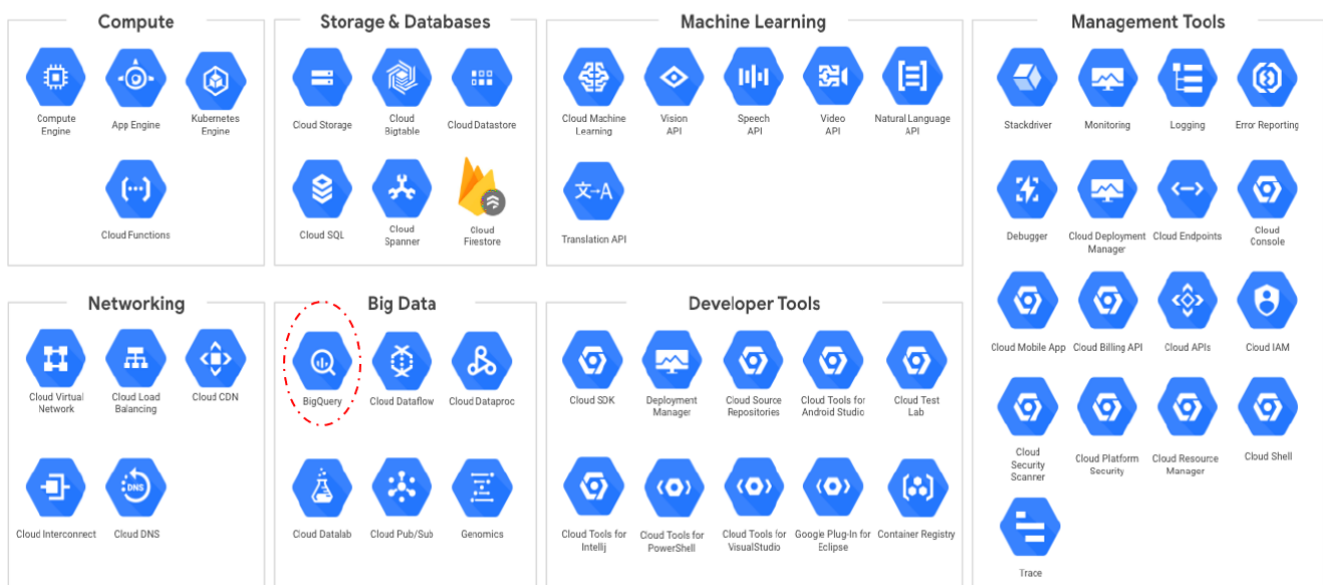
1. Google Cloud Platform Introduction
2. Google BigQuery (GBQ) Data Warehouse
3. Create GBQ Project using Public Data data-sets
4. Query Execution for 595.9 MB data-sets
5. Creation Data-sets in GBQ
6. Create table and importing data from public dataset

PRACTICUM SUPPORTS

1. Windows Operating System
2. (any) Browser Application

PRACTICUM STEPS

1) Google Cloud Platform Introduction



2) Google BigQuery (GBQ) multicloud serverless Data Warehouse

- Registration to get a GCP user id via <https://console.cloud.google.com>
- Choose "Try BigQuery Free"

3) Create GBQ Project using Public Data-sets

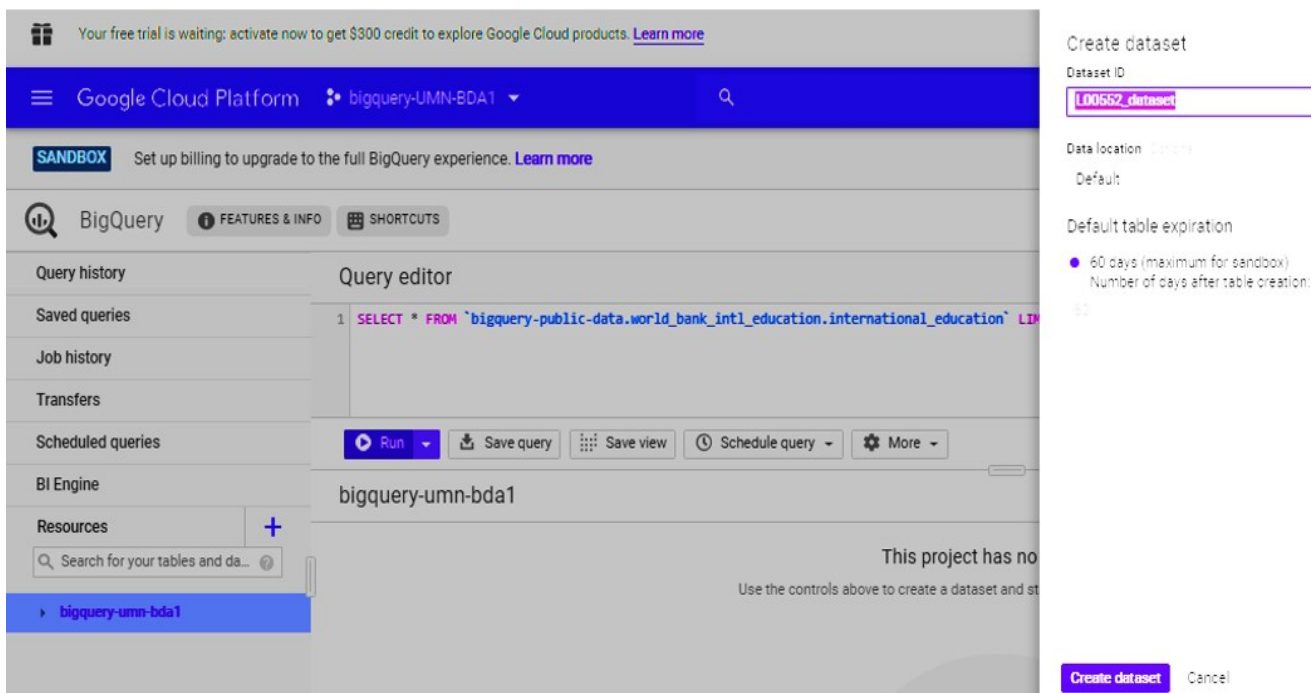
- Create a New Project "bigquery-NIM-BDA1"
- Use and choose the public data-sets dari GBQ
- The public dataset used and the appearance of its contents are in accordance with the Figure 1.1

4) Query Execution for 595.9 MB data-sets

- Using a public dataset from the World Bank, then perform a query in the query editor's query column and the resulting output is in accordance with Figure 1.2
- Congratulations! by executing the query, you have played with big data in amounts of 595.9 MB processed by 0.6 sec elapsed time.

5) Creation Data-sets in GBQ

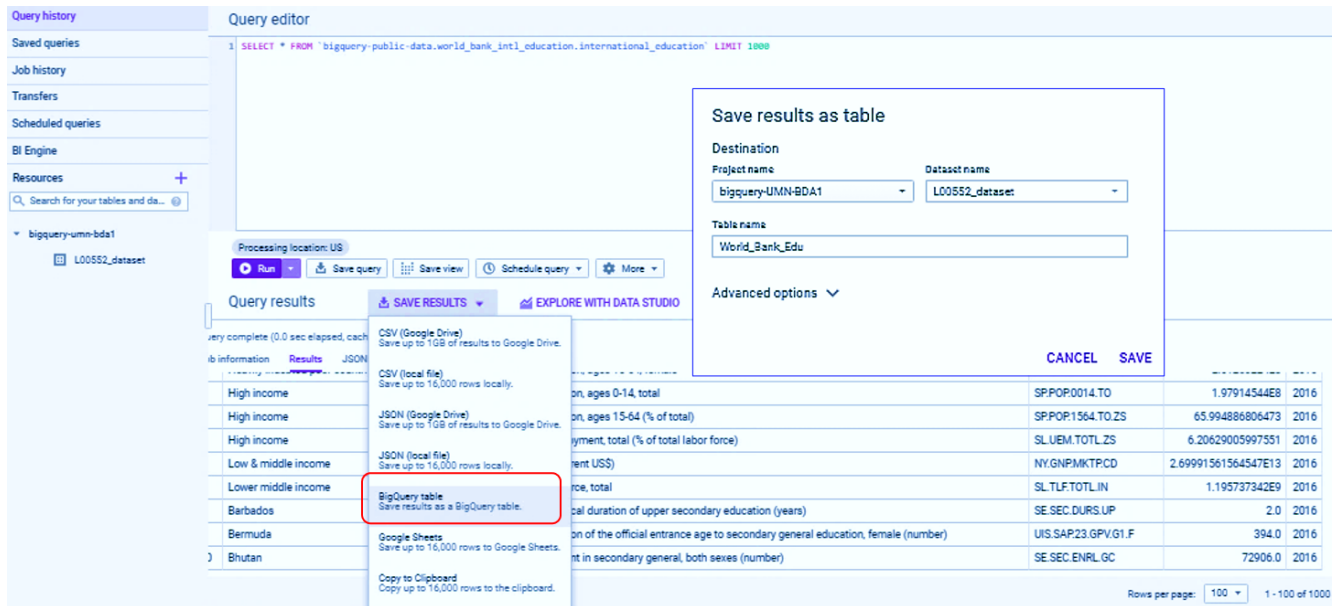
- After we try to query the table, we can store this table in our project ID so that it can be reused.
- Click the project ID bigquery-NIM-BDA1 (left hand column).
- Then click the CREATE DATASET button (right hand column), the display will appear as Figure 1.3



- Type the name of the dataset id in this case my dataset and Create dataset.
- Now we will save the world bank education data query into yourname_datasets.
- Reload your previous query using of Query History :
 - Select Query history (left column).
 - Click the query results that have a green check and Open query in editor
 - Select the BigQuery table
- The result of query history shown as Figure 1.3

6) Creation Data-sets in GBO

- Save the results in yourname_datasets and give a table name that is unique to the World_Bank_edu as follow :



The screenshot shows the Google BigQuery interface. On the left, there's a sidebar with 'Query history' and 'Resources'. The main area displays a 'Query editor' with a SQL query: `SELECT * FROM `bigquery-public-data.world_bank_intl_education.international_education` LIMIT 1000`. Below the query editor, there's a 'Query results' section showing a table of data. A 'SAVE RESULTS' dropdown menu is open, showing options like 'CSV (Google Drive)', 'JSON (Google Drive)', 'CSV (local file)', 'JSON (local file)', 'BigQuery table', 'Google Sheets', and 'Copy to Clipboard'. The 'BigQuery table' option is highlighted with a red box. A 'Save results as table' dialog box is also open, showing the 'Destination' section with 'Project name' as 'bigquery-UMN-BDA1', 'Dataset name' as 'L00552_dataset', and 'Table name' as 'World_Bank_Edu'. The dialog box has 'Advanced options' and 'CANCEL' and 'SAVE' buttons.

- Now we can see the worldbank_edu table saved in my datasets. Then try the query again, click the QUERY TABLE button (right hand column).
- The query results in a cached 0 second process. Which means it is not charged at the time of query. Although we get a 1 TB quota, we can also save by using cached (temporary table). This is very important if we have activated the billing to work on a larger project, then we can save costs
- The output of your query and exported to your table should be as Figure 1.4
- So we only need one query that is needed and then save it into the datasets that we make and then analyze it.
- Save all your screenshot(s) result as Figure 1.1 until 1.4 into 1.Tugas IS-429 Lab BDA-1 GOOGLE BIG QUERY yourname-NIM.doc and submit to e-Learning IS-429 Lab Week#1

OUTPUT/ RESULTS

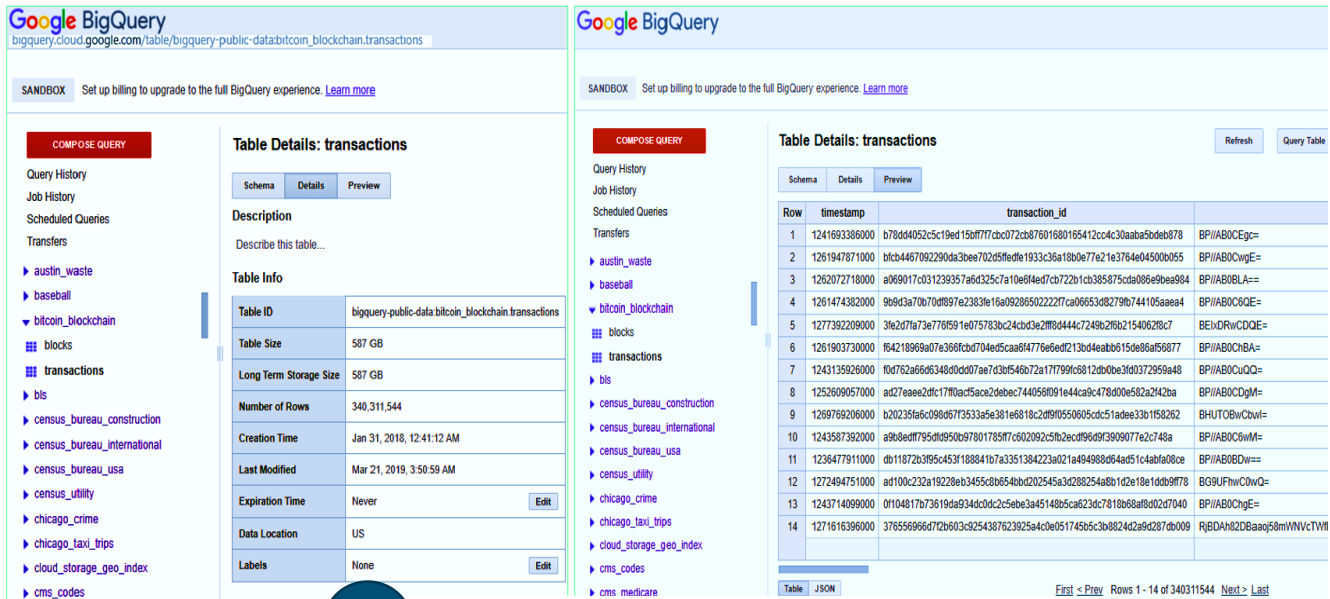


Table Details: transactions

Schema Details Preview

Description
Describe this table...

Table Info

Table ID	bigquery-public-data.bitcoin_blockchain.transactions
Table Size	587 GB
Long Term Storage Size	587 GB
Number of Rows	340,311,544
Creation Time	Jan 31, 2018, 12:41:12 AM
Last Modified	Mar 21, 2019, 3:50:59 AM
Expiration Time	Never
Data Location	US
Labels	None

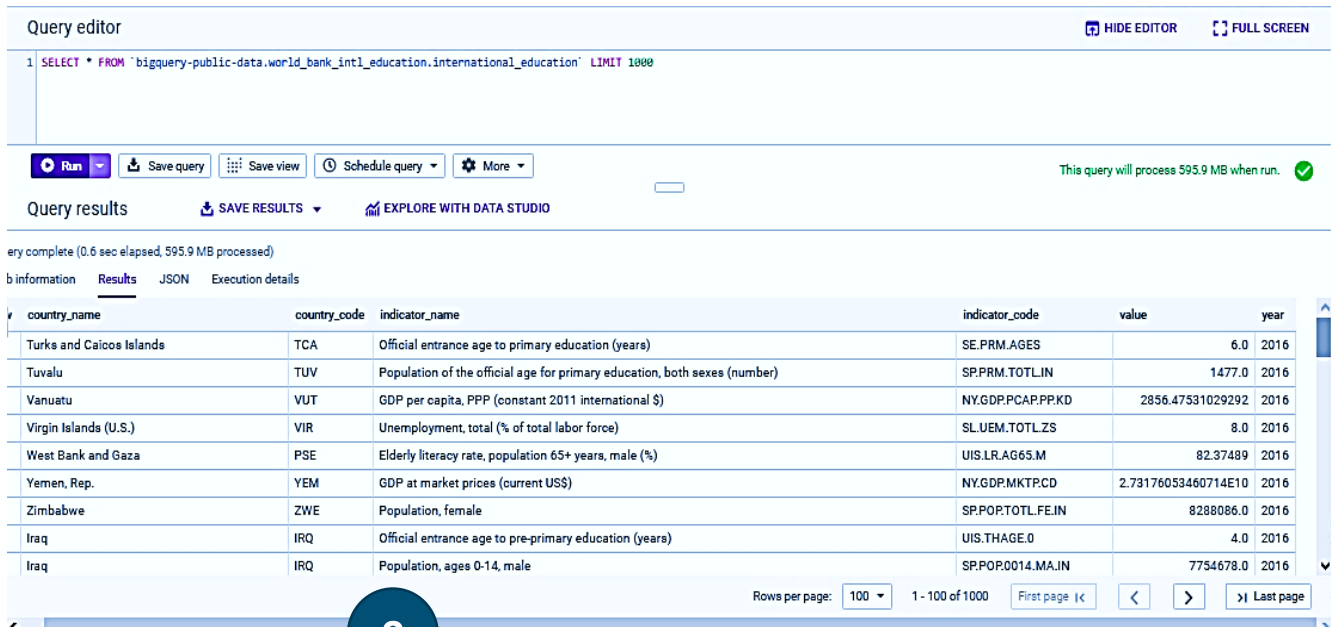
Table Details: transactions

Schema Details Preview

Row	timestamp	transaction_id	
1	1241693336000	b78dd4052c5c19ed15ef777dcb072cb87601680165412cc4c30aaba5bdeb878	BP//AB0CEg=
2	1261947871000	bfc4467092290da3bee702d5fede1933c36a18b0e77e21e3764e04500b055	BP//AB0CwgE=
3	1262072718000	a069017c031239357a6d325c7a10e04ed7cb722b1cb385875cda086e9bea984	BP//AB0BLA=
4	1261474382000	9b9d3a70b70d897e2383e1fa092865022227ca06653d8279b744105aae4	BP//AB0CQe=
5	1277392209000	3fe2d71a73e778f591e075783bc24cd3e2f80444c7249b2f0621540628c7	BE//DRwCQe=
6	1261903730000	f64218969a07e366fcb704ed5ca8f47766ed213bd4eab615de8a5f6877	BP//AB0ChBA=
7	1243135926000	f0d762a66d6348d0cd07ae7c3b546b72a177799c6812db0be3f0d372959a48	BP//AB0CwQ=
8	1252609057000	ad27eeae2dfc17f0ac5ace2debcc744056f091e44ca9c478d00e582a242ba	BP//AB0CDgM=
9	1269769206000	b20235fa6c098d57f3533a5e381e6818c2df9f0550605cd51ade913b1f58262	BHUTOBwCwM=
10	1243587392000	a9b8edf795d950b97801785f7c02092c5fb2cedf96993909077e2c746a	BP//AB0CwM=
11	1236477911000	db11872b395c45f188841b7a335138423a021a494988d54ad51c4abfa0ce	BP//AB0Dw==
12	1272494751000	ad100c232a19228eb3455c8b6540bd02545a3d288254a8b1d2e1f8e1dd89f78	B09UFWwCwQ=
13	1243714099000	0f104817b73619da934dc0dc2c5e6e3a45148b5ca623dc7818b68a0802cd7040	BP//AB0ChgE=
14	1271616396000	376556966d72b603c9254387623925a4c0e051745b5c3b8824d2a9d287db009	RjBDARH2DBaa958mVNVVtV

1

Figure 1.1 Public Data-sets



Query editor

HIDE EDITOR FULL SCREEN

```
1 SELECT * FROM `bigquery-public-data.world_bank_intl_education.international_education` LIMIT 1000
```

Run Save query Save view Schedule query More

This query will process 595.9 MB when run. ✓

Query results SAVE RESULTS EXPLORE WITH DATA STUDIO

Query complete (0.6 sec elapsed, 595.9 MB processed)

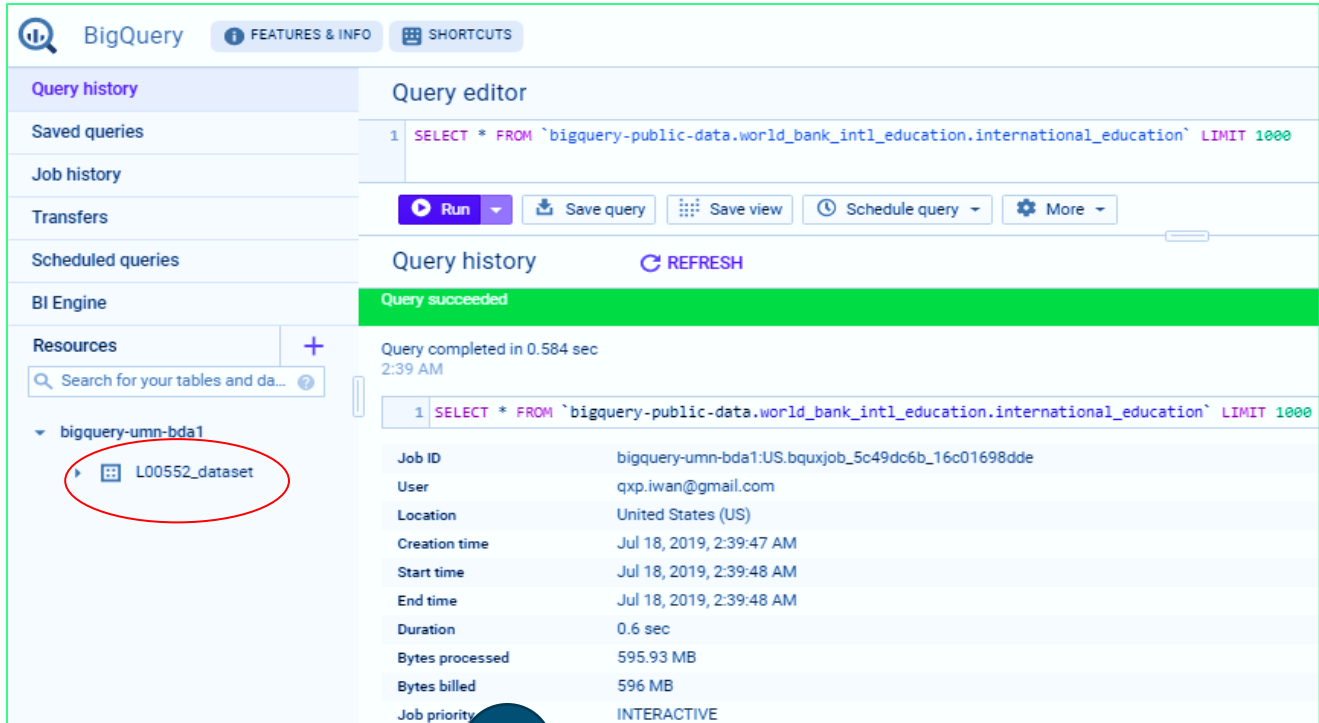
Information Results JSON Execution details

country_name	country_code	indicator_name	indicator_code	value	year
Turks and Caicos Islands	TCA	Official entrance age to primary education (years)	SE.PR.M.AGES	6.0	2016
Tuvalu	TUV	Population of the official age for primary education, both sexes (number)	SP.PR.M.TOTLIN	1477.0	2016
Vanuatu	VUT	GDP per capita, PPP (constant 2011 international \$)	NY.GDP.PCAP.PPKD	2856.47531029292	2016
Virgin Islands (U.S.)	VIR	Unemployment, total (% of total labor force)	SL.UEM.TOTL.ZS	8.0	2016
West Bank and Gaza	PSE	Elderly literacy rate, population 65+ years, male (%)	UIS.LR.AG65.M	82.37489	2016
Yemen, Rep.	YEM	GDP at market prices (current US\$)	NY.GDP.MKTP.CD	2.73176053460714E10	2016
Zimbabwe	ZWE	Population, female	SP.POP.TOTL.FE.IN	8288086.0	2016
Iraq	IRQ	Official entrance age to pre-primary education (years)	UIS.TH.AGE.0	4.0	2016
Iraq	IRQ	Population, ages 0-14, male	SP.POP.0014.MA.IN	7754678.0	2016

Rows per page: 100 1 - 100 of 1000 First page < > Last page

2

Figure 1.2 Query Execution 595.9 MB



Job ID	bigquery-umn-bda1:US.bqxjob_5c49dc6b_16c01698dde
User	qxp.iwan@gmail.com
Location	United States (US)
Creation time	Jul 18, 2019, 2:39:47 AM
Start time	Jul 18, 2019, 2:39:48 AM
End time	Jul 18, 2019, 2:39:48 AM
Duration	0.6 sec
Bytes processed	595.93 MB
Bytes billed	596 MB
Job priority	INTERACTIVE

3

Figure 1.3 Dataset & Query History

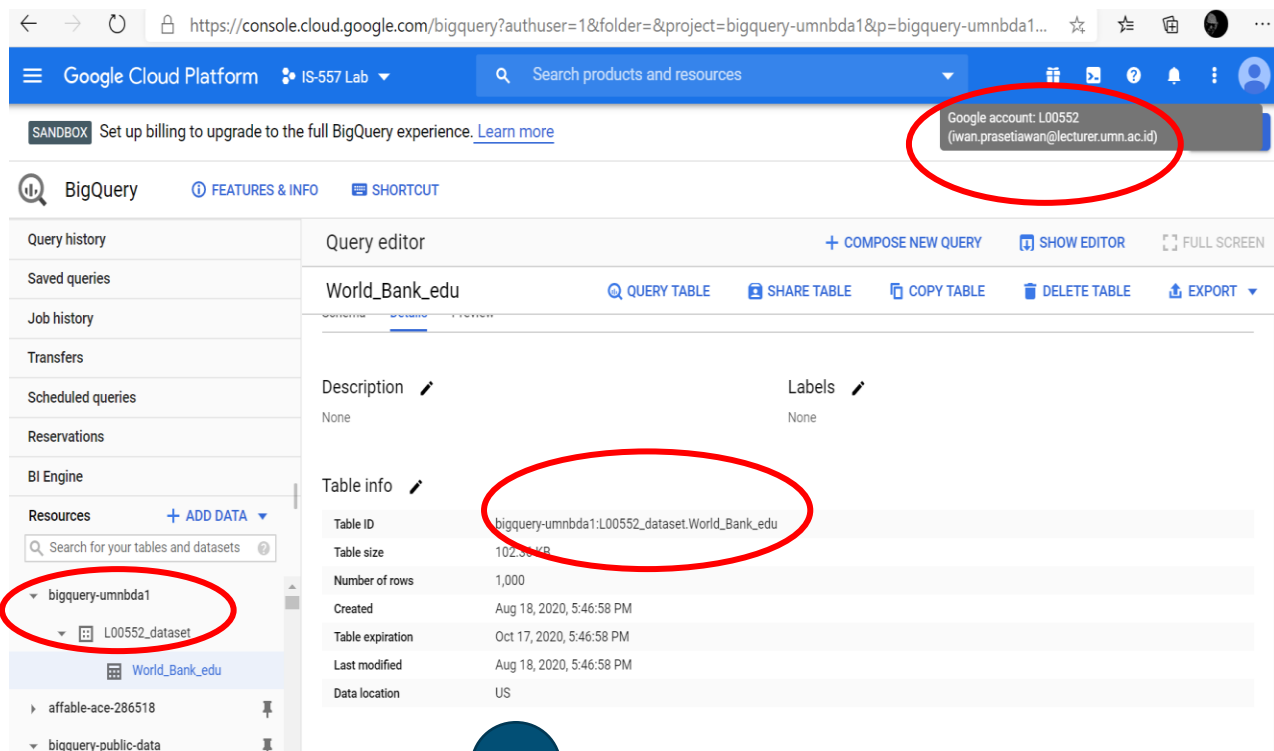


Table ID	bigquery-umn-bda1:L00552_dataset.World_Bank_edu
Table size	102.36 KB
Number of rows	1,000
Created	Aug 18, 2020, 5:46:58 PM
Table expiration	Oct 17, 2020, 5:46:58 PM
Last modified	Aug 18, 2020, 5:46:58 PM
Data location	US

4

Figure 1.4 Table Creation

REFERENCE

1. [Documentation | Google Cloud](#)
2. [Cloud Google | BigQuery](#)
3. Other additional references are excerpts from various Online Learning/websites.

The End