# Activity Exemplar: Partition data and create indexes in BigQuery

Here is a completed exemplar along with an explanation of how the exemplar fulfills the expectations for the activity.

### Assessment of Exemplar

partitions and indexes can be.

did you do well? Where can you improve? Use your answers to these questions to guide you as you continue to progress through the course. In the previous activity, you ran SQL code that created tables with partitions and indexes. Partitions and indexes help

Compare the exemplar to your completed activity. Review your work using each of the criteria in the exemplar. What

you create shortcuts to specific rows and divide large datasets into smaller, more manageable tables. By creating partitions and indexes, you can build faster and more efficient databases, making it easier to pull data when you need to analyze or visualize it. After creating the tables, you ran queries on those tables to compare their performance and demonstrate how useful

At each step, you took screenshots of the **Details** or **Execution Details** pane to compare to the following exemplar images. This will help you ensure that you completed the activity properly. It will also explain the context of why the

tables you created and the queries you ran differ from each other. By the end of this reading, you will understand how this activity demonstrates that partitions and clusters speed up queries and optimize database performance. Note that the answers for these queries might differ depending on whether you're using the sandbox or free trial/full version of BigQuery. The sandbox version might not read the full dataset, so the table size you receive might not match

the full version so you can check your work regardless of how you're using BigQuery. Explore the exemplar

the results you would get from the query in the full version. This reading explains the results for both the sandbox and

## This is the **Details** pane for the table you created without partitions or indexes. It simply describes the table size

**Table details** 

Table info

Ctd	
Created	Mar 6, 2023, 11:04:27 AM UTC-6
Last modified	Mar 6, 2023, 11:04:27 AM UTC-6
Table expiration	May 5, 2023, 12:04:27 PM UTC-5
Data location	US
Default collation	
Case insensitive	false
Description	
Labels	
Storage info @	
Number of rows	41,025
Number of rows Total logical bytes	4.37 MB
Number of rows Total logical bytes Active logical bytes	4.37 MB 4.37 MB
Number of rows Total logical bytes Active logical bytes Long term logical bytes	4.37 MB 4.37 MB 0 B
Number of rows Total logical bytes Active logical bytes Long term logical bytes Total physical bytes	4.37 MB 4.37 MB 0 B
Number of rows Total logical bytes Active logical bytes Long term logical bytes Total physical bytes Active physical bytes	4.37 MB 4.37 MB 0 B 0 B
Number of rows Total logical bytes Active logical bytes Long term logical bytes Total physical bytes	4.37 MB 4.37 MB 0 B

(4.37MB of logical and active bytes) and the number of rows (41,025).

Table info

is 0B and there is a section that includes "Table Type: Partitioned," "Partitioned by: Integer Range," "Partitioned on field: year," and "Partition filter: Not required." The partition range start (2015), end (2022) and interval is also shown.

The full version has a table size of 4.37MB, but it has the same additional partition section.

Table ID my-first-project-379816.mydataset.avocados\_partitioned Created Mar 6, 2023, 11:05:40 AM UTC-6 Last modified Mar 6, 2023, 11:05:40 AM UTC-6 Table expiration May 5, 2023, 12:05:40 PM UTC-5 Data location

Default collation  Case insensitive false  Description  Labels  Table Type Partitioned by Integer  Partitioned on field year  Partition Range Start 2015  Partition Range End 2022	er Range	
Description Labels  Table Type Partit Partitioned by Integer Partitioned on field year Partition Range Start 2015	er Range	
Labels  Table Type Partit  Partitioned by Intege  Partitioned on field year  Partition Range Start 2015	er Range	
Table Type Partit Partitioned by Intege Partitioned on field year Partition Range Start 2015	er Range	
Partitioned by Integer Partitioned on field year Partition Range Start 2015	er Range	
Partitioned by Integer Partitioned on field year Partition Range Start 2015	er Range	
Partitioned on field year Partition Range Start 2015		
Partition Range Start 2015		
Partition Range End 2022		
Partition Range Interval 1		
Partition filter Not re	equired	
Storage info		
Number of rows 41,02	e e	
Number of partitions 0	2	
Total logical bytes 4.37 l	MB	
Active logical bytes 4.371		
Long term logical bytes 0 B		
Total physical bytes 0 B		
Active physical bytes 0 B		
Long term physical 0 B bytes		
Time travel physical 0 B bytes		

May 5, 2023, 12:07:31 PM UTC-5

false

Not required

type

Table expiration Data location Default collation Case insensitive

Description Labels

Partition Range Interval Partition filter

Clustered by

344 ms

S00: Input

▶ S01: Aggregate

SHOW AVERAGE TIME SHOW MAXIMUM TIME @

Working timing

Wait:

Table Type Partitioned Partitioned by Integer Range Partitioned on field 2015 Partition Range Start Partition Range End

te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following reenshot, but the records read and records written should match with the Rows section.  Is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 025 and Records written: 3.		type	
Number of rows 41,025  Number of partitions 0  Total logical bytes 437 MB  Long term logical bytes 0B  Total physical bytes 0B  Long term logical bytes 0B  Long term physical 0B  Long	Storage info 9		
Number of partitions  O Total logical bytes  4.37 MB  Long term logical bytes  0 B  Total physical bytes  0 B  Long term logical bytes  0 B  Long term physical  bytes  Decution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the reenshots in the section.  te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following reenshot, but the records read and records written should match with the Rows section.  is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of was read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read:  O25 and Records written: 3.		41.025	
Total logical bytes 4.37 MB Active logical bytes 9.8 Total physical bytes 0.8 Total physical bytes 0.8 Long term logical bytes 0.8 Long term logical bytes 0.8 Long term physical bytes 0.8 Long term physical bytes 0.8 bytes  Prime travel physical bytes 0.8 Long term physical physics 1. Long term physical bytes 1. Lo			
Active logical bytes 4.37 MB Long term logical bytes 0.8 Active physical bytes 0.8 Active physical bytes 0.8 Active physical bytes 0.8 Time travel physical bytes 0.8 Time travel physical bytes 0.8 Execution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the reenshots in the section.  te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or or or the torun depending on differing BigQuery engine server speeds. Your screen might not match the following reenshot, but the records read and records written should match with the Rows section.  is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of two read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 0.25 and Records written: 3.			
Long term logical bytes 0 B Total physical bytes 0 B Long term physical 0 B bytes  ecution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the eenshots in the section.  te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or or or term to run depending on differing BigQuery engine server speeds. Your screen might not match the following eenshot, but the records read and records written should match with the Rows section.  is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 0.25 and Records written: 3.			
Active physical bytes 0 B Long term physical 0 B bytes  Time travel physical 0 B  ecution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the eenshots in the section.  te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following reenshot, but the records read and records written should match with the Rows section.  is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read:  Query results  A SAVEREBULTS A SEXPLORE DATA A  Conceptuation Details and the condition of the rows on the table. You'll find this in the S00:Input section, where Records read:		0 B	1
Active physical bytes 0 B Long term physical 0 B bytes  Time travel physical 0 B bytes  Time travel physical 0 B bytes  excution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't prear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the eenshots in the section.  te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following eenshot, but the records read and records written should match with the Rows section.  is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of two read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read:  O25 and Records written: 3.	Total physical bytes	0 B	
ecution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the eenshots in the section.  Ite: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following eenshot, but the records read and records written should match with the Rows section.  It is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read:  O25 and Records written: 3.		0 B	
ecution details  en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the reenshots in the section.  Ite: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following reenshot, but the records read and records written should match with the Rows section.  It is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 025 and Records written: 3.		0 B	
en, the Execution Details panes compare the query performance for each table. In the sandbox, these details won't pear for queries on the partitioned and partitioned and clustered tables. If you're using the sandbox, take note of the eenshots in the section.  Ite: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or order to run depending on differing BigQuery engine server speeds. Your screen might not match the following eenshot, but the records read and records written should match with the Rows section.  It is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 025 and Records written: 3.		0 B	
te: The Working timing section on your screen might vary in color or duration. Your SQL query might take longer or or creen to run depending on differing BigQuery engine server speeds. Your screen might not match the following eenshot, but the records read and records written should match with the Rows section.  It is is the Execution Details pane for the query on the table you created without partitions or clusters. The number of we read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 025 and Records written: 3.			nd clustered tables. If you're using the sandbox, take note of the
is is the <b>Execution Details</b> pane for the query on the table you created without partitions or clusters. The number of ws read is the total number of the rows on the table. You'll find this in the S00:Input section, where Records read: 025 and Records written: 3.  Query results  **SAVE RESULTS *** ** ** EXPLORE DATA **	orter to run depe	ending on differing BigQuery engine	server speeds. Your screen might not match the following
	is is the <b>Executio</b> ws read is the tot	on Details pane for the query on the al number of the rows on the table.	table you created without partitions or clusters. The number of
JOB INFORMATION RESULTS JSON EXECUTION DETAILS EXECUTION GRAPH PREVIEW	Query results		♣ SAVE RESULTS ▼    EXPLORE DATA ▼
	JOB INFORMATION	RESULTS JSON EXECUTION DETAILS EX	CUTION GRAPH PREVIEW

179 ms

S02: Output Records written: 3

13 ms

This is the Execution Details pane for the query on the table you created partitioned by an integer range. You'll notice that the number of records read is less. Now, Records read: 16,953 and Records written: 3. In this query, the database processes only the records from the partitions filtered by the where clause (type). When choosing a column to partition

0 ms

on, it is most effective to choose one that would frequently be used in the where clause.

324 B

13 ms

0 B @

Records read: 41025

Records read: 3 Records written: 3

≜ SAVE RESULTS ▼

Query results **▲** SAVE RESULTS ▼ JOB INFORMATION RESULTS **JSON EXECUTION DETAILS** EXECUTION GRAPH PREVIEW For help debugging or optimizing your query, check our documentation. Learn more. Bytes shuffled @ Bytes spilled to disk @ 347 ms 237 ms 324 B 0 B @ SHOW AVERAGE TIME SHOW MAXIMUM TIME \*\* S00: Input Records read: 16953 Records written: 3 S01: Aggregate Records read: 3 0 ms S02: Output Records read: 3 Compute: Records written: 3 0 ms

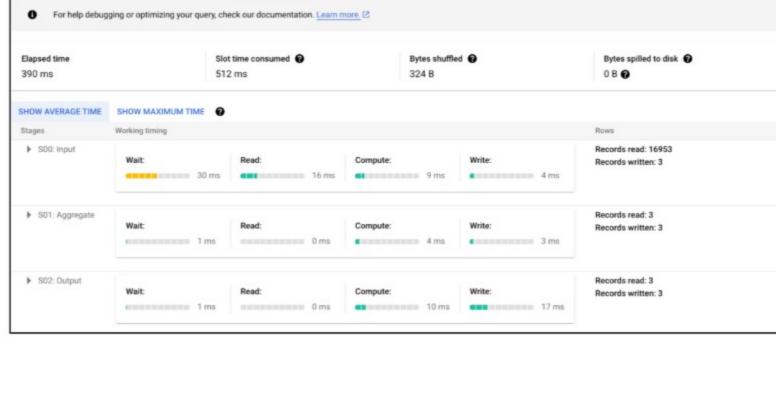
This is the Execution Details pane for the query on the table you created that is clustered by the type column. Records

partitioned one. However, the dataset you are using in this activity is too small to properly demonstrate that difference.

read: 16,953 and Records written: 3. Typically, a query on the clustered table would process fewer records than the

In other projects, you might find that clustering a table leads to a significant decrease in records read.

JOB INFORMATION **EXECUTION DETAILS** EXECUTION GRAPH PREVIEW For help debugging or optimizing your query, check our documentation. Learn more. Elapsed time Slot time consumed @ Bytes spilled to disk @ 390 ms 512 ms 324 B 0 B @



Query results

Key takeaways This activity demonstrates the impact of using partitions and indexes (known as clusters in BigQuery) in database

tables. You can use them to optimize query performance and minimize processing costs. In this exercise, applying partitions and clustering means that BigQuery can break all 41,025 records into smaller, more manageable tables to

read. The benefits of partitioning will be even more evident with larger datasets. Use this technique to optimize

Dislike

△ Like

database performance in your future projects.

Report an issue