

Assessment date

Jun 27, 2025, 3:15:00 AM

[bq-assessment-feedback@google.com](mailto:bq-assessment-feedback@google.com)



# Assessment Report: Snowflake

---

---

---

Report Completeness

93%

Longer duration of usage logs + metadata generates better details and insights

Usage timeframe:

Start Time

Dec 31, 2024, 4:00:00 PM

End Time

Jun 25, 2025, 4:59:59 PM

# Summary Highlights

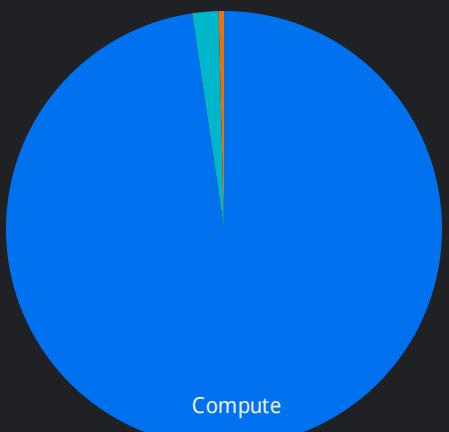
## 1. Existing system



\$XXX

Please contact Google Sales  
to unlock the numbers automatically computed for your workload.

Approximate Annual Total  
Cost\*



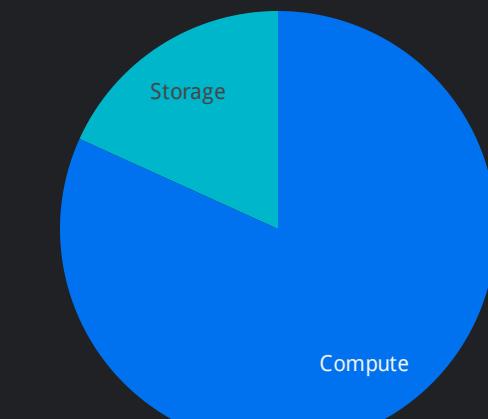
- Compute
- Storage
- Cloud Services

## 2. BigQuery steady state



\$YYY

Estimated Annual Total Cost\*



- Compute
- Storage

## 3. Migration path

Automatically Translated  
56,685

Distinct Queries  
115,616

SQL Automatically  
Translated

SQL Offline Effort

Low Impact # Patterns  
4,268

Medium Impact # Patterns  
1,314

High Impact # Patterns  
41

\* Assumptions:

Snowflake Credit Price: \$3, Storage Price \$23 TB/Month

BigQuery Enterprise Edition.

Data transfer costs not included.

Serverless features costs not included.

Based on Usage Period:

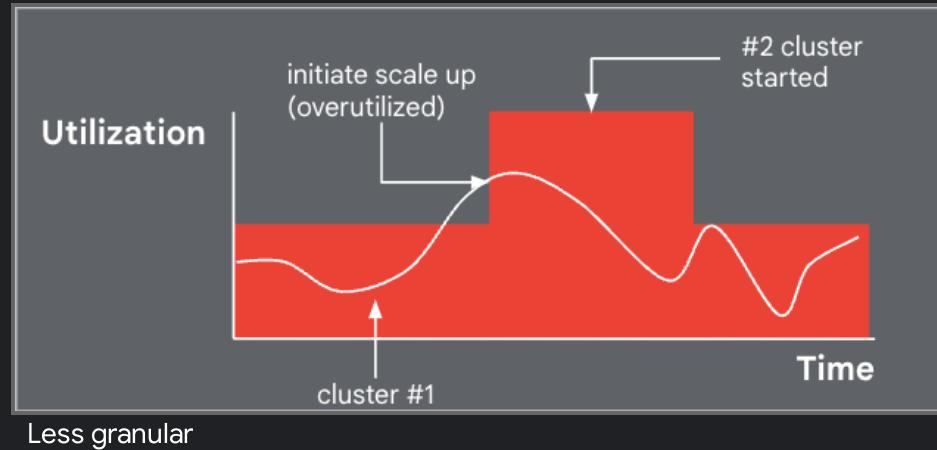
Dec 31, 2024, 3PM - Jun 25, 2025, 4PM

# Snowflake vs BigQuery Pricing Models

Snowflake

**1 Credit**

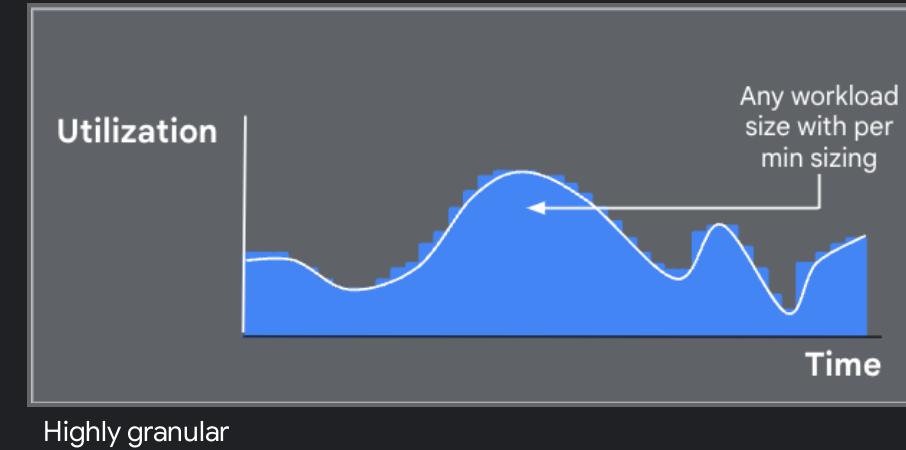
**Multi-cluster**



Compute

**50 Slots\***

**Autoscaling**



Edition

**\$2** / credit

**Standard**

**\$3** / credit

**Enterprise**

**\$4** / credit

**Business Critical**

**\$23** / TB month

**Capacity**

Storage

**Standard**

**\$2** / 50 slot hour

No commitment

**Enterprise**

**\$1.8**

3 year commitment

**\$2.4** / 50 slot hour

1 year commitment  
No commitment

**\$3**

**Enterprise Plus**

**\$3**

3 year commitment

**\$4** / 50 slot hour

1 year commitment  
No commitment

**\$5**

Documentation

[Editions features](#)

[Slots explained](#)

[Autoscaling and baseline explained](#)

\* Data is obtained from benchmarking. It is a rounded up average. Note that for ingestion the ratio is much lower.

# Total Cost of Ownership

DEMO

## Edition

BigQuery Edition: Enterprise

Commitment: 1 Yr Commit

## Compute

Based on the credits usage from Snowflake for the following period:

From  
Dec 31, 2024, 4PM

To  
Jun 25, 2025, 4PM

### Baseline

Lowest Cost Baseline Value

No data

Baseline (Commitment)

0

Annual Baseline (USD)

\$XXX

### Autoscaling

Annual  
Autoscaling Slot  
Hours

Annual Slot Hours

YYY

Annual Slot Hours

YYY

Max  
Reservation

Slots  
No data

Expected Autoscaling Cost (USD)  
\$XXX

### Total Compute

Expected Annual Compute (USD)  
\$XXX

## Storage

% of active storage

30

**Active Storage**  
(with BQ compression)

**Long-term Storage**  
(with BQ compression)

GB  
52,166.2

GB  
121,721.14

Active Storage (USD)  
\$XXX

Long Term Storage (USD)  
\$XXX

% of data loaded / changed

0

Loading (Standard storage)

GB / month  
0

Annual Ingestion Cost (USD)  
\$XXX

### Total Storage

Annual Storage Cost (USD)  
\$XXX

## Total

Total Cost (USD) \*  
\$XXX

Please contact Google Sales to unlock the numbers automatically computed for your workload.

# Automatic Translation Highlights

Automatically Translated  
56,685

Distinct Queries  
115,616

by User

Top 5

User Name	Translation Percentage ▾
ACCESS_LAYER_TEST_USER	100%
HE_ALTERYX_SYS_USER	100%
UKASTKA	100%
UDARA.GUNATHILAKE1@PEARSON.C...	100%
JUANJO.MARTINCARA@PEARSON.C...	100%

0% 20% 40% 60% 80% 100%  
1 - 5 / 187 < >

Bottom 5

User Name	Translation Percentage ▾
AHSAN.NISAR@PEARSON.COM	0%
UNAWADA	0%
PRASHANT.SINGH@PEARSON.COM	0%
UCHOUSU	0%
UBANDK2	0%

0% 20% 40% 60% 80% 100%  
1 - 5 / 187 < >

by Database

Top 5

Database Name	Translation Percentage ▾
sms_v2_ingestion	100%
bytecode	100%
pla_qa_org	100%
product_data_intelligence	100%
data_engineering_metadata	100%

0% 20% 40% 60% 80% 100%  
1 - 5 / 162 < >

Bottom 5

Database Name	Translation Percentage ▾
dac_intern_db	0%
Unknown	16.96%
salesadm_gamma	18.6%
mxci_sandbox	26.83%
mypedia	35.08%

0% 20% 40% 60% 80% 100%  
1 - 5 / 162 < >

by Error Message

Top 5

Message	Cou... ▾
BigQuery SQL does not support COPY FILE.	21,940
No target-dialect support for source-dialect-specific GET	21,902
Not yet implemented statement type GET	21,902
Positional bind variables are not yet fully supported.	18,797
An explicit cast from the input was preserved, but the expression is of an unknown or erroneous type; this is typically due to missing or incomplete metadata.	17,507

# System Overview

Usage Timeframe:

From  
Dec 31, 2024, 4PM

To  
Jun 25, 2025, 4PM

Databases

1

Tables

21,289

Queryable Data Size (TB)

216.76

Contingency Storage Size (TB)

18.3

Retained For Clone Size (TB)

7.53

Total Storage Size (TB)

242.59

QueryText

23,455

Queries per Day

11,292

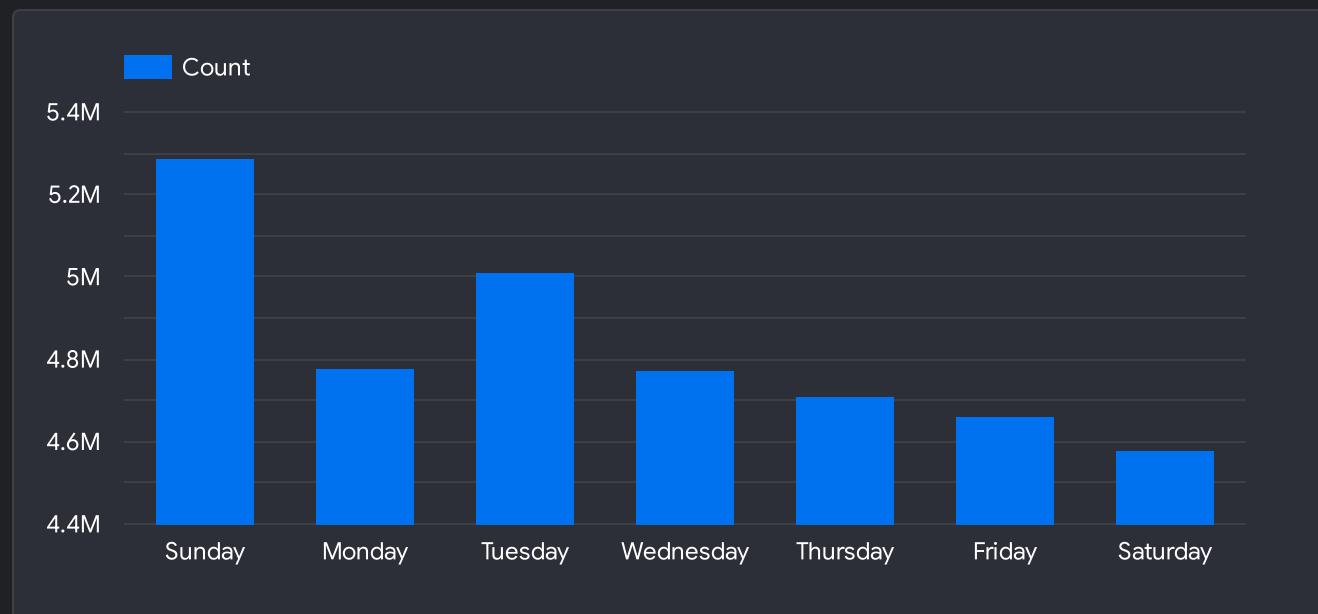
Data Scanned By Queries (GB)

11.5M

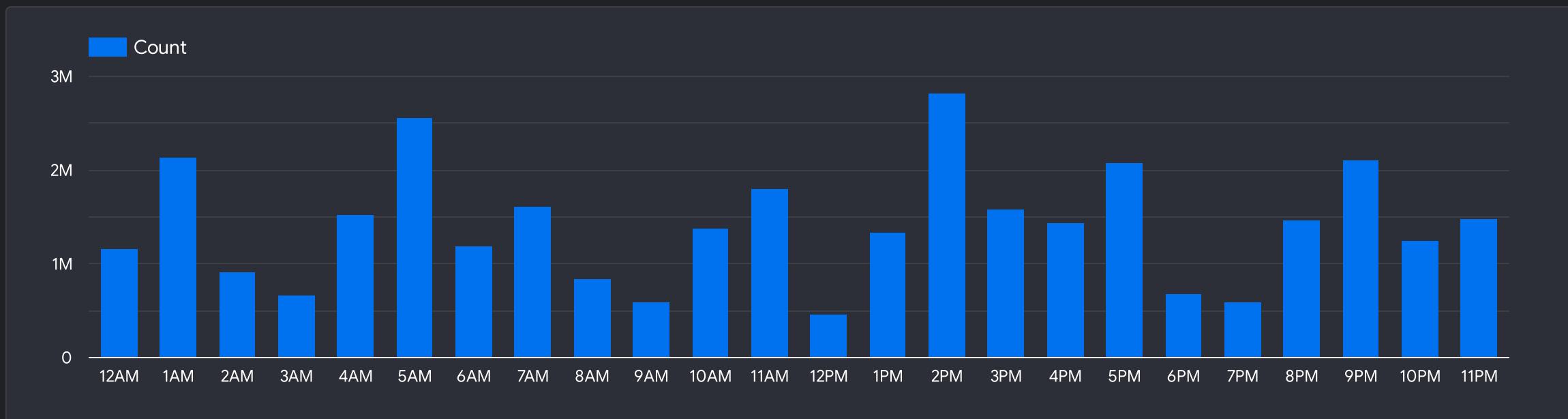
Users

170

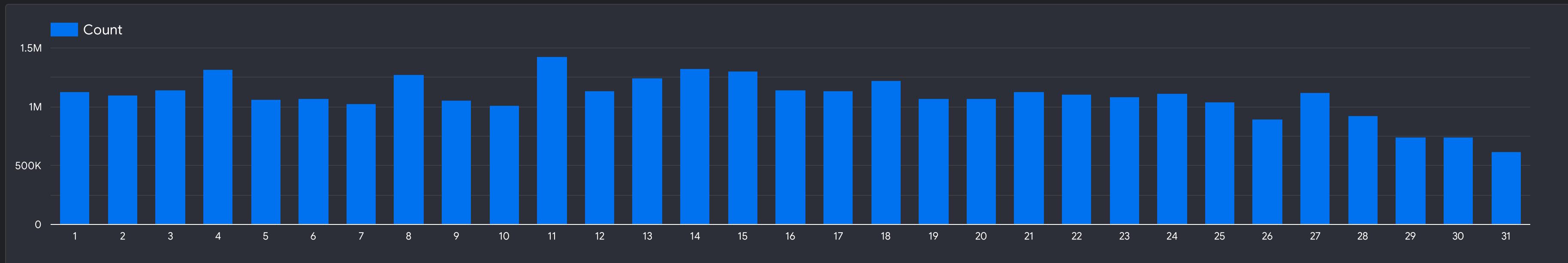
Weekday Utilization



Hourly Utilization



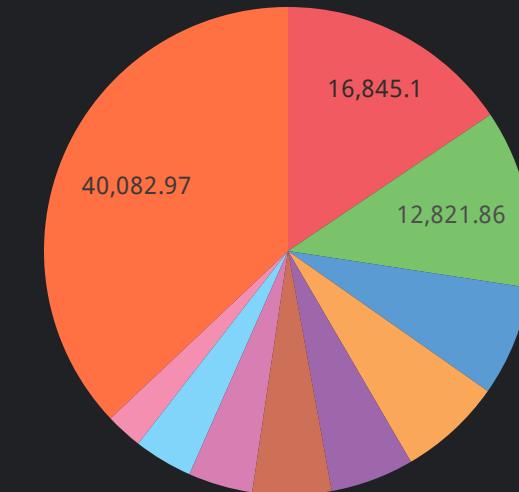
Daily Utilization



# Virtual Warehouses Overview

WarehouseName	Size	Uptime	Total Credits	Approximate Cost in Snow...	% of Total Cost
		No data			
Grand total		null	null	null	null

## Approximate Annual Cost Structure

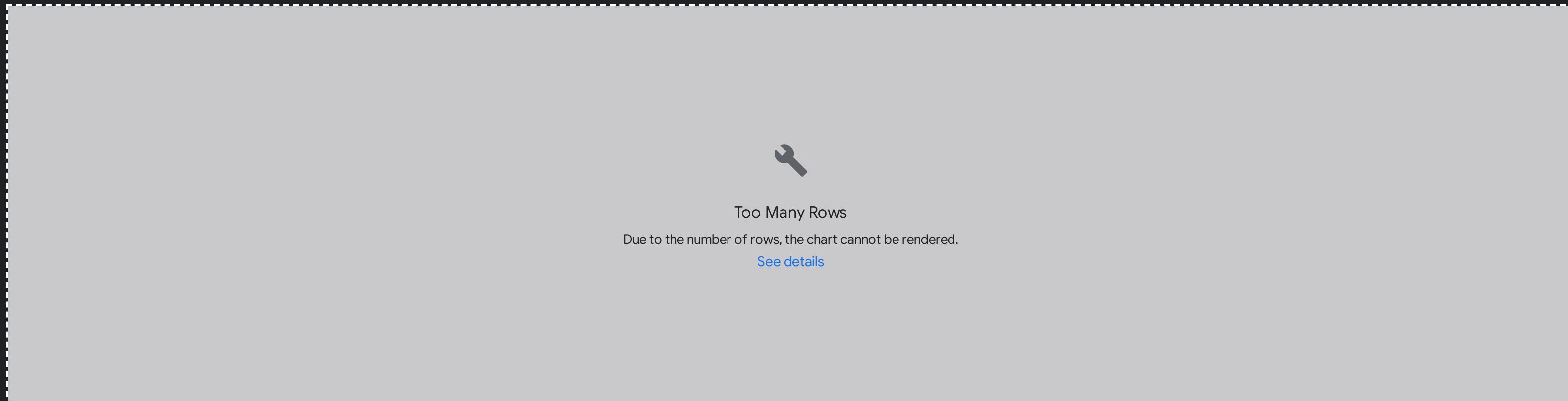


- MASTERING\_INGESTION
- MYLABS\_XL\_INGESTION
- ATTUNITY\_INGESTION\_S
- PDI\_AUTOMATION\_XL
- EVENT\_DATA\_HUB
- MXCI\_ELT\_WH
- PERFSVC\_WH
- MYLABS\_XL\_INGESTION



Warehouse Name

## Utilization Histogram - the Number of Nodes Running at a Given Time



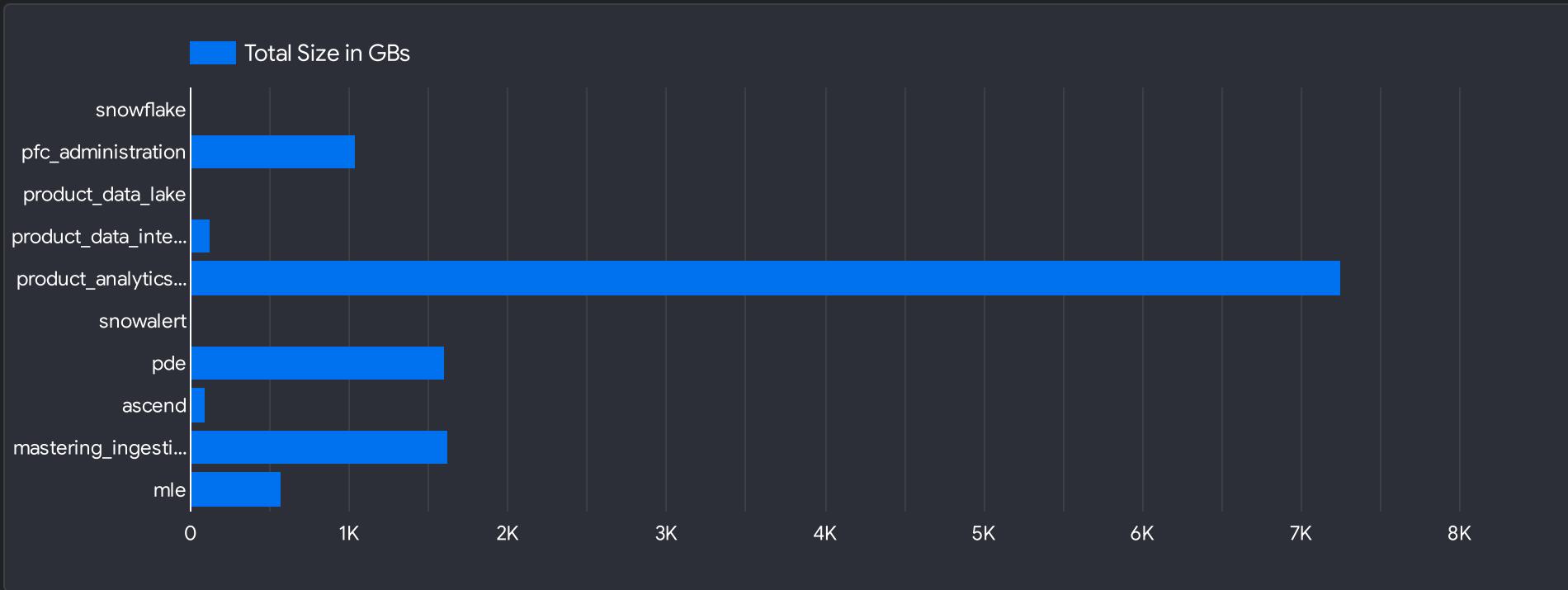
The number of nodes corresponds to the size of a virtual warehouse:

X-small	1
Small	2
Medium	4
Large	8
X-Large	16
2X-Large	32
3X-Large	64
4X-Large	128
5X-Large	256
6X-Large	512

Assumptions:

# Table Volume & Access Rate

Total (Queryable) Table Volume by Database



20 Largest Tables with Access Rate by User

FullTableName	UserName	TableSizeInMBs	UsageCount
autobahn_ingestion.raw.paf_hub_assessment_...	pdl_qe_service	16.1M	1.3K
	autobahn_sys_user	16.1M	175
	vvardh1	16.1M	35
	ueverja	16.1M	5
	thomda12	16.1M	2
autobahn_ingestion.raw.paf_hub_assessment_...	pdl_qe_service	6.6M	1.3K
	autobahn_sys_user	6.6M	175
xl_ingestion.ga.xl_154558773_ga_custom_dime...	ufox9is	9.4M	1
autobahn_ingestion.raw.pi_common_usersessi...	pdl_qe_service	2.5M	1.7K
	autobahn_sys_user	2.5M	175
autobahn_ingestion.raw.ale_ale_aleuseractivity...	pdl_qe_service	2.4M	1.7K
	autobahn_sys_user	2.4M	175
etext_ingestion.telemetry_io.etext_telemetry_hi...	everest_sys_elt_admin	1.5M	1.1K
	ufox9is	1.5M	17

Largest Tables

Database Name	SchemaName	Table Name	Size (GBs)
autobahn_ingestion	raw	paf_hub_assessment_multi_part_question...	15,708.7
etext_ingestion	firebase	event_user_analytics_268866320	12,918.5
xl_ingestion	ga	xl_154558773_ga_custom_dimension	9,219.5
naiils	experiment	tdx_from_2019	7,879.1
autobahn_ingestion	raw	paf_hub_assessment_multi_part_question...	6,457.6
naiils	experiment	kctcs_3yr_spring	6,380.1
lim_efficacy	efficacy_work	userevents_revisits_mlm	6,272
etext_ingestion	firebase	raw_analytics_268866320	6,068.2
product_data_mgmt	standard_formats	multi_part_question_user_answered_hed	3,788.9
reader_plus_ingestion	firebase	event_user_analytics_248889779	3,500.5
naiils	experiment	tdx_interactions_from_2019	3,401.1
lim_efficacy	efficacy_work	userevents_mlm	3,118
lim_efficacy	efficacy_work	userevents_revisits	2,561.9
autobahn_ingestion	raw	pi_common_usersessiontracking_1_0_0	2,422.3
autobahn_ingestion	raw	ale_ale_aleuseractivityattempt_1_0_1	2,379.2
etext_ingestion	ga4	etext_custom_field	2,244.3
etext_ingestion	telemetry_io	etext_telemetry_hit_io	2,173.4
tdx_refinement	tdx	tdx_messages_item_attempt_variables	2,053
naiils	experiment	autobahn_2223	1,991
revel_ingestion	ga	revel_86098200_ga_detail	1,895.3
revel_ingestion	ga	revel_86098200_ga_custom_dimension	1,734.7
xl_ingestion	ga	xl_154558773_ga_detail	1,722.7
etext_ingestion	telemetry	etext_telemetry_custom_fields	1,699.2
tdx_refinement	tdx	tdx_messages_user_interactions	1,592
event_data_hub_refinement	standard_formats	user_unloads_content	1,552.7

# Most frequently used tables

By Database Name

By User Name

# Table Usage

Usage timeframe:

From Dec 31, 2024, 4:00:00 PM

To Jun 25, 2025, 4:59:59 PM

Tables by Number of Users

Full Table Path	Users Count ▾
event_data_hub_refinement.standard	34
mylabs_xl_ingestion_v2.hed.home... autobahn_ingestion.autobahn_etl.au...	12
product_data_mgmt.standard_format...	10
autobahn_ingestion.raw.schemaregi...	8
prod_analytics_coe_etext.sandbox.s...	8
ies_ingestion.ies.ies_identity_profile	8
sms_v2_ingestion.smsadmin.access...	8
ies_ingestion.work.batch_stage	8
event_data_hub_refinement.standar...	7
mylabs_xl_ingestion_v2.hed.testresu...	7
event_data_hub_refinement.historic...	6
mastering_ingestion.chemistry.apite...	6
pde.he.organization	6
cortex_analyst_demo.courses_and_i...	6
naiils.he_services.subs_all_net	6
autobahn_ingestion.raw.paf_hub_as...	6
etext_ingestion.mobile_app_store.d...	5
mastering_ingestion.aandp.course	5
mastering_ingestion.biology.apitem...	5
pde.sandbox.orgservice_full_04_10_...	5
autobahn_ingestion.raw.socket_inte...	5
data_quality_sandbox.refinement.co...	5
etext_ingestion.telemetry_io.etext_t...	5
etext_ingestion.mobile_app_store.io...	5

Tables by Number of Queries

Full Table Path	Count ▾
event_data_hub_refinement.standard	2,215,226
autobahn_ingestion.autobahn_etl.au...	2,119,909
mastering_ingestion.chemistry.log_u...	253,168
mastering_ingestion.biology.log_use...	175,024
mastering_ingestion.astronomy.user...	147,818
mastering_ingestion.astronomy.apit...	145,446
mastering_ingestion.astronomy.log_...	144,996
mastering_ingestion.astronomy.assis...	144,346
mastering_ingestion.astronomy.apus...	144,302
mastering_ingestion.astronomy.user...	144,126
mastering_ingestion.astronomy.apit...	144,026
mastering_ingestion.astronomy.log_...	144,012
mastering_ingestion.astronomy.apus...	143,871
mastering_ingestion.chemistry.log_s...	143,870
mastering_ingestion.chemistry.apus...	143,806
mastering_ingestion.chemistry.assign...	143,806
mastering_ingestion.chemistry.apite...	143,576
mastering_ingestion.chemistry.log_c...	143,572
mastering_ingestion.aandp.apitemu...	143,567
mastering_ingestion.aandp.log_sub...	143,478
mastering_ingestion.aandp.assignm...	143,460
mastering_ingestion.aandp.apuserfi...	143,372
mastering_ingestion.aandp.apitemu...	143,281
mastering_ingestion.aandp.log_corr...	143,279
mastering_ingestion.chemistry.apite...	143,254

Tables by Writes

Full Table Path	Count ▾
product_data_lake.alchemy_qe.cds_...	59,845
ingestion_activation.ingestion_contr...	25,907
event_data_hub_refinement.standard	15,136
event_data_hub_refinement.standard	13,764
event_data_hub_refinement.standard	12,002
perfsvc.xl_delta_extract.audit_hed	11,978
perfsvc.xl_delta_extract.audit_school	10,330
etext_refinement.contextual_data_lo...	9,761
perfsvc.xl_delta_extract.audit_global	9,689
bytecode.refinement.job_status	9,110
event_data_hub_refinement.standard	9,100
event_data_hub_refinement.standard	8,581
ies_ingestion.qe.iesct1	8,224
event_data_hub_refinement.standard	7,972
revel_refinement.contextual_data_lo...	7,812
event_data_hub_refinement.standard	7,166
event_data_hub_refinement.standard	6,462
pde.he.organization_test_results	5,999
mylabs_xl_refinement.contextual_da...	5,893
event_data_hub_refinement.standard	5,632
event_data_hub_refinement.standard	5,532
event_data_hub_refinement.standard	5,326
event_data_hub_refinement.standard	4,928
event_data_hub_refinement.standard	4,797
event_data_hub_refinement.standard	4,602

Tables by Reads

Full Table Path	Count ▾
event_data_hub_refinement.standard	2,214,174
autobahn_ingestion.autobahn_etl.au...	2,119,855
mastering_ingestion.chemistry.log_u...	126,586
product_data_intelligence.etextmobi...	94,026
mastering_ingestion.biology.log_use...	87,513
perfsvc.xl_delta_extract.audit_hed	85,447
perfsvc.xl_delta_extract.audit_school	80,071
perfsvc.xl_delta_extract.audit_global	78,254
mastering_ingestion.astronomy.user...	73,924
mastering_ingestion.astronomy.apit...	72,739
mastering_ingestion.astronomy.log_...	72,604
mastering_ingestion.astronomy.assis...	72,279
mastering_ingestion.astronomy.apus...	72,257
mastering_ingestion.astronomy.user...	72,078
mastering_ingestion.chemistry.log_s...	72,054
mastering_ingestion.astronomy.apite...	72,028
mastering_ingestion.chemistry.assign...	72,022
mastering_ingestion.chemistry.apus...	72,022
mastering_ingestion.astronomy.log_...	72,021
mastering_ingestion.aandp.usersess...	71,953
mastering_ingestion.chemistry.apite...	71,804
mastering_ingestion.chemistry.log_c...	71,802
mastering_ingestion.aandp.apitemu...	71,802
mastering_ingestion.aandp.log_sub...	71,802
mastering_ingestion.aandp.assignm...	71,793

1 - 25 / 5362 < >

1 - 25 / 5362 < >

1 - 25 / 1853 < >

1 - 25 / 4830 < >

# Queries (last 30 days)

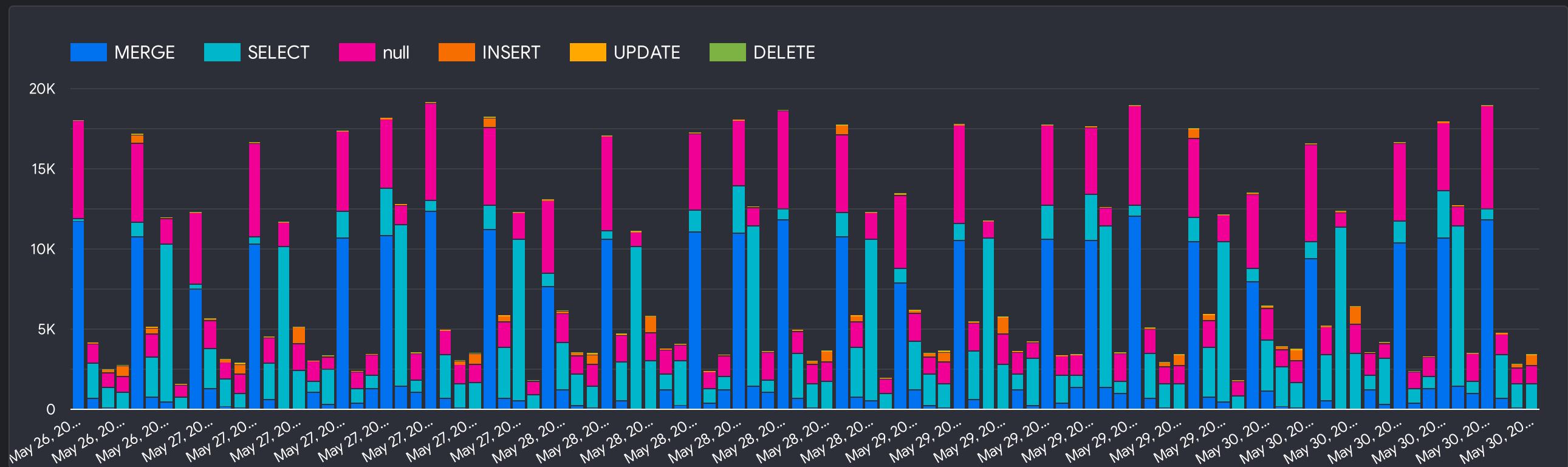
Tables Queried

By User Name

Type

Warehouse

Query Histogram by Type and Time



Query Run Frequency

QueryText	Record Count
1. MERGE INTO "HED"."TESTRESULTSDATA" ...	120
2. CALL incremental_main('IES')	120
3. COPY INTO @PLA_SMS_SUBSCRIPTION...	90
4. COPY INTO @PLA_IES_IDENTITY_PROFI...	89
5. SELECT "COURSES_AND_MODEL_DATA..."	77
6. SELECT "PSN_CODELESS_IA_USERS"."A...	60
7. EXECUTE NOTEBOOK "PROD_ANALYTI..."	54
8. MERGE INTO "HED"."HOMEWORKSESSI..."	44
9. MERGE INTO "HED"."HOMEWORKXREFE..."	42
10. execute notebook "PROD_ANALYTICS_..."	41
11. MERGE INTO "SCHOOL"."HOMEWORKS...	34

Uncategorized queries

QueryType	UserName	QueryText	Queried...	Count
No data				

Query Durations (in ms)

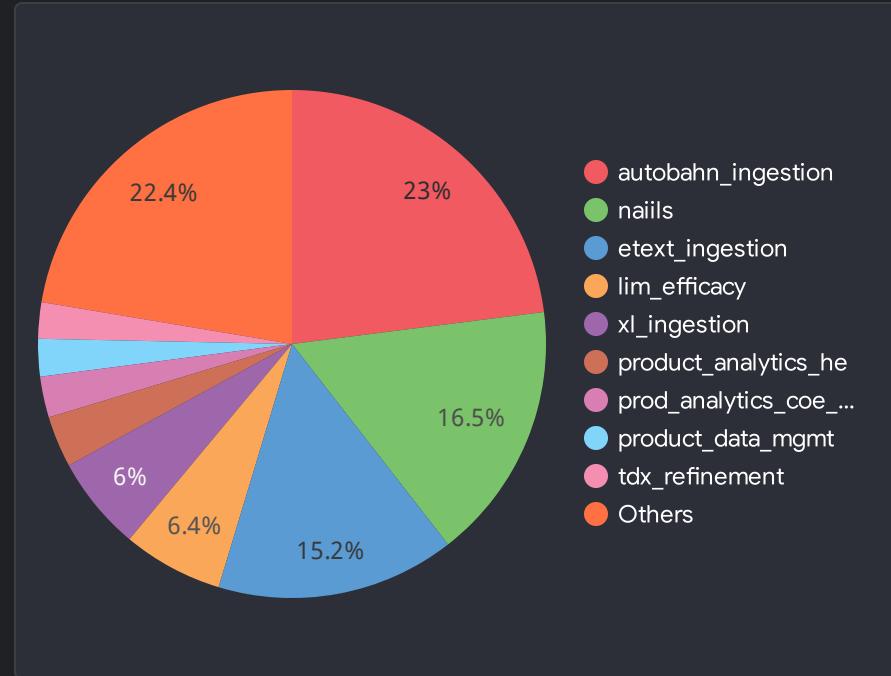
Query Text	Avg	Fastest	Slowest	Count
1. WITH active_courses AS ( SELECT CAST(course_id AS VARCHAR) AS ...	10,800,753	10,800,753	10,800,753	1
2. CALL PDI_INTERNAL.AUTOMATION.CALL_PROCEDURE_SET('DAY_ZE...')	9,453,817....	9,081,979	10,801,026	30
3. CALL PDI_DATA_PRODUCTS_NP.DAY_ZERO.UPDATE_PEARSON_PLUS...	9,451,103.9	9,078,670	10,799,109	30
4. CREATE OR REPLACE TABLE PDI_DATA_PRODUCTS_NP.DAY_ZERO.P...	9,450,150...	9,077,000	10,799,118	30
5. SELECT * FROM TDX.TDX.item_attempt_variables;	7,500,120	7,500,120	7,500,120	1
6. select acd.accesscode, accesscodebatchid, acd.accesscodetypeid, a...	7,268,465	7,268,465	7,268,465	1
7. call CPDA_INTERNAL.c_sandbox.pulse_wau()	6,803,147....	5,555,015	8,530,126	30
8. create or replace table CPDA_INTERNAL.c_sandbox.pulse_wau as ( w...	6,801,979.1	5,553,876	8,528,892	30
9. CALL cds.pearson_plus_engagement_index.procedure_etl()	6,761,798.5	3,937,489	10,800,846	4
10. call CPDA_INTERNAL.c_sandbox.okr_au()	6,557,826.7	2,990,315	9,856,484	30

Top Users

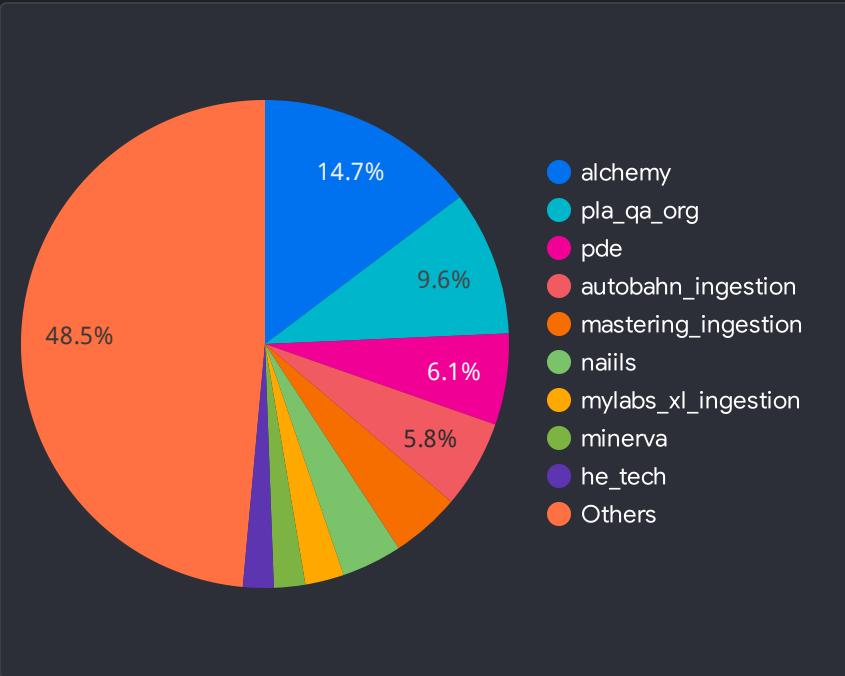
UserName	Count
1. ATTUNITY_APP	3,147,802
2. PDL_QE_SERVICE	1,207,021
3. SYSTEM	482,044
4. PERFSVC_SYS_USER	437,224
5. AUTOBAHN_SYS_USER	199,294
6. cloudbi.prod@pearson.com	161,915
7. EVEREST_SYS_elt_ADMIN	64,940
8. BACCHUS_SYS_USER	32,684
9. SNOWFLAKE_ADMIN	28,093
10. NONPROD_CLOUDBI_SYS_USER	20,073

# Databases

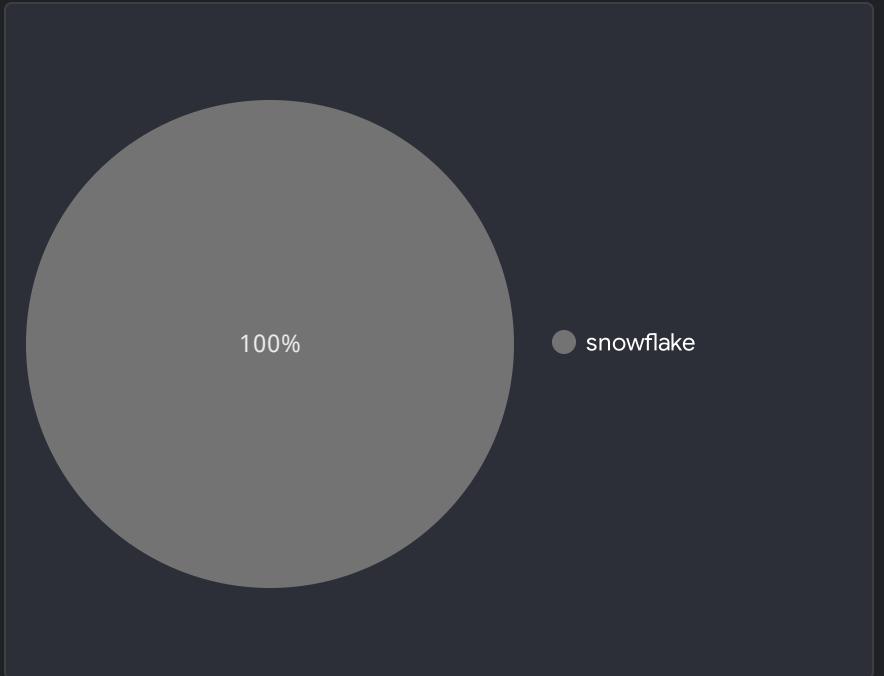
**Database by Volume**



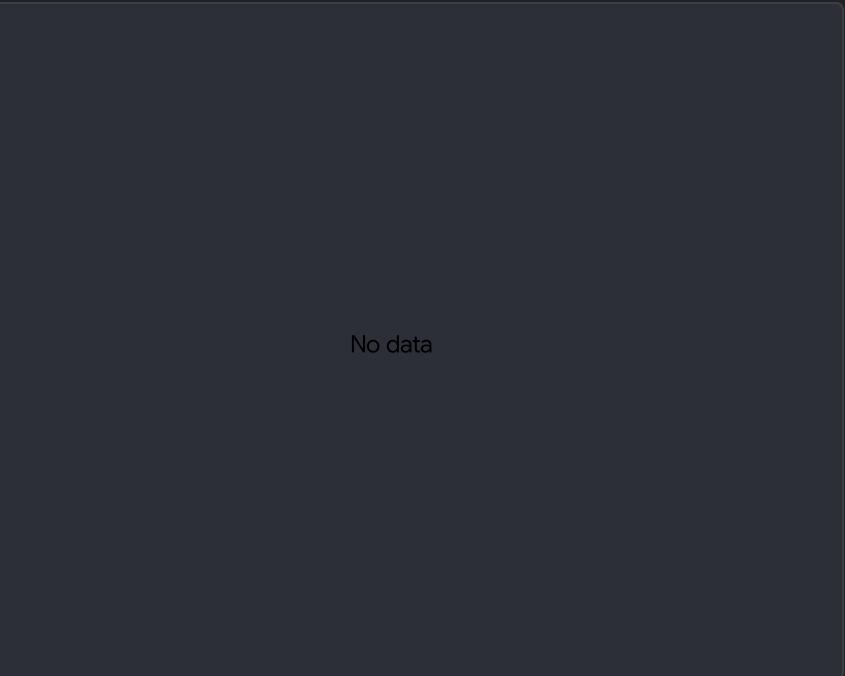
**Database by Tables**



**Database by Views**



**Database by Materialized Views**



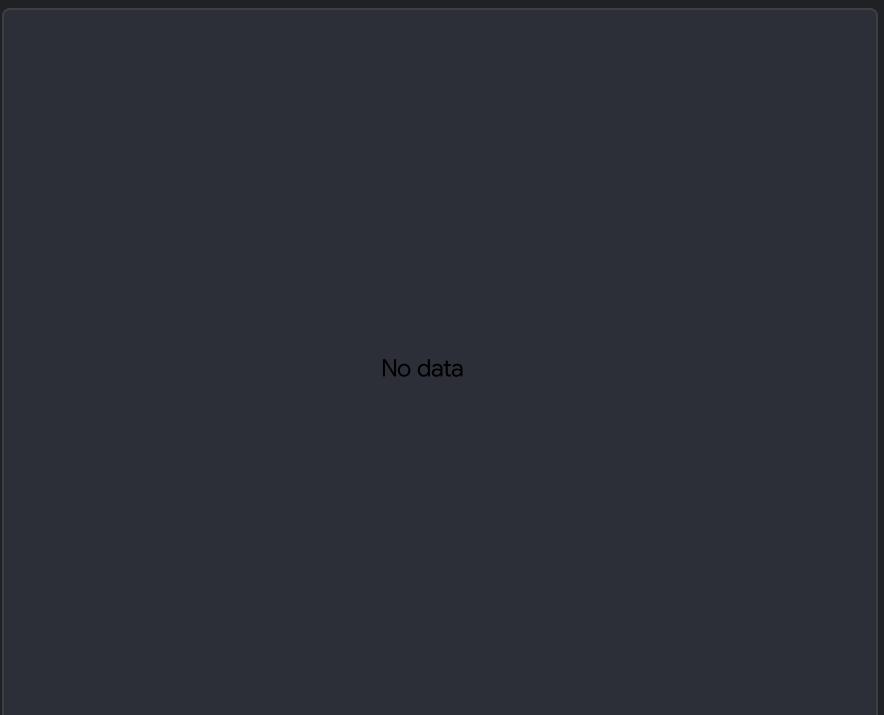
**Database by Volume\***

Database Name	Size (GBs)	Estimate (USD) ...
snowflake	0	0
pfc_administration	1,042.84	4.17
product_data_lake	3.64	0.01
product_data_intelligence	122.62	0.49
product_analytics_he	7,255.77	29.02
snowalert	0	0
pde	1,618.34	6.47
ascend	92.92	0.37
mastering_ingestion	4,797.44	19.19
mle	575.45	2.3

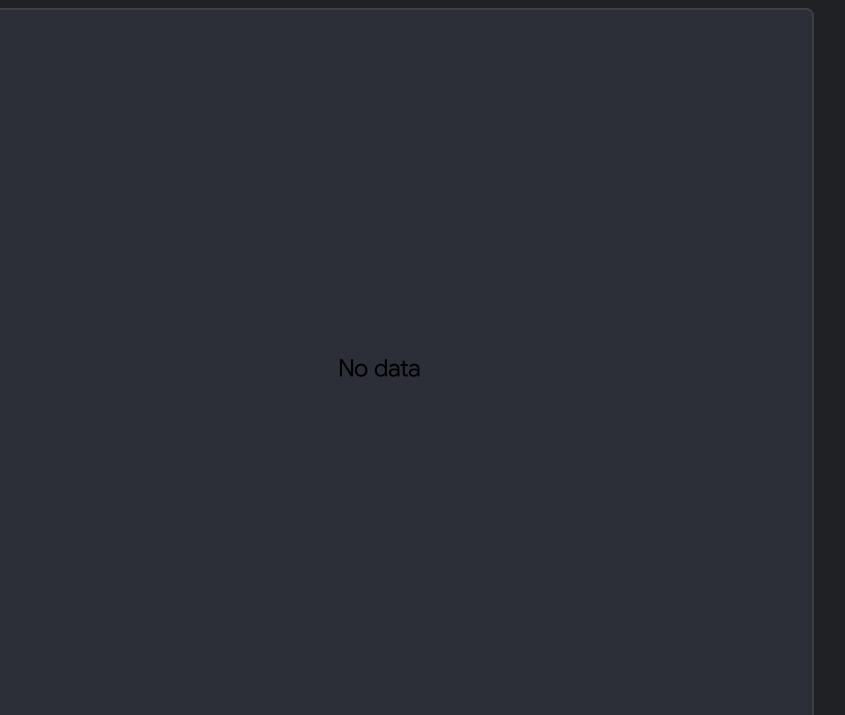
**Database by Tables**

Database Name	Number of Tables
alchemy	3,133
pla_qa_org	2,042
pde	1,288
autobahn_ingestion	1,240
mastering_ingestion	976
naiils	853
mylabs_xl_ingestion	548
minerva	441
he_tech	433
mastering_refinement	421

**Database by Writes**

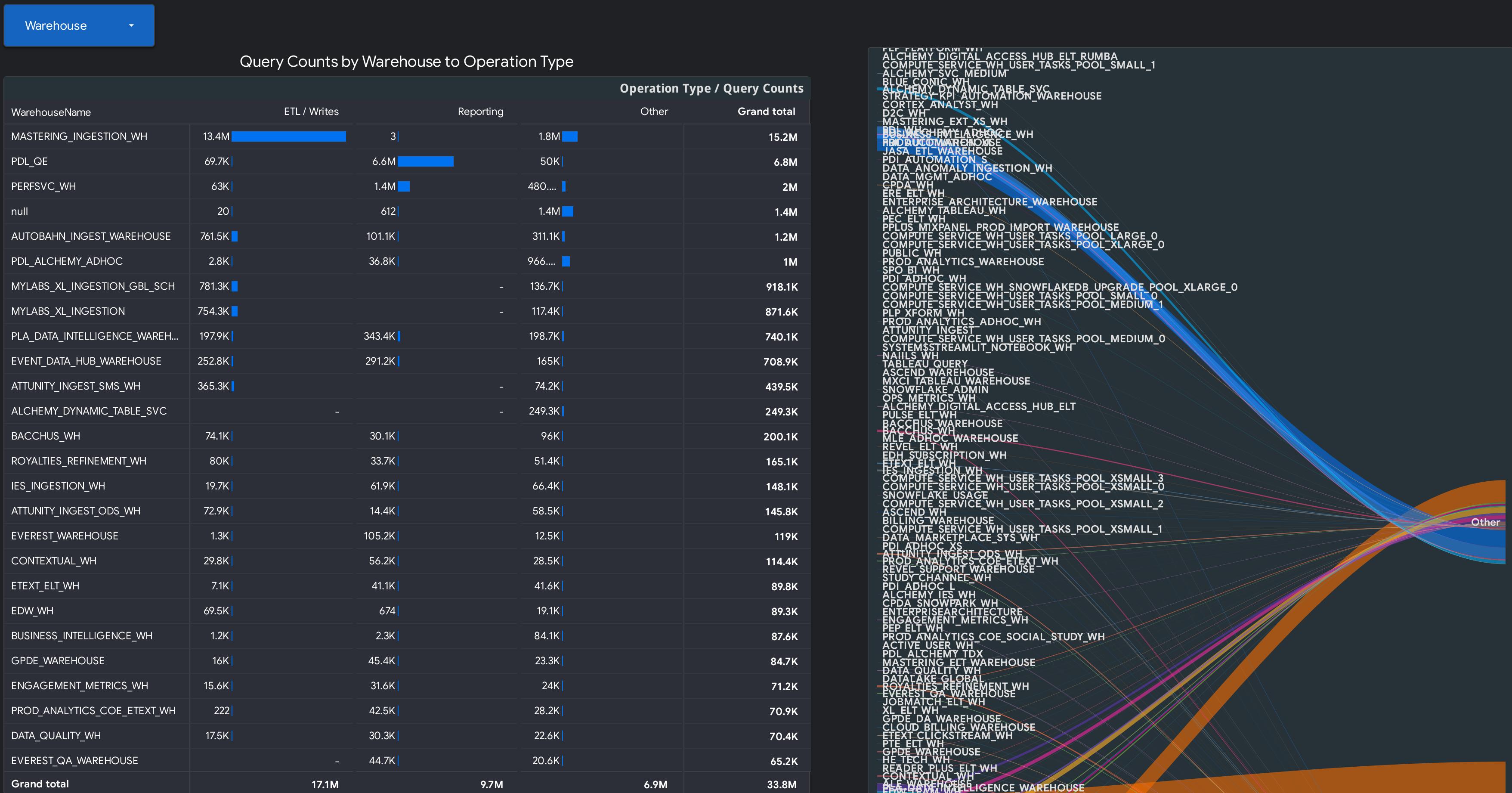


**Database by Reads**



\* Includes Contingency Bytes & Retained For Clone Bytes.

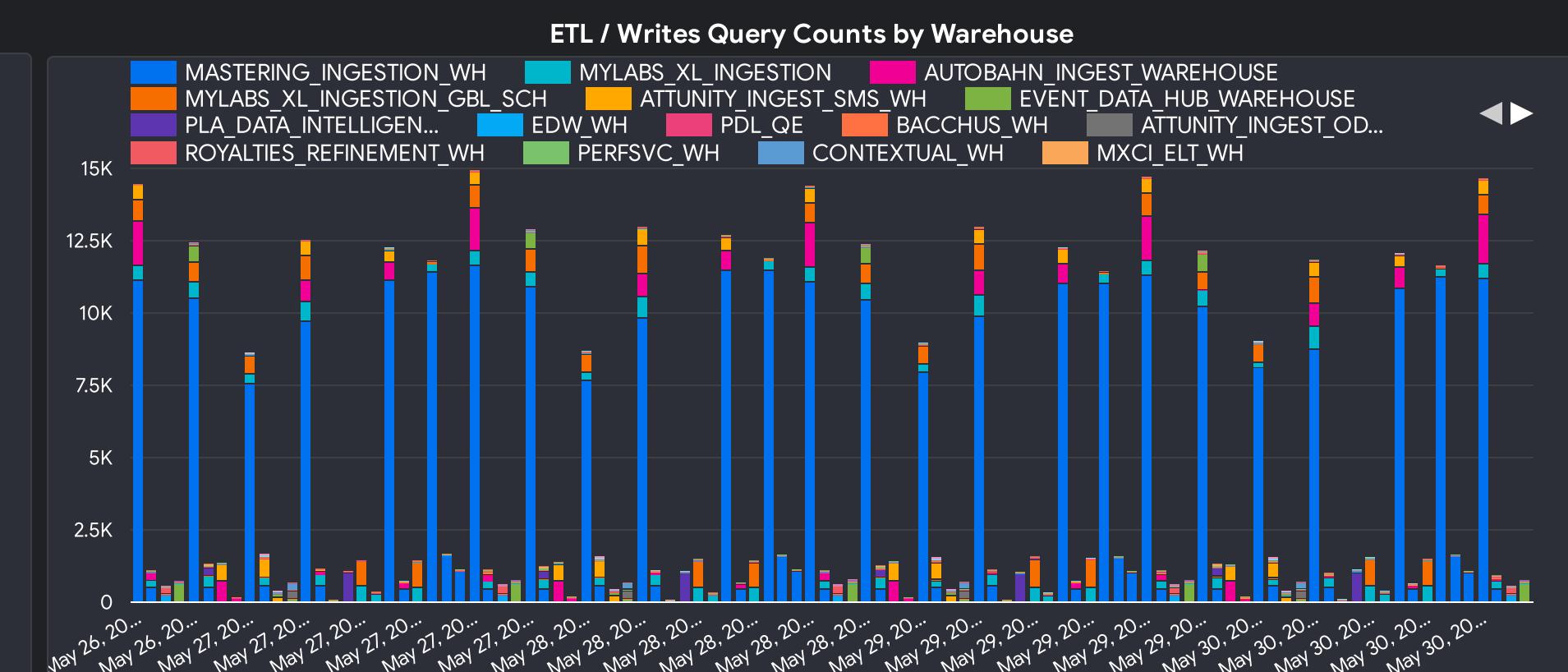
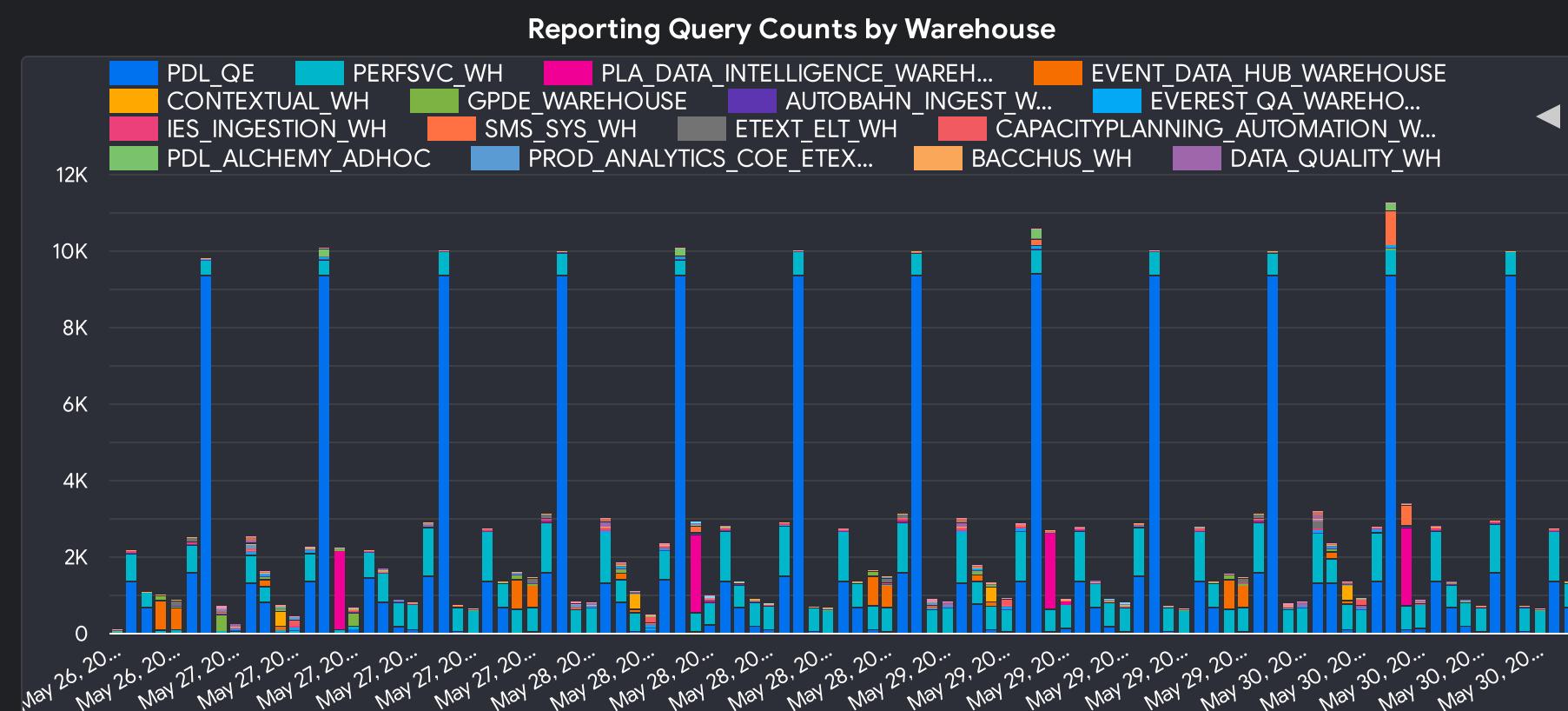
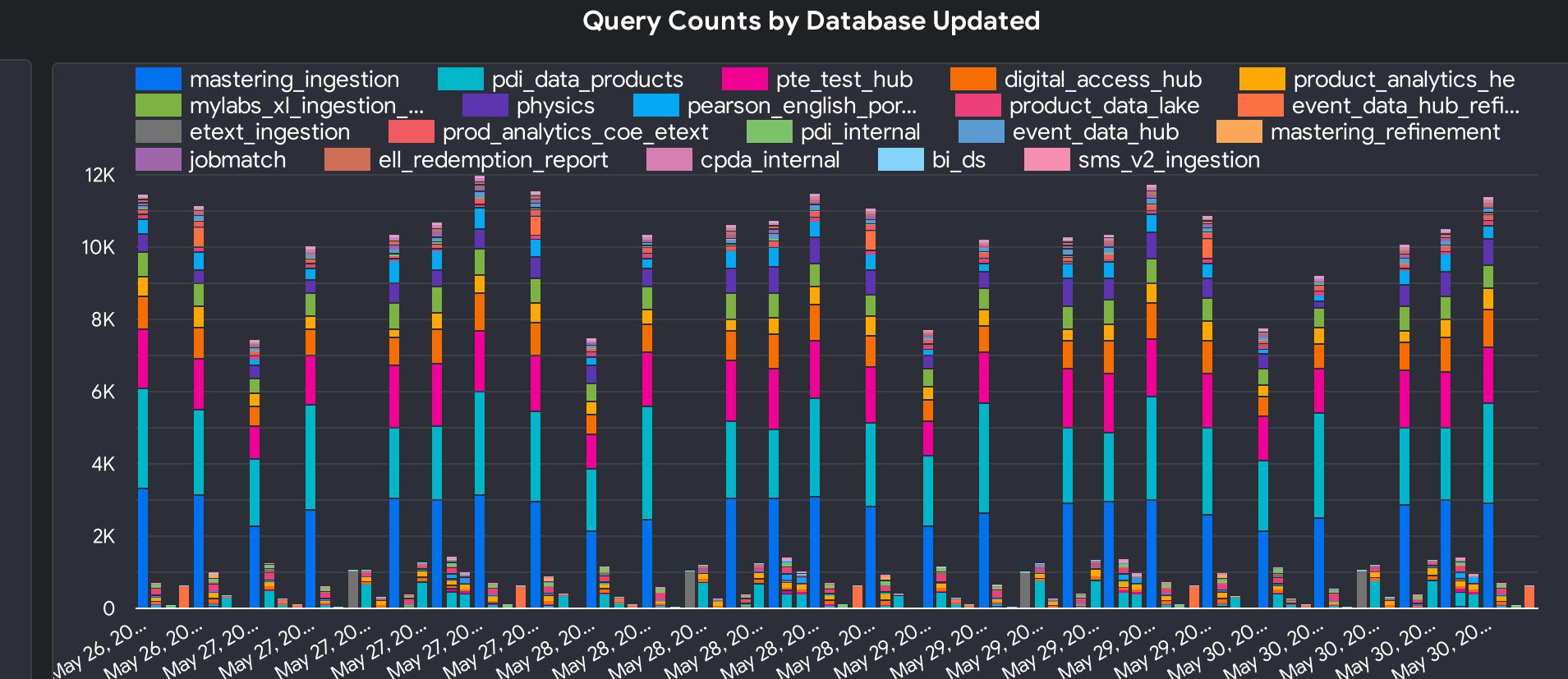
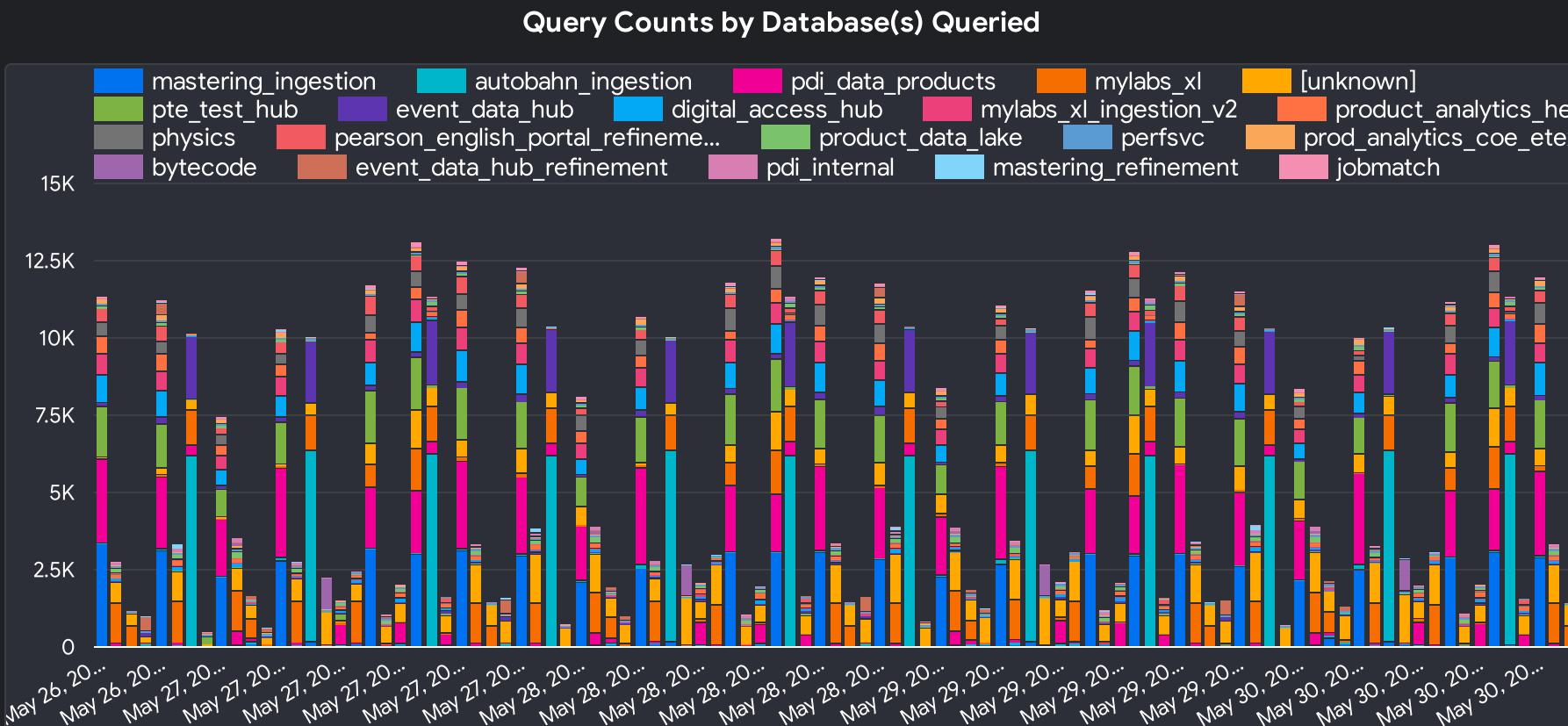
# Warehouses' query counts



# Warehouse & database interactions (last 30 days)

By User Name ▾

Warehouse ▾



# Potential system waste

## Tables With No Usage

By Database

By Schema

Tables	<b>15,927</b>
TableSizeInMBs	<b>133,234,345.81</b>

Note: External tables / pipelines are not included in the calculations.

Based on Usage Range:

Start Time	Dec 31, 2024, 4:00:00 PM
End Time	Jun 25, 2025, 4:59:59 PM

Tables w/o Usage \*

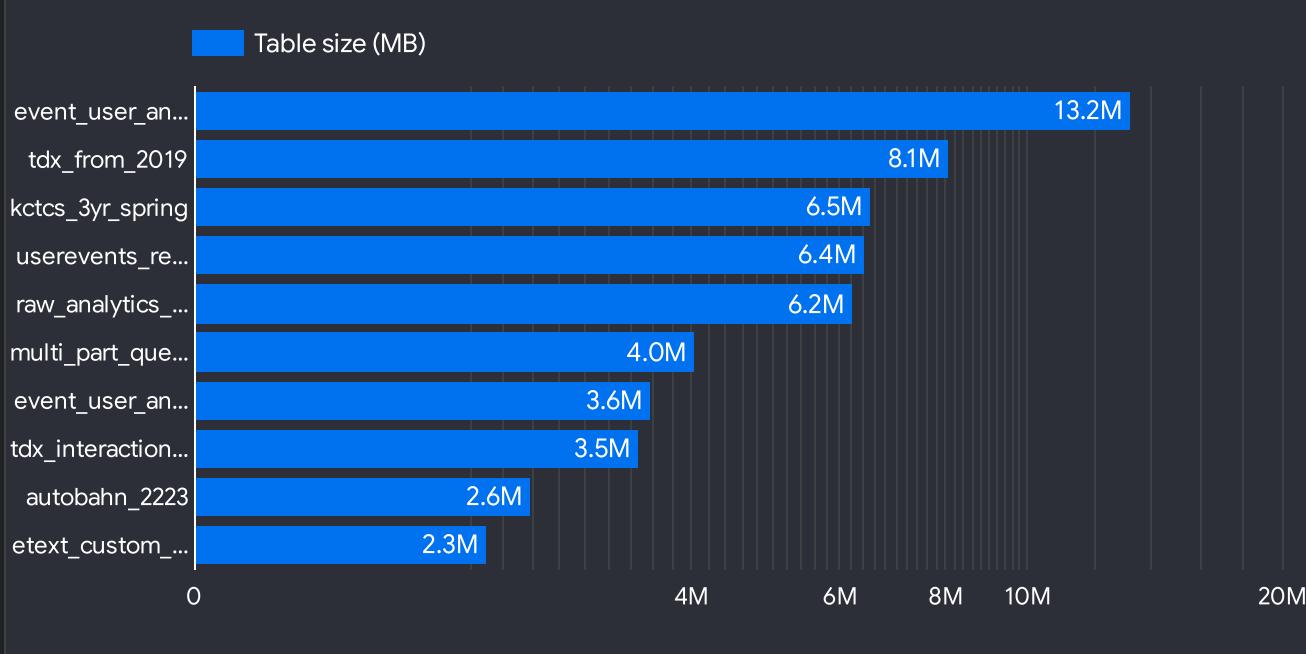
Database Name	Schema Name	Table Name	Table Size In MBs	Estimate (USD)**
etext_ingestion	firebase	event_user_analytics_26886...	13,228,557.57	25.8
naiils	experiment	tdx_from_2019	8,068,216.66	15.8
naiils	experiment	kctcs_3yr_spring	6,533,257.27	12.8
lim_efficacy	efficacy_work	userevents_revisits_mlm	6,422,529.85	12.5
etext_ingestion	firebase	raw_analytics_268866320	6,213,838.69	12.1
product_data_mgmt	standard_formats	multi_part_question_user_a...	3,879,830.56	7.6
reader_plus_ingestion	firebase	event_user_analytics_24888...	3,584,546.74	7
naiils	experiment	tdx_interactions_from_2019	3,482,722.63	6.8
etext_ingestion	ga4	etext_custom_field	2,298,144.72	4.5
etext_ingestion	telemetry_io	etext_telemetry_hit_io	2,225,585.09	4.3
naiils	experiment	autobahn_2223	2,038,748.14	4
revel_ingestion	ga	revel_86098200_ga_detail	1,940,804.91	3.8
revel_ingestion	ga	revel_86098200_ga_custo...	1,776,313.09	3.5
xl_ingestion	ga	xl_154558773_ga_detail	1,764,014.26	3.4
etext_ingestion	telemetry	etext_telemetry_custom_fiel...	1,739,982.56	3.4
naiils	experiment	wgu_asmt_homework_results	1,574,924.54	3.1
naiils	experiment	wgu_asmt_homework_query	1,483,442.25	2.9
etext_ingestion	ga4	etext_hit	1,430,361.89	2.8
naiils	experiment	fall21_spring22	1,422,391.35	2.8
alchemy	adhoc	paf_backup	1,345,150.63	2.6
etext_ingestion	telemetry	etext_telemetry_hit	1,338,006.04	2.6
prod_analytics_coe_etext	sandbox	person_etext_combine_new	1,271,077.53	2.5
reader_plus_ingestion	firebase	raw_analytics_248889779	1,255,728.41	2.5
prod_analytics_coe_etext	sandbox	person_etext_filter_email	1,251,670.99	2.4
prod_analytics_coe_etext	sandbox	etext_base	1,251,120.91	2.4
Grand total			<b>133,234,345.81</b>	<b>252.7</b>

1 - 25 / 15905 < >

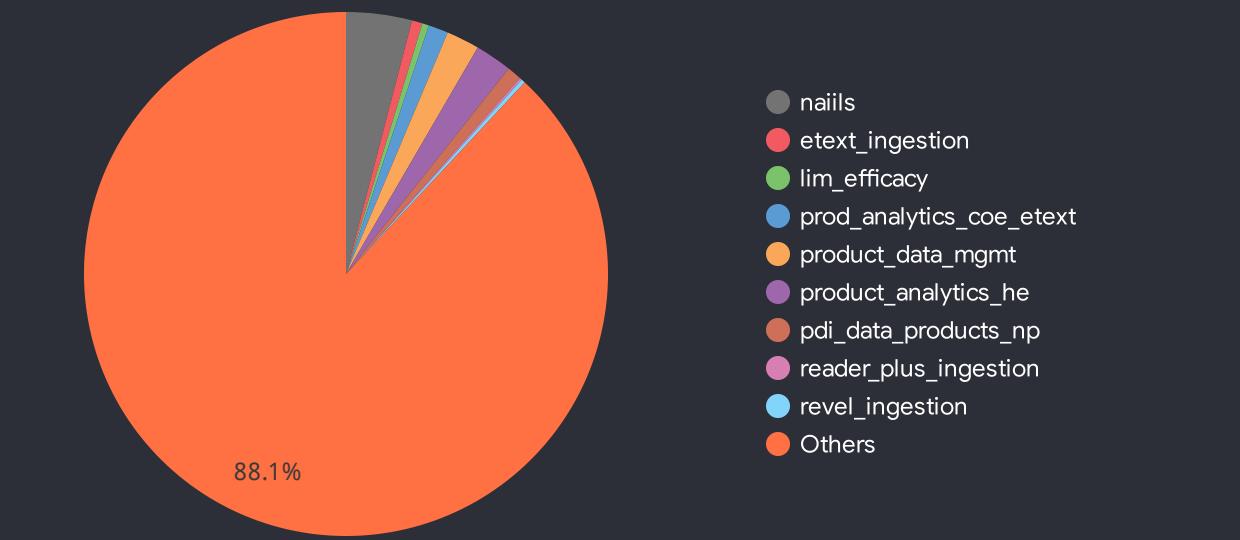
Usage vs Non Usage



Unused Tables Volume



Number of Unused Tables by Database



\* Based on the duration of the query logs window., Tables that do not have any queries.

\*\* Estimate only. The storage price is assuming more than 90 days of no usage and is assuming compression ratio of 10. First 10GB storage is free - not included in summary. Please use external documentation for more details.

# Potential system waste

# Tables With Usage but No Writes

By Database

By Schema

Tables  
19,934

TableSizeInMBs  
341,117,960.49

Note: External tables / pipelines are not included in the calculations.

Based on Usage Range:

Start Time  
Dec 31, 2024, 4:00:00 PM

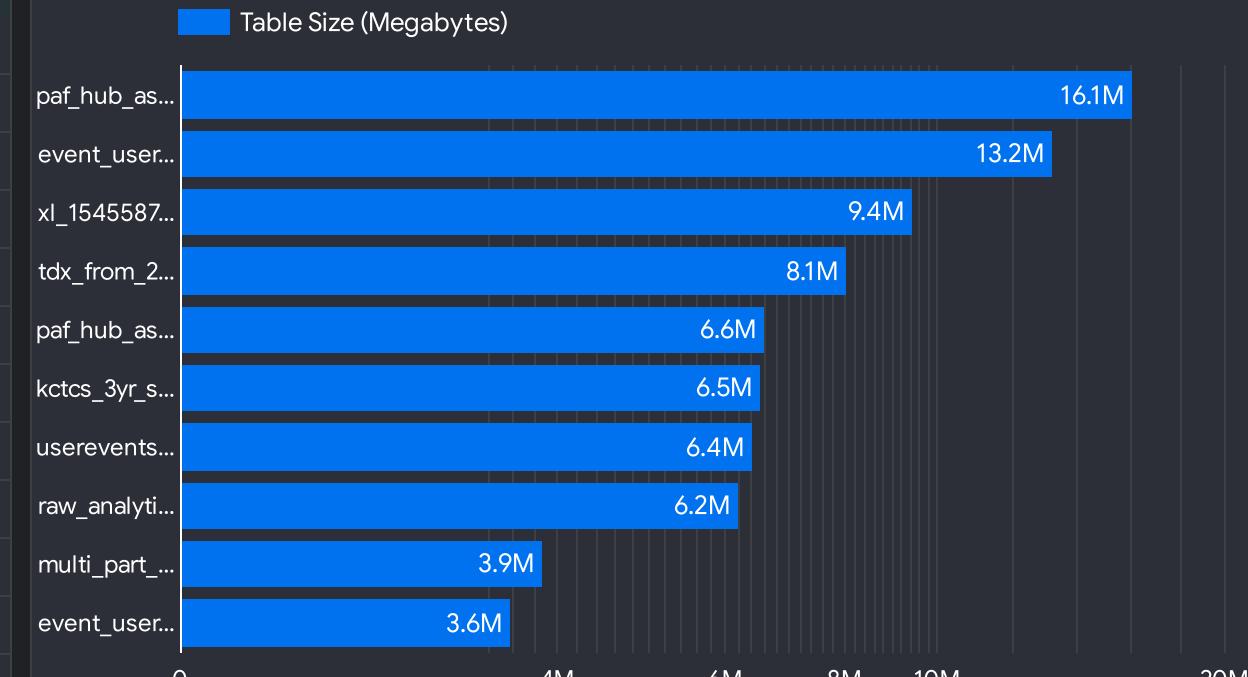
End Time  
Jun 25, 2025, 4:59:59 PM

Tables w/o Usage \*

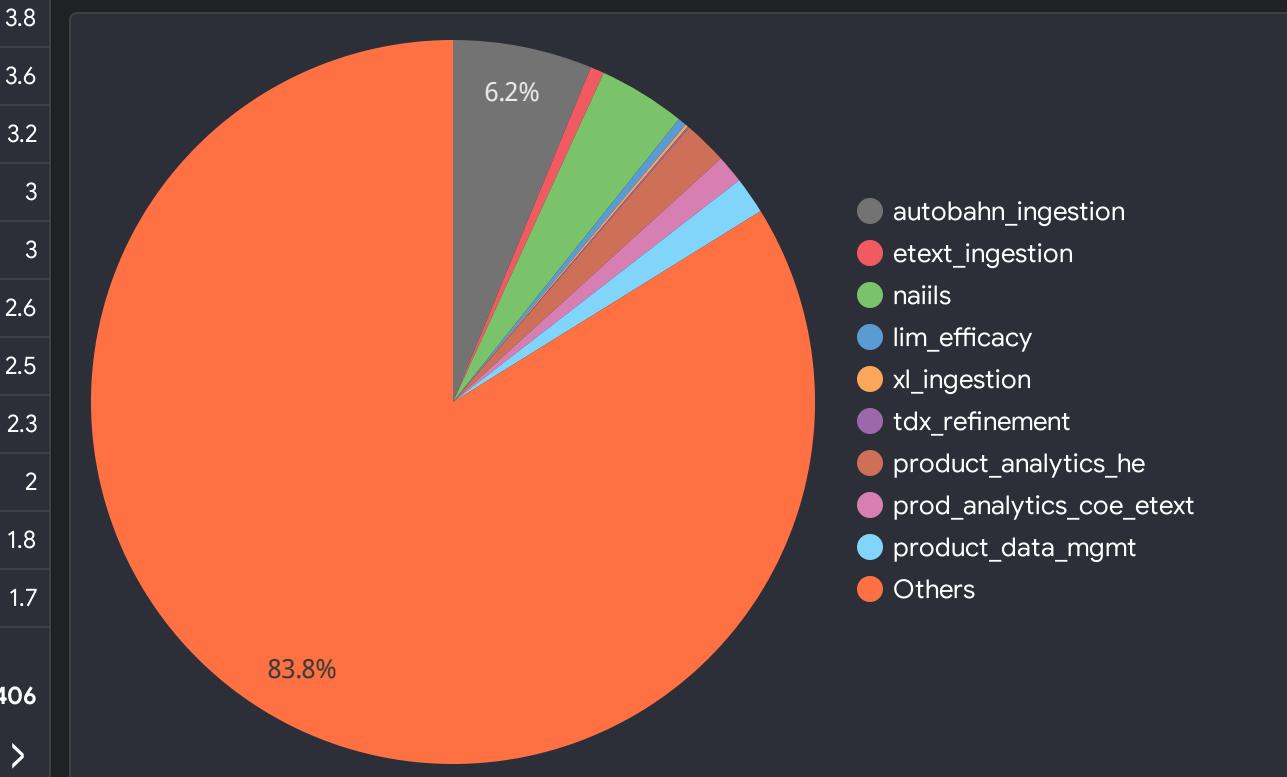
Database Name	Schema Name	Table Name	Table Size In MBs	Estimate (USD)**
autobahn_ingestion	raw	paf_hub_assessment_multi_par...	96,514,159.68	188.5
autobahn_ingestion	raw	paf_hub_assessment_multi_par...	13,225,197.46	25.8
xl_ingestion	ga	xl_154558773_ga_custom_dim...	9,440,757.31	18.4
autobahn_ingestion	raw	pi_common_usersessiontracki...	4,960,816.12	9.7
autobahn_ingestion	raw	ale_ale_aleuseractivityattempt...	4,872,624.96	9.5
autobahn_ingestion	raw	etext_common_userloadscont...	4,376,365.32	8.5
tdx_refinement	tdx	tdx_messages_main	3,732,486	7.3
lim_efficacy	efficacy_work	userevents_mlm	3,192,821.52	6.2
etext_ingestion	telemetry_io	etext_telemetry_hit_io_final	3,055,008.32	6
autobahn_ingestion	raw	etext_telemetry_usertelemetry...	2,997,364.68	5.9
lim_efficacy	efficacy_work	userevents_revisits	2,623,402.67	5.1
autobahn_ingestion	raw	etext_telemetry_usertelemetry...	2,601,759.52	5.1
autobahn_ingestion	raw	etext_common_useruploadsco...	2,555,508.92	5
tdx_refinement	tdx	tdx_messages_item_attempt_v...	2,102,287.98	4.1
autobahn_ingestion	raw	etext_telemetry_usertelemetry...	1,959,101.12	3.8
autobahn_ingestion	raw	xl_telemetry_usertelemetry_1_1...	1,849,372.22	3.6
tdx_refinement	tdx	tdx_messages_user_interactions	1,630,257.78	3.2
autobahn_ingestion	raw	pi_telemetry_usersessionenvir...	1,549,994.04	3
autobahn_ingestion	raw	xl_telemetry_usertelemetry_2_...	1,549,720.8	3
autobahn_ingestion	raw	en_common_timeontaskdailyag...	1,322,128.9	2.6
product_analytics_he	data_marts_ga	revel_dm	1,277,941.78	2.5
autobahn_ingestion	raw	xl_telemetry_usertelemetrypag...	1,178,564.82	2.3
autobahn_ingestion	raw	glp_telemetry_usertelemetrypa...	1,011,507.1	2
autobahn_ingestion	raw	globalplatformservices_gps_gl...	907,840	1.8
autobahn_ingestion	raw	ale_mastering_result_1_0_0	876,826.47	1.7
Grand total			207,883,614.68	406

1 - 25 / 4006 < >

Unused Tables Volume



Number of Unused Tables by Database



\* Based on the duration of the query logs window., Tables that do not have any queries.

\*\* Estimate only. The storage price is assuming more than 90 days of no writes and is assuming compression ratio of 10. First 10GB storage is free - not included in summary. Please use external documentation for more details.

# Queries that don't need manual intervention

## Automatic SQL Translation

Automatically Translated	56,685
Distinct Queries	115,616

Hints:  
1) Click on user or a query generalized form in the table on the right to drill down  
2) Click on focused row to reset  
3) Use +/- icons to calibrate focus

### Filters:

User Name	Queried Tables
Affected Table	Query Type

Original Query Text  
Equals ▾ Enter a value

Generalized Form  
Equals ▾ Enter a value

For the mapping between Snowflake data types and BigQuery data types, please refer to the [Snowflake SQL translation guide](#).

\* Estimate

Generalized Form	QueryType	Distinct Queries
select_(distinct_) from _ where _ = '?'	SELECT	1.4K
select _(*) from _	SELECT	1.3K
create or replace view _ select * from _	Other	1.3K
select * from _ limit 1	SELECT	944
select _(*), _ (distinct _) from _	SELECT	859
select * from _	SELECT	685
select * from _ where _ = '?'	SELECT	596
select _(_) from _	SELECT	586
truncate table _	Other	429
select _(*) from (select _(*) from _ group by _ having _(*)>1)	SELECT	409
select _(*) from (select bound_1 _(*) from _ where bound_1 between 1 and 1 and data_type('2') = 'B' group by bound_1)	SELECT	408
<b>Grand total</b>		<b>56.7K</b>

### Translated Queries

Original Query Text	Translated Query Text
1. with year_of_report as (select 2025 as year_report) select t.course_section_id, ,t.email_domain ,t.student_count from ( select a.course_section_id ,a.email_domain ,a.student_count ,a.rn_domain ,a.rank_domain ,row_number() over(partition by a.course_section_id order by a.email_domain) as rn_per_course from ( select cs.course_section_id ,ies.email_domain ,count(distinct e.person_id) as student_count ,row_number() over(partition by cs.course_section_id order by count(distinct e.person_id) desc) as rn_domain ,rank() over(partition by cs.course_section_id order by count(distinct e.person_id) desc) as rank_domain from naiils.he.services.sbs_all_net_s join event_data_hub.standard_formats.entitlement e on s.subscription_id = e.entitlement_id and e.transaction_type_code != 'Delete' left join event_data_hub.standard_formats.order_transaction ot on e.order transaction id = ot.order transaction id and ot.transaction type code	WITH YEAR_OF_REPORT AS ( SELECT NUMERIC '2025' AS YEAR_REPORT ) SELECT T.COURSE_SECTION_ID, T.EMAIL_DOMAIN, T.STUDENT_COUNT FROM ( SELECT A.COURSE_SECTION_ID, A.EMAIL_DOMAIN, A.STUDENT_COUNT, A.RN_DOMAIN, A.RANK_DOMAIN, ROW_NUMBER() OVER (PARTITION BY A.COURSE_SECTION_ID ORDER BY IF(A.EMAIL_DOMAIN IS NULL, 0, 1) DESC, A.EMAIL_DOMAIN) AS RN_PER_COURSE FROM ( SELECT CS.COURSE_SECTION_ID, IES.EMAIL_DOMAIN, CAST(COUNT(DISTINCT E.PERSON_ID) AS BIGNUMERIC) AS STUDENT_COUNT, ROW_NUMBER() OVER (PARTITION BY CS.COURSE_SECTION_ID ORDER BY IF(COUNT(DISTINCT E.PERSON_ID) IS NULL, 0, 1), CAST(COUNT(DISTINCT E.PERSON_ID) AS BIGNUMERIC) DESC) AS RN_DOMAIN,

1 - 100 / 56685 < >

BigQuery automates SQL translation from Snowflake to Standard Google SQL (free). Translation is a true compiler (vs regex matching) and achieves very high rates. Translation Accuracy will be higher if used directly with Translation Service.

# Queries requiring intervention

**Hints:**

- 1) Click on AppId, User Name, Generalized form or Complexity in this table to drill down
- 2) Click on focused row to reset

**Note:** For performance considerations, only a portion of queries can be displayed in the charts below.

# Offline SQL Translation Effort

Unique Queries  
**53,809**

## Queried Tables

## Affected Table

## Every Type

## Difficulty

**Translation offline efforts unique patterns count**

# Low Impact 4,268

## Medium Impact

1,314

High Impact  
41

## Queries suggested for reviewing

### Hints:

- 1) Click on AppId, User Name, Generalized form or Complexity in this table to drill down
- 2) Click on focused row to reset

Note: For performance considerations, only a portion of queries can be displayed in the charts below.

## Translated Queries Suggested to Review

Queried Table Name	Affected Table Name
Query Type	Difficulty

Low Difficulty  
0

Medium Difficulty  
2,491

High Difficulty  
17

Distinct Queries  
2,496

		Distinct Queries
Medium	Simple	create transient table _ clone _
		drop table _
		select _(1) from _
		select * from (select * from table(_(1,'?','?')) where 1 = 1
		select _(*) from ( select * from table(_(1,'?','?')) )
		select _(*) from _
		select _(_,_)_ from _
		insert into _ ( select _.* from _ where _='?' and _() not in (select _ from _) )
		create or replace transient table _ select _, '?' _, _, _ from _ join _ on _ = _ j...
		create or replace transient table _ select _.*_, _ from _ join _ on _ = _ join _ ...
		create or replace transient table _ select distinct _, _, _, null _, null _, _, ...
		create or replace transient table _ select distinct _, _, _, _, _, _, _, _, _, _, 1 _ fro...
		select _.* from _ where _='?' and _() not in (select _ from _)
		create or replace transient table _ select _, _, _, _, _, _ from _ union select...
		select * from _ where date_trunc('?',_) = '?' limit 1
		select _(*) from (select _(*) from _ where date_trunc('?',_) = '?' group by _ ...)
		select (_(_( cast (( _ is not null ) number ),1)),(_( distinct _)),(1) ...)
		delete from _ where date(_) < ( select date_(dateadd(_, -1, _(_)), '?') ) from...
		create or replace table _ as(select _, _, _, _(distinct _    _    _), _( '?', _()) f...
		select _, _, _, _, _, _ from (select * from table(_(1,'?','?')))
		select _(*) from _ where date_trunc('?',_) = '?' limit 1
		select _(*), _(distinct _) from _
		create or replace database _
		create or replace table _ (select _, _, case when _ in ('?', '?', '?') then '?' else...
		select _, _, _, _, _ from (select * from table(_(1,'?','?')))

QueryText	TranslatedQueryText	Message
1. with target_user as (select person_id, business_model_code, rating, category from PROD ANALYTICS COE ETEXT.SANDBOX.INMOMENT where person_id != 'userID' group by 1,2,3,4 UNION ALL select person_id, business_model_code, rating, category from PROD ANALYTICS COE ETEXT.SANDBOX.PENDO group by 1,2,3,4)	WITH TARGET_USER AS (SELECT INMOMENT.PERSON_ID, INMOMENT.BUSINESS_MODEL_CODE, INMOMENT.RATING, INMOMENT.CATEGORY FROM PROD_ANALYTICS_COE_ETEXT.SANDBOX.INMOMENT WHERE INMOMENT_PERSON_ID <> 'userID' GROUP BY 1, 2, 3, 4 UNION ALL SELECT PENDO.PERSON_ID, PENDO.BUSINESS_MODEL_CODE, PENDO.RATING, PENDO.CATEGORY FROM PROD_ANALYTICS_COE_ETEXT.SANDBOX.PENDO GROUP BY 1, 2, 3, 4)	BigQuery does not support duplicate aliases

# Avoiding Reserved Keywords - Please Review

Database ▾ Table ▾ Entity Type ▾

**Identifiers by Entity Name**

Entity Name ▾	Entity Type	Database Name	Table Name
No data			

Be aware about BigQuery [reserved keywords](#) in the source databases identifiers.

Keywords have special meaning in the Google Standard SQL language, and cannot be used as identifiers unless enclosed by backtick (`) characters.

Migration Assessment found following entities which are not handled by automatic SQL translation. They might require manual effort during migration. Consider using BigQuery [interactive SQL translation](#) to translate table DDLs with these identifiers.

BigQuery UI provides pre-validation for SQL queries, so you can check the compatibility of your query format before running it.

For the manual migration of queries, generated in non-SQL code, either enclose these identifiers by backtick characters or rename them.

## Usage of BigQuery reserved keywords (case insensitive)

Reserved Keyword	Detected Occurrences ▾
No data	

# Database Migration Coupling

# List of databases and tables accessed together in a single query

# Coupling Queries

**Query Text ▾**

1. with year\_of\_report as (select 2025 as year\_report)

```
select t.course_section_id
,t.email_domain
,t.student_count
from (
select a.course_section_id
,a.email_domain
,a.student_count
,a.m_domain
,a.rank_domain
,ROW_NUMBER() OVER(PARTITION BY a.course_section_id ORDER BY a.email_domain) AS rn_per_course
from (
select
cs.course_section_id
,ies.email_domain
,COUNT(DISTINCT e.person_id) AS student_count
,ROW_NUMBER() OVER(PARTITION BY cs.course_section_id ORDER BY COUNT(DISTINCT e.person_id) DESC) AS rn_domain
,RANK() OVER(PARTITION BY cs.course_section_id ORDER BY COUNT(DISTINCT e.person_id) DESC) AS rank_domain
from naills.he_services.subs_all_nets
JOIN event_data_hub.standard_formats.entitlement e ON s.subscription_id = e.entitlement_id AND e.transaction_type_code != 'Delete'
LEFT JOIN event_data_hub.standard_formats.order_transaction ot ON e.order_transaction_id = ot.order_transaction_id AND ot.transaction_type_code != 'Delete'
LEFT JOIN event_data_hub.standard_formats.person_course_section pc ON pc.person_id = e.person_id AND pc.transaction_type_code != 'Delete' AND pc.test_record_flag != TRUE
```

## Based on Usage Range:

Start Time  
Dec 31, 2024, 4:00:00 PM

End Time

# How your tables are updated dictates how you move them

Usage timeframe:

StartTime  
May 26, 2025, 5:00:00 PM

StartTime  
Jun 25, 2025, 4:59:58 PM

## Table Updates Schedule

UpdatedTable ▾      StartTime (Day of Week) ▾

Table Updates per Hour Based on QueryLogs (max 30 days)

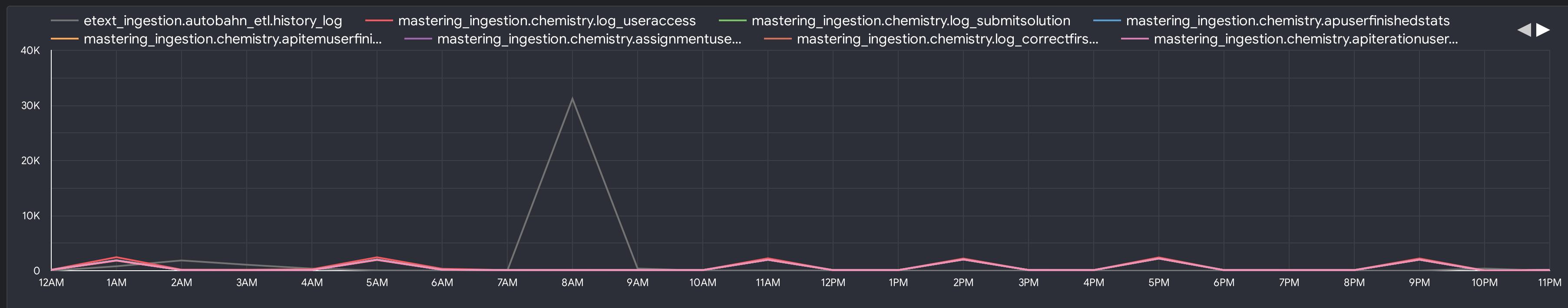
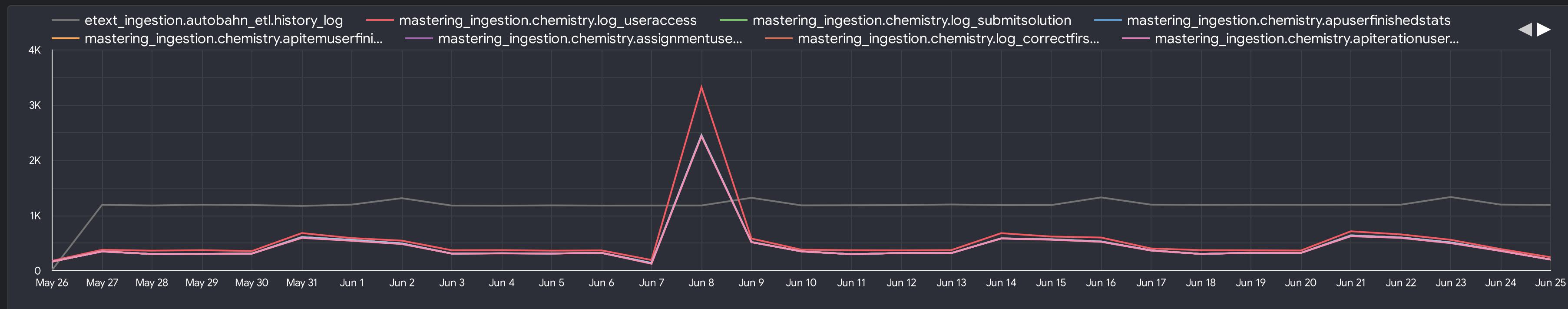


Table Updates per Day Based on QueryLogs (max 30 days)



# Demonstrating BigQuery steady state via PoC

Based on analyzed logs, these tables and queries will help demonstrate the time and cost savings, and value of BigQuery via a PoC

Row Number Limit 5

Query Type: All

## Most Frequent Queries

Query		Users	Queried Tables	Occur...
1.	copy into AUTOBAHN_INGESTION.raw.ILP_Telemetry_UserTelemetry_1_1_0 (cds_load_date, source_file, message, M_CORRELATION_ID, M_ENVIRO...	ATTUNITY_APP, AUTOBAHN_SYS_USER, U999PU, V...		2213298
2.	select count(*) as COUNT from (select namespace, message_type, version, originating_system_code from autobahn_ingestion.autobahn_etl.auto...	PDL_QE_SERVICE	autobahn_ingestion.autobahn...	2048640
3.	select namespace, message_type, version, originating_system_code from autobahn_ingestion.autobahn_etl.autobahn_ingest_metadata where lo...	PDL_QE_SERVICE	autobahn_ingestion.autobahn...	2022790
4.	alter dynamic table /* ALCHEMY.HED_HOURLY0325.ASSIGNEDTESTS = */ identifier(70837002551838) refresh at 1746094090774;	SYSTEM		1308320
5.	TRUNCATE TABLE "REPLICATE CONTROL".attrp_changes846268E2CAA50D37"	ATTUNITY_APP		1199967

1 - 5 / 5 < >

## Queries Reading Most Data

Query		Users	Queried Tables	MB scanned
1.	select * from EVENT_DATA_HUB.AUTOBAHN.PAF_HUB_ASSESSMENT_MULTI_PART_QUESTION_USER_ANSWERED_2_2 where CONTEXT_EXTENSI...	USHAHHE	autobahn_ingestion.raw.paf_hub...	4177338.3
2.	SELECT 'PRODUCT_PLATFORM_CODE' AS column_name, ROUND((COUNT(*) - COUNT(PRODUCT_PLATFORM_CODE)) * 100.0 / COUNT(*), 2) A...	EVEREST_SYS_ELT_ADMIN	etext_ingestion.telemetry_io.ete...	3980910.6
3.	CREATE OR REPLACE TABLE PROD_ANALYTICS_COE_ETEXT.SANDBOX.SPERRY_CHANNELS_RAW_DATA AS ( with channels_data as ( select a.PE...	SHAI1985, SYSTEM		2879172.5
4.	CREATE OR REPLACE TABLE PROD_ANALYTICS_COE_ETEXT.SANDBOX.SPERRY_CHANNELS_RAW_DATA AS ( with channels_data as ( select a.PE...	SYSTEM, SHAI1985		2845501.4
5.	CREATE OR REPLACE TABLE PROD_ANALYTICS_COE_ETEXT.SANDBOX.SPERRY_CHANNELS_RAW_DATA AS ( with channels_data as ( select a.PE...	SYSTEM, SHAI1985		2745668.9

1 - 5 / 5 < >

## Slowest Queries

Query		Users	Queried Tables	Duration...
1.	SELECT * FROM ( SELECT "B_C_S_E", "I_0474_BOOK_ID", "USED_IN_HW", "r_0475_BOOK_ID", "USED_IN_TEST" FROM ( SELECT * FROM ( ( SELECT ...	TESTPREP	mylabs_xl.hed.homework...	17587.321
2.	SELECT count(1) AS "COUNT(LITERAL())" FROM ( SELECT "BOOK_ID", "SHOW_TEST", "SHOW_HOMEWORK", "B_C_S_E" FROM ( SELECT "BOOK_ID..."	TESTPREP	mylabs_xl.hed.exercises	16208.445
3.	WITH NumberedRows AS ( SELECT accesscode, cashedin, cashedindate, creationdate, expirationdate, accesscodebatchid, ROW_NUMBER() OVE...	SMS_SYS_USER	sms.smsadmin.accesscode	18057.958
4.	select * from naiils.experiment.wgu_tdx_date;	U999PU, NAIILS_SYS_USER	naiils.experiment.wgu_tdx...	17944.629
5.	SELECT ( "SUBQUERY_1"."BOOKCODE" ) AS "SUBQUERY_2_COL_0" , ( "SUBQUERY_1"."CHAPTER_ID" ) AS "SUBQUERY_2_COL_1" , ( "SUBQUERY_1"...	ALE_SYS_USER	event_data_hub.autobahn...	16891.413

1 - 5 / 5 < >

## Impacted Tables

Schema	Table Name	Table Size (GB)
telemetry_io	etext_telemetry_hi...	2.9
telemetry_io	etext_telemetry_hi...	2.59
hed	testresultsdata	2.13
hed	homeworkxrefere...	2.09
raw	paf_hub_assessm...	15708.68
telemetry_io	etext_telemetry_hi...	1491.7
hed	homeworkxrefere...	14.96
hed	testresultsdata	0.33
hed	exercises	0.1
hed	homeworkresults	0.06
hed	homeworkresults	0.01
hed	exercises	0
experiment	wgu_tdx_date	0
raw	paf_hub_assessm...	0
autobahn_etl	autobahn_ingest...	0
hed	homeworkxrefere...	0
autobahn	paf_hub_assessm...	0
hed	homeworkresults	0
telemetry_io	etext_telemetry_hi...	0
smsadmin	accesscode	0
hed	testresultsdata	0

1 - 21 / 21 < >

# Demonstrating BigQuery migration plan via PoC

Row Number Limit 5

Query Type: All

# lost complex queries found in Query Logs

## Complex Queries

## Impacted Tables

Schema	Table Name	Table Size (GB)
telemetry_io	etext_telemetry_hi...	2.9
telemetry_io	etext_telemetry_hi...	2.59
ed	testresultsdata	2.13
ed	homeworkxrefere...	2.09
w	paf_hub_assessm...	15708.68
telemetry_io	etext_telemetry_hi...	1491.7
ed	homeworkxrefere...	14.96
ed	testresultsdata	0.33
ed	exercises	0.1
ed	homeworkresults	0.06
ed	homeworkresults	0.01
ed	exercises	0
xperiment	wgu_tdx_date	0
w	paf_hub_assessm...	0
utobahn_etl	autobahn_ingest_...	0
ed	homeworkxrefere...	0
utobahn	paf_hub_assessm...	0
ed	homeworkresults	0
telemetry_io	etext_telemetry_hi...	0
nsadmin	accesscode	0
ed	testresultsdata	0

# Assessment execution summary

Module Name

Status

Error Type

Report Completeness  
**93%**

## Assessment Modules

Name	Recomended Action	Completeness ▾
1. SF_M004	Extract more data	86%
2. TCO	Extract more data	92%
3. Database		95%
4. Translation		100%

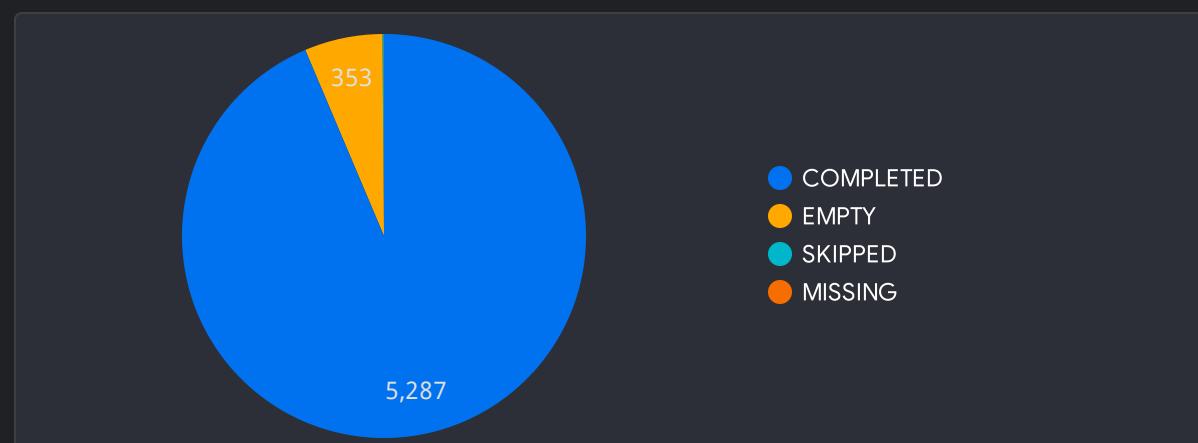
1 - 4 / 4 &lt; &gt;

Completeness defines the amount of successfully processed data that is recommended to build the assessment report. Additionally, we provide action that can be taken to improve it.

- **check processing errors** indicates errors occurred during the processing. Check the Processed Files table.

- **extract more data** indicates missing files or timeframe shorter than 2 weeks. Run the dwh-migration-dumper to extract more data.

## Input files processing status



Schema version  
**SF20240603**

## Assessment Input Data

Files Count <b>5,646</b>	Processed Data Size <b>111.54 GB</b>	Total Data Size <b>111.54 GB</b>
-----------------------------	---	-------------------------------------

**Data size** is related to the files recognized by the BigQuery migration assessment only. **SKIPPED** and **EMPTY** files that are not being processed are not included in the calculation.

## Processed Files and Errors

FilePath	EntryPath	Size	Status	Error Type	Message
1. warehouse_metering_lite.csv		0 B	MISSING	Missing input file	This file is missing in the assessment input. Run dwh-migrati...
2. warehouse_events_lite.csv		0 B	MISSING	Missing input file	This file is missing in the assessment input. Run dwh-migrati...
3. gs://snowflake-migration-prd/...	compilerworks-format.txt	18 B	SKIPPED		
4. gs://snowflake-migration-prd/...	functions-au.csv	169.8 KB	SKIPPED		
5. gs://snowflake-migration-prd/...	views-au.csv	11.6 KB	SKIPPED		
6. gs://snowflake-migration-prd/...	compilerworks-arguments.txt	356 B	SKIPPED		
7. gs://snowflake-migration-prd/...	compilerworks-version.txt	6.9 KB	SKIPPED		
8. gs://snowflake-migration-prd/...	compilerworks-metadata.yaml	8 KB	SKIPPED		
9. gs://snowflake-migration-prd/...	warehouse_load_history-au_...	271.4 KB	COMPLETED		
10. gs://snowflake-migration-prd/...	query_history-au_2025-04-14...	1.8 MB	COMPLETED		
11. gs://snowflake-migration-prd/...	query_history-au_2025-04-0...	43.7 MB	COMPLETED		
12. gs://snowflake-migration-prd/...	query_history-au_2025-05-16...	5.1 MB	COMPLETED		

1 - 20 / 5648 &lt; &gt;

## Input ZIP archives

FilePath	Completed	Empty	Failed	Skipped	All files	Processed Data
1. gs://snowflake-migration-prd/dwh-migration-snowflake-logs-20250626T143722.zip	5,280	352	0	0	5,632	100.0%
2. gs://snowflake-migration-prd/dwh-migration-snowflake-metadata.zip	10	1	0	6	17	100.0%

1 - 2 / 2 &lt; &gt;

## Usage timeframe:

Start time

Dec 31, 2024, 4:00:00PM

End time

Jun 25, 2025, 4:59:59PM