

Reports Created Colibra this targets the DQ Metastore which is Postgres.

Dataset Scans and Scores By Schema

```
%jdbc
select * from public.dataset_scan where dataset like 'public.%';
```

FINISHED

Scans By Month By Schema - 'Public'

```
%jdbc
select dataset, DATE_TRUNC('MONTH', run_id) as run_id, count(*) as Total_Scans from dataset_scan where
dataset like
'public%' group by dataset, run_id order by run_id asc
```

FINISHED

Rule Breaks Past 30 Days

```
%jdbc
select * from rule_output where dataset like 'public%' and run_id < NOW() - INTERVAL '30 DAY';
```

FINISHED

Scheduled Jobs Queue

```
%jdbc
select job_id,agent_id,dataset,run_id,status,activity,start_time from public.owlcheck_q;
```

FINISHED

Column Counts from Dataset Schema

```
%jdbc
select dataset, count(*) from dataset_schema group by dataset;
```

FINISHED

Metadata / Schema / Datatypes

```
%jdbc
select * from public.dataset_schema;
```

FINISHED

Profile Stats

```
%jdbc
select * from public.dataset_field;
```

FINISHED

Locate Similar Columns

```
%jdbc
select distinct dataset, field_nm, max_abs from dataset_field where max_abs = 'Wireless Telecommunications'
```

FINISHED

Same Column Names

```
%jdbc
select distinct dataset, field_nm from dataset_field where field_nm = 'authenticated_user'
```

FINISHED

Long Running Jobs

```
%report
select dataset,run_id,total_time from dataset_activity where total_time is not null order by total_time desc
```

FINISHED

Owlchecks by User

```
%report
select count(*) as owlchecks, username from owlcheck_q
where updt_ts < now()
group by username order by owlchecks desc
```

FINISHED

Owlcheck by User, Dataset

```
%report
select count(*), user_nm, dataset from dev.public.owl_check_history group by user_nm, dataset order by count
desc
```

FINISHED

Largest by Row Count

```
%report
select dataset,rc as row_count from dataset_scan order by rc desc
```

FINISHED

Owlchecks by Month

```
%report
with grp as (
select date_trunc('MONTH', run_id) as by_month from dataset_scan where run_id < now()
)
select count(*) as owlchecks, by_month from grp group by by_month order by by_month desc
```

Rules by User

%report

select count(*) as rules, user_nm from owl_rule group by user_nm order by rules desc

FINISHED

By Spark(Cluster) Usage

%report

select * from opt_spark order by num_executors desc

FINISHED

Breaks By Number of Records

%report

select * from public.dataset_scan where dataset like '%BUS_ID_Z%' order by rc desc

FINISHED

Queries (Rules) being Run

%report

select dataset,run_id,query,user from public.owl_check_history where dataset like '%BUS_ID_Z%'

FINISHED

Metadata Metrics for Catalog

%report

select * from dataset_field where dataset like '%BUS_ID_Z_DQ003_DR010%'

FINISHED

Failed Records for Given Rule

%report

select * from public.rule_breaks where dataset = 'BUS_ID_R_DQ003_DR01' and rule_nm = 'Exception_Records'

FINISHED