

Some useful SQL queries in Airflow

List the Airflow tables in PostgreSQL

```
SELECT
  *
FROM
  pg_catalog.pg_tables
WHERE
  schemaname != 'pg_catalog'
AND schemaname != 'information_schema'
AND tableowner = 'airflow';
```

List the task instances

```
SELECT *
FROM task_instance;
```

List the dag runs

```
SELECT *
FROM dag_run;
```

Get total completed task count

```
SELECT COUNT(1)
FROM task_instance
WHERE
  state IS NOT NULL
  AND state NOT IN ('scheduled', 'queued');
```

Get tasks started per minute for the past week

```
SELECT
  date_trunc('minute', start_date) AS d,
  count(1)
FROM task_instance
GROUP BY d
ORDER BY 1 DESC
LIMIT 24*7;
```

Get tasks finished per minute for the past week

```
SELECT
  date_trunc('minute', end_date) AS d,
  count(1)
FROM task_instance
WHERE
  state IN ('skipped', 'success', 'failed')
  AND end_date IS NOT NULL
GROUP BY d
ORDER BY 1 DESC
LIMIT 24*7;
```