DATA WAREHOUSE USING GOOGLE BIGQUERY

PEOPLE ANALYTICS

[Login to Google BigQuery Studio using your google workspace account.](#_vnnp0z5s09z4)

[Create a project](#_ad0aphqqeuid)

[Create a dataset](#_9ipoyaw9up8v)

[Create table](#_kqyzwyb8ma11)

[Diagram](#_m9rukrv6rqby)

[Operations](#_3f5ed55uj9se)

Google BigQuery is a data warehouse tool which in turn allows SQL operations.

<https://cloud.google.com/bigquery/docs/introduction>

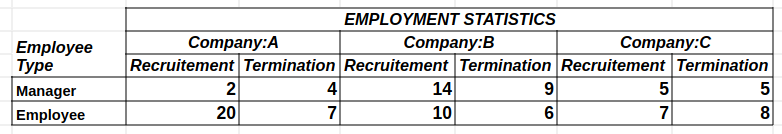
# Login to Google BigQuery Studio using your google workspace account.

<https://console.cloud.google.com/bigquery>

Note that SANDBOX is a way to utilize Google BigQuery without any enrollment as such. Here, your work/data is retained but queries are in the session only and not allowed to persist with cloud storage. Free for Beginners with limited access as such.

A data analyst is planning to record information shown, after certain processing steps, into a data warehouse. Provide schemas for data warehouse (multidimensional modeling), to be able to store using RDBMS. Narrate OLAP operations on this newly created data warehouse.

**Employment statistics in the year of 2024**



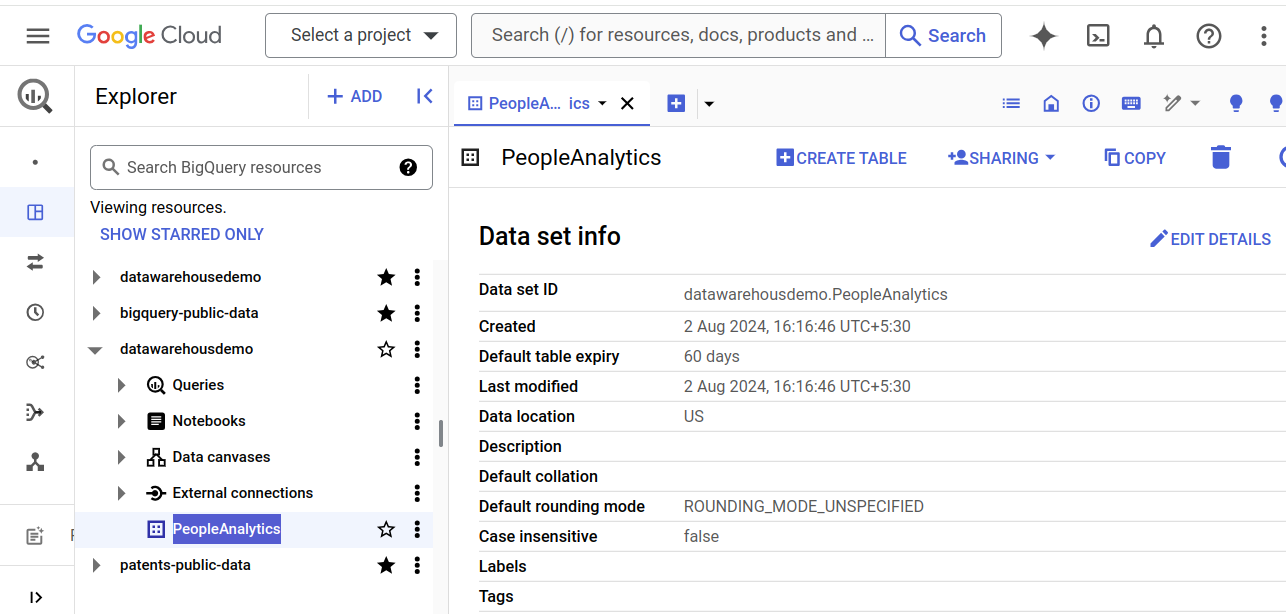
**P.S. Note that the companies already had employees in the year of 1993**

# Create a project

datawarehousedemo

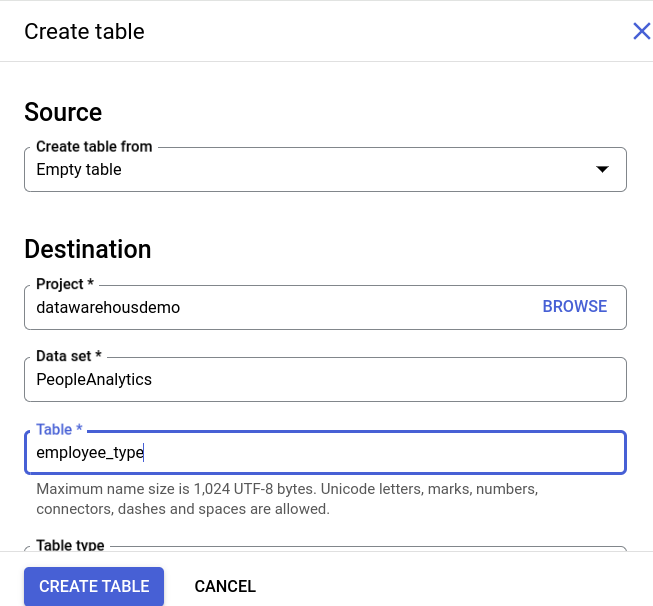
## Create a dataset

datawarehousdemo.PeopleAnalytics



### Create table

1. **Employee\_type (X axis)**



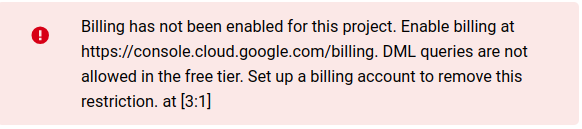
While trying to insert into empty table using SQL ‘insert’ limitation of not being able to perform DML was thrown.

insert into

datawarehousdemo.PeopleAnalytics.employee\_type

(employee\_type\_id, employee\_type\_name, employee\_type\_description)

values(1,'Manager','Employee who is a manager by job title');



Hence

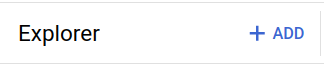
| employee\_type\_id | employee\_type\_name | employee\_type\_description |
| --- | --- | --- |
| 1 | Manager | Employee who is a manager by job title |
| 2 | GeneralEmployee | Regular employee who may be assigned any given task per day. |

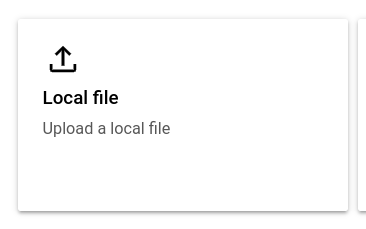
employee\_type.csv

employee\_type\_id,employee\_type\_name,employee\_type\_description

1,Manager,"Employee who is a manager by job title."

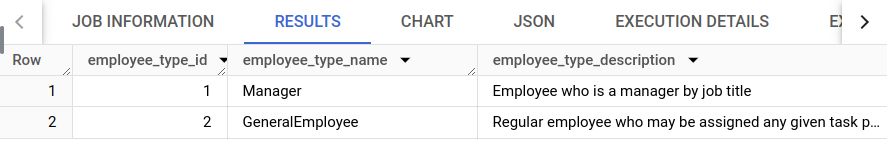
2,GeneralEmployee,"Regular employee who may be assigned any given task per day."





Upload csv file content and verify the records in employee\_type table.

select \* from datawarehousdemo.PeopleAnalytics.employee\_type;



1. **Company\_type (Y axis)**

company\_type\_id,company\_type\_name,company\_type\_description

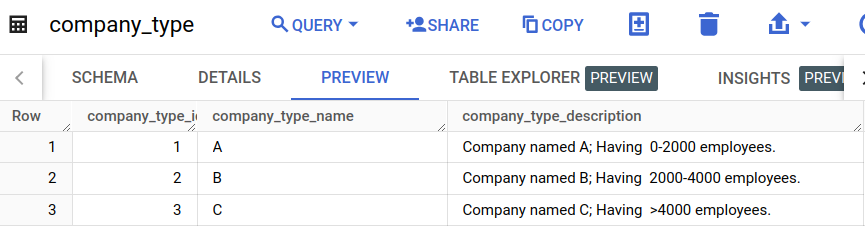
1,A,"Company named A; Having 0-2000 employees."

2,B,"Company named B; Having 2000-4000 employees."

3,C,"Company named C; Having >4000 employees."

Utilise “+ADD” button to create table and insert data from file same single step. Verify with the preview tab or SQL results way.



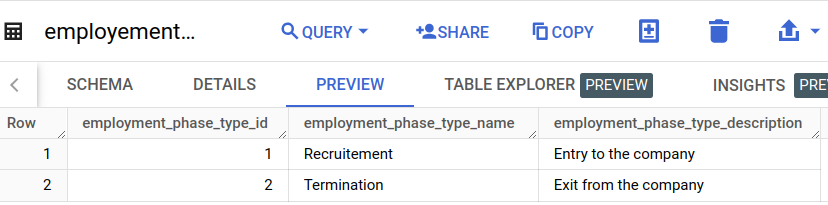


1. **Employement\_phase\_type (Z axis)**

employment\_phase\_type\_id,employment\_phase\_type\_name,employment\_phase\_type\_description

1,Recruitement,"Entry to the company"

2,Termination,"Exit from the company"



1. **employee\_in\_out\_facts**

employee\_type\_id,company\_type\_id,employment\_phase\_type\_id,count\_of\_employees

1,1,1,2

1,1,2,4

1,2,1,14

1,2,2,9

1,3,1,5

1,3,2,5

2,1,1,20

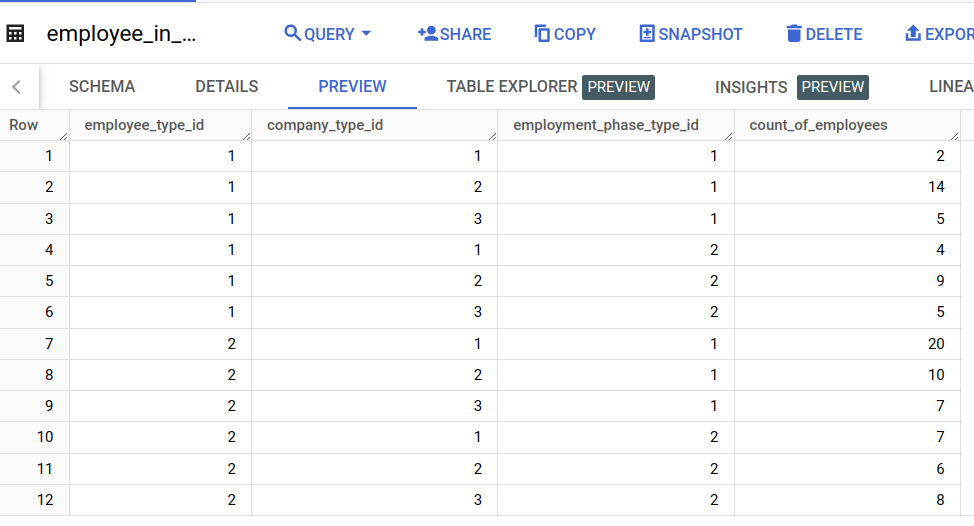
2,1,2,7

2,2,1,10

2,2,2,6

2,3,1,7

2,3,2,8



# Diagram



# Operations

1. How many managers were recruited?

select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=1;

21

1. How many employees were terminated in type A companies?

select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where company\_type\_id=1 and employment\_phase\_type\_id=2;

11

P.s. Considering employees both manager and non-manager.

1. What was the overall increase in recruitment in type A companies?

As we only have data available for a single year i.e. 2024. We don’t have anything to compare with the recruitment in type A companies. We may be able to use baseline with either type B companies or type C companies.

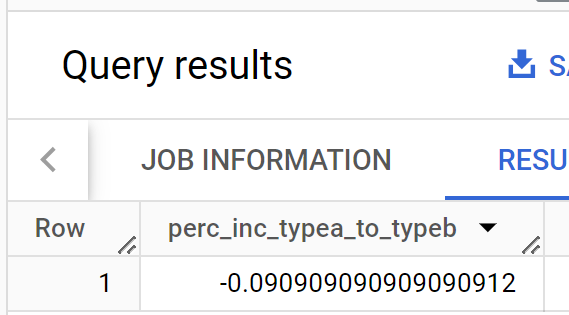
Let’s find out the overall increase in recruitment in type A companies to type B companies.

select (inner\_table.typeacomp-inner\_table.typebcomp)/inner\_table.typeacomp as perc\_inc\_typea\_to\_typeb from

(select (select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where company\_type\_id=1 and employment\_phase\_type\_id=1) as typeacomp, (select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where company\_type\_id=2 and employment\_phase\_type\_id=1) as typebcomp) as inner\_table;



select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where company\_type\_id=1 and employment\_phase\_type\_id=1;

22

select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where company\_type\_id=2 and employment\_phase\_type\_id=1;

24

(22-24)/22 =-0.0909

The negative symbol shows that there was actually a percentage decrease of company types A to company types B overall recruitment.

-9.1 %

1. If the total number of managers at the beginning of 2024 was 65, then what was the percentage increase in it at the end of 2024?

Newly recruited managers

select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=1 ;

21

Managers terminated from the job/post

select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=2 ;

18

Hence, there were 21-18=3 only 3 managers were added to the existing count that is 65 at the beginning of 2024.

Percentage increase

(68-65)/65 = 0.046

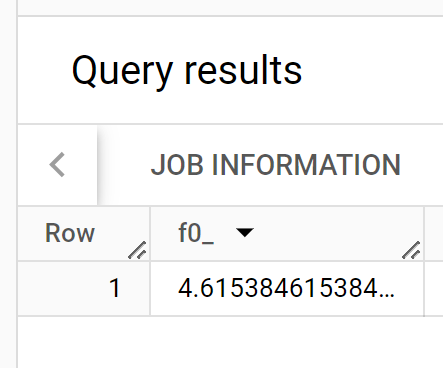
4.6 %

select

(((((select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=1))

-((select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts where employee\_type\_id=1 and employment\_phase\_type\_id=2))+65)-65)/65)\*100;



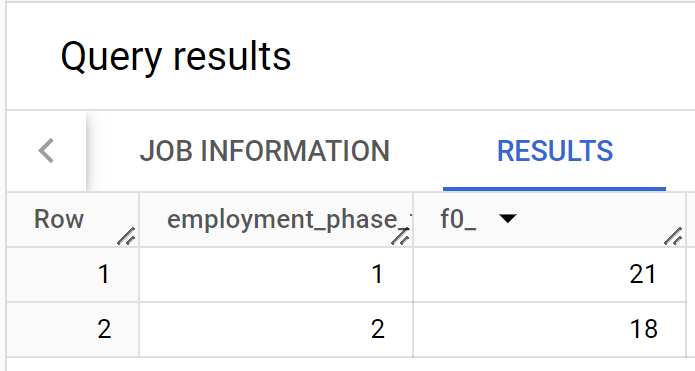
1. Net employment about managers overall across all companies

21-18=3

select employment\_phase\_type\_id, sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1

group by employment\_phase\_type\_id



select (select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=1)-

(select sum(count\_of\_employees) from datawarehousdemo.PeopleAnalytics.employee\_in\_out\_facts

where employee\_type\_id=1 and employment\_phase\_type\_id=2) as net\_employment;

