

**DATA WAREHOUSE USING GOOGLE BIGQUERY**

*PEOPLE ANALYTICS*

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

[Create a project](#)

[Create a dataset](#)

[Create table](#)

[Diagram](#)

[Operations](#)

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

Employee Type	EMPLOYMENT STATISTICS					
	Company:A		Company:B		Company:C	
	Recruitment	Termination	Recruitment	Termination	Recruitment	Termination
Manager	2	4	14	9	5	5
Employee	20	7	10	6	7	8

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

## Create a project

datawarehousedemo

## Create a dataset

datawarehousedemo.PeopleAnalytics

The screenshot shows the Google Cloud BigQuery Explorer interface. On the left, the 'Explorer' sidebar lists resources under the 'datawarehousedemo' project. The 'PeopleAnalytics' dataset is selected. The main panel displays the 'Data set info' for 'PeopleAnalytics'.

Data set info	
Data set ID	datawarehousedemo.PeopleAnalytics
Created	2 Aug 2024, 16:16:46 UTC+5:30
Default table expiry	60 days
Last modified	2 Aug 2024, 16:16:46 UTC+5:30
Data location	US
Description	
Default collation	
Default rounding mode	ROUNDING_MODE_UNSPECIFIED
Case insensitive	false
Labels	
Tags	

## Create table

### 1) Employee\_type (X axis)

Create table ✕

### Source

Create table from  
Empty table ▼

### Destination

Project \*  
datawarehousdemo BROWSE

Data set \*  
PeopleAnalytics

Table \*  
employee\_type

Maximum name size is 1,024 UTF-8 bytes. Unicode letters, marks, numbers, connectors, dashes and spaces are allowed.

Table type

CREATE TABLE

CANCEL

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');
```

5



Billing has not been enabled for this project. Enable billing at <https://console.cloud.google.com/billing>. DML queries are not allowed in the free tier. Set up a billing account to remove this restriction. at [3:1]

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."

Explorer

+ ADD



**Local file**

Upload a local file

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

```
select * from datawarehousdemo.PeopleAnalytics.employee_type;
```

	JOB INFORMATION	RESULTS	CHART	JSON	EXECUTION DETAILS	ED
Row	employee_type_id	employee_type_name	employee_type_description			
1	1	Manager	Employee who is a manager by job title			
2	2	GeneralEmployee	Regular employee who may be assigned any given task p...			

## 2) 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.

Explorer

+ ADD

company_type		QUERY	SHARE	COPY				
	SCHEMA	DETAILS	PREVIEW	TABLE EXPLORER	PREVIEW	INSIGHTS	PREVIEW	
Row	company_type_id	company_type_name	company_type_description					
1	1	A	Company named A; Having 0-2000 employees.					
2	2	B	Company named B; Having 2000-4000 employees.					
3	3	C	Company named C; Having >4000 employees.					

### 3) Employment\_phase\_type (Z axis)

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

- 1,Recrutement,"Entry to the company"
- 2,Termination,"Exit from the company"

employment...

QUERY

SHARE

COPY

<

SCHEMA

DETAILS

PREVIEW

TABLE EXPLORER

PREVIEW

INSIGHTS

PRE

Row	employment_phase_type_id	employment_phase_type_name	employment_phase_type_description
1	1	Recrutement	Entry to the company
2	2	Termination	Exit from the company

### 4) 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



employee\_in\_...

QUERY

SHARE

COPY

SNAPSHOT

DELETE

EXPORT

<

SCHEMA

DETAILS

PREVIEW

TABLE EXPLORER

PREVIEW

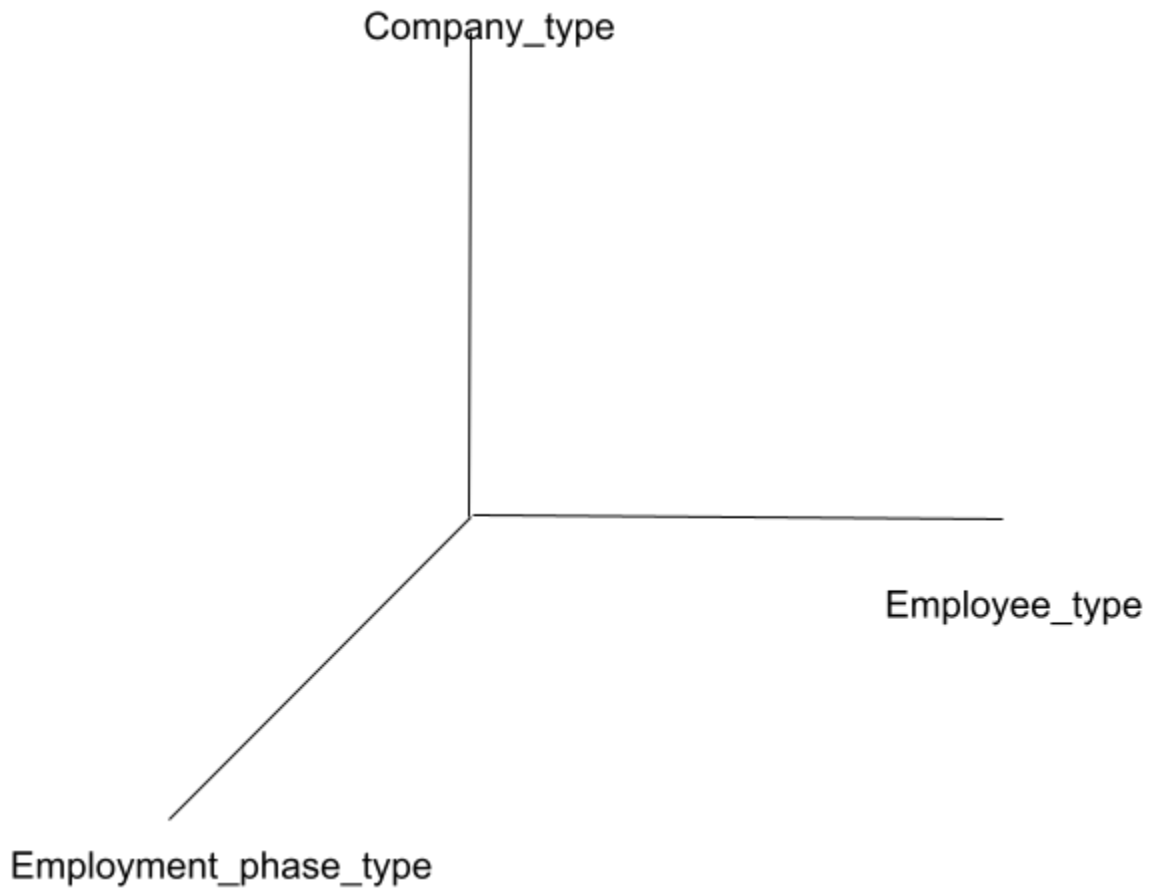
INSIGHTS

PREVIEW

LINEA

Row	employee_type_id	company_type_id	employment_phase_type_id	count_of_employees
1	1	1	1	2
2	1	2	1	14
3	1	3	1	5
4	1	1	2	4
5	1	2	2	9
6	1	3	2	5
7	2	1	1	20
8	2	2	1	10
9	2	3	1	7
10	2	1	2	7
11	2	2	2	6
12	2	3	2	8

# Diagram



## Operations

(1) How many managers were recruited?

```
select sum(count_of_employees) from  
datawarehousedemo.PeopleAnalytics.employee_in_out_facts  
where employee_type_id=1 and employment_phase_type_id=1;
```

21

(2) 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.

(3) 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;
```

Query results		Download	Share
<	JOB INFORMATION	RESULTS	
Row	perc_inc_typea_to_typeb		
1	-0.0909090909090912		

```
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 %

- (4) 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;

```

Query results	
<div> <div>&lt;</div> <div>JOB INFORMATION</div> </div>	
Row	f0_ ▼
1	4.615384615384...

(5) 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

```

Query results		
<div> <div>&lt;</div> <div>JOB INFORMATION</div> <div>RESULTS</div> </div>		
Row	employment_phase	f0_ ▼
1	1	21
2	2	18

```

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;
```

## Query results

 JOB INFORMATION		
Row	net_employment	
1	3	