

# Tasks

## 1. Find the number of rows in the table?

The screenshot shows the DBeaver 21.3.5 interface. On the left, the Database Navigator shows the 'dataanalyticsbootcamp' database with a table 'payroll\_csv' containing 107M rows. The main editor displays a SQL script with multiple queries, including a final query to count the rows: `select count(*) from project.payroll_csv`. The Results panel at the bottom shows the output of this query: a single row with the value 1,161,078.

```
--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate), "year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate)
--from project.payroll_csv
--select max(salary_rate)
--from project.payroll_csv
--where agency_code =56
--select max(salary_rate)
--from project.payroll_csv
--where agency_code =868
select count(*)
from project.payroll_csv
```

Grid	123 count	Value
1	1,161,078	1161078

## 2. What are the columns that are there in the table?

The screenshot shows the DBeaver 21.3.5 interface with the 'payroll\_csv' table selected. The left pane shows the table's structure with columns: year (int4), agency\_code (int4), employee\_name (varchar(32)), agency\_name (varchar(32)), title\_code (varchar(8)), pay\_class (varchar(8)), salary\_rate\_str (varchar(16)), and salary\_rate (int4). The main editor shows the table's data, with columns year, agency\_code, employee\_name, and agency\_name highlighted. The Results panel at the bottom shows the first 26 rows of data.

Grid	123 year	123 agency_code	asc employee_name	asc agency_name	Value
1	2,014	868	B J SANDIFORD	DEPARTMENT OF CITYWIDE	69
2	2,014	868	C A WIGFALL	DEPARTMENT OF CITYWIDE	
3	2,014	69	A E A-AWOSOGBA	HRA/DEPARTMENT OF SOC	
4	2,014	868	K D AABY	DEPARTMENT OF CITYWIDE	
5	2,014	56	I D AADIL	POLICE DEPARTMENT	
6	2,014	69	M AAKIRI	HRA/DEPARTMENT OF SOC	
7	2,014	464	A AALAI	CUNY QUEENSBOROUGH C	
8	2,014	998	A V AALEVIK	N.Y.C. TRANSIT AUTHORITY	
9	2,014	998	M AAMIR	N.Y.C. TRANSIT AUTHORITY	
10	2,014	826	M R AARABI	DEPARTMENT OF ENVIRON	
11	2,014	827	M A AARLEV	DEPARTMENT OF SANITATI	
12	2,014	69	A AARON	HRA/DEPARTMENT OF SOC	
13	2,014	996	A AARON	N.Y.C. HOUSING AUTHORIT	
14	2,014	826	A L AARON	DEPARTMENT OF ENVIRON	
15	2,014	998	A AARON	N.Y.C. TRANSIT AUTHORITY	
16	2,014	998	A A AARON	N.Y.C. TRANSIT AUTHORITY	
17	2,014	996	G AARON	N.Y.C. HOUSING AUTHORIT	
18	2,014	125	H AARON	DEPARTMENT FOR THE AGI	
19	2,014	465	J L AARON	CUNY KINGSBOROUGH CO	
20	2,014	69	M AARON	HRA/DEPARTMENT OF SOC	
21	2,014	868	R L AARON	DEPARTMENT OF CITYWIDE	
22	2,014	816	S A AARON	DEPARTMENT OF HEALTH /	
23	2,014	72	S AARON	DEPARTMENT OF CORRECT	
24	2,014	820	S AARON	OFFICE OF ADMINISTRATIV	
25	2,014	57	W L AARON	FIRE DEPARTMENT	
26	2,014	998	H AARONS	N.Y.C. TRANSIT AUTHORITY	

### 3. What are the data types of columns in the table?

The screenshot shows the DBeaver 21.3.5 interface with the 'payroll\_csv' table selected. The left sidebar displays the database structure, including the 'project' schema and the 'payroll\_csv' table. The main pane shows the table's properties, including the 'Columns' tab which lists the following columns and their data types:

Column Name	#	Data type
123 year	1	int4
123 agency_code	2	int4
noc employee_name	3	varchar(32)
noc agency_name	4	varchar(32)
noc title_code	5	varchar(8)
noc pay_class	6	varchar(8)
noc salary_rate_str	7	varchar(16)
123 salary_rate	8	int4

### 4. Find the employee that got paid the highest in each of the calendar year? (i.e., 2014, 2015 and so on)

The screenshot shows the DBeaver 21.3.5 interface with a SQL script executed. The script is as follows:

```
--select max(salary_rate), year  
--from project.payroll_csv  
--group by "year"  
--select avg(salary_rate), "year"  
--from project.payroll_csv  
--group by "year"  
--select max(salary_rate), "year"  
--from project.payroll_csv  
--group by "year"  
--select max(salary_rate)  
--from project.payroll_csv  
--select max(salary_rate)  
--from project.payroll_csv  
--where agency_code = 56  
--select max(salary_rate)  
--from project.payroll_csv  
--where agency_code = 868  
--select count(*)  
--from project.payroll_csv  
select max(salary_rate), "year"  
from project.payroll_csv  
group by "year"
```

The results of the query are displayed in a grid:

max	year
300,000	2,014
350,000	2,015
350,000	2,016
350,000	2,017

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: dataanalyticsbootcamp - database-1.c7973a0i8c.us-east-1.amazonaws.com

- dataanalyticsbootcamp
  - Schemas
    - lecture
    - practice
    - project
      - Tables
        - payroll\_csv (107M)
          - Columns
            - 123 year (int4)
            - 123 agency\_code (int4)
            - abc employee\_name (varchar(32))
            - abc agency\_name (varchar(32))
            - abc title\_code (varchar(8))
            - abc pay\_class (varchar(8))
            - abc salary\_rate\_str (varchar(16))
            - 123 salary\_rate (int4)
          - Constraints
            - Foreign Keys
            - Indexes
            - Dependencies
            - References

SQL Editor:

```

--select max(salary_rate), year
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate), "year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate)
--from project.payroll_csv
--select max(salary_rate)
--from project.payroll_csv
--where agency_code =56
--select max(salary_rate)
--from project.payroll_csv
--where agency_code =868
--select count(*)
--from project.payroll_csv
select employee_name, year, salary_rate
from project.payroll_csv
where salary_rate = 300000 and year = 2014
  
```

Results 31 | Results 32 | Results 33 | Results 34 | Results 35 | Results 36 | payroll\_csv 37 | "30"

Grid:

abc employee_name	123 year	123 salary_rate
E J MARTI	2,014	300,000
C BIANCO	2,014	300,000

2 row(s) fetched - 284ms (7ms fetch), on Mar 24, 23:18:13

5.What is the average salary rate in each calendar year?

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: dataanalyticsbootcamp - database-1.c7973a0i8c.us-east-1.amazonaws.com

- dataanalyticsbootcamp
  - Schemas
    - lecture
    - practice
    - project
      - Tables
        - payroll\_csv (107M)
          - Columns
            - 123 year (int4)
            - 123 agency\_code (int4)
            - abc employee\_name (varchar(32))
            - abc agency\_name (varchar(32))
            - abc title\_code (varchar(8))
            - abc pay\_class (varchar(8))
            - abc salary\_rate\_str (varchar(16))
            - 123 salary\_rate (int4)
          - Constraints
            - Foreign Keys
            - Indexes
            - Dependencies
            - References

SQL Editor:

```

--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
select avg(salary_rate),"year"
from project.payroll_csv
group by "year"
  
```

Results 7 | payroll\_csv 8 | "2"

Grid:

123 avg	123 year
54,143.6491969256	2,014
52,384.2557466581	2,015
56,299.9700791411	2,016
55,708.2475474628	2,017

4 Rows: 1

6.What is the max salary rate in each calendar year?(i.e.. 2014, 2015 and so on)

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: project > Tables > payroll\_csv (107M)

Columns:

- 123 year (int4)
- 123 agency\_code (int4)
- 123 employee\_name (varchar(32))
- 123 agency\_name (varchar(32))
- 123 title\_code (varchar(8))
- 123 pay\_class (varchar(8))
- 123 salary\_rate\_str (varchar(16))
- 123 salary\_rate (int4)

Script:

```
--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
select max(salary_rate), "year" |
from project.payroll_csv
group by "year"
```

Results:

123 max	123 year
300,000	2,014
350,000	2,015
350,000	2,016
350,000	2,017

Value: 300000

7.What is the max salary rate among all calendar years?

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: project > Tables > payroll\_csv (107M)

Columns:

- 123 year (int4)
- 123 agency\_code (int4)
- 123 employee\_name (varchar(32))
- 123 agency\_name (varchar(32))
- 123 title\_code (varchar(8))
- 123 pay\_class (varchar(8))
- 123 salary\_rate\_str (varchar(16))
- 123 salary\_rate (int4)

Script:

```
--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
select max(salary_rate), "year"
from project.payroll_csv
group by "year"
```

Results:

123 max	123 year
350,000	2,017

Value: 350000

8.Find the max salary paid by agency 56 in any of the calendar years?

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: dataanalyticsbootcamp - database-1.c71973a0i8cl.us-east-1.amazonaws.com

- dataanalyticsbootcamp
  - Schemas
    - lecture
    - practice
    - project
      - payroll\_csv (107M)
        - Columns
          - year (int4)
          - agency\_code (int4)
          - employee\_name (varchar(32))
          - agency\_name (varchar(32))
          - title\_code (varchar(8))
          - pay\_class (varchar(8))
          - salary\_rate\_str (varchar(16))
          - salary\_rate (int4)
        - Constraints
        - Foreign Keys
        - Indexes
        - Dependencies
        - References

SQL Editor:

```
--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate), "year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate)
--from project.payroll_csv
select max(salary_rate)
from project.payroll_csv
where agency_code =56
```

Results 1:

123 max	Value
1	226,366

1 row(s) fetched - 154ms, on Mar 23, 22:42:34

EST en\_US Writable Smart Insert 12 : 1 : 299 Sel: 0 | 0

9. Find the max salary paid by agency 868 in any of the calendar years?

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-13

Database Navigator: dataanalyticsbootcamp - database-1.c71973a0i8cl.us-east-1.amazonaws.com

- dataanalyticsbootcamp
  - Schemas
    - lecture
    - practice
    - project
      - payroll\_csv (107M)
        - Columns
          - year (int4)
          - agency\_code (int4)
          - employee\_name (varchar(32))
          - agency\_name (varchar(32))
          - title\_code (varchar(8))
          - pay\_class (varchar(8))
          - salary\_rate\_str (varchar(16))
          - salary\_rate (int4)
        - Constraints
        - Foreign Keys
        - Indexes
        - Dependencies
        - References

SQL Editor:

```
--select max(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select avg(salary_rate),"year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate), "year"
--from project.payroll_csv
--group by "year"
--select max(salary_rate)
--from project.payroll_csv
--select max(salary_rate)
--from project.payroll_csv
--where agency_code =56
select max(salary_rate)
from project.payroll_csv
where agency_code =868
```

Results 1:

123 max	Value
1	226,366

1 row(s) fetched - 331ms, on Mar 23, 22:48:56

EST en\_US Writable Smart Insert 15 : 24 : 401 Sel: 0 | 0

10. Which title code has max average salary in any of the calendar years?

The screenshot shows the DBeaver 21.3.5 interface. On the left, the Database Navigator shows the 'dataanalyticsbootcamp' database with a table 'payroll\_csv' (107M). The main editor displays the following SQL query:

```
select title_code, avg(salary_rate) as max_avg_salary
from project.payroll_csv
group by title_code
order by max_avg_salary desc
limit 1
```

The 'Results' panel shows one row:

title_code	max_avg_salary
12707	318,644.5

The status bar indicates '1 row(s) fetched - 446ms (1ms fetch), on Mar 27, 22:31:58'.

11. Count the number of records that are present for each agency code

The screenshot shows the DBeaver 21.3.5 interface. The main editor displays the following SQL query:

```
select agency_code, count(agency_code) as num_of_records
from project.payroll_csv
group by agency_code
```

The 'Results' panel shows 14 rows of data:

agency_code	num_of_records
2	2,030
3	3,317
4	440
8	150
9	1,813
10	215
11	231
12	278
13	243
14	182

The status bar indicates '149 row(s) fetched - 262ms (1ms fetch), on Mar 28, 00:05:34'.

12. Count the number of records that are present for each agency name



DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-14

Database Navigator: dataanalyticsbootcamp - database-1.c7i973aoi8cl.us-e

dataanalyticsbootcamp

- Schemas
  - lecture
  - practice
  - project
- Tables
  - payroll\_csv 107M
- Columns
  - 123 year (int4)
  - 123 agency\_code (int4)
  - asc employee\_name (varchar(32))
  - asc agency\_name (varchar(32))
  - asc title\_code (varchar(8))
  - asc pay\_class (varchar(8))
  - asc salary\_rate\_str (varchar(16))
  - 123 salary\_rate (int4)
- Constraints
- Foreign Keys
- Indexes
- Dependencies
- References
- Partitions
- Triggers
- Rules
- Views
  - Materialized Views
- Indexes
- Functions
- Sequences
- Data types
- Aggregate functions
- Event Triggers
- Extensions

SQL Editor:

```
--select title_code , avg(salary_rate) as max_avg_salary
--from project.payroll_csv
--group by title_code
--order by max_avg_salary desc
--limit 1
select agency_name , count(agency_name) as num_of_records
from project.payroll_csv
group by agency_name
```

Results 2:

asc agency_name	num_of_records
1 ADMINISTRATION FOR CHILDRE	25,917
2 BOARD OF CORRECTIONS	90
3 BOARD OF ELECTIONS	3,317
4 BOROUGH PRESIDENT-BRONX	231
5 BOROUGH PRESIDENT-BROOKLYN	278
6 BOROUGH PRESIDENT-MANHATTA	215
7 BOROUGH PRESIDENT-QUEENS	243
8 BOROUGH PRESIDENT-RICHMOND	182
9 BUSINESS INTEGRITY COMMISS	338
10 CAMPAIGN FINANCE BOARD	440
11 CITY CLERK	301

149 row(s) fetched - 318ms, on Mar 28, 00:03:30

### 13.Count the number of records that are present for each pay class

DBeaver 21.3.5 - <dataanalyticsbootcamp> Script-14

Database Navigator: dataanalyticsbootcamp - database-1.c7i973aoi8cl.us-e

dataanalyticsbootcamp

- Schemas
  - lecture
  - practice
  - project
- Tables
  - payroll\_csv 107M
- Columns
  - 123 year (int4)
  - 123 agency\_code (int4)
  - asc employee\_name (varchar(32))
  - asc agency\_name (varchar(32))
  - asc title\_code (varchar(8))
  - asc pay\_class (varchar(8))
  - asc salary\_rate\_str (varchar(16))
  - 123 salary\_rate (int4)
- Constraints
- Foreign Keys
- Indexes
- Dependencies
- References
- Partitions
- Triggers
- Rules
- Views
  - Materialized Views
- Indexes
- Functions
- Sequences
- Data types
- Aggregate functions
- Event Triggers
- Extensions

SQL Editor:

```
--select title_code , avg(salary_rate) as max_avg_salary
--from project.payroll_csv
--group by title_code
--order by max_avg_salary desc
--limit 1
select pay_class , count(pay_class) as num_of_records
from project.payroll_csv
group by pay_class
```

Results 3:

asc pay_class	num_of_records
1 0	6
2 1	50,583
3 2	2,895
4 3	76,115
5 4	90,222
6 A	436,066
7 B	13,336
8 BW	6,727
9 D	235,175
10 E	5,372

25 row(s) fetched - 362ms, on Mar 28, 00:04:24