HUMAN RESOURCES DATA ANALYSIS WITH SQL

August 2024

Presented By
MUHAMMAD ABBAS

Human Resources Data Analysis With SQL

Total Files:

```sql
-- Total Data Files

select *from departments
select *from dept_emp
select *from dept_manager
select *from employees
select *from salaries
select *from titles
```

Data Cleaning:

```sql
-- Data Cleaning & Preparation

exec sp_rename 'dbo.departments.["dept_no"]', 'dept_no', 'column'
exec sp_rename 'dbo.departments.["dept_name"]', 'dept_name', 'column'
exec sp_rename 'dbo.employees.sex', 'gender', 'column'

alter table salaries alter column salary bigint
alter table employees alter column birth_date date
alter table employees alter column hire_date date
```

Data Analysis & Visualization:

Query:

```
-- Determine how many employees work in each department

select
dept_name as 'Department Name',
count('dept_no') as 'Total Employees'
from departments left join dept_emp on
departments.dept_no=dept_emp.dept_no
group by dept_name
```

Output:

	Department Name	Total Employees
1	Customer Service	23580
2	Development	85707
3	Finance	17346
4	Human Resources	17786
5	Marketing	20211
6	Production	73485
7	Quality Management	20117
8	Research	21126
9	Sales	52245

Query:

```
-- Identify the top 10 highest paid employees in the company

select top 10
first_name as 'First Name',
last_name as 'Last Name',
dept_name as 'Department Name',
salary as Salary
from employees join salaries
on employees.emp_no=salaries.emp_no
join dept_emp on
dept_emp.emp_no=salaries.emp_no
join departments on
departments.dept_no=dept_emp.dept_no
order by salary desc
```

Output:

	First Name	Last Name	Department Name	Salary
1	Charmane	Griswold	Sales	129492
2	Slavian	Peac	Sales	127238
3	Nahid	Varker	Sales	127041
4	Poorav	Esposito	Sales	126703
5	Juichirou	Thambidurai	Sales	125469
6	Toshimo	Reghbati	Sales	124357
7	Radoslaw	Pfau	Sales	123749
8	Tsutomu	Alameldin	Sales	123668
9	Martine	Heydon	Sales	123477
10	Chinhyun	Klyachko	Sales	122726

Query:

```
-- Calculate the average salary for each departments

select
dept_name as 'Department Name',
avg(salary) as 'Average Salary'
from dept_emp join salaries
on dept_emp.emp_no=salaries.emp_no
join departments on
departments.dept_no=dept_emp.dept_no
group by dept_name
```

Output:

	Department Name	Average Salary
1	Quality Management	46456
2	Marketing	61095
3	Research	48850
4	Finance	59533
5	Production	48760
6	Development	48697
7	Human Resources	44678
8	Sales	69832
9	Customer Service	47998

Query:

```sql
-- Compare the average salary of male and female employees

select
gender as Gender,
avg(salary) as 'Average Salary'
from employees join salaries
on employees.emp_no=salaries.emp_no
group by gender
```

Output:

	Gender	Average Salary
1	M	52982
2	F	52953

Query:

```sql
-- compare the how many employees work in company basis of genders

select
gender as Gender,
count(gender) as 'Employee Count'
from employees
group by gender
```

Output:

	Gender	Employee Count
1	M	179973
2	F	120051

Query:

```sql
-- Analyze the employees in company on the basis of their ages

select
datediff(year, birth_date,hire_date) as Age,
count(datediff(year, birth_date,hire_date)) as 'Count Employees'
from employees
group by datediff(year, birth_date,hire_date)
order by 'Count Employees' desc
```

Output:

	Age	Count Employees
1	33	22560
2	32	22212
3	31	21311
4	30	20695
5	34	20210
6	29	19418
7	28	18100
8	35	17281
9	27	16391
10	26	14875
11	36	14819
12	25	12784
13	37	12432
14	24	10591
15	38	10177
16	39	8224
17	23	8149
18	40	6493

19	22	5876
20	41	5022
21	42	3713
22	21	2896
23	43	2507
24	44	1632
25	45	904
26	46	407
27	20	222
28	47	123

Query:

```
-- identify the employees who are not cureently assigned to any department

select
count(dept_no) as 'Not Assigned Employees'
from dept_emp
where dept_no is null
```

Output:

	Not Assigned Employees
1	0

Query:

```sql
-- Caculate the ratio of managers to employees in each department

select
dept_name as 'Department Name',
count(emp_no) as 'Total Managers'
from departments join dept_manager
on departments.dept_no=dept_manager.dept_no
group by dept_name
```

Output:

	Department Name	Total Managers
1	Customer Service	4
2	Development	2
3	Finance	2
4	Human Resources	2
5	Marketing	2
6	Production	4
7	Quality Management	4
8	Research	2
9	Sales	2

Query:

```sql
-- Analyze how many employees hired in each year

select
year(hire_date) as Years,
count(emp_no) as  'Total Employees'
from employees
group by year(hire_date)
order by year(hire_date)
```

Output:

	Years	Total Employees
1	1985	35316
2	1986	36150
3	1987	33501
4	1988	31436
5	1989	28394
6	1990	25610
7	1991	22568
8	1992	20402
9	1993	17772
10	1994	14835
11	1995	12115
12	1996	9574
13	1997	6669
14	1998	4155
15	1999	1514
16	2000	13

Query:

```sql
-- Calculate the total salary expence for each department

select
dept_name as 'Department Name',
sum(salary) as 'Salaries Expances'
from departments join dept_emp
on departments.dept_no=dept_emp.dept_no
join salaries
on dept_emp.emp_no=salaries.emp_no
group by dept_name
order by sum(salary)
```

Output:

	Department Name	Salaries Expances
1	Human Resources	794654412
2	Quality Management	934555467
3	Research	1032009108
4	Finance	1032668378
5	Customer Service	1131808649
6	Marketing	1234809182
7	Production	3583161729
8	Sales	3648379483
9	Development	4173702462

Query:

```
-- Calculate the number of employees by team member distribution

select
title as 'Employee Distribution',
count(emp_title_id) as Salaries
from titles join employees
on titles.title_id=employees.emp_title_id
group by title
order by Salaries desc
```

Output:

	Employee Distribution	Average Salary
1	Senior Staff	58550
2	Staff	58465
3	Manager	51531
4	Technique Leader	48582
5	Assistant Engineer	48564
6	Engineer	48535
7	Senior Engineer	48506

Query:

```sql
-- Calculate the number of employees as gender by team member distribution

select
title as 'Employee Distribution',
sum(case when gender='M' then 1 else 0 end) as Male,
sum(case when gender='F' then 1 else 0 end) as Female
from titles join employees
on titles.title_id=employees.emp_title_id
group by title
```

Output:

	Employee Distribution	Male	Female
1	Assistant Engineer	3502	2333
2	Engineer	28340	18963
3	Manager	11	13
4	Senior Engineer	58608	39139
5	Senior Staff	15937	10646
6	Staff	64534	42850
7	Technique Leader	9041	6107

Query:

```sql
-- Analyze the department managers by gender

select
gender as Gender,
count(dept_no) as 'Managers Counts'
from employees join dept_manager
on employees.emp_no=dept_manager.emp_no
group by gender
```

Output:

	Gender	Managers Counts
1	F	13
2	M	11

Query:

```
-- Calculate the total salary expence for by team member distribution

select
title as 'Employee Distribution',
sum(salary) as Salary
from titles join employees
on titles.title_id=employees.emp_title_id
join salaries
on salaries.emp_no=employees.emp_no
group by title
order by sum(salary) desc
```

Output:

	Employee Distribution	Salary
1	**Staff**	**6278246672**
2	Senior Engineer	4741394167
3	Engineer	2295867023
4	Senior Staff	1556439241
5	Technique Leader	735933710
6	Assistant Engineer	283373475
7	Manager	1236745

Query:

```
-- Calculate the average salary for employee based on hired year

select
year(hire_date) as 'Hired Year',
avg(salary) as Salary
from employees join salaries
on employees.emp_no = salaries.emp_no
group by year(hire_date)
order by 'Hired Year' asc
```

Output:

	Hired Year	Salary
1	1985	53010
2	1986	53032
3	1987	52971
4	1988	52834
5	1989	53050
6	1990	52972
7	1991	52876
8	1992	52902
9	1993	53152
10	1994	52994
11	1995	52843
12	1996	52934
13	1997	52888
14	1998	53098
15	1999	53098
16	2000	52292

Query:

```sql
-- Analyze how the company monthly salary has changed over the time

select
month(hire_date) as 'Hired Month',
avg(salary) as Salary
from employees join salaries
on employees.emp_no = salaries.emp_no
where year(hire_date) =1999
group by
month(hire_date)
order by 'Hired Month' asc
```

Output:

	Hired Month	Salary
1	1	53194
2	2	52739
3	3	52683
4	4	54556
5	5	52975
6	6	51450
7	7	52709
8	8	56039
9	9	51396
10	10	50259
11	11	54674
12	12	55618

Query:

```sql
-- Analyze how many employees were hired each month in a specific year

select
month(hire_date) as 'Hired Month',
count(emp_no) as 'Total Employees'
from employees
where year(hire_date)=1995
group by month(hire_date)
order by 'Hired Month' asc
```

Output:

	Hired Month	Total Employees
1	1	1160
2	2	988
3	3	1113
4	4	1057
5	5	1040
6	6	981
7	7	1017
8	8	1013
9	9	943
10	10	943
11	11	907
12	12	953

Query:

```sql
-- Determine the distribution of employees birthdays by month

select
month(birth_date) as 'Birth Month',
count(emp_no) as 'Birthdays In Each Months'
from employees
group by month(birth_date)
order by 'Birth Month' asc
```

Output:

	Birth Month	Birthdays In Each Months
1	1	25412
2	2	23483
983	3	25649
4	4	24631
5	5	25113
6	6	24712
7	7	25698
8	8	25262
9	9	24720
10	10	25518
11	11	24500
12	12	25326