

In [ ]:

In [7]: `import pandas as pd`

```
file_path=(r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\attendance.xls")
df = pd.read_excel(file_path)
```

```
import pandas as pd
file_path = (r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\departments.xlsx")
df = pd.read_excel(file_path)
```

In [8]: `import pandas as pd`

```
file_path = (r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\employees.xls")
df = pd.read_excel(file_path)
```

In [17]: `import pandas as pd`

```
from sqlalchemy import create_engine
engine = create_engine("mysql+pymysql://root:root@localhost/projects")
files={
    "employees": r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\employee",
    "departments": r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\depart",
    "attendance": r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\attenda"
}
for table, path in files.items():
    df = pd.read_excel(path)
    df.columns = df.columns.str.replace(r'^[0-9a-zA-Z]+', '_', regex=True)
    df.to_sql(table, con=engine, if_exists="replace", index=False)
    print(f"Loaded {table} into database with cleaned column name")
```

Loaded employees into database with cleaned column name

Loaded departments into database with cleaned column name

Loaded attendance into database with cleaned column name

In [12]: `%load_ext sql`

```
%sql mysql+mysqlconnector://root:root@localhost/test
```

The sql extension is already loaded. To reload it, use:

```
%reload_ext sql
```

In [13]: `%%sql`

```
use projects
```

```
* mysql+mysqlconnector://root:***@localhost/test
```

0 rows affected.

Out[13]: `[]`In [27]: `%%sql`

```
select * from attendance
```

```
* mysql+mysqlconnector://root:***@localhost/test
```

30 rows affected.

Out[27]:

EmployeeID	Date	Status
110	2024-12-04 00:00:00	Remote
107	2024-12-04 00:00:00	Present
104	2024-12-10 00:00:00	Absent
105	2024-12-10 00:00:00	Absent
103	2024-12-08 00:00:00	Present
102	2024-12-05 00:00:00	Absent
104	2024-12-08 00:00:00	Present
102	2024-12-03 00:00:00	Absent
110	2024-12-05 00:00:00	Present
106	2024-12-07 00:00:00	Remote
102	2024-12-08 00:00:00	Absent
108	2024-12-06 00:00:00	Present
107	2024-12-09 00:00:00	Remote
101	2024-12-02 00:00:00	Absent
103	2024-12-07 00:00:00	Absent
110	2024-12-06 00:00:00	Absent
108	2024-12-07 00:00:00	Absent
110	2024-12-08 00:00:00	Absent
101	2024-12-01 00:00:00	Absent
104	2024-12-07 00:00:00	Absent
109	2024-12-09 00:00:00	Present
101	2024-12-04 00:00:00	Absent
101	2024-12-08 00:00:00	Absent
109	2024-12-02 00:00:00	Present
105	2024-12-10 00:00:00	Present
103	2024-12-03 00:00:00	Present
103	2024-12-04 00:00:00	Remote
106	2024-12-05 00:00:00	Present
108	2024-12-05 00:00:00	Absent
104	2024-12-09 00:00:00	Remote

In [25]:

```
%%sql
select * from departments
```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
4 rows affected.

Out[25]: **DepartmentID DepartmentName Head**

201	HR	Sophia
202	Finance	Liam
203	IT	Olivia
204	Sales	Mason

In [82]: **%%sql**  
**select \* from employees**

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
10 rows affected.

Out[82]: **EmployeeID Name Department Salary JoiningDate ManagerID**

101	Alice	HR	45000	2020-01-15 00:00:00	201
102	Bob	Finance	55000	2019-03-10 00:00:00	202
103	Charlie	IT	60000	2021-06-20 00:00:00	203
104	David	IT	75000	2018-07-01 00:00:00	203
105	Eva	Sales	50000	2022-02-12 00:00:00	204
106	Frank	Finance	65000	2017-11-23 00:00:00	202
107	Grace	Sales	48000	2021-12-01 00:00:00	204
108	Hannah	HR	47000	2019-09-15 00:00:00	201
109	Ian	IT	80000	2023-04-01 00:00:00	203
110	Jack	Sales	52000	2022-08-19 00:00:00	204

In [17]: **%%sql**  
*/\*Which department has the highest number of employees?\*/*  
**select**  
department,  
**count(name) as total\_emp**  
**from employees**  
**group by department**  
**order by total\_emp desc**  
**limit 1**

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
1 rows affected.

Out[17]: **department total\_emp**

Sales	3
-------	---

In [18]: **%%sql**  
*/\*Aggregate by Department*  
*👉 Find the average number of working days (Present) for employees in each depart*  
**with present\_days as(**  
    **select**  
        e.department,

```

        a.employeeid,
        count(*) as total_present_days
    from attendance a
    join employees e
        on a.employeeid = e.employeeid
    where status = "present"
    and year(a.date)=2024
    group by e.department,a.employeeid
)
select
    department,
    avg(total_present_days) as avg_present_days
from present_days
group by department

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test

4 rows affected.

Out[18]: **department avg\_present\_days**

IT	1.6667
Sales	1.0000
Finance	1.0000
HR	1.0000

In [119...

```

%%sql
select
    e.department,
    a.employeeid,
    count(*) as total_present_days
from attendance a
join employees e
    on a.employeeid = e.employeeid
where status = "present"
and year(a.date)=2024
group by e.department,a.employeeid

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test

8 rows affected.

Out[119... **department employeeid total\_present\_days**

IT	103	2
IT	104	1
Sales	105	1
Finance	106	1
Sales	107	1
HR	108	1
IT	109	2
Sales	110	1

In [120...

```

%%sql
/*Aggregate by Department

```

👉 Find the average number of present days for employees in each department \*/

```

with present_avg as(
select
count(status)as total,
Status
from attendance
group by status
)
select
avg(total) as avg_present,status
from present_avg
group by status

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
3 rows affected.

Out[120... **avg\_present** **status**

5.0000	Remote
10.0000	Present
15.0000	Absent

In [96]: **%%sql**  

```

select
count(status)as total,
Status
from attendance
group by status

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
3 rows affected.

Out[96]: **total** **Status**

5	Remote
10	Present
15	Absent

In [55]: **%%sql**  

```

/*Find the total number of employees in each department.*/
select
count(name) as total_employee,
department
from employees
group by department

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
4 rows affected.

Out[55]: **total\_employee** **department**

2	HR
2	Finance
3	IT
3	Sales

```
In [81]: %%sql
/*Manager-Employee Relation
👉 Get the List of departments along with their manager name.*/
select
count(e.name) as total_employee,
d.departmentname,
d.head as manager_name
from employees e
join departments d
on d.departmentid = e.managerid
group by d.head,d.departmentname
```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
4 rows affected.

Out[81]: **total\_employee** **departmentname** **manager\_name**

2	HR	Sophia
2	Finance	Liam
3	IT	Olivia
3	Sales	Mason

In [ ]:

```
In [61]: %%sql
/*Attendance Analysis
👉 Find the total number of absent days for each employee in 2024*/
select
count(status) as total_days,
status
from attendance
where status = "absent"
```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
1 rows affected.

Out[61]: **total\_days** **status**

15	Absent
----	--------

```
In [59]: %%sql
/*Filter + Join
👉 Get the names of employees who belong to the "HR" department.*/
select
name,
department
from employees
where department = "HR"
```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
2 rows affected.

Out[59]: **name** **department**

Alice	HR
Hannah	HR

In [60]: `%%sql`  
*/\*Employee Details*  
 🖱️ *List all employees along with their department name.\*/*  
`select`  
  
     name,  
  
     department  
  
`from` employees

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
 10 rows affected.

Out[60]: **name department**

Alice	HR
Bob	Finance
Charlie	IT
David	IT
Eva	Sales
Frank	Finance
Grace	Sales
Hannah	HR
Ian	IT
Jack	Sales

In [56]: `%%sql`  
*/\*Basic Count*  
 🖱️ *Find the total number of departments in the company.\*/*  
`select`  
     count(department) as total\_dept,  
     department  
`from` employees  
`group by` department

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
 4 rows affected.

Out[56]: **total\_dept department**

2	HR
2	Finance
3	IT
3	Sales

In [49]: `%%sql`  
*/\*Total number of working days (Status = "Present") for each employee \*/*  
`select`

```

year(a.Date) as year,
month(a.Date) as month,
a.employeeid,
e.name,
e.department,
a.Status
from employees e
  join attendance a
on a.employeeid = e.employeeid
  join departments d
on d.departmentid = e.managerid
where status = "present"

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
10 rows affected.

Out[49]:

year	month	employeeid	name	department	Status
2024	12	108	Hannah	HR	Present
2024	12	106	Frank	Finance	Present
2024	12	109	Ian	IT	Present
2024	12	109	Ian	IT	Present
2024	12	104	David	IT	Present
2024	12	103	Charlie	IT	Present
2024	12	103	Charlie	IT	Present
2024	12	110	Jack	Sales	Present
2024	12	107	Grace	Sales	Present
2024	12	105	Eva	Sales	Present

In [38]:

```

%%sql
/*Total number of working days (Status = "Present") for each employee in January
select
  e.employeeid,
  e.name,
  e.department,
  count(*) as total_prestnt_days
from employees e
  join attendance a
on a.employeeid = e.employeeid
  join departments d
on d.departmentid = e.managerid
where a.status = "present"
  and year(a.date)=2024
  and month(a.date)=1
group by e.employeeid,e.name,e.department;

```

\* mysql+mysqlconnector://root:\*\*\*@localhost/test  
0 rows affected.



Out[38]: **employeeid name department total\_prestnt\_days**

```
import pandas as pd from sqlalchemy import create_engine engine =
create_engine("mysql+pymysql://root:root@localhost/projects") folder = r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR
DATA SET\hr_data_set" queries = { "open attendance table": ""select * from attendance"", "open departments
table": ""select * from departments"", "open employees table": ""select * from employees"", "Employee Details List all
employees along with their department name": ""select name, department from employees"", "Total number of working
days Status Present for each employee in January": ""select e.employeeid, e.name, e.department, count(*) as
total_prestnt_days from employees e join attendance a on a.employeeid = e.employeeid join departments d on d.departmentid
= e.managerid where a.status = 'present' and year(a.date)=2024 and month(a.date)=1 group by
e.employeeid,e.name,e.department "", "Total number of working days Status Present for each employee": ""select
year(a.Date) as year, month(a.Date) as month, a.employeeid, e.name, e.department, a.Status from employees e join attendance
a on a.employeeid = e.employeeid join departments d on d.departmentid = e.managerid where status = 'present' "", "Basic
Count Find the total number of departments in the company": ""select count(department) as total_dept, department from
employees group by department "", " Get the names of employees who belong to the HR department": ""select name,
department from employees where department = 'HR' "", " Find the total number of absent days for each employee": ""select
count(status) as total_days, status from attendance where status = 'absent' "", "Manager-Employee Relation Get the list of
departments along with their manager name": ""select count(e.name) as total_employee, d.departmentname, d.head as
manager_name from employees e join departments d on d.departmentid = e.managerid group by d.head,d.departmentname "",
"Find the total number of employees in each department": ""select count(name) as total_employee, department from
employees group by department "", "Find the number of working days Present for employees in each department": ""select
e.department, a.employeeid, count(*) as total_present_days from attendance a join employees e on a.employeeid =
e.employeeid where status = "present" and year(a.date)=2024 group by e.department,a.employeeid "", "Aggregate by
Department Find the average number of present days for employees in each department": ""with present_avg as( select
count(status)as total, Status from attendance group by status ) select avg(total) as avg_present,status from present_avg group
by status "", "Aggregate by Department Find the average number of working days Present for employees in each
department": ""with present_days as( select e.department, a.employeeid, count(*) as total_present_days from attendance a
join employees e on a.employeeid = e.employeeid where status = "present" and year(a.date)=2024 group by
e.department,a.employeeid ) select department, avg(total_present_days) as avg_present_days from present_days group by
department "" } for name, query in queries.items(): try: df = pd.read_sql(query, engine) file_path = f"{folder}\\{name}.csv"
df.to_csv(file_path, index=False) print(f"✅ Saved {file_path}") except Exception as e: print(f"❌ Error in query '{name}':
{e}")
```

```
In [6]: import pandas as pd
from sqlalchemy import create_engine

engine = create_engine("mysql+pymysql://root:root@localhost/projects")
folder = r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr_data_set"

queries = {
    "open attendance table": ""select * from attendance"",
    "open departments table": ""select * from departments"",
    "open employees table": ""select * from employees"",

    "Employee Details List all employees along with their department name": ""
        select
            name,
            department
        from employees
    "",

    "Total number of working days Status Present for each employee in January":
        select
            e.employeeid,
            e.name,
            e.department,
            count(*) as total_present_days
        from employees e
        join attendance a on a.employeeid = e.employeeid
        join departments d on d.departmentid = e.managerid
        where a.status = 'present'
```

```
        and year(a.date) = 2024
        and month(a.date) = 1
    group by e.employeeid, e.name, e.department
""",

"Total number of working days Status Present for each employee": ""
    select
        year(a.Date) as year,
        month(a.Date) as month,
        a.employeeid,
        e.name,
        e.department,
        a.Status
    from employees e
    join attendance a on a.employeeid = e.employeeid
    join departments d on d.departmentid = e.managerid
    where status = 'present'
""",

"Basic Count Find the total number of departments in the company": ""
    select
        count(department) as total_dept,
        department
    from employees
    group by department
""",

"Get the names of employees who belong to the HR department": ""
    select
        name,
        department
    from employees
    where department = 'HR'
""",

"Find the total number of absent days for each employee": ""
    select
        employeeid,
        count(*) as total_absent_days
    from attendance
    where status = 'absent'
    group by employeeid
""",

"Manager-Employee Relation Get the list of departments along with their mana
    select
        count(e.name) as total_employee,
        d.departmentname,
        d.head as manager_name
    from employees e
    join departments d on d.departmentid = e.managerid
    group by d.head, d.departmentname
""",

"Find the total number of employees in each department": ""
    select
        count(name) as total_employee,
        department
    from employees
    group by department
```

```

"""
,

"Find the number of working days Present for employees in each department":
    select
        e.department,
        a.employeeid,
        count(*) as total_present_days
    from attendance a
    join employees e on a.employeeid = e.employeeid
    where status = 'present'
        and year(a.date) = 2024
    group by e.department, a.employeeid
"""

"Aggregate by Department Find the average number of present days for employee
with present_days as (
    select
        e.department,
        a.employeeid,
        count(*) as total_present_days
    from attendance a
    join employees e on a.employeeid = e.employeeid
    where status = 'present'
        and year(a.date) = 2024
    group by e.department, a.employeeid
)
select
    department,
    avg(total_present_days) as avg_present_days
from present_days
group by department
"""

}

for name, query in queries.items():
    try:
        df = pd.read_sql(query, engine)
        file_path = f"{folder}\\{name}.csv"
        df.to_csv(file_path, index=False)
        print(f"✅ Saved {file_path}")
    except Exception as e:
        print(f"❌ Error in query '{name}': {e}")

```

- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\open attendance table.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\open departments table.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\open employees table.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Employee Details List all employees along with their department name.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Total number of working days Status Present for each employee in January.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Total number of working days Status Present for each employee.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Basic Count Find the total number of departments in the company.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Get the names of employees who belong to the HR department.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Find the total number of absent days for each employee.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Manager-Employee Relation Get the list of departments along with their manager name.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Find the total number of employees in each department.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Find the number of working days Present for employees in each department.csv
- ✓ Saved C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set\Aggregate by Department Find the average number of present days for employees in each department.csv

```
In [4]: import os

# Check if folder exists, if not create it
if not os.path.exists(folder):
    os.makedirs(folder)
    print(f"📁 Created folder: {folder}")
else:
    print(f"📁 Folder already exists: {folder}")
```

📁 Folder already exists: C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set

```
In [5]: import pandas as pd
import os
from sqlalchemy import create_engine

engine = create_engine("mysql+pymysql://root:root@localhost/projects")

folder = r"C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr_data_set"

# ✓ Ensure folder exists
if not os.path.exists(folder):
    os.makedirs(folder)
    print(f"📁 Created folder: {folder}")
else:
    print(f"📁 Folder already exists: {folder}")
```

📁 Folder already exists: C:\Users\abdul\OneDrive\Desktop\PROJECTS\HR DATA SET\hr\_data\_set

In [ ]: