DS1_C4_S6_Practice

In [1]:

```
import pandas as pd
import warnings
warnings.filterwarnings('ignore')
```

In [ ]:

```
pip install mysql-connector-python
```

In [ ]:

```
import mysql.connector as sql
```

In [ ]:

```
db = sql.connect(host='localhost',user='root', password= 'password', database='hr')
```

In [ ]:

```
cursor = db.cursor()
```

Task 1

In [6]:

```
sq = 'select * from hr.countries;'
cursor.execute(sq)
data=cursor.fetchall()
data
```

Out[6]:

```
[('AR', 'Argentina', 2),
 ('AU', 'Australia', 3),
 ('BE', 'Belgium', 1),
 ('BR', 'Brazil', 2),
 ('CA', 'Canada', 2),
 ('CH', 'Switzerland', 1),
 ('CN', 'China', 3),
 ('DE', 'Germany', 1),
 ('DK', 'Denmark', 1),
 ('EG', 'Egypt', 4),
 ('FR', 'France', 1),
 ('HK', 'HongKong', 3),
 ('IL', 'Israel', 4),
 ('IN', 'India', 3),
 ('IT', 'Italy', 1),
 ('JP', 'Japan', 3),
 ('KW', 'Kuwait', 4),
 ('MX', 'Mexico', 2),
 ('NG', 'Nigeria', 4),
 ('NL', 'Netherlands', 1),
 ('SG', 'Singapore', 3),
 ('UK', 'United Kingdom', 1),
 ('US', 'United States of America', 2),
 ('ZM', 'Zambia', 4),
 ('ZW', 'Zimbabwe', 4)]
```

In [7]:

```
sq1 = 'select * from hr.regions;'
cursor.execute(sq1)
data1=cursor.fetchall()
data1
```

Out[7]:

```
[(1, 'Europe'), (2, 'Americas'), (3, 'Asia'), (4, 'Middle East and Africa')]
```

In [8]:

```
sq3 = """select country_id,country_name,region_name from countries inner join regions on re
cursor. execute(sq3)
data2 = cursor.fetchall()
data2
```

Out[8]:

```
[('BE', 'Belgium', 'Europe'),
 ('CH', 'Switzerland', 'Europe'),
 ('DE', 'Germany', 'Europe'),
 ('DK', 'Denmark', 'Europe'),
 ('FR', 'France', 'Europe'),
 ('IT', 'Italy', 'Europe'),
 ('NL', 'Netherlands', 'Europe'),
 ('UK', 'United Kingdom', 'Europe'),
 ('AR', 'Argentina', 'Americas'),
 ('BR', 'Brazil', 'Americas'),
 ('CA', 'Canada', 'Americas'),
 ('MX', 'Mexico', 'Americas'),
 ('US', 'United States of America', 'Americas'),
 ('AU', 'Australia', 'Asia'),
 ('CN', 'China', 'Asia'),
 ('HK', 'HongKong', 'Asia'),
 ('IN', 'India', 'Asia'),
 ('JP', 'Japan', 'Asia'),
 ('SG', 'Singapore', 'Asia'),
 ('EG', 'Egypt', 'Middle East and Africa'),
 ('IL', 'Israel', 'Middle East and Africa'),
 ('KW', 'Kuwait', 'Middle East and Africa'),
 ('NG', 'Nigeria', 'Middle East and Africa'),
 ('ZM', 'Zambia', 'Middle East and Africa'),
 ('ZW', 'Zimbabwe', 'Middle East and Africa')]
```

Task 2

In [9]:

```
sq4 = """select employee_id,first_name,last_name, country_name from countries inner join lo
departments on departments.location_id = locations.location_id inner join employees on depa
cursor. execute(sq4)
data3 = cursor.fetchall()
data3
```

Out[9]:

```
[(103, 'Alexander', 'Hunold', 'United States of America'),
 (104, 'Bruce', 'Ernst', 'United States of America'),
 (105, 'David', 'Austin', 'United States of America'),
 (106, 'Valli', 'Pataballa', 'United States of America'),
 (107, 'Diana', 'Lorentz', 'United States of America'),
 (120, 'Matthew', 'Weiss', 'United States of America'),
 (121, 'Adam', 'Fripp', 'United States of America'),
 (122, 'Payam', 'Kaufling', 'United States of America'),
 (123, 'Shanta', 'Vollman', 'United States of America'),
 (124, 'Kevin', 'Mourgos', 'United States of America'),
 (125, 'Julia', 'Nayer', 'United States of America'),
 (126, 'Irene', 'Mikkilineni', 'United States of America'),
 (127, 'James', 'Landry', 'United States of America'),
 (128, 'Steven', 'Markle', 'United States of America'),
 (129, 'Laura', 'Bissot', 'United States of America'),
 (130, 'Mozhe', 'Atkinson', 'United States of America'),
 (131, 'James', 'Marlow', 'United States of America'),
 (132, 'TJ', 'Olson', 'United States of America'),
```

In [10]:

```
data4=pd.read_sql_query('select employee_id,first_name,last_name, country_name from countri
data4
```

Out[10]:

	employee_id	first_name	last_name	country_name
0	103	Alexander	Hunold	United States of America
1	104	Bruce	Ernst	United States of America
2	105	David	Austin	United States of America
3	106	Valli	Pataballa	United States of America
4	107	Diana	Lorentz	United States of America
...	...	...	...	...
101	175	Alyssa	Hutton	United Kingdom
102	176	Jonathon	Taylor	United Kingdom
103	177	Jack	Livingston	United Kingdom
104	179	Charles	Johnson	United Kingdom
105	204	Hermann	Baer	Germany

106 rows × 4 columns

In [11]:

```
group = data4.groupby(['country_name'])['employee_id'].count()
group
```

Out[11]:

```
country_name
Canada                         2
Germany                        1
United Kingdom                35
United States of America      68
Name: employee_id, dtype: int64
```

Task 3

In [12]:

```
sq5 = """select employee_id, concat(first_name,last_name) , country_name, city from countri
departments on departments.location_id = locations.location_id inner join employees on depa
cursor. execute(sq5)
data5 = cursor.fetchall()
data5
```

Out[12]:

```
[(103, 'AlexanderHunold', 'United States of America', 'Southlake'),
 (104, 'BruceErnst', 'United States of America', 'Southlake'),
 (105, 'DavidAustin', 'United States of America', 'Southlake'),
 (106, 'ValliPataballa', 'United States of America', 'Southlake'),
 (107, 'DianaLorentz', 'United States of America', 'Southlake'),
 (120, 'MatthewWeiss', 'United States of America', 'South San Francisc
o'),
 (121, 'AdamFripp', 'United States of America', 'South San Francisco'),
 (122, 'PayamKaufling', 'United States of America', 'South San Francisc
o'),
 (123, 'ShantaVollman', 'United States of America', 'South San Francisc
o'),
 (124, 'KevinMourgos', 'United States of America', 'South San Francisc
o'),
 (125, 'JuliaNayer', 'United States of America', 'South San Francisco'),
 (126, 'IreneMikkilineni', 'United States of America', 'South San Francis
co'),
 (127, 'JamesLandry', 'United States of America', 'South San Francisco'),
```

In [13]:

```
data5=pd.read_sql_query('select employee_id, concat(first_name,last_name) , country_name, c
data5
```

Out[13]:

	employee_id	concat(first_name,last_name)	country_name	city
0	103	AlexanderHunold	United States of America	Southlake
1	104	BruceErnst	United States of America	Southlake
2	105	DavidAustin	United States of America	Southlake
3	106	ValliPataballa	United States of America	Southlake
4	107	DianaLorentz	United States of America	Southlake
...	...	...	...	...
101	175	AlyssaHutton	United Kingdom	Oxford
102	176	JonathonTaylor	United Kingdom	Oxford
103	177	JackLivingston	United Kingdom	Oxford
104	179	CharlesJohnson	United Kingdom	Oxford
105	204	HermannBaer	Germany	Munich

106 rows × 4 columns

In [14]:

```
pvt = pd.pivot_table(data5,index=['country_name'],values=['employee_id'], aggfunc= 'count')
pvt
```

Out[14]:

country_name	employee_id
Canada	2
Germany	1
United Kingdom	35
United States of America	68

In [15]:

```
import matplotlib.pyplot as plt
```

In [31]:

```
plt.pie(pvt['employee_id'],labels=pvt.index,autopct='%.2f%%')
plt.title('Pie Chart')
plt.show()
```