Read the following data set:

https://archive.ics.uci.edu/ml/machine-learning-databases/adult/
Rename the columns as per the description from this file:
https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.names
Task:
Create a sql db from adult dataset and name it sqladb

1. Select 10 records from the adult sqladb

In [1]:
```
import pandas as pd
```

In [2]:
```
col_names = ['age','workclass','fnlwgt','education','education-num','marital-stat
adult_data = pd.read_csv("https://archive.ics.uci.edu/ml/machine-learning-databas
adult_data.head(10)
```

Out[2]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	se
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Mal
1	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Mal
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Mal
3	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Mal
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Femal
5	37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	White	Femal
6	49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family	Black	Femal
7	52	Self-emp-not-inc	209642	HS-grad	9	Married-civ-spouse	Exec-managerial	Husband	White	Mal
8	31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family	White	Femal
9	42	Private	159449	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Mal

2. Show me the average hours per week of all men who are working in private sector

In [13]:
```
adult_data[adult_data["workclass"]==" Private"].groupby("workclass")["hours-per-w
```

Out[13]:
```
workclass
 Private    40.267096
Name: hours-per-week, dtype: float64
```

3. Show me the frequency table for education, occupation and relationship, separately

In [16]:
```
pd.crosstab(index=adult_data["education"],columns="count")
```

Out[16]:

col_0	count
education	
10th	933
11th	1175
12th	433
1st-4th	168
5th-6th	333
7th-8th	646
9th	514
Assoc-acdm	1067
Assoc-voc	1382
Bachelors	5355
Doctorate	413
HS-grad	10501
Masters	1723
Preschool	51
Prof-school	576
Some-college	7291

In [17]:
```
pd.crosstab(index=adult_data["occupation"],columns="count")
```

Out[17]:

occupation	col_0 count
?	1843
Adm-clerical	3770
Armed-Forces	9
Craft-repair	4099
Exec-managerial	4066
Farming-fishing	994
Handlers-cleaners	1370
Machine-op-inspct	2002
Other-service	3295
Priv-house-serv	149
Prof-specialty	4140
Protective-serv	649
Sales	3650
Tech-support	928
Transport-moving	1597

In [18]:
```
pd.crosstab(index=adult_data["relationship"],columns="count")
```

Out[18]:

relationship	col_0 count
Husband	13193
Not-in-family	8305
Other-relative	981
Own-child	5068
Unmarried	3446
Wife	1568

4. Are there any people who are married, working in private sector and having a masters degree

In [28]:
```
ata["marital-status"].str.startswith(" Married")) & (adult_data["workclass"]==" Pr
```

Out[28]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race
5	37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	White
87	33	Private	202051	Masters	14	Married-civ-spouse	Prof-specialty	Husband	White
100	76	Private	124191	Masters	14	Married-civ-spouse	Exec-managerial	Husband	White
188	31	Private	99928	Masters	14	Married-civ-spouse	Prof-specialty	Wife	White
198	35	Private	138992	Masters	14	Married-civ-	Prof-	Other-	White

5. What is the average, minimum and maximum age group for people working in different sectors

In [29]:
```
adult_data.groupby("workclass")["age"].agg([pd.np.min, pd.np.max, pd.np.mean])
```

Out[29]:

workclass	amin	amax	mean
?	17	90	40.960240
Federal-gov	17	90	42.590625
Local-gov	17	90	41.751075
Never-worked	17	30	20.571429
Private	17	90	36.797585
Self-emp-inc	17	84	46.017025
Self-emp-not-inc	17	90	44.969697
State-gov	17	81	39.436055
Without-pay	19	72	47.785714

6. Calculate age distribution by country

In [33]:
```
pd.crosstab(index=adult_data["native-country"],columns=adult_data["age"])
```

Out[33]:

age	17	18	19	20	21	22	23	24	25	26	...	80	81	82	83	84	85
native-country																	
?	2	8	5	10	11	12	6	14	11	18	...	0	2	0	0	0	0
Cambodia	0	1	0	0	0	0	0	0	1	0	...	0	0	0	0	0	0
Canada	2	1	1	2	0	1	3	3	5	2	...	1	0	0	0	0	0
China	0	0	0	0	0	1	2	2	1	0	...	0	0	0	0	0	0
Columbia	0	1	0	0	3	0	2	0	4	3	...	0	0	0	0	0	0
Cuba	0	0	0	0	2	0	3	0	1	0	...	0	0	1	0	0	0
Dominican-Republic	0	1	0	0	3	3	4	3	2	3	...	0	0	0	0	0	0
Ecuador	0	0	0	0	1	0	1	2	1	0	...	0	0	0	0	0	0
El-Salvador	2	0	4	5	2	2	6	7	8	3	...	0	0	0	0	0	0
England	1	0	2	1	1	1	2	3	1	1	...	0	0	0	0	0	0
France	0	0	0	1	0	0	0	1	0	0	...	0	0	0	0	0	0
Germany	0	1	0	5	3	4	1	2	0	3	...	0	0	0	0	0	0
Greece	0	0	0	0	0	2	1	1	0	0	...	0	0	0	0	0	0
Guatemala	0	0	3	3	2	3	4	1	6	2	...	0	0	0	0	0	0
Haiti	1	0	1	1	0	2	2	2	0	0	...	0	0	0	0	0	0
Holand-Netherlands	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
Honduras	0	1	1	0	0	0	0	0	0	1	...	0	0	0	0	0	0
Hong	0	0	1	0	0	0	0	0	0	2	...	0	0	0	0	0	0
Hungary	0	0	0	0	0	0	0	1	0	1	...	0	1	0	0	0	0
India	1	1	0	1	1	2	1	2	3	3	...	0	0	0	0	0	0
Iran	0	0	0	0	0	2	1	1	0	0	...	0	0	0	0	0	0
Ireland	0	0	0	0	0	0	2	0	0	0	...	0	0	0	0	0	0
Italy	0	0	1	1	1	1	0	1	2	0	...	0	0	0	0	0	0
Jamaica	0	1	1	0	3	3	2	4	4	3	...	0	0	0	0	0	0
Japan	0	0	2	0	0	0	0	1	1	2	...	0	0	0	0	0	0
Laos	0	0	2	0	0	0	1	1	0	0	...	0	0	0	0	0	0
Mexico	6	6	14	18	24	24	32	25	39	22	...	0	1	0	0	0	0
Nicaragua	0	0	2	4	1	0	0	0	1	1	...	0	0	0	0	0	0
Outlying-US(Guam-USVI-etc)	0	0	0	0	1	0	1	0	1	1	...	0	0	0	0	0	0
Peru	2	1	0	1	1	1	1	1	1	1	...	0	0	0	0	0	0
Philippines	1	2	2	2	1	4	6	5	3	6	...	0	0	0	0	0	0
Poland	1	0	0	0	1	0	2	0	3	2	...	0	1	0	0	0	1
Portugal	0	0	1	0	1	1	1	1	0	1	...	0	0	0	0	0	0
Puerto-Rico	1	0	3	2	1	1	5	1	1	3	...	0	0	0	0	0	0
Scotland	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0

age	17	18	19	20	21	22	23	24	25	26	...	80	81	82	83	84	85
native-country																	
South	0	0	3	2	1	5	3	3	0	3	...	0	0	0	0	0	0
Taiwan	0	0	0	1	1	2	1	2	3	2	...	0	0	0	0	0	0
Thailand	0	0	2	0	0	1	0	1	0	1	...	0	0	0	0	0	0
Trinadad&Tobago	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
United-States	374	524	659	689	651	683	771	705	734	694	...	21	15	11	6	10	2
Vietnam	0	0	2	3	3	3	10	2	3	1	...	0	0	0	0	0	0
Yugoslavia	0	0	0	1	0	1	0	0	1	0	...	0	0	0	0	0	0

42 rows × 73 columns

7. Compute a new column as 'Net-Capital-Gain' from the two columns 'capital-gain' and 'capital-loss'

In [36]:
```
adult_data["Net-Capital-Gain"] = adult_data["capital-gain"] - adult_data["capital
adult_data.head()
```

Out[36]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	se
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Mal
1	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Mal
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Mal
3	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Mal
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Femal

In [ ]:
```
```