practise questions

In [ ]:
```
a={"name":["madhu","kusum","kinshuk","ankit","shruti"],
"2014":[100.5,150.8,200.9,30000,40000],
"2015":[12000,18000,22000,30000,45000],
"2016":[20000,50000,70000,100000,125000],
"2017":[50000,60000,70000,80000,90000]}
```

In [ ]:
```
import pandas as pd
df=pd.DataFrame(a)
```

In [ ]:
```
df.set_index("name",inplace=True)
df
```

Out[ ]:

name	2014	2015	2016	2017
madhu	100.5	12000	20000	50000
kusum	150.8	18000	50000	60000
kinshuk	200.9	22000	70000	70000
ankit	30000.0	30000	100000	80000
shruti	40000.0	45000	125000	90000

In [ ]:
```
#row labels
df.index
```

Out[ ]:
```
Index(['madhu', 'kusum', 'kinshuk', 'ankit', 'shruti'], dtype='object', name='name')
```

In [ ]:
```
#column labels
df.columns
```

Out[ ]:
```
Index(['2014', '2015', '2016', '2017'], dtype='object')
```

In [ ]:
```
#data types of eac columns
df.dtypes
```

Out[ ]:
```
2014    float64
2015      int64
2016      int64
2017      int64
dtype: object
```

In [ ]:
```
#dimension
df.ndim
```

Out[ ]:
```
2
```

In [ ]:
```
#shape
df.shape
```

Out[ ]:
```
(5, 4)
```

In [ ]:
```
#size
df.size
```

Out[ ]:
```
20
```

In [ ]:
```
#last two rows
df.iloc[-2:]
```

Out[ ]:

name	2014	2015	2016	2017
ankit	30000.0	30000	100000	80000
shruti	40000.0	45000	125000	90000

In [ ]:
```
df.tail(2)
```

Out[ ]:

name	2014	2015	2016	2017
ankit	30000.0	30000	100000	80000
shruti	40000.0	45000	125000	90000

In [ ]:
```
#first two columns
df.iloc[:,:2]
```

Out[ ]:

name	2014	2015
madhu	100.5	12000
kusum	150.8	18000
kinshuk	200.9	22000
ankit	30000.0	30000
shruti	40000.0	45000

In [ ]:
```
#creating a dictionary and and using the dict to create another dataframe
#and checking the the dataframe is empty
b={"name":["madhu","kusum","kinshuk","ankit","shruti"],
"2018":[160000,110000,500000,340000,900000]}
df1=pd.DataFrame(b)
df1.set_index("name",inplace=True)
df1
```

Out[ ]:

name	2018

name	2018
madhu	160000
kusum	110000
kinshuk	500000
ankit	340000
shruti	900000

In [ ]:
```
#join the two tables
df2=df.join(df1,on="name",how="left")
df2
```

Out[ ]:

name	2014	2015	2016	2017	2018
madhu	100.5	12000	20000	50000	160000
kusum	150.8	18000	50000	60000	110000
kinshuk	200.9	22000	70000	70000	500000
ankit	30000.0	30000	100000	80000	340000
shruti	40000.0	45000	125000	90000	900000

In [ ]:
```
#transpose the dataframe
df2.transpose()
```

Out[ ]:

name	madhu	kusum	kinshuk	ankit	shruti
2014	100.5	150.8	200.9	30000.0	40000.0
2015	12000.0	18000.0	22000.0	30000.0	45000.0
2016	20000.0	50000.0	70000.0	100000.0	125000.0
2017	50000.0	60000.0	70000.0	80000.0	90000.0
2018	160000.0	110000.0	500000.0	340000.0	900000.0

In [ ]:
```
#sales made in the year 2017
df2["2017"]
```

Out[ ]:
```
name
madhu      50000
kusum      60000
kinshuk    70000
ankit      80000
shruti     90000
Name: 2017, dtype: int64
```

In [ ]:
```
#sales made by mathu and ankit in the year 2017 and 2018
df2.loc[["madhu","ankit"],["2017","2018"]]
```

Out[ ]:

name	2017	2018
madhu	50000	160000
ankit	80000	340000

In [ ]:
```
#sales made by shruti in 2016
df2.loc["shruti","2016"]
```

Out[ ]:
```
125000
```

In [ ]:
```
#adding a row with index "sumeet"
df2.loc["sumeet"]=[196.2,37800,52000,78438,38852]
df2
```

Out[ ]:

name	2014	2015	2016	2017	2018
madhu	100.5	12000.0	20000.0	50000.0	160000.0
kusum	150.8	18000.0	50000.0	60000.0	110000.0
kinshuk	200.9	22000.0	70000.0	70000.0	500000.0
ankit	30000.0	30000.0	100000.0	80000.0	340000.0
shruti	40000.0	45000.0	125000.0	90000.0	900000.0
sumeet	196.2	37800.0	52000.0	78438.0	38852.0

In [ ]:
```
#deleting the 2014 column
df2.drop("2014",axis=1,inplace=True)
```

In [ ]:
```
#deleting the row with index "kinshuk"
df2.drop("kinshuk",axis=0,inplace=True)
```

In [ ]:
```
#changing the sales person name in index
df2.rename(index={"ankit":"vivaan","madhu":"shailesh"},inplace=True)
df2
```

Out[ ]:

name	2015	2016	2017	2018
shailesh	12000.0	20000.0	50000.0	160000.0
kusum	18000.0	50000.0	60000.0	110000.0
vivaan	30000.0	100000.0	80000.0	340000.0
shruti	45000.0	125000.0	90000.0	900000.0
sumeet	37800.0	52000.0	78438.0	38852.0

In [ ]:
```
#updating the sale made by shailesh in 2018 to 100000
```

```
df2.at["shailesh","2018"]=100000
df2
```

Out[ ]:

name	2015	2016	2017	2018
shailesh	12000.0	20000.0	50000.0	100000.0
kusum	18000.0	50000.0	60000.0	110000.0
vivaan	30000.0	100000.0	80000.0	340000.0
shruti	45000.0	125000.0	90000.0	900000.0
sumeet	37800.0	52000.0	78438.0	38852.0

In [ ]:

```
#saving the dataframe to csv file without index and column labels
df2.to_csv("salesfigures.csv",index=False,header=False)
```

In [ ]:

```
#reading the saved csv file(since i saved the file without header iam using header=N
df3=pd.read_csv("salesfigures.csv",header=None)
df3
```

Out[ ]:

	0	1	2	3
0	12000.0	20000.0	50000.0	100000.0
1	18000.0	50000.0	60000.0	110000.0
2	30000.0	100000.0	80000.0	340000.0
3	45000.0	125000.0	90000.0	900000.0
4	37800.0	52000.0	78438.0	38852.0

In [ ]:

```
#updating the column and  row labels
df3.rename(columns={0:"2015",1:"2016",2:"2017",3:"2018"},inplace=True)
df3.rename(index={0:"shailesh",1:"kusum",2:"vivaan",3:"shruti",4:"sumeet"},inplace=T
df3
```

Out[ ]:

	2015	2016	2017	2018
shailesh	12000.0	20000.0	50000.0	100000.0
kusum	18000.0	50000.0	60000.0	110000.0
vivaan	30000.0	100000.0	80000.0	340000.0
shruti	45000.0	125000.0	90000.0	900000.0
sumeet	37800.0	52000.0	78438.0	38852.0