In [ ]:
```
Python pandas
Nurshanov Dias
IT3-2208
```

In [177…
```
import pandas as pd

data = {
    '2014': [100.5, 150.9, 200.6, 30000, 40000],
    '2015': [12000, 18000, 22000, 30000, 45000],
    '2016': [20000, 50000, 70000, 100000, 125000],
    '2017': [50000, 60000, 80000, 90000, 90000]
}
sales_persons = ['Zack', 'Yasmin', 'Xander', 'Walker', 'Vanda']

Sales = pd.DataFrame(data, index=sales_persons)
Sales
```

Out[177…

	2014	2015	2016	2017
Zack	100.5	12000	20000	50000
Yasmin	150.9	18000	50000	60000
Xander	200.6	22000	70000	80000
Walker	30000.0	30000	100000	90000
Vanda	40000.0	45000	125000	90000

In [178…
```
print('2a) Display the row labels of Sales.')

print(Sales.index)
```

```
2a) Display the row labels of Sales.
Index(['Zack', 'Yasmin', 'Xander', 'Walker', 'Vanda'], dtype='object')
```

In [179…
```
print('2b) Display the column labels of Sales.')

print(Sales.columns)
```

```
2b) Display the column labels of Sales.
Index(['2014', '2015', '2016', '2017'], dtype='object')
```

In [180…
```
print('2c) Display the data types of each column of Sales. --- attribute dataframe_name.dtype')

Sales.dtypes
```

```
2c) Display the data types of each column of Sales. --- attribute dataframe_name.dtype
```

Out[180…
```
2014    float64
2015      int64
2016      int64
2017      int64
dtype: object
```

In [181…
```
print('2d) Display the dimensions, shape, size and values of Sales.')

{
    "dimensions": Sales.ndim,
    "shape": Sales.shape,
    "size": Sales.size,
    "values": Sales.values,
}
```

```
2d) Display the dimensions, shape, size and values of Sales.
```

Out[181…
```
{'dimensions': 2,
 'shape': (5, 4),
 'size': 20,
 'values': array([[1.005e+02, 1.200e+04, 2.000e+04, 5.000e+04],
        [1.509e+02, 1.800e+04, 5.000e+04, 6.000e+04],
        [2.006e+02, 2.200e+04, 7.000e+04, 8.000e+04],
        [3.000e+04, 3.000e+04, 1.000e+05, 9.000e+04],
        [4.000e+04, 4.500e+04, 1.250e+05, 9.000e+04]])}
```

In [182…
```
print('2e) Display the last two rows of Sales.')

Sales.tail(2)
```

```
2e) Display the last two rows of Sales.
```

Out[182…

	2014	2015	2016	2017
Walker	30000.0	30000	100000	90000
Vanda	40000.0	45000	125000	90000

In [183…
```
print('2f) Display the first two columns of Sales.')

Sales.head(2)
```

2f) Display the first two columns of Sales.

Out[183…

	2014	2015	2016	2017
Zack	100.5	12000	20000	50000
Yasmin	150.9	18000	50000	60000

In [184…
```
print('2g) Create a dictionary using the following data. Use this dictionary to create a DataFrame Sales2.')

data = {
    '2018': [
        160000,
        110000,
        500000,
        340000,
        900000,
    ]
}

Sales2 = pd.DataFrame(data, index=sales_persons)
Sales2
```

2g) Create a dictionary using the following data. Use this dictionary to create a DataFrame Sales2.

Out[184…

	2018
Zack	160000
Yasmin	110000
Xander	500000
Walker	340000
Vanda	900000

In [185…
```
print('2h) Check if Sales2 is empty or it contains data.')

Sales2.empty
```

2h) Check if Sales2 is empty or it contains data.

Out[185… False

In [186…
```
print('3a) Append the DataFrame Sales2 to the DataFrame Sales.')

Sales = Sales.join(Sales2)
Sales
```

3a) Append the DataFrame Sales2 to the DataFrame Sales.

Out[186…

	2014	2015	2016	2017	2018
Zack	100.5	12000	20000	50000	160000
Yasmin	150.9	18000	50000	60000	110000
Xander	200.6	22000	70000	80000	500000
Walker	30000.0	30000	100000	90000	340000
Vanda	40000.0	45000	125000	90000	900000

In [187…
```
print('3b) Change the DataFrame Sales such that it becomes its transpose.')

Sales = Sales.T
Sales
```

3b) Change the DataFrame Sales such that it becomes its transpose.

Out[187…

	Zack	Yasmin	Xander	Walker	Vanda
2014	100.5	150.9	200.6	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	80000.0	90000.0	90000.0
2018	160000.0	110000.0	500000.0	340000.0	900000.0

In [188…
```
print('3c) Display the sales made by all sales persons in the year 2017.')
```

```
Sales.loc['2017']
```

3c) Display the sales made by all sales persons in the year 2017.

Out[188…
```
Zack      50000.0
Yasmin    60000.0
Xander    80000.0
Walker    90000.0
Vanda     90000.0
Name: 2017, dtype: float64
```

In [189…
```
print('3d) Display the sales made by Yasmin and Vanda in the year 2017 and 2018.')

Sales.loc[['2017', '2018'],['Yasmin', 'Vanda']]
```

3d) Display the sales made by Yasmin and Vanda in the year 2017 and 2018.

Out[189…

	Yasmin	Vanda
2017	60000.0	90000.0
2018	110000.0	900000.0

In [190…
```
print('3e) Display the sales made by Xander 2016.')

Sales.loc['2016','Xander']
Sales
```

3e) Display the sales made by Xander 2016.

Out[190…

	Zack	Yasmin	Xander	Walker	Vanda
2014	100.5	150.9	200.6	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	80000.0	90000.0	90000.0
2018	160000.0	110000.0	500000.0	340000.0	900000.0

In [191…
```
print('3f) Add data to Sales for salesman Ursula where the sales made are [196.2, 37800, 52000, 78438,38852] in

Sales['Ursula'] = [196.2, 37800.0, 52000.0, 78438.0, 38852.0]
Sales
```

3f) Add data to Sales for salesman Ursula where the sales made are [196.2, 37800, 52000, 78438,38852] in the yea
rs [2014, 2015, 2016, 2017,2018] respectively.

Out[191…

	Zack	Yasmin	Xander	Walker	Vanda	Ursula
2014	100.5	150.9	200.6	30000.0	40000.0	196.2
2015	12000.0	18000.0	22000.0	30000.0	45000.0	37800.0
2016	20000.0	50000.0	70000.0	100000.0	125000.0	52000.0
2017	50000.0	60000.0	80000.0	90000.0	90000.0	78438.0
2018	160000.0	110000.0	500000.0	340000.0	900000.0	38852.0

In [192…
```
print('3g) Delete the data for the year 2014 from the DataFrame Sales.')

Sales = Sales.drop(index=['2014'])
Sales
```

3g) Delete the data for the year 2014 from the DataFrame Sales.

Out[192…

	Zack	Yasmin	Xander	Walker	Vanda	Ursula
2015	12000.0	18000.0	22000.0	30000.0	45000.0	37800.0
2016	20000.0	50000.0	70000.0	100000.0	125000.0	52000.0
2017	50000.0	60000.0	80000.0	90000.0	90000.0	78438.0
2018	160000.0	110000.0	500000.0	340000.0	900000.0	38852.0

In [193…
```
print('3h) Delete the data for sales man Zack from the DataFrame Sales.')

Sales = Sales.drop(columns=['Zack'])
Sales
```

3h) Delete the data for sales man Zack from the DataFrame Sales.

Out[193…

	Yasmin	Xander	Walker	Vanda	Ursula
2015	18000.0	22000.0	30000.0	45000.0	37800.0
2016	50000.0	70000.0	100000.0	125000.0	52000.0
2017	60000.0	80000.0	90000.0	90000.0	78438.0
2018	110000.0	500000.0	340000.0	900000.0	38852.0

In [194…

```
print('3i) Change the name of the salesperson Vanda to Vision and Yasmin to Aladdin.')

Sales.rename(columns={'Vanda': 'Vision', 'Yasmin': 'Aladdin'}, inplace=True)
Sales
```

3i) Change the name of the salesperson Vanda to Vision and Yasmin to Aladdin.

Out[194…

	Aladdin	Xander	Walker	Vision	Ursula
2015	18000.0	22000.0	30000.0	45000.0	37800.0
2016	50000.0	70000.0	100000.0	125000.0	52000.0
2017	60000.0	80000.0	90000.0	90000.0	78438.0
2018	110000.0	500000.0	340000.0	900000.0	38852.0

In [195…

```
print('3j) Update the sale made by Xander in 2018 to 100000.')

Sales['Xander']['2018']=100000
Sales
```

3j) Update the sale made by Xander in 2018 to 100000.

Out[195…

	Aladdin	Xander	Walker	Vision	Ursula
2015	18000.0	22000.0	30000.0	45000.0	37800.0
2016	50000.0	70000.0	100000.0	125000.0	52000.0
2017	60000.0	80000.0	90000.0	90000.0	78438.0
2018	110000.0	100000.0	340000.0	900000.0	38852.0

In [196…

```
print('3k) Write the values of DataFrame Sales to a comma separated file SalesFigures.csv on the disk. Do not w
output_path = 'SalesFigures.csv'
Sales.to_csv(output_path, index=False, header=False)

output_path
```

3k) Write the values of DataFrame Sales to a comma separated file SalesFigures.csv on the disk. Do not write the row labels and column labels.

Out[196… 'SalesFigures.csv'

In [197…

```
print('3l) Read the data in the file SalesFigures.csv into a DataFrame SalesRetrieved and Display it.Now update
SalesRetrieved to be the same as that of Sales.')
SalesRetrieved = pd.read_csv(output_path, header=None)

SalesRetrieved.columns = Sales.columns
SalesRetrieved.index = Sales.index

SalesRetrieved
```

3l) Read the data in the file SalesFigures.csv into a DataFrame SalesRetrieved and Display it.Now update the row labels and column labels of SalesRetrieved to be the same as that of Sales.

Out[197…

	Aladdin	Xander	Walker	Vision	Ursula
2015	18000.0	22000.0	30000.0	45000.0	37800.0
2016	50000.0	70000.0	100000.0	125000.0	52000.0
2017	60000.0	80000.0	90000.0	90000.0	78438.0
2018	110000.0	100000.0	340000.0	900000.0	38852.0