Lab – 27

Student Name: Navneet P

Student ID: AF0411619

Topic: Pandas Pivot Table

Pivot table in Pandas:

A **pivot table** in **Pandas** is a powerful tool for data summarization and transformation, used to group and aggregate data based on one or more columns. It's particularly useful for reshaping data for easier analysis, especially when dealing with large datasets.

pd.pivot_table(data, values=None, index=None, columns=None, aggfunc='mean', fill_value=None, margins=False, margins_name='All', dropna=True)

Dataset used: salesdata.csv

Preview:

	Region	Manager	SalesMan	Item	Units	Unit_price	Sale_amt
0	East	Martha	Alexander	Television	95	1198.0	113810.0
1	Central	Hermann	Shelli	Home Theater	50	500.0	25000.0
2	Central	Hermann	Luis	Television	36	1198.0	43128.0
3	Central	Timothy	David	Cell Phone	27	225.0	6075.0
4	West	Timothy	Stephen	Television	56	1198.0	67088.0
5	East	Martha	Alexander	Home Theater	60	500.0	30000.0
6	Central	Martha	Steven	Television	75	1198.0	89850.0
7	Central	Hermann	Luis	Television	90	1198.0	107820.0
8	West	Douglas	Michael	Television	32	1198.0	38336.0

Question 1: Region-wise, Manager-wise and Salesman wise - To Find Total Sales Amount.

Solution:

Question 1

```
# Import necessary libraries
import pandas as pd
import numpy as np

# Read the sales data from a CSV file into a DataFrame
df = pd.read_csv("salesdatanew.csv")

# Print a message indicating the result being displayed
print("Total Sales amount Region, Manager, and Salesman-wise")

# Create a pivot table to calculate the total sales amount (Sale_amt) for each
combination of Region, Manager, and Salesman
# 'index' specifies that we want to group by the "Region", "Manager", and "SalesMan"
columns
# 'values' specifies that we want to perform aggregation on the "Sale_amt" column
# 'aggfunc=np.sum' calculates the total sales amount for each unique combination of
Region, Manager, and Salesman
df4 = pd.pivot_table(df, index=["Region", "Manager", "SalesMan"], values=
["Sale_amt"], aggfunc=np.sum)

# Print the resulting pivot table
print(df4)
```

Output:

```
Total Sales amount Region,Manager and Salesman-wise
                          Sale_amt
Region  Manager SalesMan
Central Douglas John       124016.0
        Hermann Luis       206373.0
                Shelli      33698.0
                Sigal      125037.5
        Marth   Steven      14000.0
        Martha  Steven     185690.0
        Timothy David      140955.0
East    Douglas Karen       48204.0
        Martha  Alexander  236703.0
                Diana       36100.0
West    Douglas Michael     66836.0
        Timothy Stephen     88063.0
```

Question 2: Item-Wise - To find total Units

Solution:

Question 1

```
# Import necessary libraries
import pandas as pd
import numpy as np

# Read the sales data from a CSV file into a DataFrame
df = pd.read_csv("salesdatanew.csv")

# Print a message indicating the result being displayed
print("Total units Item-wise\n")

# Create a pivot table to aggregate the total units sold per item
# 'index' specifies that we want to group by the "Item" column
# 'values' specifies that we want to perform aggregation on the "Units" column
# 'aggfunc=np.sum' indicates that we are summing the "Units" for each unique "Item"
df1 = pd.pivot_table(df, index=["Item"], values=["Units"], aggfunc=np.sum)

# Print the resulting pivot table
print(df1)
```

Output:

```
Total units Item-wise

              Units
Item
Cell Phone      278
Desk             10
Home Theater    722
Television      716
Video Games     395
```

Question 3: Region-Wise, Item Wise - To find Sales Amt

Solution:

Question 2

```
# Import necessary libraries
import pandas as pd
import numpy as np

# Read the sales data from a CSV file into a DataFrame
df = pd.read_csv("salesdatanew.csv")

# Print a message indicating the result being displayed
print("Total Sales amount Region and Item-wise")

# Create a pivot table to aggregate the total sales amount (Sale_amt) per region and
item
# 'index' specifies that we want to group by both "Region" and "Item" columns
# 'values' specifies that we want to perform aggregation on the "Sale_amt" column
# 'aggfunc=np.sum' indicates that we are summing the "Sale_amt" for each unique
combination of "Region" and "Item"
df2 = pd.pivot_table(df, index=["Region", "Item"], values=["Sale_amt"],
aggfunc=np.sum)

# Print the resulting pivot table
print(df2)
```

Output:

```
Total Sales amount Region and Item-wise
                    Sale_amt
Region  Item
Central Cell Phone     6075.0
        Desk            875.0
        Home Theater  212000.0
        Television    596604.0
        Video Games    14215.5
East    Cell Phone     39375.0
        Home Theater  117000.0
        Television    155740.0
        Video Games     8892.0
West    Cell Phone     17100.0
        Desk            375.0
        Home Theater   32000.0
        Television    105424.0
```

Question 4: Manager-Wise - To Find Mean Sale Amt

Solution:

Question 4

```
# Import necessary libraries
import pandas as pd
import numpy as np

# Read the sales data from a CSV file into a DataFrame
df = pd.read_csv("salesdatanew.csv")

# Print a message indicating the result being displayed
print("Average Sales amount Manager-wise")

# Create a pivot table to calculate the average sales amount (Sale_amt) per manager
# 'index' specifies that we want to group by the "Manager" column
# 'values' specifies that we want to perform aggregation on the "Sale_amt" column
# 'aggfunc=np.mean' calculates the mean (average) sales amount for each manager
# '.round(2)' rounds the result to 2 decimal places for better readability
df3 = pd.pivot_table(df, index=["Manager"], values=["Sale_amt"],
aggfunc=np.mean).round(2)

# Print the resulting pivot table
print(df3)
```

Output:

```
Average Sales amount Manager-wise
          Sale_amt
Manager
Douglas   29882.00
Hermann   30425.71
Marth     14000.00
Martha    35268.69
Timothy   25446.44
Total Sales amount Region,Manager and Salesman-wise
```