Exercise 3 – Queries

File reading and DataFrame allocation:

```
import pandas as pd
path_file = "CarSalesDataForReports.xlsx"
df_Clients = pd.read_excel(path_file, sheet_name = "Clients")
df_Stock = pd.read_excel(path_file, sheet_name = "Stock")
df_InvoiceLines = pd.read_excel(path_file, sheet_name = "InvoiceLines")
df_Invoices = pd.read_excel(path_file, sheet_name = "Invoices")
df_Colors = pd.read_excel(path_file, sheet_name = "Colors")
```

1. **Create a query that returns the top 3 car brands most sold (i.e., having the most car brand sales) during first and third quarter of year 2015.**

Answer:

```
df_Quarter1_Quest1 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20150101) & (df_Invoices.InvoiceDateKey <= 20150331)]
df_Quarter3_Quest1 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20150701) & (df_Invoices.InvoiceDateKey <= 20150931)]

df_Buy_Cars = pd.merge(df_InvoiceLines , df_Stock,how="inner",on="StockID")
df_Buy_Cars_Quarter1 = pd.merge(df_Buy_Cars, df_Quarter1_Quest1,how="inner",on="InvoiceID")["Make"].value_counts().head(3).rename_axis('Make').reset_index(name='Amount')
df_Buy_Cars_Quarter3 = pd.merge(df_Buy_Cars, df_Quarter3_Quest1,how="inner",on="InvoiceID")["Make"].value_counts().head(3).rename_axis('Make').reset_index(name='Amount')


print("****  Car Brands Most Sold Quarter 1 For 2015 ****\n")
print(df_Buy_Cars_Quarter1)
print("\n")
print("****  Car Brands Most Sold Quarter  3 For 2015 ****\n")
print(df_Buy_Cars_Quarter3)
```

```
****  Car Brands Most Sold Quarter 1 For 2015 ****

          Make  Amount
0       Jaguar      16
1  Aston Martin      14
2   Rolls Royce       8


****  Car Brands Most Sold Quarter  3 For 2015 ****

          Make  Amount
0       Jaguar      14
1  Aston Martin      10
2   Rolls Royce       8
```

2. **Create a query that shows the top 3 most sold car colors (i.e., having the most car color sales) for each quarter for the years 2012,2013,2014,2015.**

Answer:

```
df_Quarter1_2012 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20120101) & (df_Invoices.InvoiceDateKey <= 20120331)]
df_Quarter2_2012 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20120401) & (df_Invoices.InvoiceDateKey <= 20120631)]
df_Quarter3_2012 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20120701) & (df_Invoices.InvoiceDateKey <= 20120931)]
df_Quarter4_2012 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20121001) & (df_Invoices.InvoiceDateKey <= 20121231)]

df_Quarter1_2013 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20130101) & (df_Invoices.InvoiceDateKey <= 20130331)]
df_Quarter2_2013 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20130401) & (df_Invoices.InvoiceDateKey <= 20130631)]
df_Quarter3_2013 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20130701) & (df_Invoices.InvoiceDateKey <= 20130931)]
df_Quarter4_2013 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20131001) & (df_Invoices.InvoiceDateKey <= 20131231)]

df_Quarter1_2014 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20140101) & (df_Invoices.InvoiceDateKey <= 20140331)]
df_Quarter2_2014 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20140401) & (df_Invoices.InvoiceDateKey <= 20140631)]
df_Quarter3_2014 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20140701) & (df_Invoices.InvoiceDateKey <= 20140931)]
df_Quarter4_2014 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20141001) & (df_Invoices.InvoiceDateKey <= 20141231)]

df_Quarter1_2015 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20150101) & (df_Invoices.InvoiceDateKey <= 20150331)]
df_Quarter2_2015 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20150401) & (df_Invoices.InvoiceDateKey <= 20150631)]
df_Quarter3_2015 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20150701) & (df_Invoices.InvoiceDateKey <= 20150931)]
df_Quarter4_2015 = df_Invoices[(df_Invoices.InvoiceDateKey >= 20151001) & (df_Invoices.InvoiceDateKey <= 20151231)]
```

```
print("**** Best selling colors 2012 ****\n")

df_Quarter1_2012_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter1_2012,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 1 2012 Best selling colors##\n")
print(df_Quarter1_2012_Colors)
print("\n")

df_Quarter2_2012_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter2_2012,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 2 2012 Best selling colors##\n")
print(df_Quarter2_2012_Colors)
print("\n")

df_Quarter3_2012_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter3_2012,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 3 2012 Best selling colors##\n")
print(df_Quarter3_2012_Colors)
print("\n")

df_Quarter4_2012_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter4_2012,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 4 2012 Best selling colors##\n")
print(df_Quarter4_2012_Colors)
print("\n")

print("**** Best selling colors 2013 ****\n")

df_Quarter1_2013_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter1_2013,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 1 2013 Best selling colors##\n")
print(df_Quarter1_2013_Colors)
print("\n")

df_Quarter2_2013_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter2_2013,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 2 2013 Best selling colors##\n")
print(df_Quarter2_2013_Colors)
print("\n")

df_Quarter3_2013_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter3_2013,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 3 2013 Best selling colors##\n")
print(df_Quarter3_2013_Colors)
print("\n")

df_Quarter4_2013_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter4_2013,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 4 2013 Best selling colors##\n")
print(df_Quarter4_2013_Colors)
print("\n")

print("**** Best selling colors 2014 ****\n")

df_Quarter1_2014_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter1_2014,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 1 2014 Best selling colors##\n")
print(df_Quarter1_2014_Colors)
print("\n")

df_Quarter2_2014_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter2_2014,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 2 2014 Best selling colors##\n")
print(df_Quarter2_2014_Colors)
print("\n")

df_Quarter3_2014_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter3_2014,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 3 2014 Best selling colors##\n")
print(df_Quarter3_2014_Colors)
print("\n")

df_Quarter4_2014_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter4_2014,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 4 2014 Best selling colors##\n")
print(df_Quarter4_2014_Colors)
print("\n")

print("**** Best selling colors 2015 ****\n")

df_Quarter1_2015_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter1_2015,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 1 2015 Best selling colors##\n")
print(df_Quarter1_2015_Colors)
print("\n")
```

```
df_Quarter2_2015_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter2_2015,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 2 2015 Best selling colors##\n")
print(df_Quarter2_2015_Colors)
print("\n")

df_Quarter3_2015_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter3_2015,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 3 2015 Best selling colors##\n")
print(df_Quarter3_2015_Colors)
print("\n")

df_Quarter4_2015_Colors = pd.merge(df_Colors, pd.merge(df_Buy_Cars, df_Quarter4_2015,how="inner",on="InvoiceID"),how="inner",on="ColorID")["Color"].value_counts().head(3).rename_axis('Color').reset_index(name='Amount')
print("##Quarter 4 2015 Best selling colors##\n")
print(df_Quarter4_2015_Colors)
print("\n")
print("\n")
```

```
                    Color  Amount
0         Canary Yellow       6
1                   Red       4
2  British Racing Green       2


##Quarter 3 2014 Best selling colors##

           Color  Amount
0  Canary Yellow       4
1           Blue       4
2            Red       2


##Quarter 4 2014 Best selling colors##

                    Color  Amount
0         Canary Yellow       9
1                Silver       6
2  British Racing Green       2


**** Best selling colors 2015 ****

##Quarter 1 2015 Best selling colors##

                 Color  Amount
0  British Racing Green      10
1               Silver       8
2          Dark Purple       8


##Quarter 2 2015 Best selling colors##

           Color  Amount
0            Red      22
1  Canary Yellow       8
2          Black       8


##Quarter 3 2015 Best selling colors##

           Color  Amount
0  Canary Yellow      10
1         Silver       8
2    Dark Purple       8


##Quarter 4 2015 Best selling colors##

    Color  Amount
0     Red      14
1  Silver      10
2    Blue      10
```

✓ 0 s se ejecutó 17:10