Aim:- To perform OLAP operations on a data warehouse to analyze sales data stored in Firestore.

Software Used: IDLE

Theory :

Code-

Use service key from Practical 6 for extraction and analysisCode [serviceAccountKey.json file]

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from google.cloud import firestore
import os

# Set the path to your service account key file
os.environ["GOOGLE_APPLICATION_CREDENTIALS"] = "serviceAccountKey.json"

# Initialize Firestore DB
db = firestore.Client()

# Fetch data from Firestore
sales_ref = db.collection('sales')
docs = sales_ref.stream()

# Convert Firestore documents to DataFrame
data = []
for doc in docs:
```

```python
    data.append(doc.to_dict())
df = pd.DataFrame(data)

# Roll-Up: Aggregate data by product
df_rollup = df.groupby('Product').agg({'Quantity': 'sum', 'Price': 'sum'}).reset_index()
df_rollup['Total_Sales'] = df_rollup['Quantity'] * df_rollup['Price']
print("Roll-Up (Aggregate by Product):")
print(df_rollup)

# Drill-Down: Break down data by region
df_drilldown = df.groupby(['Region', 'Product']).agg({'Quantity': 'sum', 'Price': 'sum'}).reset_index()
df_drilldown['Total_Sales'] = df_drilldown['Quantity'] * df_drilldown['Price']
print("Drill-Down (Break Down by Region):")
print(df_drilldown)

# Slice: Analyze data for a specific product (e.g., 'Widget')
df_slice = df[df['Product'] == 'Widget']
print("Slice (Data for Product 'Widget'):")
print(df_slice)

# Dice: Analyze a sub-cube (e.g., products 'Widget' and 'Gadget' in 'Region1')
df_dice = df[(df['Region'] == 'Region1') & (df['Product'].isin(['Widget', 'Gadget']))]
print("Dice (Products 'Widget' and 'Gadget' in 'Region1'):")
print(df_dice)

# Visualization
plt.figure(figsize=(12, 6))

# Total Sales by Product (Roll-Up)
plt.subplot(1, 2, 1)
sns.barplot(x='Product', y='Total_Sales', data=df_rollup)
plt.title('Total Sales by Product')

# Total Sales by Region and Product (Drill-Down)
plt.subplot(1, 2, 2)
sns.barplot(x='Product', y='Total_Sales', hue='Region', data=df_drilldown)
plt.title('Total Sales by Region and Product')

plt.tight_layout()
plt.show()
```

Output-

```
Roll-Up (Aggregate by Product):
   Product  Quantity  Price  Total_Sales
0  Gadget          8  100.0        800.0
1   Gizmo          4   75.0        300.0
2  Widget         10   75.0        750.0
Drill-Down (Break Down by Region):
    Region  Product  Quantity  Price  Total_Sales
0  Region1   Gadget         7   50.0        350.0
1  Region1   Widget         8   50.0        400.0
2  Region2   Gadget         1   50.0         50.0
3  Region2    Gizmo         4   75.0        300.0
4  Region2   Widget         2   25.0         50.0
Slice (Data for Product 'Widget'):
   Product  OrderID  Price   Region  Quantity        Date
0  Widget         1   25.0  Region1         5  2024-01-01
1  Widget         2   25.0  Region1         3  2024-01-03
3  Widget         4   25.0  Region2         2  2024-01-02
Dice (Products 'Widget' and 'Gadget' in 'Region1'):
   Product  OrderID  Price   Region  Quantity        Date
0  Widget         1   25.0  Region1         5  2024-01-01
1  Widget         2   25.0  Region1         3  2024-01-03
2  Gadget         3   50.0  Region1         7  2024-01-04
```


Total Sales by Product

Total Sales by Region and Product