Applied A.I. Solutions
Foundations of Data Management

Lab Exercises

Group-10 members

1. Goyal, Vinayak
2. Bhasgauri, Harshal Shashikant
3. Sebastian, Arun
4. ., Himani
5. Singh, Satyajeet
6. Trongkitroongruang, Kajhonprom
7. Cheng, Qianfan

Operational Sales Report: Analysis of Regional in the U.S. in January 2021

Period	Region	State	City	Category	Total Orders	Total Returns	Total Sales	Sales KPI	Total Profits	Total COGS
2021-01	Central	Illinois	Aurora	Furniture	1	0	69.375	69	-47.18	116.555
				Office Supplies	3	0	309.6	103	96.56	213.04
				Technology	1	0	2003.168	2003	250.4	1752.768
			Chicago	Office Supplies	4	0	46.64	12	-10.64	57.28
		Indiana	Richmond	Furniture	1	0	18.96	19	8.53	10.43
				Office Supplies	4	0	100.02	25	28.43	71.59
				Technology	3	0	239.72	80	74.22	165.5
		Iowa	Des Moines	Furniture	1	0	34.58	35	14.52	20.06
				Office Supplies	4	0	98.64	25	39.67	58.97
				Technology	1	0	207	207	51.75	155.25
			Marion	Furniture	1	0	14.91	15	4.62	10.29
				Office Supplies	3	0	121.29	40	55.25	66.04
		Kansas	Wichita	Office Supplies	1	0	279.9	L	137.15	142.75
		Michigan	Detroit	Furniture	1	0	210.98	211	21.1	189.88
				Technology	1	0	3059.982	3060	680	2379.982
			Jackson	Furniture	1	0	302.67	303	72.64	230.03
				Office Supplies	4	0	5603.95	1401	2578.58	3025.37
				Technology	3	0	1135.913	379	313.18	822.733
		Missouri	Springfield	Furniture	1	0	212.94	213	53.24	159.7
				Office Supplies	3	0	4406.39	1469	153.41	4252.98
		Texas	Austin	Office Supplies	2	0	37.06	19	-59.08	96.14
				Technology	2	0	110.576	55	27.7	82.876
			Dallas	Office Supplies	1	0	760.98	761	-1141.47	1902.45
			El Paso	Furniture	1	0	913.43	913	-169.64	1083.07
				Office Supplies	3	0	409.32	136	18.3	391.02
			Houston	Office Supplies	1	0	18.16	18	1.82	16.34
			Huntsville	Furniture	2	0	452.164	226	-214.02	666.184
				Office Supplies	5	0	502.766	101	-176.94	679.706
			Keller	Office Supplies	1	0	6	6	2.1	3.9
		Wisconsin	Franklin	Office Supplies	1	0	3.6	4	1.73	1.87
	East	Connecticut	Waterbury	Office Supplies	1	0	3.52	4	1.02	2.5
		District of Colum	Washington	Furniture	1	0	37.68	38	15.83	21.85
				Office Supplies	1	0	40.08	40	19.24	20.84
		Massachusetts	Quincy	Office Supplies	1	0	12.7	13	5.84	6.86
		New York	New York City	Furniture	1	0	207.846	208	2.31	205.536
				Office Supplies	1	0	5.22	5	2.4	2.82
				Technology	2	0	587.85	294	193.33	394.52
		Ohio	Kent	Office Supplies	1	0	14.016	14	1.75	12.266
				Technology	2	0	179.958	90	-36	215.958
			Lorain	Furniture	1	0	48.896	49	8.56	40.336
		Pennsylvania	Philadelphia	Furniture	2	0	919.239	460	-56.99	976.229
				Office Supplies	5	0	2259.49	452	95.34	2164.15
				Technology	1	0	429.6	430	-93.08	522.68
		Vermont	Burlington	Office Supplies	4	0	637.18	159	196.31	440.87
	South	Alabama	Hoover	Office Supplies	2	0	22.63	11	7.62	15.01
			Tuscaloosa	Office Supplies	1	0	33.74	34	15.52	18.22
		Florida	Miami	Furniture	1	0	419.136	419	-68.11	487.246
			Ormond Beach	Office Supplies	1	0	2.808	3	-1.97	4.778
		Georgia	Columbus	Furniture	1	0	62.72	63	24.46	38.26
				Technology	1	0	2939.93	2940	764.38	2175.55
			Smyrna	Office Supplies	1	0	5.67	6	0.11	5.56
		North Carolina	Charlotte	Office Supplies	2	0	383.992	192	1.84	382.152
			Jacksonville	Office Supplies	3	0	66.258	22	-26.35	92.608
				Technology	2	0	703.62	352	-27.14	730.76
		Tennessee	Johnson City	Office Supplies	4	4	63.102	16	12.93	50.172
				Technology	1	1	111.984	112	7	104.984
	West	Arizona	Tucson	Office Supplies	1	0	4.938	5	-3.62	8.558
				Technology	1	0	95.984	96	12	83.984
		California	Costa Mesa	Furniture	1	1	37.74	38	12.83	24.91
				Technology	1	1	239.97	240	26.4	213.57
			Long Beach	Office Supplies	4	0	332.8	83	80.69	252.11
			Los Angeles	Furniture	3	0	902.022	301	192.37	709.652
				Office Supplies	2	0	176.3	88	75.92	100.38
				Technology	2	0	177.366	89	15.86	161.506
			Rancho Cucamonga	Office Supplies	1	0	38.88	39	18.66	20.22
			San Francisco	Furniture	1	1	120.784	121	-13.59	134.374
				Office Supplies	11	2	2364.598	215	763.52	1601.078
				Technology	1	0	359.976	360	130.49	229.486
			San Jose	Office Supplies	2	0	302.644	151	96.67	205.974
				Technology	1	0	110.352	110	8.28	102.072
		Colorado	Aurora	Office Supplies	1	0	168.624	169	14.75	153.874
				Technology	1	0	169.064	169	-14.79	183.854
		Montana	Great Falls	Office Supplies	3	0	1189.43	396	75.57	1113.86
				Technology	1	0	2999.95	3000	1379.98	1619.97
		Washington	Seattle	Furniture	2	0	977.96	489	99.07	878.89
				Office Supplies	7	0	441.372	63	81.79	359.582
				Technology	3	0	871.09	290	155.51	715.58

This research focuses on a detailed analysis of sales throughout four key regions in the United States: Central, East, South, and West. We will investigate the fundamental challenges in each operational city, evaluating sales and return rates for 2021.

Using the findings from our analysis, it is clear that the South region, notably Johnson City, Tennessee, is a major source of concern. The product category "office supply" receives 4 returns, which corresponds exactly to the total number of faulty orders.

Upon dissecting the profit-loss metrics, challenges emerge across all regions:

1. Central: Illinois

2. East: Ohio and Pennsylvania

3. South: Florida and North Carolina

4. West: Arizona, California, and Colorado.

The most prevalent product category associated with these losses is **Office Supply**. Addressing these issues has become critical for increasing our profitability.

KPI Performance

We set a 10% growth goal for each year, which means sales of current year need be greater than the sales of previous year times 1.1.

KPI performance: Sales(current year) >= Sales(previous year) * (1 + 10%)

Sales = Unit Price of Product * quantity * (1 – discount)

COGS = Sales – ProfitsSales KPI = sales / orders

Our operational analytics suggest a return rate of roughly 6.45% 6.45% for the given month, implying 6 to 7 returns for every 100 purchases. This stresses the need of scrutinizing sectors such as the South region, particularly the Office Supply category. Our profit margin is roughly 16.24 % 16.24%, which means that for every $1 earned in sales, we net approximately $0.162 in profit. The average transaction value, as represented by the Sales KPI, is roughly $326.26, serving as a baseline for our average order value across all sectors and geographies.

Potential Future Directions

Logistics Redesign: Our logistics might use a new coat of paint. Consider route optimization, cost-cutting, and on-time delivery. Modern logistical tools could be the solution.

Improving Our Quality: The frequent "Office Supply" bug screams poor quality. It's past time we evaluated our quality control procedures, scrutinized our suppliers, and increased product inspections.

Warehouse Work: Time is of the essence, and our warehouses cannot afford to be late. We're talking about high-quality warehouse systems, process automation, and regular employee upskilling.

Solutions that can be implemented include improving our logistical operations, ensuring on-time delivery, increasing order correctness, and improving the efficiency of our warehouses and shipping protocols.

Executive Report: Analysis in the U.S. 2021 vs 2020

Region	Sales 2020 ($)	Sales 2021 ($)	Sales 2021 vs 2020 (%)	KPI Performance	Profit 2021 ($)	Profit 2021 vs 2020 (%)	COGS 2020 (%)	COGS 2021 (%)
Central	147426.9134	147100.5877	-0.22%	Below Target	7550.78	-62.05%	86.50%	94.87%
East	180673.1328	213083.3194	17.94%	Above Target	33230.46	64.99%	88.85%	84.40%
South	93618.3163	122905.197	31.28%	Above Target	8848.89	-50.01%	81.09%	92.80%
West	187479.5583	250118.9997	33.41%	Above Target	43809.04	82.15%	87.17%	82.48%

This research provides a comparative examination of sales in four key regions of the United States: Central, East, South, and West. To determine the performance trend, we compared sales, profit, and COGS from 2021 to statistics from 2020.

1. **Central Region**

While sales fell moderately by 0.22% from 2020, profit fell by a more significant 62.05%. The cost of goods sold (COGS) grew from 86.5% in 2020 to 94.87% in 2021. This region's performance is categorized as **"Below Target."**

2. **East Region**

Sales increased by 17.94% in 2021 over 2020. Profit increased significantly by 64.99%. COGS declined from 88.85% in 2020 to 84.4% in 2021, suggesting improved cost management and trend maintenance. This region's performance is rated **"Above Target."**

3. **South Region**

Sales in the South region increased by 31.28% from 2020. Profit, on the other hand, fell by 50.01%. The cost of goods sold increased from 81.09% in 2020 to 92.8% in 2021. Despite the drop in profits, the region's performance is rated **"Above Target."**

4. **West Region**

The West region saw the greatest increase in sales, with a 33.41% gain. Profits increased by 82.15% as well. COGS has decreased from 87.17% in 2020 to 82.48% in 2021. The performance of this region is also classified as **"Above Target."**

In Summary, while sales have increased across the board, profitability in the Central region is a source of worry. Except for the Central and South, most regions were able to maintain or cut their COGS. Strategic measures are required to assure long-term growth and profitability, particularly in regions that are falling short of their targets.

Appendix I

```
import csv
import decimal
import os
from sqlalchemy.orm import Session, aliased
import sqlalchemy as sa

from src.constants import ROOT_DIR
from src.database import engine
from src.models import mapped_models as mm

engine = engine.sql_engine()
er_filename = os.path.join(ROOT_DIR, 'data', 'executive_report.csv')
opr_filename = os.path.join(ROOT_DIR, 'data', 'operational_report.csv')

class ReportBase():
    engine = None

    def __init__(self, engine):
        self.engine = engine


    def gen_report(self, data=[], headers=[], csv_file=''):
        # writing to csv file
        with open(csv_file, 'w') as csvfile:
            # creating a csv writer object
            csvwriter = csv.writer(csvfile)
            csvwriter.writerow(headers)
            # writing the data rows
            csvwriter.writerows(data)

class ExecutiveReport(ReportBase):
    '''
    generate executive report
    '''
    def query_orders(self, session, start_date, end_date):
        return (session.query(sa.extract('year', mm.Order.order_date).label('year'),
                        mm.Region.name.label('region'),
                        sa.func.sum(mm.ProductOrder.order_price * mm.ProductOrder.quantity * (1 -
mm.ProductOrder.order_discount)).label('order_sales'),
                        sa.func.sum(mm.ProductOrder.order_profit).label('order_profits'),
                        sa.func.sum(mm.ProductOrder.order_price * mm.ProductOrder.quantity * (1 -
mm.ProductOrder.order_discount) - mm.ProductOrder.order_profit).label('order_cogs'),
                        )
                .select_from(mm.ProductOrder)
                .join(mm.Order, mm.ProductOrder.order_id == mm.Order.id)
```

```
            .join(mm.Customer, mm.Customer.id == mm.Order.customer_id)

            # .join(mm.AddressCustomer, mm.Customer.id == mm.AddressCustomer.customer_id)
            .join(mm.Address, mm.Order.address_id == mm.Address.id)
            .join(mm.City, mm.Address.city_id == mm.City.id)
            .join(mm.State, mm.City.state_id == mm.State.id)
            .join(mm.Region, mm.Region.id == mm.State.region_id)
            # .join(mm.Product, mm.Product.id == mm.ProductOrder.product_id, isouter=True)
            .filter(mm.Order.order_date.between(start_date, end_date))
            .group_by(sa.extract('year', mm.Order.order_date),
                      mm.Region.name,
                      )
            .order_by('year', 'region')
            .all())



class OperationalReport(ReportBase):
    '''
    generate operational report
    '''
    def query_orders(self, session, start_date, end_date):
        category = aliased(mm.Category)
        category_p = aliased(mm.Category)
        return (session.query(
                       sa.func.date_format(mm.Order.order_date, '%Y-%m').label('year_and_month'),
                      #   mm.ProductOrder,
                      mm.Region.name.label('region'),
                      mm.State.name.label('state'),
                      mm.City.name.label('city'),
                      category_p.name.label('category'),
                      sa.func.count(mm.Order.id).label('sum_orders'),
                      sa.func.sum(sa.case (
                         (mm.Order.status_id == 2, 1),
                         else_=0
                      ) ).label('sum_returned'),
                      sa.func.sum(mm.ProductOrder.order_price * mm.ProductOrder.quantity * (1 -
mm.ProductOrder.order_discount)).label('sum_sales'),
                      sa.func.sum(mm.ProductOrder.order_profit).label('sum_profits'),
                      sa.func.sum(mm.ProductOrder.order_price * mm.ProductOrder.quantity * (1 -
mm.ProductOrder.order_discount) - mm.ProductOrder.order_profit).label('sum_cogs'),
                      )
            .select_from(mm.ProductOrder)
            .join(mm.Order, mm.ProductOrder.order_id == mm.Order.id)
            .join(mm.Product, mm.ProductOrder.product_id == mm.Product.id)
            .join(category, category.id == mm.Product.category_id)
            .join(category_p, category_p.id == category.parent_id)
            .join(mm.Customer, mm.Customer.id == mm.Order.customer_id)
```

```
            .join(mm.Address, mm.Order.address_id == mm.Address.id)
            .join(mm.City, mm.Address.city_id == mm.City.id)
            .join(mm.State, mm.City.state_id == mm.State.id)
            .join(mm.Region, mm.Region.id == mm.State.region_id)

            # .join(mm.Product, mm.Product.id == mm.ProductOrder.product_id, isouter=True)
            .filter(mm.Order.order_date.between(start_date, end_date))
            .group_by(sa.func.date_format(mm.Order.order_date, '%Y-%m'),
                      mm.Region.name,
                      mm.State.name,
                      mm.City.name,
                      category_p.name,
                      )
            .order_by('year_and_month', 'region', 'state', 'city', 'category')
            .all()
        )

if __name__ == '__main__':
    kpi_performance = 0.1 # 10% increase of sales yearly
    er = ExecutiveReport(engine)
    er_data = []
    er_headers = ['Region', 'Sales 2020 ($)', 'Sales 2021 ($)', 'Sales 2021 vs 2020 (%)', 'KPI Performance',
            'Profit 2021 ($)', 'Profit 2021 vs 2020 (%)',
            'COGS 2020 (%)', 'COGS 2021 (%)']
    with Session(bind=engine) as session:
        q_2020 = er.query_orders(session=session, start_date='2020-01-01', end_date='2020-12-31')
        q_2021 = er.query_orders(session=session, start_date='2021-01-01', end_date='2021-12-31')
    for i in range(len(q_2021)):
        er_data.append((q_2021[i][1],
                q_2020[i][2],
                q_2021[i][2],
                f'{round(((q_2021[i][2] - q_2020[i][2])/q_2020[i][2]) * 100, 2)}%',
                'Above Target' if q_2021[i][2] >= (q_2020[i][2] * (1 + decimal.Decimal(kpi_performance))) else 'Below Target',
                q_2021[i][3],
                f'{round(((q_2021[i][3] - q_2020[i][3])/q_2020[i][3]) * 100, 2)}%',
               #  q_2021[i][2] - q_2021[i][3],
               #  f'{round(((q_2021[i][2] - q_2021[i][3] - (q_2020[i][2] - q_2020[i][3]))/(q_2020[i][2] - q_2020[i][3])) * 100, 2)}%',
               f'{round(((q_2020[i][2] - q_2020[i][3])/q_2020[i][2]) * 100, 2)}%',
               f'{round(((q_2021[i][2] - q_2021[i][3])/q_2021[i][2]) * 100, 2)}%',
                ))
    # generate executive report
    er.gen_report(er_data, er_headers, er_filename)

    opr = OperationalReport(engine)
    opr_headers = ['Period', 'Region', 'State', 'City', 'Category',
            'Total Orders', 'Total Returns',
            'Total Sales', 'Total Profits',
            'Total COGS']
```

```
with Session(bind=engine) as session:
    q = opr.query_orders(session=session, start_date='2021-01-01', end_date='2021-03-31')
# generate operational report
opr.gen_report(q, opr_headers, opr_filename)
```