Assignment 2

# PART A – Coffee Sales

**A. Just for Starters – SQL Questions**

A1.

1. **CREATE** **TABLE** TempA1 **AS**
2. **SELECT** StateName, AVG(AreaSales) **as** StateAvg
3. **FROM** (
4. **SELECT** S.StateName, A.AreaID, SUM(F.ActSales) **as** AreaSales
5. **FROM** States S, AreaCode A, FactCoffee F
6. **WHERE** S.StateID = A.StateID AND A.AreaID = F.AreaID AND EXTRACT(year **from** F.FactDate) = 2013
7. **GROUP** **BY** S.StateName, A.AreaID
8. )
9. **GROUP** **BY** StateName;
11. **SELECT** A.AreaID, S.StateName, ROUND(T.StateAvg,2) **as** StateAvg, SUM(F.ActSales) **as** AreaSales
12. **FROM** AreaCode A, States S, FactCoffee F, TempA1 T
13. **WHERE** A.StateID = S.StateID AND A.AreaID = F.AreaID AND S.StateName = T.StateName
14. **GROUP** **BY** A.AreaID, S.StateName, StateAvg, ROUND(T.StateAvg,2)
15. **HAVING** SUM(F.ActSales) > 1.1\*T.StateAvg
16. **ORDER** **BY** S.StateName;

|  |  |  |  |
| --- | --- | --- | --- |
| AREAID | STATENAME | STATEAVG | AREASALES |
| 209 | California | 1901.92 | 5543 |
| 213 | California | 1901.92 | 6741 |
| 310 | California | 1901.92 | 3823 |
| 323 | California | 1901.92 | 2754 |
| 415 | California | 1901.92 | 6148 |
| 530 | California | 1901.92 | 5780 |
| 559 | California | 1901.92 | 2104 |
| 562 | California | 1901.92 | 6106 |
| 626 | California | 1901.92 | 3412 |
| 650 | California | 1901.92 | 6616 |

The table shows the first 10 records from the result (area codes with sales more than 10% the average sales of all area codes within that state for the year 2013).

A2.

1. **SELECT** P.ProdName, SUM(F.ActSales) **as** TotalSales, ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)),2) **as** ProfMargPerc
2. **FROM** ProdCoffee P, FactCoffee F
3. **WHERE** P.ProductID = F.ProductID
4. **GROUP** **BY** P.ProdName
5. **HAVING** ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)),2) > 15
6. **ORDER** **BY** TotalSales **DESC**;

|  |  |  |
| --- | --- | --- |
| PRODNAME | TOTALSALES | PROFMARGPERC |
| Lemon | 95926 | 31.14 |
| Colombian | 128311 | 43.49 |
| Caffe Mocha | 84904 | 20.82 |
| Decaf Espresso | 78162 | 37.74 |
| Chamomile | 75578 | 36.03 |
| Darjeeling | 73151 | 39.72 |
| Earl Grey | 66772 | 36.19 |
| Decaf Irish Cream | 62248 | 22.47 |
| Caffe Latte | 35899 | 31.69 |
| Mint | 35710 | 17.23 |

The table shows the first 10 records from the result (products with profit margins as percentage of sales of at least 15%).

A3.

1. **SELECT** \* **FROM** (
2. **SELECT** A.AreaID, P.ProdLine, SUM(F.ActProfit) **as** AreaProfit
3. **FROM** AreaCode A, ProdCoffee P, FactCoffee F
4. **WHERE** A.AreaID = F.AreaID AND P.ProductID = F.ProductID AND EXTRACT(year **from** F.FactDate) = 2012
5. **GROUP** **BY** A.AreaID, P.ProdLine
6. )
7. PIVOT (
8. SUM(AreaProfit)
9. **FOR** ProdLine IN ('Beans' **as** BeansProf, 'Leaves' **as** LeavesProf)
10. )
11. **WHERE** LeavesProf > 2\*BeansProf;

|  |  |  |
| --- | --- | --- |
| AREAID | BEANSPROF | LEAVESPROF |
| 619 | -56 | 147 |
| 607 | -208 | 563 |
| 712 | 241 | 1818 |
| 330 | 51 | 227 |
| 518 | -284 | 889 |
| 661 | 140 | 440 |
| 937 | 95 | 315 |
| 315 | 75 | 267 |
| 505 | 93 | 237 |
| 347 | -232 | 261 |

The table shows the first 10 records from the result (AreaIDs where total profit from leaves in 2012 are two times greater than that from beans.

**B. DECLINING PROFITS:**

B1.

1. **CREATE** **TABLE** TempAB **AS**
2. **SELECT** \* **FROM** (
3. **SELECT** AreaID, StateName, Y2012, Y2013, Y2012-Y2013 **as** ProfDec
4. **FROM**(
5. **SELECT** \* **FROM**(
6. **SELECT** A.AreaID, S.StateName, EXTRACT(year **from** F.FactDate) **as** ProfYear, SUM(F.ActProfit) **As** AreaProf
7. **FROM** AreaCode A, FactCoffee F, States S
8. **WHERE** A.AreaID = F.AreaID and S.StateID = A.StateID
9. **GROUP** **BY** A.AreaID, S.StateName, EXTRACT(year **from** F.FactDate)
10. )
11. PIVOT (
12. Sum(AreaProf)
13. **FOR** ProfYear IN (2012 **as** Y2012, 2013 **as** Y2013)
14. ))
15. **WHERE** 100\*(Y2013-Y2012)/abs(Y2012) <> 0
16. **ORDER** **BY** ProfDec **DESC**)
17. **WHERE** ROWNUM <= 5;

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| AREAID | STATENAME | Y2012 | Y2013 | PROFDEC |
| 845 | New York | 1023 | -1059 | 2082 |
| 508 | Massachusetts | 1499 | 368 | 1131 |
| 626 | California | 953 | -111 | 1064 |
| 712 | Iowa | 2059 | 1020 | 1039 |
| 631 | New York | 1559 | 749 | 810 |

The table shows the entire result (top 5 area codes with declining profits and how much the profits declined)

B2.

1. **SELECT** AreaID, StateName, ProdName, Y2012 - Y2013 **as** ProfDec
2. **FROM** (
3. **SELECT** \* **FROM** (
4. **SELECT** T.AreaID, T.StateName, P.ProdName, EXTRACT(year **from** F.FactDate) **as** Year, SUM(F.ActProfit) **as** ProdProfit
5. **FROM** TempAB T, ProdCoffee P, FactCoffee F
6. **WHERE** T.AreaID = F.AreaID AND P.ProductID = F.ProductID
7. **GROUP** **BY** T.AreaID, T.StateName, P.ProdName, EXTRACT(year **from** F.FactDate)
8. )
9. PIVOT (
10. SUM(ProdProfit)
11. **FOR** Year IN (2012 **as** Y2012, 2013 **as** Y2013)
12. )
13. )
14. **WHERE** Y2012 - Y2013 > 0
15. **ORDER** **BY** AreaID, ProfDec **DESC**;

|  |  |  |  |
| --- | --- | --- | --- |
| AREAID | STATENAME | PRODNAME | PROFDEC |
| 508 | Massachusetts | Regular Espresso | 24 |
| 508 | Massachusetts | Lemon | 9 |
| 508 | Massachusetts | Darjeeling | 1 |
| 626 | California | Decaf Irish Cream | 51 |
| 631 | New York | Regular Espresso | 242 |
| 712 | Iowa | Decaf Irish Cream | 31 |
| 845 | New York | Caffe Mocha | 765 |
| 845 | New York | Mint | 418 |

The table shows all products with declining profits in the area codes from B1 along with how much their profits have declined. For AreaID 508, Regular Espresso had significantly higher profit decline than the other products. For AreaID 626, Decaf Irish Cream was the only product with a profit decline. For AreaID 631, Regular Espresso was the only product with a profit decline. For AreaID 712, Decaf Irish Cream was the only product with a profit decline. For AreaID 845, Caffe Mocha had a significantly higher profit decline than other products, but Mint also saw a significant decline.

**C. BUDGETED Numbers:**

C1.

1. **SELECT** \* **FROM** (
2. **SELECT** S.StateName, SUM(F.ActProfit) - SUM(F.BudProfit) **as** ProfDiff, SUM(F.ActSales) - SUM(F.BudSales) **as** SalesDiff
3. **FROM** States S, FactCoffee F, AreaCode A
4. **WHERE** S.StateID = A.StateID AND A.AreaID = F.AreaID
5. **GROUP** **BY** S.StateName
6. **HAVING** SUM(F.ActProfit) - SUM(BudProfit) > 0 AND SUM(F.ActSales) - SUM(F.BudSales) > 0
7. **ORDER** **BY** ProfDiff **DESC**)
8. **WHERE** ROWNUM <= 5;

|  |  |  |
| --- | --- | --- |
| STATENAME | PROFDIFF | SALESDIFF |
| Iowa | 3872 | 7750 |
| California | 3405 | 6212 |
| Massachusetts | 1842 | 2965 |
| New York | 1656 | 7272 |
| Florida | 510 | 4843 |

The table shows the top 5 states for the year 2012 that have substantially higher actual numbers relative to budgeted numbers for profits and sales.

C2.

1. **SELECT** T.StateName, A.AreaID, SUM(F.BudProfit) **as** BudProfit, SUM(F.ActProfit) - SUM(F.BudProfit) **as** ProfDiff,
2. SUM(F.BudSales) **as** BudSales, SUM(F.ActSales) - SUM(F.BudSales) **as** SalesDiff,
3. ROUND(100\*(SUM(F.ActProfit) - SUM(F.BudProfit))/SUM(F.BudProfit), 2) **as** ProfDiffPerc,
4. ROUND(100\*(SUM(F.ActSales) - SUM(F.BudSales))/SUM(F.BudSales), 2) **as** SalesDiffPerc
5. **FROM** TempAC T, AreaCode A, FactCoffee F, States S
6. **WHERE** T.StateName = S.StateName AND S.StateID = A.StateID AND A.AreaID = F.AreaID
7. **GROUP** **BY** T.StateName, A.AreaID
8. **HAVING** 100\*(SUM(F.ActProfit) - SUM(F.BudProfit))/SUM(F.BudProfit) > 20
9. AND 100\*(SUM(F.ActSales) - SUM(F.BudSales))/SUM(F.BudSales) > 15
10. **ORDER** **BY** StateName;

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| STATENAME | AREAID | BUDPROFIT | PROFDIFF | BUDSALES | SALESDIFF | PROFDIFFPERC | SALESDIFFPERC |
| California | 310 | 1090 | 396 | 3120 | 703 | 36.33 | 22.53 |
| California | 323 | 760 | 365 | 2330 | 424 | 48.03 | 18.2 |
| California | 408 | 320 | 65 | 1340 | 207 | 20.31 | 15.45 |
| California | 510 | 620 | 256 | 1490 | 382 | 41.29 | 25.64 |
| California | 619 | 640 | 233 | 1560 | 393 | 36.41 | 25.19 |
| California | 760 | 730 | 198 | 1670 | 486 | 27.12 | 29.1 |
| Iowa | 515 | 3860 | 926 | 10200 | 1613 | 23.99 | 15.81 |
| Iowa | 563 | 4490 | 1196 | 11390 | 2036 | 26.64 | 17.88 |
| Iowa | 641 | 3810 | 1051 | 10100 | 1894 | 27.59 | 18.75 |
| Massachusetts | 857 | 1620 | 554 | 2800 | 490 | 34.2 | 17.5 |

The table shows the first 10 records (out of 12) from the result. The 12 AreaIDs in the result are the area codes within the 5 states from C1 that beat budgeted sales and profits significantly. I chose to define significant as actual profit exceeding budget by at least 20% and actual sales exceeding budget by at least 15%.

**D. PRODUCT Related:**

D1.

1. **SELECT** StateMkt, ProdName, ProfInc, MktRank
2. **FROM** (
3. **SELECT** StateMkt, ProdName, Y2013 - Y2012 **as** ProfInc, RANK() OVER (PARTITION **BY** StateMkt
4. **ORDER** **BY** Y2013 - Y2012 **DESC**) **as** MktRank
5. **FROM** (
6. **SELECT** S.StateMkt, P.ProdName, EXTRACT(year **from** F.FactDate) **as** Year, SUM(F.ActProfit) **AS** Profit
7. **FROM** States S, ProdCoffee P, FactCoffee F, AreaCode A
8. **WHERE** S.StateID = A.StateID AND P.ProductID = F.ProductID AND A.AreaID = F.AreaID
9. **GROUP** **BY** S.StateMkt, P.ProdName, EXTRACT(YEAR **FROM** F.FactDate)
10. )
11. PIVOT (
12. SUM(Profit)
13. **FOR** Year in (2012 **as** Y2012, 2013 **as** Y2013)
14. )
15. )
16. **WHERE** MktRank <= 3;

|  |  |  |  |
| --- | --- | --- | --- |
| STATEMKT | PRODNAME | PROFINC | MKTRANK |
| Central | Caffe Mocha | 2704 | 1 |
| Central | Chamomile | 2677 | 2 |
| Central | Darjeeling | 1979 | 3 |
| East | Colombian | 4998 | 1 |
| East | Regular Espresso | 1881 | 2 |
| East | Lemon | 1436 | 3 |
| South | Colombian | 1615 | 1 |
| South | Decaf Espresso | 1088 | 2 |
| South | Caffe Mocha | 970 | 3 |
| West | Lemon | 2413 | 1 |

The table shows the first 10 records (out of 12) from the result. The full result shows the top 3 products in each market with the greatest increase in profits.

D2.

1. **CREATE** **TABLE** TempAD **AS**
2. **SELECT** StateMkt, ProdType, SalesInc, MktRank
3. **FROM** (
4. **SELECT** StateMkt, ProdType, Y2013 - Y2012 **AS** SalesInc, RANK() OVER (PARTITION **BY** StateMkt
5. **ORDER** **BY** Y2013 - Y2012 **DESC**) **AS** MktRank
6. **FROM** (
7. **SELECT** S.StateMkt, P.ProdType, EXTRACT(YEAR **FROM** F.FactDate) **AS** Year, SUM(F.ActSales) **AS** Sales
8. **FROM** States S, ProdCoffee P, FactCoffee F, AreaCode A
9. **WHERE** S.StateID = A.StateID AND P.ProductID = F.ProductID AND A.AreaID = F.AreaID
10. **GROUP** **BY** S.StateMkt, P.ProdType, EXTRACT(YEAR **FROM** F.FactDate)
11. )
12. PIVOT (
13. SUM(Sales)
14. **FOR** Year IN (2012 **AS** Y2012, 2013 **AS** Y2013)
15. )
16. )
17. **WHERE** MktRank <=2;

|  |  |  |  |
| --- | --- | --- | --- |
| STATEMKT | PRODTYPE | SALESINC | MKTRANK |
| Central | Tea | 1474 | 1 |
| Central | Coffee | 1464 | 2 |
| East | Coffee | 1161 | 1 |
| East | Espresso | 1053 | 2 |
| South | Espresso | 974 | 1 |
| South | Coffee | 699 | 2 |
| West | Tea | 1544 | 1 |
| West | Espresso | 1508 | 2 |

The table shows the top 2 product types in each market that have the greatest increase in sales.

D3.

1. **SELECT** StateMkt, ProdType, ProdName, ProdSales2013 - ProdSales2012 **AS** ProdSalesInc
2. **FROM** (
3. **SELECT** T.StateMkt, T.ProdType, P.ProdName, EXTRACT(YEAR **FROM** F.FactDate) **AS** Year, SUM(F.ActSales) **AS** ProdSales
4. **FROM** TempAD T, ProdCoffee P, FactCoffee F, AreaCode A, States S
5. **WHERE** T.StateMkt = S.StateMkt AND T.ProdType = P.ProdType AND P.ProductID = F.ProductID AND
6. S.StateID = A.StateID AND A.AreaID = F.AreaID
7. **GROUP** **BY** T.StateMkt, T.ProdType, P.ProdName, EXTRACT(YEAR **FROM** F.FactDate)
8. )
9. PIVOT (
10. SUM(ProdSales)
11. **FOR** Year IN (2013 **AS** ProdSales2013, 2012 **AS** ProdSales2012)
12. )
13. **ORDER** **BY** StateMkt, ProdType, ProdSalesInc **DESC**;

|  |  |  |  |
| --- | --- | --- | --- |
| STATEMKT | PRODTYPE | PRODNAME | PRODSALESINC |
| Central | Coffee | Colombian | 605 |
| Central | Coffee | Decaf Irish Cream | 553 |
| Central | Coffee | Amaretto | 306 |
| Central | Tea | Earl Grey | 725 |
| Central | Tea | Darjeeling | 632 |
| Central | Tea | Green Tea | 117 |
| East | Coffee | Colombian | 975 |
| East | Coffee | Decaf Irish Cream | 126 |
| East | Coffee | Amaretto | 60 |
| East | Espresso | Regular Espresso | 531 |

The table shows the first 10 records from the result. The overall result shows increase in sales for each product within the product types D2. For all of the product types, there are one or two products that have a significantly greater increase in sales than the other products. For example, for the Coffee product type in the Central market, Colombian and Decaf Irish Cream have seen a significantly greater increase in sales than Amaretto, the third product.

**E. MARKETING EXPENSES (LOWEST):**

E1.

1. **SELECT** \* **FROM** (
2. **SELECT** S.StateName, ROUND(100\*(SUM(F.ActMarkCost)/SUM(F.ActSales)), 2) **AS** MktPercSales
3. **FROM** States S, FactCoffee F, AreaCode A
4. **WHERE** S.StateID = A.StateID AND F.AreaID = A.AreaID
5. **GROUP** **BY** S.StateName
6. **ORDER** **BY** MktPercSales)
7. **WHERE** ROWNUM <= 5;

|  |  |
| --- | --- |
| STATENAME | MKTPERCSALES |
| Massachusetts | 11.44 |
| Texas | 12.69 |
| Illinois | 12.82 |
| Iowa | 13.87 |
| Colorado | 14.13 |

The table shows the top 5 states with the lowest market expenses as a percentage of their sales.

E2.

1. **CREATE** **TABLE** TempAE **AS**
2. **SELECT** \* **FROM** (
3. **SELECT** S.StateName, ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)), 2) **AS** ProfPercSales
4. **FROM** States S, FactCoffee F, AreaCode A
5. **WHERE** S.StateID = A.StateID AND F.AreaID = A.AreaID
6. **GROUP** **BY** S.StateName
7. **ORDER** **BY** ProfPercSales **DESC**)
8. **WHERE** ROWNUM <= 5;

|  |  |
| --- | --- |
| STATENAME | PROFPERCSALES |
| Massachusetts | 54.87 |
| Illinois | 44.1 |
| Texas | 42.14 |
| Iowa | 40.57 |
| Colorado | 36.83 |

The table shows the top 5 states with the highest profits as a percentage of sales. The states are the same as the 5 with the lowest market expenses as a percentage of sales, but they are ranked differently. Illinois has higher profits as a percentage of sales than Texas but also higher market expenses as a percent of sales.

E3.

1. **SELECT** StateName, ProdName, MktExp **FROM**(
2. **SELECT** T.StateName, P.ProdName, SUM(F.ActMarkCost) **AS** MktExp, RANK() OVER (PARTITION **BY** T.StateName
3. **ORDER** **BY** SUM(F.ActMarkCost)) **AS** StateRank
4. **FROM** TempAE T, ProdCoffee P, FactCoffee F, States S, AreaCode A
5. **WHERE** T.StateName = S.StateName AND S.StateID = A.StateID AND A.AreaID = F.AreaID AND P.ProductID = F.ProductID
6. **GROUP** **BY** T.StateName, P.ProdName
7. **ORDER** **BY** StateName, MktExp)
8. **WHERE** StateRank <= 3;

|  |  |  |
| --- | --- | --- |
| STATENAME | PRODNAME | MKTEXP |
| Colorado | Earl Grey | 400 |
| Colorado | Mint | 414 |
| Colorado | Green Tea | 448 |
| Illinois | Earl Grey | 504 |
| Illinois | Darjeeling | 512 |
| Illinois | Lemon | 526 |
| Iowa | Decaf Espresso | 8 |
| Iowa | Caffe Mocha | 124 |
| Iowa | Amaretto | 126 |
| Massachusetts | Green Tea | 138 |

The table shows the top 10 records from the result. The result shows the 3 products in each market with the lowest marketing expenses.

**F. MARKETING EXPENSES (HIGHEST):**

F1.

1. **CREATE** **TABLE** TempAF **AS**
2. **SELECT** \* **FROM** (
3. **SELECT** S.StateName, ROUND(100\*(SUM(F.ActMarkCost)/SUM(F.ActSales)), 2) **AS** MktPercSales,
4. ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)), 2) **AS** ProfPercSales
5. **FROM** States S, FactCoffee F, AreaCode A
6. **WHERE** S.StateID = A.StateID AND F.AreaID = A.AreaID
7. **GROUP** **BY** S.StateName
8. **ORDER** **BY** MktPercSales **DESC**)
9. **WHERE** ROWNUM <= 5;

|  |  |  |
| --- | --- | --- |
| STATENAME | MKTPERCSALES | PROFPERCSALES |
| Nevada | 20.04 | 17.65 |
| Wisconsin | 19.81 | 26.31 |
| New Mexico | 19.14 | 5.03 |
| Washington | 18.61 | 29.3 |
| New York | 18.41 | 28.36 |

This table shows the 5 states with the highest marketing expenses as a percentage of sales along with their profits as a percentage of sales. To determine if these marketing expenses are justified, we will compare the profits as a percentage of sales to the average for all states. If the state profits as a percentage of sales is higher than the average, we will say that the expenses are justified.

1. **SELECT** AVG(ProfPercSales) **AS** AvgStateProfPercSales **FROM** (
2. **SELECT** S.StateName, ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)), 2) **AS** ProfPercSales
3. **FROM** States S, FactCoffee F, AreaCode A
4. **WHERE** S.StateID = A.StateID AND A.AreaID = F.AreaID
5. **GROUP** **BY** S.StateName);

|  |
| --- |
| AVGSTATEPROFPERCSALES |
| 30.0165 |

None of the states listed above have profits as a percentage as sales as high as the average for all states. By the criteria defined above, this means that the marketing expenses are not justified.

F2.

1. **SELECT** \* **FROM** (
2. **SELECT** T.StateName, A.AreaID, ROUND(100\*(SUM(F.ActMarkCost))/SUM(F.ActSales), 2) **AS** MktPercSales,
3. ROUND(100\*(SUM(F.ActProfit)/SUM(F.ActSales)),2) **AS** ProfPercSales
4. **FROM** TempAF T, AreaCode A, FactCoffee F, States S
5. **WHERE** T.StateName = S.StateName AND S.StateID = A.StateID AND A.AreaID = F.AreaID
6. **GROUP** **BY** T.StateName, A.AreaID
7. **ORDER** **BY** StateName, MktPercSales **DESC**
8. )
9. **WHERE** ProfPercSales < 25;

|  |  |  |  |
| --- | --- | --- | --- |
| STATENAME | AREAID | MKTPERCSALES | PROFPERCSALES |
| Nevada | 775 | 20.65 | 14.31 |
| Nevada | 702 | 19.42 | 21.06 |
| New Mexico | 505 | 19.14 | 5.03 |
| New York | 914 | 28.71 | -4.05 |
| New York | 845 | 25.34 | -0.67 |
| New York | 212 | 24.07 | 17.37 |
| New York | 347 | 23.25 | 17.43 |
| New York | 607 | 19.95 | 23.01 |
| Wisconsin | 608 | 23.3 | 22.92 |

The table shows all area codes within the 5 states from F1 that have lower than 25% profits as a percentage of sales. It is clear that a few of these area codes spend too much on marketing expenses relative to others. For example, AreaID 914 and AreaID845 both spend more than 25% of their sales on marketing and have negative profits.

**G. STRATEGY**

G1. To evaluate which stores to close, the first thing I did was create a new table that contains only the stores that have had below average profits for both 2012 and 2013 when compared to the profits of all stores. Next, I looked at the change in sales from 2012 to 2013 for these locations and found the stores with the largest percent decrease in sales from 2012 to 2013

1. -- Find AreaIDs with below avereage profits in 2012 and 2013
2. **CREATE** **TABLE** TempAG1 **AS**
3. **SELECT** A.AreaID, S.StateName
4. **FROM** AreaCode A, States S, FactCoffee F
5. **WHERE** A.StateID = S.StateID AND A.AreaID = F.AreaID AND EXTRACT(YEAR **FROM** F.FactDate) = 2013
6. **GROUP** **BY** A.AreaID, S.StateName
7. **HAVING** SUM(F.ActProfit) <
8. (**SELECT** AVG(SUM(ActProfit)) **AS** AvgProf
9. **FROM** FactCoffee
10. **WHERE** EXTRACT(YEAR **FROM** FactDate) = 2013
11. **GROUP** **BY** AreaID)
12. **INTERSECT**
13. **SELECT** A.AreaID, S.StateName
14. **FROM** AreaCode A, States S, FactCoffee F
15. **WHERE** A.StateID = S.StateID AND A.AreaID = F.AreaID AND EXTRACT(YEAR **FROM** F.FactDate) = 2012
16. **GROUP** **BY** A.AreaID, S.StateName
17. **HAVING** SUM(F.ActProfit) <
18. (**SELECT** AVG(SUM(ActProfit)) **AS** AvgProf
19. **FROM** FactCoffee
20. **WHERE** EXTRACT(YEAR **FROM** FactDate) = 2012
21. **GROUP** **BY** AreaID);
23. -- Find AreaIDs with largest percent decrease in sales
24. **SELECT** AreaID, StateName, Sales2012, Sales2013, Sales2013 - Sales2012 **AS** SalesDiff,
25. ROUND(100\*(Sales2013 - Sales2012)/Sales2012,2) **AS** PercDiff
26. **FROM** (
27. **SELECT** T.AreaID, T.StateName, EXTRACT(YEAR **FROM** F.FactDate) **AS** Year, SUM(F.ActSales) **AS** Sales
28. **FROM** TempAG1 T, FactCoffee F, AreaCode A
29. **WHERE** T.AreaID = A.AreaID AND A.AreaID = F.AreaID
30. **GROUP** **BY** T.AreaID, T.StateName, EXTRACT(YEAR **FROM** F.FactDate))
31. PIVOT (
32. SUM(Sales)
33. **FOR** Year IN (2013 **AS** Sales2013, 2012 **AS** Sales2012))
34. **ORDER** **BY** PercDiff;

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| AREAID | STATENAME | SALES2012 | SALES2013 | SALESDIFF | PERCDIFF |
| 831 | California | 1796 | 322 | -1474 | -82.07 |
| 214 | Texas | 1337 | 310 | -1027 | -76.81 |
| 936 | Texas | 982 | 245 | -737 | -75.05 |
| 210 | Texas | 1213 | 382 | -831 | -68.51 |
| 972 | Texas | 1490 | 562 | -928 | -62.28 |
| 740 | Ohio | 2326 | 878 | -1448 | -62.25 |
| 832 | Texas | 1156 | 449 | -707 | -61.16 |
| 956 | Texas | 1201 | 488 | -713 | -59.37 |
| 407 | Florida | 1027 | 424 | -603 | -58.71 |
| 863 | Florida | 1474 | 673 | -801 | -54.34 |

This table shows the 10 locations with the largest percentage sales decrease from 2012 to 2013 out of the stores with below average profits for both 2012 and 2013. If we are looking for potential stores to close, these would be the first that I would look at.

G2.

1. **SELECT** AreaID, StateName, Sales2012, Sales2013, ROUND(100\*(Sales2013 - Sales2012)/Sales2012,2) **AS** PercDiff
2. **FROM** (
3. **SELECT** A.AreaID, S.StateName, EXTRACT(YEAR **FROM** F.FactDate) **AS** Year, SUM(F.ActSales) **AS** TotSales
4. **FROM** AreaCode A, States S, FactCoffee F
5. **WHERE** A.StateID = S.StateID AND A.AreaID = F.AreaID
6. **GROUP** **BY** A.AreaID, S.StateName, EXTRACT(YEAR **FROM** F.FactDate))
7. PIVOT (
8. SUM(TotSales)
9. **FOR** Year IN (2013 **AS** Sales2013, 2012 **AS** Sales2012))
10. **WHERE** Sales2012 **IS** NOT NULL AND Sales2013 **IS** NOT NULL
11. **ORDER** **BY** PercDiff **DESC**;

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| AREAID | STATENAME | SALES2012 | SALES2013 | PERCDIFF |
| 323 | California | 298 | 2456 | 724.16 |
| 352 | Florida | 159 | 1281 | 705.66 |
| 510 | California | 234 | 1638 | 600 |
| 432 | Texas | 318 | 1549 | 387.11 |
| 806 | Texas | 190 | 861 | 353.16 |
| 914 | New York | 781 | 3026 | 287.45 |
| 817 | Texas | 318 | 1066 | 235.22 |
| 530 | California | 1332 | 4448 | 233.93 |
| 774 | Massachusetts | 1225 | 3922 | 220.16 |
| 619 | California | 468 | 1485 | 217.31 |

The table shows the first 10 records from the result. The entire result shows every location’s sales in 2012, sales in 2013 and the percent difference. These are the 10 area codes with the largest percent increase in sales from 2012 to 2013. If the firm is looking to expand, I would recommend looking into these locations first.

# PART B – Office Products

QUESTION 1:

1. **CREATE** **TABLE** Managers (
2. RegID       NUMBER,
3. Region      VARCHAR2(10 BYTE),
4. RegManager  VARCHAR2(10 BYTE),
5. **PRIMARY** **KEY** (RegID),
6. **CONSTRAINT** ch\_reg **CHECK** (Region IN ('East', 'South', 'Central', 'West')));
8. **CREATE** **TABLE** Products (
9. ProdID        NUMBER,
10. ProdName      VARCHAR2(100 BYTE),
11. ProdCat       VARCHAR2(30 BYTE),
12. ProdSubCat    VARCHAR2(30 BYTE),
13. ProdCont      VARCHAR2(20 BYTE),
14. ProdUnitPrice NUMBER(7,2),
15. ProdMargin    NUMBER(5,3),
16. **PRIMARY** **KEY** (ProdID),
17. **CONSTRAINT** ch\_cat **CHECK** (ProdCat IN ('Technology', 'Furniture', 'Office Supplies')),
18. **CONSTRAINT** ch\_cont **CHECK** (ProdCont IN ('Jumbo Drum', 'Medium Box', 'Jumbo Box', 'Wrap Bag',
19. 'Large Box', 'Small Box', 'Small Pack')));
21. **CREATE** **TABLE** Orders (
22. OrderID   NUMBER,
23. Status    VARCHAR2(10 BYTE),
24. **PRIMARY** **KEY** (OrderID));
26. **CREATE** **TABLE** Customers (
27. CustID      NUMBER,
28. CustName    VARCHAR2(35 BYTE),
29. CustReg     NUMBER(1,0),
30. CustState   VARCHAR2(20 BYTE),
31. CustCity    VARCHAR2(20 BYTE),
32. CustZip     NUMBER(5,0),
33. CustSeg     VARCHAR2(15 BYTE),
34. **PRIMARY** **KEY** (CustID),
35. **FOREIGN** **KEY** (CustReg) **REFERENCES** Managers(RegID) **ON** **DELETE** **CASCADE**,
36. **CONSTRAINT** ch\_seg **CHECK** (CustSeg IN ('Home Office', 'Corporate', 'Small Business', 'Consumer')));
38. **CREATE** **TABLE** OrderDet (
39. OrderID       NUMBER,
40. CustID        NUMBER,
41. ProdID        NUMBER,
42. OrdPriority   VARCHAR2(15 BYTE),
43. OrdDiscount   NUMBER(3,2),
44. OrdShipMode   VARCHAR2(15 BYTE),
45. OrdDate       **DATE**,
46. OrdShipDate   **DATE**,
47. OrdShipCost   NUMBER(5,2),
48. OrdQty        NUMBER,
49. OrdSales      NUMBER(8,2),
50. **PRIMARY** **KEY** (OrderID, CustID, ProdID),
51. **FOREIGN** **KEY** (OrderID) **REFERENCES** Orders(OrderID),
52. **FOREIGN** **KEY** (CustID) **REFERENCES** Customers(CustID),
53. **FOREIGN** **KEY** (ProdID) **REFERENCES** Products(ProdID),
54. **CONSTRAINT** ch\_priority **CHECK** (OrdPriority IN ('Low', 'Medium', 'High', 'Critical', 'Not Specified')),
55. **CONSTRAINT** ch\_mode **CHECK** (OrdShipMode IN ('Regular Air', 'Delivery Truck', 'Express Air')));

The above script was used to create the tables. Once the tables were created, the data was loaded to them from the different Excel spreadsheets.

QUESTION 2: ORDER Cancellations

a)

1. **SELECT** ROUND(R.ReturnCount/O.TotOrders,4) **AS** FracReturned
2. **FROM** (
3. **SELECT** COUNT(OrderID) **AS** ReturnCount
4. **FROM** Orders
5. **WHERE** Status LIKE 'Returned') R,
6. (**SELECT** COUNT(OrderID) **AS** TotOrders
7. **FROM** Orders) O;

|  |
| --- |
| FRACRETURNED |
| 0.0093 |

A fraction of about 0.0093, or 0.93%, of the orders were cancelled.

b)

1. **SELECT** SUM(D.OrdSales) **AS** SalesReturned
2. **FROM** Orders O, OrderDet D
3. **WHERE** O.OrderID = D.OrderID AND O.Status LIKE 'Returned';

|  |
| --- |
| SALESRETURNED |
| 308455.12 |

The total sales from cancelled orders was $308,455.12.

c)

1. **SELECT** CustName, RetOrders **FROM** (
2. **SELECT** C.CustName, COUNT(O.OrderID) RetOrders, RANK() OVER (**ORDER** **BY** COUNT(O.OrderID) **DESC**) **AS** Rank
3. **FROM** Orders O, OrderDet D, Customers C
4. **WHERE** O.OrderID = D.OrderID AND D.CustID = C.CustID AND O.Status LIKE 'Returned'
5. **GROUP** **BY** C.CustName)
6. **WHERE** Rank <= 5;

|  |  |
| --- | --- |
| CUSTNAME | RETORDERS |
| Hazel Jennings | 7 |
| Keith Marsh | 5 |
| Jenny Gold | 5 |
| Maxine Collier Grady | 4 |
| Gordon Brandt | 4 |
| Leigh Burnette Hurley | 4 |
| Billy Perry Browning | 4 |
| Constance Flowers | 4 |

The table shows the result of the 8 customers with the most cancelled orders. There are 8 rather than 5 because 5 customers are tied for the fourth most cancelled orders.

QUESTION 3: CUSTOMER Related

a)

1. **SELECT** CustName, TotSales **FROM** (
2. **SELECT** C.CustName, SUM(D.OrdSales) **AS** TotSales, RANK() OVER (**ORDER** **BY** SUM(D.OrdSales) **DESC**) **AS** Rank
3. **FROM** Customers C, OrderDet D
4. **WHERE** C.CustID = D.CustID
5. **GROUP** **BY** C.CustName)
6. **WHERE** Rank <= 10;

|  |  |
| --- | --- |
| CUSTNAME | TOTSALES |
| Gordon Brandt | 123745.62 |
| Glen Caldwell | 89269.7 |
| Rosemary O'Brien | 86540.75 |
| Leigh Burnette Hurley | 83651.7 |
| Kristine Connolly | 81296.39 |
| Nina Horne Kelly | 78243.6 |
| Neal Wolfe | 69118 |
| Priscilla Kane | 61610.6 |
| Dana Teague | 61298.98 |
| Kim Weiss | 58947.41 |

The table shows the entire result (top 10 customers in terms of total sales/revenues generated).

b)

1. -- Percent of customers who buy all 3 product categories
2. **SELECT** ROUND(100\*(AllCatsCount/TotCount),2) **AS** AllCatPerc
3. **FROM**(
4. **SELECT** COUNT(CatsPurchased) **AS** AllCatsCount
5. **FROM** (
6. **SELECT** CustName, COUNT(ProdCat) **AS** CatsPurchased
7. **FROM** (
8. **SELECT** C.CustName, P.ProdCat
9. **FROM** Customers C, Products P, OrderDet O
10. **WHERE** C.CustID = O.CustID AND O.ProdID = P.ProdID
11. **GROUP** **BY** C.CustName, P.ProdCat
12. **ORDER** **BY** C.CustName)
13. **GROUP** **BY** CustName
14. **ORDER** **BY** CatsPurchased)
15. **WHERE** CatsPurchased = 3),
16. (**SELECT** COUNT(CatsPurchased) **AS** TotCount
17. **FROM** (
18. **SELECT** CustName, COUNT(ProdCat) **AS** CatsPurchased
19. **FROM** (
20. **SELECT** C.CustName, P.ProdCat
21. **FROM** Customers C, Products P, OrderDet O
22. **WHERE** C.CustID = O.CustID AND O.ProdID = P.ProdID
23. **GROUP** **BY** C.CustName, P.ProdCat
24. **ORDER** **BY** C.CustName)
25. **GROUP** **BY** CustName
26. **ORDER** **BY** CatsPurchased));

|  |
| --- |
| ALLCATPERC |
| 24.24 |

The result of this query is the percentage of customers who buy products from all 3 categories. Less than 25% of customers fit this criteria, meaning more than 75% do not buy products from all 3 categories. Clearly there is some potential for this 75% to buy other product categories.

1. -- Percent of sales in each category for each customer
2. **SELECT** X.CustName, Y.ProdCat, Y.CatSales, ROUND(100\*(Y.CatSales/X.CustSales),2) **AS** CatPerc
3. **FROM** (
4. **SELECT** C.CustName, SUM(D.OrdSales) **AS** CustSales
5. **FROM** Customers C, OrderDet D
6. **WHERE** C.CustID = D.CustID
7. **GROUP** **BY** C.CustName) X,
8. (**SELECT** C.CustName, P.ProdCat, SUM(D.OrdSales) **AS** CatSales
9. **FROM** Products P, OrderDet D, Customers C
10. **WHERE** C.CustID = D.CustID AND P.ProdID = D.ProdID
11. **GROUP** **BY** C.CustName, P.ProdCat) Y
12. **WHERE** X.CustName = Y.CustName AND ROUND(100\*(Y.CatSales/X.CustSales),2) < 10
13. **ORDER** **BY** CustName;

|  |  |  |  |
| --- | --- | --- | --- |
| CUSTNAME | PRODCAT | CATSALES | CATPERC |
| Aaron Davies Bruce | Office Supplies | 83.82 | 3.51 |
| Aaron Riggs | Office Supplies | 127.86 | 3.4 |
| Adam G Sawyer | Office Supplies | 292.18 | 9.43 |
| Adam G Sawyer | Furniture | 26.68 | 0.86 |
| Adam Saunders Gray | Office Supplies | 256.72 | 8.02 |
| Adam Saunders Gray | Technology | 233.98 | 7.31 |
| Albert Frost | Technology | 739.9 | 7.56 |
| Albert Frost | Furniture | 342.62 | 3.5 |
| Alex Harding | Furniture | 1348.73 | 7.19 |
| Alex Watkins | Office Supplies | 19.12 | 2.17 |

The table shows the first 10 records from the result. The entire result shows the total amount spent by a customer on a product category and the percentage of their total spending on that category if the percentage is less than 10%. Clearly there are numerous customers who spend a very low percentage on one or more categories. This could indicate that there is potential for these customers to spend more on these product categories.

QUESTION 4:

a)

1. **SELECT** SUM(ActPrice) - SUM(TheorPrice) **AS** TotalDiff **FROM** (
2. **SELECT** D.OrderID, ((P.ProdUnitPrice\*D.OrdQty)\*(1-D.OrdDiscount)+OrdShipCost) **AS** TheorPrice, D.OrdSales **AS** ActPrice
3. **FROM** OrderDet D, Products P
4. **WHERE** D.ProdID = P.ProdID);

|  |
| --- |
| TOTALDIFF |
| -21791.2566 |

The result shows the total difference between actual sales and theoretical sales for all orders. It appears that their actual sales are about $21,791.26 less than they theoretically should be.

b)

1. **SELECT** Region, RegManager, ROUND(AVG(ActPrice),2) **AS** AvgAct, ROUND(AVG(TheorPrice),2) **AS** AvgTheor,
2. ROUND(100\*(AVG(ActPrice) - AVG(TheorPrice))/AVG(TheorPrice),2) **AS** AvgPercDiff
3. **FROM** (
4. **SELECT** D.OrderID, M.Region, M.RegManager, ((P.ProdUnitPrice\*D.OrdQty)\*(1-D.OrdDiscount)+OrdShipCost) **AS** TheorPrice, D.OrdSales **AS** ActPrice
5. **FROM** OrderDet D, Products P, Managers M, Customers C
6. **WHERE** D.ProdID = P.ProdID AND M.RegID = C.CustReg AND C.CustID = D.CustID)
7. **GROUP** **BY** Region, RegManager
8. **ORDER** **BY** AvgPercDiff;

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| REGION | REGMANAGER | AVGACT | AVGTHEOR | AVGDIFF |
| West | William | 1041.41 | 1046.82 | -0.52 |
| South | Sam | 808.55 | 810.25 | -0.21 |
| Central | Chris | 866.49 | 867.77 | -0.15 |
| East | Erin | 1042.9 | 1044.01 | -0.11 |

The result shows the average actual price of an order, the average theoretical price of an order, and the percent difference for each region and manager. It appears that all four managers are generally pricing less than the theoretical price, but William in the West region is doing it significantly more than the others.

QUESTION 5: Product Related Questions

a)

1. **SELECT** ProdName
2. **FROM** Products
3. **WHERE** REGEXP\_LIKE (ProdName, '\d');

|  |
| --- |
| PRODNAME |
| APC 7 Outlet Network SurgeArrest Surge Protector |
| AT&T 1430 2.4GHz Analog Phone w/Caller ID |
| AT&T 2230 Dual Handset Phone With Caller ID/Call Waiting |
| AT&T Black Trimline Phone, Model 210 |
| Atlantic Metals Mobile 2-Shelf Bookcases, Custom Colors |
| Atlantic Metals Mobile 3-Shelf Bookcases, Custom Colors |
| Atlantic Metals Mobile 5-Shelf Bookcases, Custom Colors |
| Avanti 1.7 Cu. Ft. Refrigerator |
| Avanti 4.4 Cu. Ft. Refrigerator |
| Avery 05222 Permanent Self-Adhesive File Folder Labels for Typewriters, on Rolls, White, 250/Roll |

The table shows the first 10 records from the result. The entire result shows all products with digits in their name.

b)

1. **SELECT** Rank, ProdName, TotSales **FROM** (
2. **SELECT** P.ProdName, SUM(D.OrdSales) **AS** TotSales, RANK() OVER (**ORDER** **BY** SUM(D.OrdSales) **DESC**) **AS** Rank
3. **FROM** Products P, OrderDet D
4. **WHERE** P.ProdID = D.ProdID AND EXTRACT(YEAR **FROM** D.OrdDate) = 2011
5. **GROUP** **BY** P.ProdName)
6. **WHERE** Rank <=5;

|  |  |  |
| --- | --- | --- |
| RANK | PRODNAME | TOTSALES |
| 1 | Canon imageCLASS 2200 Advanced Copier | 64713.42 |
| 2 | Canon PC1080F Personal Copier | 60380.82 |
| 3 | Sharp AL-1530CS Digital Copier | 58482.4 |
| 4 | Bretford CR8500 Series Meeting Room Furniture | 58279.87 |
| 5 | Global Troy™ Executive Leather Low-Back Tilter | 48602.86 |

The result shows the top 5 selling products and their total sales in the year 2011.

c)

1. **SELECT** Rank, ProdName, TotMargin **FROM** (
2. **SELECT** P.ProdName, SUM(D.OrdSales\*P.ProdMargin) **AS** TotMargin, RANK() OVER (**ORDER** **BY** SUM(D.OrdSales\*P.ProdMargin) **DESC**) **AS** RANK
3. **FROM** Products P, OrderDet D
4. **WHERE** P.ProdID = D.ProdID
5. **GROUP** **BY** P.ProdName)
6. **WHERE** Rank <= 10;

|  |  |  |
| --- | --- | --- |
| RANK | PRODNAME | TOTMARGIN |
| 1 | Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish | 117920.993 |
| 2 | Global Troy™ Executive Leather Low-Back Tilter | 116415.384 |
| 3 | Bretford CR8500 Series Meeting Room Furniture | 71134.7789 |
| 4 | Chromcraft Bull-Nose Wood 48" x 96" Rectangular Conference Tables | 63456.0774 |
| 5 | Bretford CR4500 Series Slim Rectangular Table | 54207.7652 |
| 6 | Canon PC1080F Personal Copier | 51466.385 |
| 7 | Hon 2090 “Pillow Soft” Series Mid Back Swivel/Tilt Chairs | 48947.0592 |
| 8 | BoxOffice By Design Rectangular and Half-Moon Meeting Room Tables | 48811.1858 |
| 9 | Lexmark 4227 Plus Dot Matrix Printer | 45328.3435 |
| 10 | Hon Multipurpose Stacking Arm Chairs | 44422.7107 |

The result shows the top 10 products with the greatest total profit margin along with their total margins.

d)

1. **SELECT** Rank, ProdName, TotSales **FROM** (
2. **SELECT** P.ProdName, SUM(D.OrdSales) **AS** TotSales, RANK() OVER (**ORDER** **BY** SUM(D.OrdSales)) **AS** Rank
3. **FROM** Products P, OrderDet D
4. **WHERE** P.ProdID = D.ProdID
5. **GROUP** **BY** P.ProdName)
6. **WHERE** Rank <= 5;

|  |  |  |
| --- | --- | --- |
| RANK | PRODNAME | TOTSALES |
| 1 | Alliance Rubber Bands | 7.43 |
| 2 | \*Staples\* Packaging Labels | 11.71 |
| 3 | Blackstonian Pencils | 13.18 |
| 4 | Avery 482 | 16.67 |
| 5 | Sony IBM Color Diskettes, 25/Pack | 18.17 |

The result shows the worst 5 products in terms of sales along with their total sales.