Group Assignment

Group - 3: Chunlu Zhu, Na Qian, Pragati Koladiya, Sijie Wang
Fall A 2020
ALY 6030
Group Assignment
Last Updated: 10/05/2020

AIM:
To address the given case in context of the AW database.

Case #4:

- An executive states that R&D has been increasing its cost by 15% every quarter and believes that the cost is associated with Accessory product churn.
- The executive believes that the churn is unnecessary and is losing money for the company.
- What data can you provide to support or refute the executive's claim?

Is R&D increasing its cost by 15% every quarter?

First, we are going to check the cost that R&D spends in each quarter. We choose fiscal year and fiscal quarter in this case, because we are analysing the Financial statements. They are issued through fiscal years. Also, we have three conditions: i) expenses are actual; ii) department is R&D department; iii) account type is expenditure in the income statements. We ranked the data for the further analysis (cost increase by percent for every quarter). Finally, we put the data into the table named “R_D_Cost”.

Input Query:

```
-- ---------------------------------------------------------------------------
/*First, we are going to check the cost that R&D spends in each quarter. */
-- ---------------------------------------------------------------------------

CREATE TABLE R_D_Cost
SELECT
    B.year_quarter,
    B.Total_cost_R_D,
    RANK() OVER (ORDER BY B.year_quarter ASC) AS rank_year
FROM
    (SELECT
        CONCAT(A.FiscalYear, '-' , A.FiscalQuarter) AS year_quarter,
        A.Total_cost_R_D
    FROM
        (SELECT
            t.FiscalYear,
            t.FiscalQuarter,
            SUM(f.Amount) AS Total_cost_R_D
        FROM FactFinance AS f
        INNER JOIN DimTime AS t
            ON f.TimeKey = t.TimeKey
        INNER JOIN DimAccount AS a
            ON f.AccountKey = a.AccountKey
        INNER JOIN DimScenario AS s
            ON f.ScenarioKey = s.ScenarioKey
        INNER JOIN DimDepartmentGroup AS d
            ON f.DepartmentGroupKey = d.DepartmentGroupKey
        WHERE s.ScenarioName = "Actual" AND d.DepartmentGroupName = "Research and Development" AND a.AccountType = "Expenditures"
        GROUP BY t.FiscalYear, t.FiscalQuarter
        ) AS A
    ) AS B;
```

The result shows below:

year_quarter	Total_cost_R_D	rank_year
2018-1	2607840.901977539	1
2018-2	3523345.3420386314	2
2018-3	3262258.2403535843	3
2018-4	3976494.7061138153	4
2019-1	6655741.095246315	5
2019-2	5345352.074413776	6
2019-3	4331353.764190316	7
2019-4	5040620.10611546	8
2020-1	8472490.706979752	9
2020-2	6742537.53653717	10
2020-3	5740958.618612289	11
2020-4	6772867.801546097	12

- We got the cost that R&D spent in every quarter. Now we need to know the growth rate of the R&D cost.

```
-- ------------------------------------------------------------------------------
/* Second, check whether R&D has been increasing its cost by 15% every quarter*/
-- ------------------------------------------------------------------------------

CREATE TABLE R_D_Percent_cost
SELECT
    C.year_quarter,
    CONCAT(ROUND(100 * (C.Total_cost_R_D - C.Total_cost_lag)/C.Total_cost_lag, 0), ' ' , "%") AS Cost_percent_increasing
FROM
    (SELECT
        A.Total_cost_R_D AS Total_cost_lag,
        A.year_quarter AS year_quarter_lag,
        B.Total_cost_R_D AS Total_cost_R_D,
        B.year_quarter AS year_quarter
    FROM R_D_Cost AS A
    INNER JOIN R_D_Cost AS B
        ON A.rank_year +1 = B.rank_year
    ) AS C;
```

We self-join the table that we got from above and then calculate the growth rate. The result is shown below:

year_quarter	Cost_percent_increasing
2018-2	35 %
2018-3	-7 %
2018-4	22 %
2019-1	67 %
2019-2	-20 %
2019-3	-19 %
2019-4	16 %
2020-1	68 %
2020-2	-20 %
2020-3	-15 %
2020-4	18 %

- We can see that the growth rate for R&D cost for each quarter is not always increasing. The growth rates are fluctuating over the quarters. But in the 4th quarter of 2020, it increases by 18%. Therefore, the executive statement that R&D has been increasing its cost by 15% every quarter is not correct.

Does the company have accessory product churn?

To answer this question, we need to understand the definition of accessory product churn. The accessory product churn is the speed at which products are introduced and then retired. From an executive point of view, he thinks accessory products were already retired. In other words, he is under the impression that the accessory products don't exist. In this case, we check this statement from DimProductCategory, DimProductSubCategory, DimProduct, FactInternetSales, and DimTime table. We select product category names as accessories and status as current. Note: Data is saved in the Accessory_Info table.

```
-- ------------------------------------------------
/* Check total accessory profit for each quarter */
-- ------------------------------------------------

CREATE TABLE Accessory_Info
SELECT
    CONCAT(A.FiscalYear, "-", A.FiscalQuarter) AS year_quarter,
    A.Accessory_Cost,
    A.Accessory_Profit
FROM
    (SELECT
        t.FiscalYear,
        t.FiscalQuarter,
        SUM(i.TotalProductCost) AS Accessory_Cost,
        SUM(i.SalesAmount - i.TotalProductCost) AS Accessory_Profit
    FROM DimProductCategory AS p
    INNER JOIN DimProductSubcategory AS s
        ON p.ProductCategoryKey = s.ProductCategoryKey
    INNER JOIN DimProduct AS product
        ON s.ProductSubcategoryKey = product.ProductSubcategoryKey
    INNER JOIN FactInternetSales AS i
        ON i.ProductKey = product.ProductKey
    INNER JOIN DimTime AS t
        ON i.OrderDateKey = t.TimeKey
    WHERE p.EnglishProductCategoryName = "Accessories" AND product.Status = "Current"
    GROUP BY t.FiscalYear, t.FiscalQuarter
    ) AS A;
```

The final result is as below:

year_quarter	Accessory_Cost	Accessory_Profit
2020-1	44134.00	74422.00
2020-2	65087.00	109762.00
2020-3	64578.00	108793.00
2020-4	74310.00	125229.00
2021-1	12548.00	21151.00

- The result table shows that accessory products have existed since the first quarter of 2020. Moreover, the accessory profit is increasing over the quarters. We select the current status of accessory products, so we can prove that the executive statement of accessory product churn is incorrect.

Is R&D cost associated with accessory product churn?

We join Accessory_Info, R_D_Cost, and R_D_Percent_Cost tables.

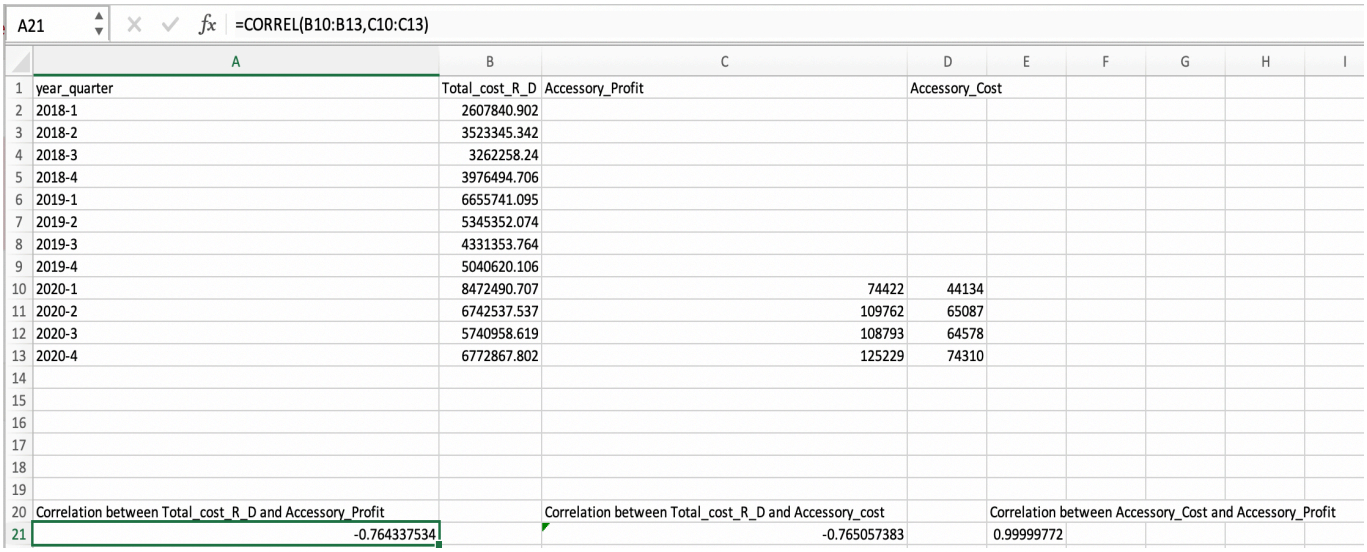
```
-- ------------------------------------------------------
/*Join Accessory_Info, R_D_Cost, and R_D_Percent_cost */
-- ------------------------------------------------------

SELECT
    a.year_quarter,
    a.Accessory_Cost,
    a.Accessory_Profit,
    r.Total_cost_R_D,
    p.Cost_percent_increasing
FROM Accessory_Info AS a
INNER JOIN R_D_Cost AS r
    ON a.year_quarter = r.year_quarter
INNER JOIN R_D_Percent_cost AS p
    ON a.year_quarter = p.year_quarter;
```

The result shows below:

year_quarter	Accessory_Cost	Accessory_Profit	Total_cost_R_D	Cost_percent_increasi...
2020-1	44134.00	74422.00	8472490.706979752	68 %
2020-2	65087.00	109762.00	6742537.53653717	-20 %
2020-3	64578.00	108793.00	5740958.618612289	-15 %
2020-4	74310.00	125229.00	6772867.801546097	18 %

There are only 4 quarters in 2020. To see the relationship between accessory cost, accessory profit, and total cost of R&D, we have built the following correlation on excel:



The correlation between total cost R&D and accessory profit is -0.7643, which means that R&D cost is increasing, while accessory profit is decreasing. It seems like this evidence proves the executive's statement, since his statement is that when accessory profit is churn (accessory profit is decreasing), R&D cost is increasing. We disagree with this statement, there are few reasons:

1. We only calculate the correlation between accessory profit and R&D cost but not causation between accessory profit and R&D cost. In fact, we cannot control other variables affecting R&D cost to calculate causation between these two.
2. The data points are only 4 quarters in 2020. We cannot make the conclusion from a few data points.
3. We only can get the R&D total cost, but we cannot get the R&D accessory cost specifically.

Based on these reasons, we cannot conclude that R&D cost is associated with accessory product churn.

Whether churn is unnecessary and is losing money for the company?

Since we know that accessory products still exist in the company, we disagree with this statement.

In conclusion, we refute the executive's claim, because 1) the growth rate for R&D is not always increasing from 2018 to 2020; 2) accessory products still exist in the company; 3) we cannot find a causation relationship between accessory profit and R&D cost.

Case #5:

- Your manager has asked you to prepare a revenue forecast. Select the data that is necessary to predict revenue 12 months into the future.

To get the revenue forecasting table, we need to get the time on a monthly basis, we choose to use the calendar year in this case. For that we need a monthly basis table and we only have calendar year and calendar month in the DimTime table. We join DimTime with FactInternetSales, because we want to predict sales revenue. Note that we use “case when” to convert all English month names to numerical data.

```
-- ----------------------------------------------------
/*Get the table for time-series revenue forecasting*/
-- ----------------------------------------------------

SELECT
    CONCAT(A.CalendarYear, "-", A.new_month) AS YearMonth,
    SUM(A.SalesAmount) AS Total_SalesAmount
FROM
    (SELECT
        t.CalendarYear,
        i.SalesAmount,
        CASE
            WHEN t.EnglishMonthName = "January" THEN "01"
            WHEN t.EnglishMonthName = "February" THEN "02"
            WHEN t.EnglishMonthName = "March" THEN "03"
            WHEN t.EnglishMonthName = "April" THEN "04"
            WHEN t.EnglishMonthName = "May" THEN "05"
            WHEN t.EnglishMonthName = "June" THEN "06"
            WHEN t.EnglishMonthName = "July" THEN "07"
            WHEN t.EnglishMonthName = "August" THEN "08"
            WHEN t.EnglishMonthName = "September" THEN "09"
            WHEN t.EnglishMonthName = "October" THEN "10"
            WHEN t.EnglishMonthName = "November" THEN "11"
            ELSE "12"
        END AS new_month
    FROM DimTime AS t
    INNER JOIN FactInternetSales AS i
        ON i.OrderDateKey = t.TimeKey
    ) AS A
GROUP BY YearMonth;
```

The result shows below:

YearMonth	Total_SalesAmount
2017-07	473360.00
2017-08	506161.00
2017-09	473910.00
2017-10	513297.00
2017-11	543960.00
2017-12	755480.00
2018-01	596710.00
2018-02	550779.00
2018-03	644096.00
2018-04	663652.00
2018-05	673516.00
2018-06	676723.00
2018-07	500365.00
2018-08	545999.00
2018-09	350462.00
2018-10	415378.00
2018-11	335084.00
2018-12	577296.00
2019-01	438842.00
2019-02	489060.00
2019-03	485545.00
2019-04	506372.00
2019-05	562732.00
2019-06	554754.00
2019-07	886596.00
2019-08	847259.00
2019-09	1010089.00
2019-10	1080282.00
2019-11	1196818.00
2019-12	1731548.00
2020-01	1340067.00
2020-02	1462283.00
2020-03	1480736.00
2020-04	1608526.00
2020-05	1878067.00
2020-06	1949123.00
2020-07	50783.00

- The table above can be used to forecast the revenue from August 2020 to August 2021(12 months).