

Evaluación Práctica en SQL sobre el Caso de SmartDesk

MASTER BIG DATA, DATA SCIENCE Y BUSINESS ANALYTICS

Cristian Galán Amaro

Preguntas a resolver:

1. Análisis completo de ventas y beneficio por producto

```
-- 1. Análisis completo de ventas y beneficio por producto

SELECT
    CATEGORY,
    SUM(PRODUCT) AS PRODUCTO,
    SUM(MAINTENANCE) AS MANTENIMIENTO,
    SUM(PARTS) AS PARTES,
    SUM(SUPPORT) AS SOPORTE,
    SUM(PROFIT) AS BENEFICIO_TOTAL,
    ROUND(AVG(PROFIT), 2) AS BENEFICIO_PROMEDIO,
    SUM(UNITS_SOLD) AS UNIDADES_VENDIDAS
FROM SALES
WHERE (YEAR=2020) AND (ACCOUNT='Abbot Industries')
GROUP BY CATEGORY;
```

CATEGORY	PRODUCTO	MANTENIMIENTO	PARTES	SOPORTE	BENEFICIO_TOTAL	BENEFICIO_PROMEDIO	UNIDADES_VENDIDAS
Breakroom	611058.00	134433.00	0.00	0.00	298196.00	298196.00	1745.88
Chairs	2300563.00	347800.00	0.00	0.00	953411.00	953411.00	6573.04
Desks	1160080.00	267600.00	0.00	0.00	513965.00	513965.00	3314.51
Electronics	457900.00	235900.00	0.00	200000.00	312830.00	312830.00	1308.29

En este apartado vemos las cuentas de Abbot Industries en el año 2020 agrupadas por la categoría de sus productos, podemos apreciar como esta empresa apenas vende productos por partes y únicamente ejerce algún tipo de soporte en la categoría de Electronics. Una propuesta podría ser especializarse en el mercado de sillas y escritorios que son los que generan mas beneficios.

2. Cálculo de pronóstico total y beneficio esperado

```
-- 2. Cálculo de pronóstico total y beneficio esperado

SELECT
    COALESCE(F.CATEGORY, S.CATEGORY) AS CATEGORIA,
    SUM(CASE WHEN QUARTER='2020 Q1' THEN PROFIT ELSE 0 END) AS BENEFICIO_Q1_2020,
    SUM(CASE WHEN QUARTER='2021 Q3' THEN PROFIT ELSE 0 END) AS BENEFICIO_Q3_2021,
    SUM(CASE WHEN F.YEAR='2022' THEN FORECAST_PROFIT ELSE 0 END) AS BENEFICIO_ESPERADO_2022,
    IFNULL (MIN(OPPORTUNITY_AGE),0) AS MAS_RECIENTE,
    IFNULL (MAX(OPPORTUNITY_AGE),0) AS MAS_ANTIGUA

FROM SALES AS S FULL JOIN FORECASTS AS F
ON S.YEAR=F.YEAR AND S.CATEGORY=F.CATEGORY
GROUP BY CATEGORIA
ORDER BY CATEGORIA;
```

CATEGORIA	BENEFICIO_Q1_2020	BENEFICIO_Q3_2021	BENEFICIO_ESPERADO_2022	MAS_RECIENTE	MAS_ANTIGUA
Breakroom	2077967.00	5107911.00	33107609.00	2	378
Chairs	1148111.00	188800.00	4343857.00	3	173
Desks	1114668.00	1934225.00	8681586.00	3	383
Electronics	881830.00	1324853.00	9929645.00	3	383

En este caso volvemos a agrupar nuestros datos por categoría, pero realizamos un filtrado en tres momentos del tiempo distintos, observamos que los beneficios obtenidos en el tercer trimestre de 2021 son mucho mayores que los del primer trimestre de 2020, esto puede deberse a diversos factores, por ejemplo, la expansión de las empresas o en cambio, puede que esté relacionado con algún derivado de la estacionalidad, estos datos no se pueden comparar con los obtenidos en el beneficio esperado de 2022 ya que el tiempo de las series elegidas no es el mismo.

3. Comparación de ventas, unidades vendidas y beneficio entre industrias en APAC y EMEA

```
-- 3. Comparación de ventas, unidades vendidas y beneficio entre industrias en APAC y EMEA

SELECT
    REGION,
    INDUSTRY,
    COUNTRY,
    SUM(PRODUCT) AS PRODUCTO,
    SUM(PROFIT) AS BENEFICIO,
    ROUND(AVG(PROFIT),2) AS BENEFICIO_PROMEDIO,
    SUM(UNITS_SOLD) AS UNIDADES_VENDIDAS

FROM SALES AS S INNER JOIN ACCOUNTS AS A
ON S.ACCOUNT=A.ACCOUNT
WHERE REGION='APAC' OR REGION='EMEA'
GROUP BY INDUSTRY, COUNTRY, REGION
ORDER BY REGION, INDUSTRY, COUNTRY;
```

REGION	INDUSTRY	COUNTRY	PRODUCTO	BENEFICIO	BENEFICIO_PROMEDIO	UNIDADES_VENDIDAS
APAC	Consulting	Australia	5978048.00	2787133.00	1393566.50	17080.14
APAC	Consulting	Korea	3475223.00	1846700.00	615566.67	9929.21
APAC	Entertainment and Media	Singapore	3321068.00	1293408.00	1293408.00	9488.77
APAC	Finance	Japan	2089093.00	1218312.00	406104.00	5968.83
APAC	Home Services	Singapore	1000000.00	1008000.00	1008000.00	2857.14
APAC	Hospitality	Korea	913562.00	541313.00	541313.00	2610.18
APAC	Technology	Australia	739677.00	923910.00	923910.00	2113.36
EMEA	Automotive	Germany	1601599.00	927994.00	927994.00	4576.00
EMEA	Automotive	Spain	2583739.00	1044523.00	348174.33	7382.12
EMEA	Banking	Spain	1070574.00	448602.00	448602.00	3058.78
EMEA	Biotech and Pharmaceutical	Austria	1184944.00	788272.00	788272.00	3385.55
EMEA	Consulting	France	1904646.00	911156.00	303718.67	5441.85
EMEA	Consulting	Germany	1761773.00	1242617.00	621308.50	5033.64
EMEA	Entertainment and Media	France	6546290.00	3453479.00	493354.14	18703.69
EMEA	Entertainment and Media	Germany	1552644.00	591247.00	591247.00	4436.13
EMEA	Entertainment and Media	Italy	1865296.00	867680.00	289226.67	5329.42
EMEA	Entertainment and Media	Spain	973692.00	539583.00	539583.00	2781.98
EMEA	Entertainment and Media	Switzerland	3068546.00	1442272.00	480757.33	8767.27
EMEA	Finance	France	1660147.00	758989.00	252996.33	4743.28
EMEA	Finance	Germany	1329741.00	528239.00	528239.00	3799.26
EMEA	Finance	Portugal	2318719.00	963449.00	321149.67	6624.91
EMEA	Finance	Switzerland	694658.00	366344.00	366344.00	1984.74
EMEA	Home Services	France	2715395.00	1365097.00	341274.25	7758.27
EMEA	Hospitality	Sweden	749523.00	380303.00	380303.00	2141.49
EMEA	Insurance	Germany	2436850.00	1032793.00	344264.33	6962.43
EMEA	Law	France	1883462.00	799097.00	399548.50	5381.32
EMEA	Manufacturing	France	4529601.00	2078402.00	519600.50	12941.72
EMEA	Real Estate	Austria	1994003.00	1065342.00	355114.00	5697.15
EMEA	Retail	Austria	2884399.00	1040105.00	1040105.00	8241.14
EMEA	Retail	France	2394083.00	1276800.00	319200.00	6840.23
EMEA	Retail	Germany	9475797.00	4485663.00	640809.00	27073.70
EMEA	Retail	Italy	737309.00	359807.00	359807.00	2106.60
EMEA	Technology	Austria	634183.00	332639.00	332639.00	1811.95
EMEA	Technology	France	586264.00	248331.00	248331.00	1675.04
EMEA	Technology	Germany	4912602.00	2458930.00	351275.71	14036.00
EMEA	Technology	Portugal	3066914.00	1593759.00	796879.50	8762.61
EMEA	Technology	Sweden	500000.00	182900.00	182900.00	1428.57

El análisis en esta ocasión se centra en dos regiones, APAC y EMEA, y se agrupan los resultados por industria y país, este tipo de filtrado nos puede enseñar que tipo de países son los más punteros en un determinado tipo de industria en su región. Por ejemplo, en APAC en Consulting el país puntero es Australia, mientras que en EMEA es Alemania. Esto podría ser determinante para las empresas que buscan un lugar donde expandirse de manera internacional.

4. Beneficio por tipo de empresa

```
-- 4. Beneficio por tipo de empresa DUDA CON SI SE REFIERE A BENEFICIO DE LA PRIMERA TABLA

SELECT
  INDUSTRY,
  SUM(FORECAST_PROFIT) AS INGRESO_FUTURO_TOTAL,
  CASE
    WHEN INGRESO_FUTURO_TOTAL >= 500000 AND INGRESO_FUTURO_TOTAL <= 1000000 THEN 'Beneficio normal'
    WHEN INGRESO_FUTURO_TOTAL > 1000000 THEN 'Beneficio alto'
    ELSE 'No hay registros'
  END AS TIPO_DE_BENEFICIO

FROM FORECASTS AS F INNER JOIN ACCOUNTS AS A
ON (F.ACCOUNT=A.ACCOUNT)
WHERE F.ACCOUNT IN (SELECT ACCOUNT
  FROM FORECASTS
  WHERE FORECAST_PROFIT > 500000)

GROUP BY INDUSTRY;
```

INDUSTRY	INGRESO_FUTURO_TOTAL	TIPO_DE_BENEFICIO
Manufacturing	1423198.00	Beneficio alto
Biotech and Pharmaceutical	1395858.00	Beneficio alto
Entertainment and Media	4533561.00	Beneficio alto
Healthcare	4008251.00	Beneficio alto
Law	2777974.00	Beneficio alto
Technology	3517510.00	Beneficio alto
Consulting	6925531.00	Beneficio alto
Retail	5264537.00	Beneficio alto
Real Estate	1315380.00	Beneficio alto
Home Services	1925380.00	Beneficio alto
Banking	1678289.00	Beneficio alto
Automotive	603000.00	Beneficio normal
Finance	2148635.00	Beneficio alto
Hospitality	3377578.00	Beneficio alto

El filtrado de datos se vuelve a agrupar por industria, pero en este caso se busca obtener una tabla que exponga si los beneficios van a ser altos o normales según una serie de criterios, estableciendo además un mínimo. Esto puede ser muy útil a la hora de establecer intervalos para simplificar la lectura de datos, lo que permite una mayor eficiencia en la toma de decisiones. En este caso particular nos permite apreciar que una industria se prevé que se quede rezagada en el año 2022, Automotive.

5. Beneficio acumulado por trimestre particionado por industria

```
-- 5. Beneficio acumulado por trimestre particionado por industria

SELECT
    INDUSTRY,
    QUARTER_OF_YEAR,
    SUM(PROFIT) AS BENEFICIO_TOTAL,
    SUM(SUM(PROFIT)) OVER (PARTITION BY INDUSTRY ORDER BY QUARTER_OF_YEAR) AS BENEFICIO_ACUMULADO,
    SUM(SUM(FORECAST_PROFIT)) OVER (PARTITION BY INDUSTRY ORDER BY QUARTER_OF_YEAR) AS
FORECAST_ACUMULADO,
    MIN(OPPORTUNITY_AGE) AS MAS_RECIENTE,
    MAX(OPPORTUNITY_AGE) AS MAS_ANTIGUA

FROM ACCOUNTS AS A
INNER JOIN SALES AS S
ON (A.ACCOUNT=S.ACCOUNT)
INNER JOIN FORECASTS AS F
ON (A.ACCOUNT=F.ACCOUNT)
GROUP BY INDUSTRY, QUARTER_OF_YEAR
ORDER BY INDUSTRY, QUARTER_OF_YEAR;
```

INDUSTRY	QUARTER_OF_YEAR	BENEFICIO_TOTAL	BENEFICIO_ACUMULADO	FORECAST_ACUMULADO	MAS_RECIENTE	MAS_ANTIGUA
Banking	Q4	897204.00	897204.00	1158180.00	3	20
Biotech and Pharmaceutical	Q1	280500.00	280500.00	50000.00	7	7
Biotech and Pharmaceutical	Q3	788272.00	1068772.00	645858.00	5	5
Biotech and Pharmaceutical	Q4	819000.00	1887772.00	1445858.00	18	18
Consulting	Q1	1015202.00	1015202.00	932677.00	4	44
Consulting	Q3	1706623.00	2721825.00	3603225.00	13	21
Consulting	Q4	4943287.00	7665112.00	6650495.00	4	21
Education	Q1	539202.00	539202.00	401987.00	11	45
Education	Q2	1516400.00	2055602.00	1205961.00	11	45
Education	Q4	1672000.00	3727602.00	1607948.00	11	45
Entertainment and Media	Q1	367200.00	367200.00	291513.00	3	3
Entertainment and Media	Q2	3145442.00	3512642.00	2688882.00	3	343
Entertainment and Media	Q3	2026516.00	5539158.00	3663522.00	4	343
Entertainment and Media	Q4	1106528.00	6645686.00	4351535.00	3	343
Finance	Q1	4760012.00	4760012.00	3066957.00	3	383
Finance	Q3	631756.00	5391768.00	3350964.00	129	378
Finance	Q4	1277978.00	6669746.00	4670964.00	12	17
Healthcare	Q2	748304.00	748304.00	959532.00	19	19
Healthcare	Q4	420266.00	1168570.00	1919064.00	19	19
Home Services	Q1	675526.00	675526.00	437651.00	19	21
Home Services	Q2	1418134.00	2093660.00	1598922.00	4	49
Home Services	Q3	886918.00	2980578.00	2150193.00	4	49
Home Services	Q4	687290.00	3667868.00	2701464.00	4	49
Hospitality	Q1	380303.00	380303.00	335561.00	49	49
Hospitality	Q4	1431012.00	1811315.00	2464504.00	14	143
Insurance	Q1	568620.00	568620.00	693941.00	5	19
Insurance	Q3	620052.00	1188672.00	1520337.00	5	250
Insurance	Q4	453000.00	1641672.00	1990078.00	5	5
Law	Q1	4628922.00	4628922.00	4534676.00	10	10
Law	Q2	2484000.00	7112922.00	6665676.00	10	10
Law	Q4	1485004.00	8597926.00	7402650.00	3	14
Manufacturing	Q1	10762538.00	10762538.00	6289070.00	3	204
RealEstate	Q1	1794900.00	1794900.00	684558.00	17	34
RealEstate	Q2	518684.00	2313584.00	908837.00	17	34
Retail	Q1	474181.00	474181.00	471747.00	4	20
Retail	Q2	1966351.00	2440532.00	998898.00	20	33
Retail	Q3	541400.00	2981932.00	2858898.00	4	25
Retail	Q4	3414018.00	6395950.00	7403898.00	20	44
Technology	Q1	1201521.00	1201521.00	2156287.00	17	44
Technology	Q2	597376.00	1798897.00	2462183.00	4	5
Technology	Q3	968436.00	2767333.00	4646883.00	24	24
Technology	Q4	1780577.00	4547910.00	6950475.00	5	77

En este caso volvemos a agrupar por industria fijándonos en el beneficio dependiendo del trimestre del año, lo que nos podría hacer identificar algún tipo de estacionalidad. Además, nos fijaremos también en el beneficio acumulado y el Forecast acumulado, por último, podremos evaluar la diferencia entre el tiempo que está abierta la oportunidad más reciente y la mas antigua. Identificamos, por ejemplo, que en “Home Services” el segundo trimestre del año provoca casi el doble de beneficio que cada uno de los otros tres, por lo que podría ser recomendable en las empresas que tengan empleados en esta industria contratar más personal en esta época o en su defecto aumentar las horas extras.

Caso Práctico: Análisis Libre

1. Introducción y análisis exploratorio

En la reunión del viernes 15 de noviembre se pidió al equipo de analistas de SmartDesk las siguientes tareas por parte de los distintos departamentos:

1. En primer lugar, el departamento de Recursos Humanos expuso la necesidad de una reestructuración profunda en el área de “Executives”, este área esta formado por los ejecutivos de cuentas que gestionan la relación con los clientes de la empresa. Esta reestructuración va a provocar que algunos ejecutivos promocionen de puesto y sueldo, otros sean despedidos y otros se queden como están durante el próximo ejercicio.

Para ello, nos piden que marquemos unos criterios justos en los que los promocionados deben oscilar entre 2 y 5, los que continúan entre 10 y 15 y los despedidos no pueden ser menos de 12.

2. En segundo lugar, el departamento de publicidad ha propuesto en la reunión, que se desarrolle una campaña de publicidad específica en un país de los que ya se opera y nos piden que hagamos una propuesta en base a los datos de beneficio e ingresos esperados.

2. Análisis SQL

A través del análisis SQL de nuestra base de datos vamos a procurar dar respuesta a las peticiones de los departamentos antes descritas. Estas son las consultas SQL realizadas:

```
-- CASO PRACTICO.
-- A. Beneficio segun el ejecutivo que gestiona la relacion con el cliente.

SELECT
  A.ACCOUNT_EXECUTIVE,
  SUM(CASE WHEN S.YEAR = 2019 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2019,
  SUM(CASE WHEN S.YEAR = 2020 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2020,
  SUM(CASE WHEN S.YEAR = 2021 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2021,
  IFNULL(SUM(F.FORECAST_PROFIT), 0) AS INGRESOS_ESPERADOS_2022,
  ROUND(
    CASE
      WHEN BENEFICIO_2021 = 0 AND INGRESOS_ESPERADOS_2022 > 0 THEN 100
      WHEN BENEFICIO_2021 = 0 AND INGRESOS_ESPERADOS_2022 = 0 THEN 0
      ELSE
        (INGRESOS_ESPERADOS_2022 - BENEFICIO_2021) / BENEFICIO_2021 * 100
    END, 1) AS VARIACION_PORCENTUAL_21_22,

  CASE
    WHEN VARIACION_PORCENTUAL_21_22 >= 50 AND (BENEFICIO_2019+BENEFICIO_2020+BENEFICIO_2021) >= 3000000
  THEN 'PROMOCIONA'
    WHEN VARIACION_PORCENTUAL_21_22 >= 50 OR (BENEFICIO_2019+BENEFICIO_2020+BENEFICIO_2021) >= 3000000
  THEN 'CONTINUA'
    ELSE 'DESPEDIDO'
  END AS RECURSOS_HUMANOS

FROM ACCOUNTS AS A
INNER JOIN SALES AS S
  ON A.ACCOUNT = S.ACCOUNT
LEFT JOIN FORECASTS AS F
  ON A.ACCOUNT = F.ACCOUNT

GROUP BY A.ACCOUNT_EXECUTIVE
ORDER BY A.ACCOUNT_EXECUTIVE;
```



```
-- B. Países donde la empresa debe fomentar su crecimiento.

SELECT
  COUNTRY,
  SUM(CASE WHEN S.YEAR = 2019 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2019,
  SUM(CASE WHEN S.YEAR = 2020 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2020,
  SUM(CASE WHEN S.YEAR = 2021 THEN S.PROFIT ELSE 0 END) AS BENEFICIO_2021,
  IFNULL(SUM(F.FORECAST_PROFIT), 0) AS INGRESOS_ESPERADOS_2022,
  CASE
    WHEN (BENEFICIO_2019 <= BENEFICIO_2020 AND BENEFICIO_2020 <= BENEFICIO_2021)
    THEN 'APTO'
    ELSE 'NO APTO'
  END AS CRECIMIENTO_19_20_21,
  CASE
    WHEN (INGRESOS_ESPERADOS_2022 <= BENEFICIO_2021)
    THEN 'APTO'
    ELSE 'NO APTO'
  END AS PREVISION_CRECIMIENTO_21_22,
  CASE
    WHEN (CRECIMIENTO_19_20_21='APTO') AND (PREVISION_CRECIMIENTO_21_22='APTO')
    THEN 'ACTIVAR CAMPAÑA'
    ELSE 'CAMPAÑA DENEGADA'
  END AS PUBLICIDAD

FROM ACCOUNTS AS A
LEFT JOIN SALES AS S
  ON A.ACCOUNT = S.ACCOUNT
LEFT JOIN FORECASTS AS F
  ON A.ACCOUNT = F.ACCOUNT
GROUP BY A.COUNTRY
ORDER BY A.COUNTRY;
```

3. Reflexión y sugerencia de estrategias

Para dar respuesta a las preguntas y establecer nuestras recomendaciones vamos a exponer los resultados de nuestro análisis.

1. En el caso de la petición del departamento de Recursos Humanos, el equipo de análisis de datos va a tener en cuenta dos criterios: En primer lugar, que el ejecutivo en su previsión para el año 2022 obtenga una tasa de variación con respecto al beneficio de 2021 de al menos el 50%. En segundo lugar, que la suma de los beneficios de los años 2019, 2020 y 2021 sea de al menos 3 millones de euros.

Si el ejecutivo es capaz de cumplir ambos requisitos promocionara de sueldo, si solo es capaz de cumplir uno de ellos, continuará en su puesto y si ninguno de los dos se cumple deberá ser despedido.

Esta ha sido la salida enviada al departamento de RRHH:

ACCOUNT_EXECUTIVE	RECURSOS_HUMANOS
AsherOchoa	CONTINUA
BarrettGriffin	DESPEDIDO
CalebCombs	DESPEDIDO
CalistaJohns	CONTINUA
CarynBright	CONTINUA
ChavaGutierrez	CONTINUA
ColeHartman	CONTINUA
DawnBoyer	DESPEDIDO
DonovanGardner	CONTINUA
EmilyWolfe	DESPEDIDO
GraceBurris	CONTINUA
GrayJoyner	PROMOCIONA
HamishBlake	DESPEDIDO
HarrisonYang	DESPEDIDO
HavivaLester	DESPEDIDO
HollyKirby	DESPEDIDO
HolmesMeadows	CONTINUA
IngaLancaster	PROMOCIONA
IraBuchanan	CONTINUA
KaiRogers	CONTINUA
LeandraEllison	PROMOCIONA
LeslieGentry	DESPEDIDO
LesterBrock	DESPEDIDO
MarthaDickerson	DESPEDIDO
MeganRosa	CONTINUA
MufutauSmith	DESPEDIDO
NasimMayo	DESPEDIDO
RamaFranco	CONTINUA
RileyWiley	DESPEDIDO
TeaganFrank	CONTINUA
WingRoberts	CONTINUA
XanderReed	DESPEDIDO
ZacheryWhitney	PROMOCIONA

- En el caso del departamento de publicidad nuestra propuesta se ha basado en encontrar un país con una expansión constante, es decir que año tras año haya dejado en la empresa un beneficio mayor y que en la previsión de 2022 no siga esta tendencia creciente y, por lo tanto, necesite un impulso.

Al aunar esas dos condiciones el país propuesto para la campaña de publicidad ha sido Italia. Esta ha sido la salida enviada al departamento de publicidad:

COUNTRY	CRECIMIENTO_19_20_21	PREVISION_CRECIMIENTO_21_22	PUBLICIDAD
Argentina	NO APTO	NO APTO	CAMPAÑA DENEGADA
Australia	NO APTO	APTO	CAMPAÑA DENEGADA
Austria	NO APTO	NO APTO	CAMPAÑA DENEGADA
France	NO APTO	NO APTO	CAMPAÑA DENEGADA
Germany	APTO	NO APTO	CAMPAÑA DENEGADA
India	APTO	NO APTO	CAMPAÑA DENEGADA
Italy	APTO	APTO	ACTIVAR CAMPAÑA
Japan	APTO	NO APTO	CAMPAÑA DENEGADA
Korea	NO APTO	NO APTO	CAMPAÑA DENEGADA
Mexico	NO APTO	NO APTO	CAMPAÑA DENEGADA
Portugal	NO APTO	NO APTO	CAMPAÑA DENEGADA
Singapore	NO APTO	NO APTO	CAMPAÑA DENEGADA
Spain	APTO	NO APTO	CAMPAÑA DENEGADA
Sweden	NO APTO	NO APTO	CAMPAÑA DENEGADA
Switzerland	APTO	NO APTO	CAMPAÑA DENEGADA
United Kingdom	APTO	NO APTO	CAMPAÑA DENEGADA
United States	NO APTO	NO APTO	CAMPAÑA DENEGADA

