

Project SQL from Womenderful Bootcamp
(Menganalisis Database Situs Fashion)
by Puri Tri Handayani

Terdapat 3 tabel pada database situs Fashion, yaitu:

1. **items** (id, name, gender, price, cost)
2. **sales_records** (id, user_id, item_id, purchased_at)
3. **users** (id, name, gender, age)

Pada tabel-tabel tersebut terdapat:

- **30 item terdaftar**
- **910 sales record** pada Juli 2018
- **325 user terdaftar**

Permintaan data no. 1:

Tim keuangan meminta angka total untuk pendapatan (revenue), laba dan persentase laba website pembelian untuk mencari tau gambaran besar dari jumlah penjualan dan laba yang didapatkan website tersebut.

Persentase laba dapat dikalkulasikan dengan formula sebagai berikut:

Total persentase laba = total laba / total pendapatan * 100

Transformasi total persentase dari integer menjadi desimal dengan formula:

Persentase laba= 1.0 * total laba / total pendapatan * 100

Code:

```
1 SELECT SUM(items.price) AS "total_revenue", SUM(items.price - items.cost) AS "total_laba",  
2 1.0 * SUM(items.price - items.cost) / SUM(items.price) * 100 AS "persentase_laba"  
3 FROM sales_records  
4 JOIN items  
5 ON sales_records.item_id = items.id;  
6
```

Output:

total_revenue	total_laba	persentase_laba
46779	31428	67.1839928172898

Analisis Hasil:

Hasilnya diperoleh total revenue dari seluruh item yang terjual sebesar **46779**, total keuntungan/laba sebesar **31428** dan persentase laba sebesar **67,18%**. Analisis lanjutan dapat dilihat item mana yang paling berpengaruh terhadap revenue dan mana yang tidak, serta item mana yang memiliki persentase laba terbesar dan terkecil.

Permintaan data no. 2:

Manakah 5 item yang berpengaruh paling besar terhadap revenue?

Code:

```
1 SELECT items.name, COUNT(*) AS "total_penjualan", items.price,
2    | | items.price * COUNT(*) AS "total_pemasukan"
3 FROM sales_records
4 JOIN items
5 ON sales_records.item_id = items.id
6 GROUP BY items.name, items.price
7 ORDER BY COUNT(*) * items.price DESC
8 LIMIT 5;
```

Output:

name	total_penjualan	price	total_pemasukan
jaket kulit	27	320	8640
mantel merah	25	270	6750
celana denim	61	60	3660
sepatu lari	35	75	2625
rok	52	40	2080

Analisis Hasil:

Output yang dihasilkan adalah 5 item yang menghasilkan pendapatan tertinggi secara berurutan (jaket kulit, mantel merah, celana denim, sepatu lari dan rok). Item teratas adalah "jaket kulit". Meskipun jumlah penjualannya tidak terlalu besar tapi harga satuannya sangat tinggi, yang membuatnya menjadi item teratas. Sebaliknya, “sepatu lari” dan “rok” memiliki harga yang lebih rendah tetapi menghasilkan keuntungan besar karena jumlah penjualan yang besar.

Permintaan data no. 3:

Berapakah laba dan persentase laba dari 5 item penghasil revenue teratas?

Code:

```
1 SELECT items.name, (items.price - items.cost) AS "laba",
2    1.0 * (items.price - items.cost) / (items.price) * 100 AS "persentase_laba",
3    COUNT(*) * (items.price - items.cost) AS "total_laba"
4 FROM sales_records
5 JOIN items
6 ON sales_records.item_id = items.id
7 WHERE items.id IN (
8     SELECT items.id
9     FROM sales_records
```

```
10 JOIN items
11 ON sales_records.item_id = items.id
12 GROUP BY items.id
13 ORDER BY COUNT(*) * items.price DESC
14 LIMIT 5
15 )
16 GROUP BY items.name, laba, persentase_laba
17 ORDER BY COUNT(*) * items.price DESC;
```

Output:

name	laba	persentase_laba	total_laba
jaket kulit	222	69.375	5994
mantel merah	195	72.22222222222222	4875
celana denim	34	56.66666666666667	2074
...
sepatu lari	47	62.66666666666667	1645
rok	32	80.0	1664

Analisis Hasil:

Di antara 5 item yang menghasilkan pendapatan teratas, "jaket kulit", "mantel merah", dan "celana denim" memiliki tingkat keuntungan yang lebih tinggi dan 2 lainnya memiliki tingkat keuntungan yang lebih rendah daripada yang ada di seluruh situs.

Permintaan data no. 4:

Manakah 5 item yang memberikan penghasilan laba tertinggi?

Code:

```
1 SELECT items.name, (items.price - items.cost) AS "laba",
2    1.0 * (items.price - items.cost) / (items.price) * 100 AS "persentase_laba",
3    COUNT(*) * (items.price - items.cost) "total_laba"
4 FROM sales_records
5 JOIN items
6 ON sales_records.item_id = items.id
7 GROUP BY items.name, laba, persentase_laba
8 ORDER BY total_laba DESC
9 LIMIT 5;
```

Output:

name	laba	persentase_laba	total_laba
jaket kulit	222	69.375	5994
mantel merah	195	72.22222222222222	4875
celana denim	34	56.66666666666667	2074
...
sweater	32	94.1176470588235	1856
rok	32	80.0	1664

Analisis Hasil:

Output yang dihasilkan adalah 5 item yang menghasilkan laba tertinggi secara berurutan (**jaket kulit, mantel merah, celana denim, sweater dan rok**). Dari 5 item tersebut, 4 item juga menempati posisi item pendapatan tertinggi. Sekarang dapat memutuskan untuk tetap menjual 5 item penghasil pendapatan tertinggi karena juga menguntungkan. Tapi **terkadang ada barang yang laku tapi untungnya sedikit**, sehingga dapat disimpulkan bahwa mungkin harus **mengubah taktik penjualan**. Item "sweater" menempati posisi No. 4 sebagai penghasil keuntungan. Tetapi persentase keuntungannya paling tinggi yaitu 94,12%. Artinya, kita bisa coba mendiskon barang

(masih menguntungkan) dan meningkatkan kuantitas penjualan untuk memaksimalkan total keuntungan.

Permintaan data no. 5:

Setelah mengetahui item-item yang paling berpengaruh terhadap pendapatan dan laba, Agar dapat mengalokasikan biaya marketing secara efektif, manakah 5 item yang paling berpengaruh terhadap penghasil pendapatan terendah?

Code:

```
1 SELECT items.name, COUNT(*) AS "jumlah_penjualan", items.price,  
2     || items.price * COUNT(*) AS "total_pendapatan"  
3 FROM sales_records  
4 JOIN items  
5 ON sales_records.item_id = items.id  
6 GROUP BY items.name, items.price  
7 ORDER BY total_pendapatan ASC  
8 LIMIT 5;  
9
```

Output:

name	jumlah_penjualan	price	total_pendapatan
kaos kaki	18	10	180
kaos abu-abu	9	23	207
topi	20	15	300
syal	17	18	306
sandal	21	18	378

Analisis Hasil:

Output yang dihasilkan adalah 5 item yang menghasilkan pendapatan terendah secara berurutan (kaos kaki, kaos abu-abu, topi, syal dan sandal). Untuk kuantitas penjualan kecil (terutama "kaos abu-abu"), stoknya dapat dikurangi untuk menghemat biaya penyimpanan.

Permintaan data no. 6:

Manakah 5 item dengan penghasilan laba terendah?

Code:

```
1 SELECT items.name, (items.price - items.cost) AS "laba",  
2     1.0 * (items.price - items.cost) / (items.price) * 100 AS "persentase_laba",  
3     COUNT(*) * (items.price - items.cost) "total_laba"  
4 FROM sales_records  
5 JOIN items  
6 ON sales_records.item_id = items.id  
7 GROUP BY items.name, laba, persentase_laba  
8 ORDER BY total_laba ASC  
9 LIMIT 5;
```

Output:

name	laba	persentase_laba	total_laba
kaos abu-abu	13	56.5217391304348	117
kaos kaki	8	80.0	144
kaos biru tua	11	55.0	220
syal	13	72.2222222222222	221
topi	12	80.0	240

Analisis Hasil:

Output yang dihasilkan adalah 5 item yang menghasilkan laba terendah secara berurutan (kaos abu-abu, kaos kaki, kaos biru tua, syal dan topi). Persentase laba dari “kaos abu-abu” dan “kaos biru tua” sangat rendah, sehingga perlu diambil langkah untuk barang-barang tersebut. Persentase laba dari “kaos kaki” tidak buruk meskipun keuntungannya sangat kecil. Langkah yang dapat dilakukan mungkin dapat memberikan diskon dengan penjualan 3 pasang item.

Permintaan data no. 7:

Berapa total laba dan persentase laba berdasarkan gender?

Code:

```
1 SELECT items.gender, SUM(items.price - items.cost) AS "total_laba", 1.0 * SUM(items.price - items.cost) / SUM(items.price) * 100 AS "persentase_laba"
2 FROM sales_records
3 JOIN items
4 ON sales_records.item_id = items.id
5 GROUP BY items.gender;
```

Output:

gender	total_laba	persentase_laba
0	3644	60.5818786367415
1	18450	70.8362128541811
2	9334	63.4189427911401

Analisis hasil:

Output pada kolom gender, 0 berarti netral, 1 berarti wanita dan 2 berarti pria. Kategori “wanita” menghasilkan laba dan persentase laba tertinggi yaitu sebesar 18450 dan 70,8%.

Permintaan data no. 8:

Berapakah jumlah pengguna aktif (pengguna yang membeli item setidaknya sekali) dan persentasenya?

Persentase pengguna aktif dapat dihitung dengan formula:

$(\text{jumlah pengguna aktif}) / (\text{jumlah total pengguna}) * 100$

Code:

```
1 SELECT COUNT(DISTINCT(sales_records.user_id)) AS "jumlah_pengguna_aktif", 1.0 * 100 * COUNT(DISTINCT(sales_records.user_id)) / (  
2     SELECT COUNT(*)  
3     FROM users  
4 )  
5     AS "persentase_pengguna_aktif"  
6 FROM sales_records  
7 JOIN users  
8 ON sales_records.user_id = users.id;  
9
```

Output:

jumlah_pengguna_aktif	persentase_pengguna_aktif
283	87.0769230769231

Analisis Hasil:

Jumlah pengguna aktif yaitu sebanyak **283** dari 325 pengguna yang mendaftar disitus belanja, atau sebesar **87,08%**. Untuk mendapatkan lebih banyak keuntungan, dapat dilakukan pendekatan terhadap pengguna yang tidak aktif, misalnya dengan mengirim email yang menarik pengguna untuk berbelanja atau membuat pengguna yang sudah aktif untuk berbelanja lebih sering atau membeli barang yang lebih mahal.

Permintaan data no. 9:

Berapakah frekuensi rata-rata pembelian, pengeluaran rata-rata, dan pengeluaran rata-rata per pembelian pengguna aktif?

- Rumus frekuensi rata-rata pembelian:
 $1.0 * (\text{Jumlah total pembelian di situs}) / (\text{Jumlah pengguna aktif})$
- Rumus Pengeluaran rata-rata:
 $1.0 * (\text{Jumlah harga item yang terjual}) / (\text{Jumlah pengguna aktif})$
- Rumus Pengeluaran rata-rata per pembelian:
 $1.0 * (\text{Belanja rata-rata}) / (\text{Frekuensi rata-rata pembelian})$

Code:

```
1 SELECT 1.0 * COUNT(*)/COUNT(DISTINCT(sales_records.user_id)) AS "frek_mean_pembelian",  
2     1.0 * SUM(items.price) / COUNT(DISTINCT(sales_records.user_id)) AS "mean_pengeluaran",  
3     1.0 * (1.0 * SUM(items.price) / COUNT(DISTINCT(sales_records.user_id))) /  
4     (1.0 * COUNT(*)/COUNT(DISTINCT(sales_records.user_id))) AS mean_pengeluaran_per_pembelian  
5 FROM sales_records  
6 JOIN items  
7 ON sales_records.item_id = items.id;
```

Output:

frek_mean_pembelian	mean_pengeluaran	mean_pengeluaran_per_pembelian
3.21554770318021	165.296819787986	51.4054945054945

Analisis hasil:

Frekuensi rata-rata pembelian yaitu 3.22 kali, pengeluaran rata-rata sebesar \$165,3, dan pengeluaran rata-rata per pembelian dalam sebulan sebesar \$51,4. Sebagai analisis lanjutan dapat dilihat list pengguna aktif yang paling sering berbelanja dan menghabiskan lebih dari harga rata-rata.

Permintaan data no. 10:

Dapatkan list users yang melakukan pembelian lebih dari rata-rata!

Code:

```
1 SELECT users.name, COUNT(sales_records.user_id) AS "jumlah_pembelian"
2 FROM sales_records
3 JOIN users
4 ON sales_records.user_id = users.id
5 GROUP BY users.name
6 HAVING jumlah_pembelian > (
7     SELECT 1.0 * COUNT(*)/COUNT(DISTINCT(sales_records.user_id)) AS "frek_mean_pembelian"
8     FROM sales_records
9     JOIN users
10    ON sales_records.user_id = users.id
11 )
12 ORDER BY jumlah_pembelian DESC;
```

Output:

name	jumlah_pembelian
Hobie	14
Burch	12
Sam	12

dst.

Analisis hasil:

Untuk customer yang sering melakukan pembelian diatas rata-rata, dapat diberikan **diskon spesial** supaya mereka tetap setia berbelanja.