**ASSIGNMENT CHAPTER 1**

**SQL FOUNDATION**

Nama : Agnes Monica Puspitaningtyas

Kompi : 6

SC ID : SC60619

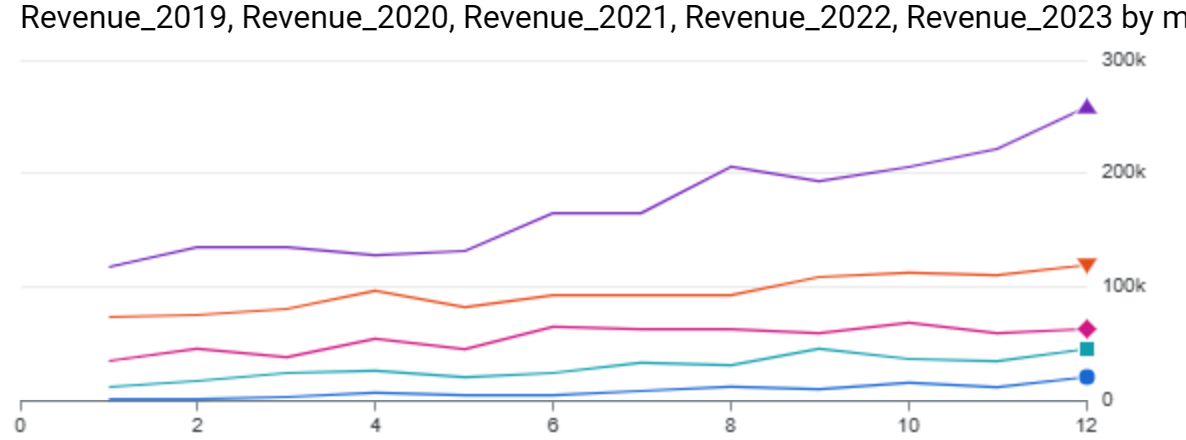
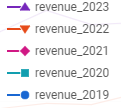
Analisis data memiliki peran yang penting untuk mengevaluasi kebijakan bisnis yang akan dijalankan. Tujuan dari analisis data adalah untuk menemukan informasi penting dan relevan dari data untuk mengambil keputusan terbaik dan informatif. Sebagai *data analyst*, kita diharuskan untuk fokus pada tujuan, yaitu membantu perusahaan membuat keputusan berdasarkan data yang ada. Untuk dapat menganalisis data, diperlukan sebuah *tools* untuk membantu pekerjaan seorang *data analyst*, yaitu SQL. Dalam hal ini, saya akan menganalisis sebuah dataset *e-commerce* The Look untuk mengetahui performa penjualan dan identifikasi pelanggan beserta perilakunya dalam melakukan transaksi.

The Look merupakan situs *E-commerce* fiksi yang diciptakan oleh Looker team. Dataset ini memuat data dan tabel berupa:

|  |  |
| --- | --- |
| **Nama Tabel** | **Nama Kolom** |
| users | id, first\_name, last\_name, email, age, gender, state, street\_address, postal\_code, city, country, latitude, longitude, traffic source, created\_at |
| order\_items | id, order\_id, user\_id, product\_id, inventory\_item\_id, status, created\_at, shipped\_at, delivered\_at, returned\_at, sale\_price |
| distribution\_center | id, name, latitude, longitude |
| inventory\_items | id, product\_id, created\_at, sold\_at, cost, product\_category, product\_name, product\_brand, product\_retail\_price, product\_department, product\_sku, product\_distribution\_center\_id |
| products | id, cost, category, name, brand, retail\_price, department, sku, distribution\_center\_id |
| orders | order\_id, user\_id, status, gender, created\_at, returned\_at, shipped\_at, delivered\_at, num\_of\_item |
| events | id, user\_id, sequence\_number, session\_id, created\_at, ip\_address, city, state, postal\_code, browser, traffic\_source, url, event\_type |

Berikut *insight* yang didapatkan:

1. **Bagaimana pendapatan *e-commerce* pada setiap tahun?**

1. Pendapatan pada tahun dari tahun 2019 hingga 2023 terus mengalami peningkatkan. Hal ini menunjukkan hasil yang signifikan perusahaan dalam mengembangkan bisnisnya sehingga mengalami kemajuan yang pesat.
2. Pendapatan melonjak pada bulan Juli dan Desember. Hal ini dikarenakan Juli merupakan musim panas di beberapa wilayah sehingga orang cenderung lebih banyak berbelanja. Desember merupakan akhir tahun dimana banyak orang sedang menikmati liburan Natal dan tahun baru sehingga orang cenderung berbelanja pula pada bulan ini. Selain itu, promosi dan diskon menjadi faktor konsumen menjadi lebih sering melakukan transaksi.

*Call to action*: perusahaan perlu menyesuaikan strategi pemasaran dengan memberikan promosi dan diskon, meningkatkan persediaan barang, dan meningkatkan interaksi dengan pelanggan seperti melakukan *live* agar membangun koneksi yang lebih dekat dengan audiens, serta membuat konten yang bersifat menarik dan persuasif untuk dibagikan ke media sosial.

Query:

SELECT

  EXTRACT(MONTH FROM oi.Created\_at) AS months,

  ROUND(SUM(CASE WHEN EXTRACT(YEAR FROM oi.created\_at) = 2019 THEN oi.sale\_price \* o.num\_of\_item ELSE 0 END), 2) AS Revenue\_2019,

  ROUND(SUM(CASE WHEN EXTRACT(YEAR FROM oi.created\_at) = 2020 THEN oi.sale\_price \* o.num\_of\_item ELSE 0 END), 2) AS Revenue\_2020,

  ROUND(SUM(CASE WHEN EXTRACT(YEAR FROM oi.created\_at) = 2021 THEN oi.sale\_price \* o.num\_of\_item ELSE 0 END), 2) AS Revenue\_2021,

  ROUND(SUM(CASE WHEN EXTRACT(YEAR FROM oi.created\_at) = 2022 THEN oi.sale\_price \* o.num\_of\_item ELSE 0 END), 2) AS Revenue\_2022,

  ROUND(SUM(CASE WHEN EXTRACT(YEAR FROM oi.created\_at) = 2023 THEN oi.sale\_price \* o.num\_of\_item ELSE 0 END), 2) AS Revenue\_2023

FROM bigquery-public-data.thelook\_ecommerce.order\_items oi

INNER JOIN `bigquery-public-data.thelook\_ecommerce.orders` AS o ON oi.order\_id = o.order\_id

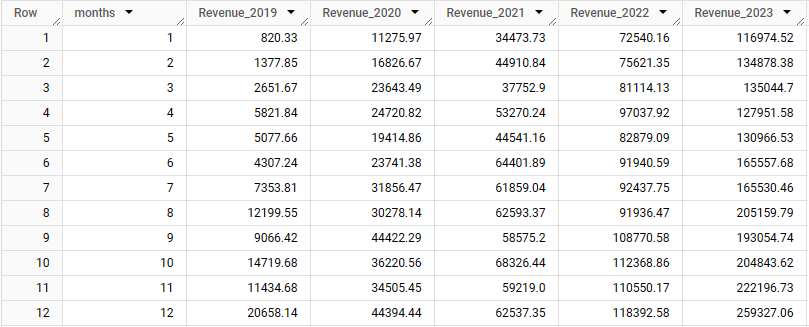
WHERE oi.status = 'Complete'

  AND EXTRACT(YEAR FROM oi.Created\_at) BETWEEN 2019 AND 2023

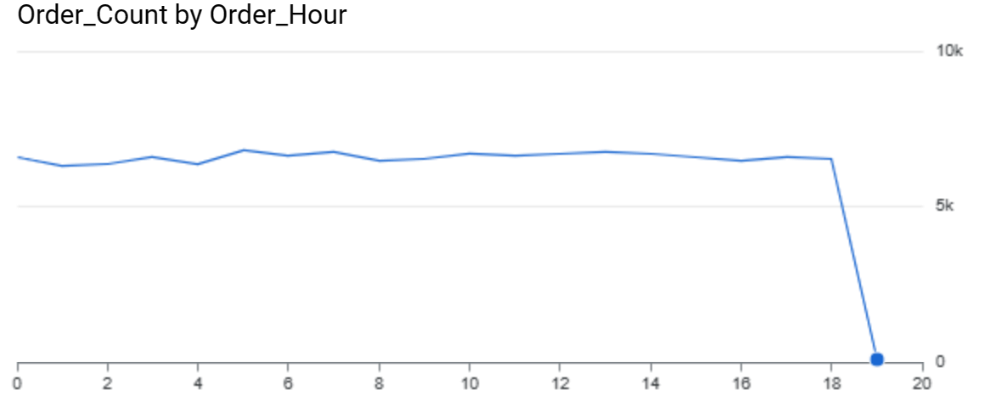
GROUP BY months

ORDER BY months

Output:



1. **Kapan waktu pemesanan yang paling sibuk?**



1. Waktu pemesanan rata-rata konstan, tetapi pelanggan paling banyak melakukan transaksi pada saat jam 5 pagi sebanyak 6.779 pesanan dan 1 siang sebanyak 6.772 pesanan. Sementara itu, waktu paling sedikit bertransaksi adalah pukul 7 malam.
2. Hal dikarenakan kebiasan pelanggan yang lebih aktif berbelanja pada saat pagi hari sebelum memulai hari kerja atau jam istirahat makan siang. Sementara itu, pada pukul 7 malam sebagian besar orang sedang beristirahat setelah hari kerja sehingga aktivitas pembelian *online* cenderung menurun.

*Call to action*: Perusahaan dapat mengoptimalkan strategi promosi pada jam-jam puncak, seperti memberikan diskon/promosi dan melakukan *live* pada jam tersebut.

* Query:

SELECT

  EXTRACT(HOUR FROM created\_at) AS Order\_Hour,

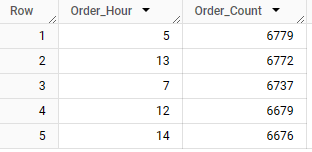
  COUNT(order\_id) AS Order\_Count

FROM bigquery-public-data.thelook\_ecommerce.orders

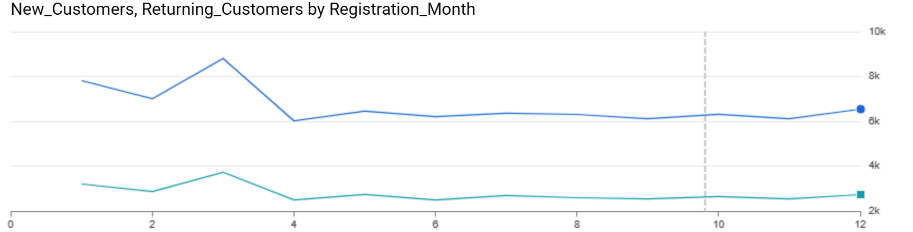
GROUP BY order\_hour

ORDER BY order\_count DESC

* Output:



1. **Bagaimana tren pelanggan baru yang melakukan transaksi di bulan yang sama?**

1. Bulan dengan registrasi pengguna paling banyak berada pada bulan Maret, yaitu sebanyak 7.800 pengguna dan yang melakukan transaksi sebanyak 3.707 pelanggan. Sementara itu, bulan dengan registrasi pengguna paling sedikit yaitu bulan April sebanyak 6.140 pengguna dan yang melakukan transaksi sebanyak 2.497 pelanggan.
2. Hal ini dikarenakan pada bulan Maret terdapat promosi secara besar-besaran untuk pengguna yang baru melakukan transaksi. Ketika seluruh kebutuhannya serta hasrat untuk berbelanja sudah terpenuhi pada bulan tersebut, animo pelanggan turun pada bulan setelahnya karena pada bulan April perusahaan tidak lagi melakukan promosi serupa.

*Call to action*: perusahaan dapat terus memanfaatkan strategi promosi pada bulan Maret untuk menarik lebih banyak pelanggan baru. Namun, perusahaan perlu mengadakan promosi yang berkelanjutan pada bulan setelahnya agar dapat meningkatkan aktivitas dan transaksi pelanggan. Perusahaan perlu membangun komunikasi dan melakukan penawaran yang disesuaikan dengan preferensi pelanggan.

* Query:

SELECT

  EXTRACT(MONTH FROM u.created\_at) AS Registration\_Month,

  COUNT(DISTINCT u.id) AS New\_Customers,

  COUNT(DISTINCT CASE WHEN EXTRACT(MONTH FROM o.created\_at) = EXTRACT(MONTH FROM delivered\_at) THEN user\_id END) AS Returning\_Customers

FROM bigquery-public-data.thelook\_ecommerce.orders AS o

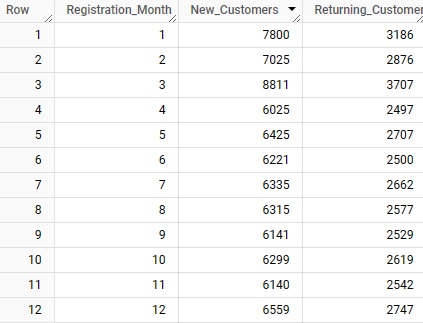
INNER JOIN bigquery-public-data.thelook\_ecommerce.users AS u

ON o.user\_id = u.id

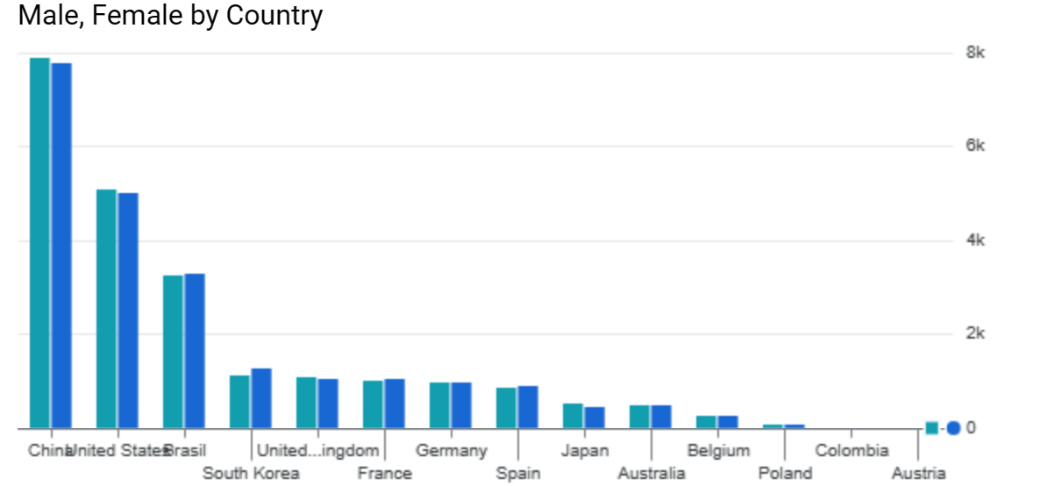
GROUP BY Registration\_Month

ORDER BY Registration\_Month

* Output:



1. **Bagaimana profil pelanggan pada setiap negara?**



1. Pelanggan paling banyak berasal dari negara China dengan total sebanyak 15.665 pelanggan yang terdiri dari 7.887 pelanggan laki-laki dan 7.778 pelanggan perempuan. Sementara itu, pelanggan paling sedikit berasal dari negara Austria dengan total sebanyak 2 pelanggan laki-laki.
2. Secara keseluruhan, proporsi pelanggan antara laki-laki dan wanita pada setiap negara hampir sama.

*Call to action*: Perusahaan dapat mengembangkan strategi pemasaran yang unik dan menarik berdasarkan preferensi dan kebutuhan antara laki-laki dan perempuan.

Query:

SELECT

  DISTINCT Country,

  SUM(CASE WHEN u.gender = 'M' THEN 1 else null END) AS Male,

  SUM(CASE WHEN u.gender = 'F' THEN 1 else null END) as Female,

  COUNT(oi.user\_id) AS Total\_Pelanggan

FROM bigquery-public-data.thelook\_ecommerce.users AS u

INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items AS oi

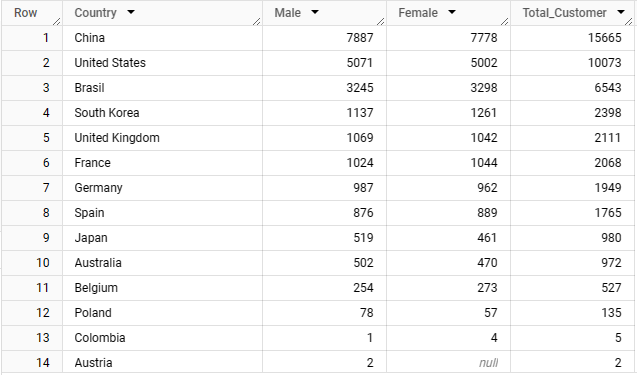
ON u.id = oi.user\_id

WHERE oi.status = 'Complete'

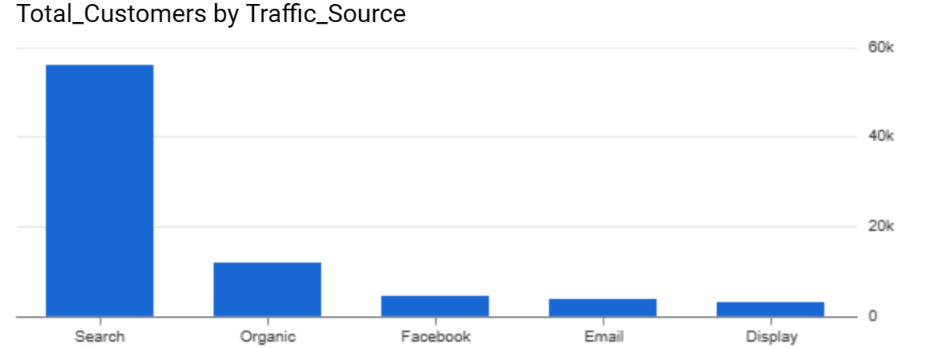
GROUP BY Country

ORDER BY Total\_Pelanggan DESC

Output:



1. **Apa sumber *traffic* yang lebih banyak diakses oleh pelanggan?**



1. Pelanggan paling banyak datang ke situs web melalui mesin pencarian, seperti Google, Mozila Firefox, Bing, dll sebanyak 56.009 pelanggan. Kemudian, diikuti oleh organik, yang merupakan pencarian yang muncul secara alami di mesin pencari sebanyak 12.080, diikuti oleh Facebook sebanyak 4.751, Email sebanyak 4.003, dan *display* (iklan) sebanyak 3.223.
2. Hal ini dikarenakan karena mesin pencarian menjadi alat utama yang digunakan oleh banyak orang untuk mencari informasi dan produk secara *online* sehingga membuat situs web lebih mudah ditemukan dan diakses oleh pengguna yang ingin mencari informasi.

*Call to action*: Perusahaan mengoptimalkan situs web untuk mesin pencari, meningkatkan kehadiran di Facebook, dan mulai menyusun strategi pemasaran Email secara efektif untuk memaksimalkan *traffic.*

Query:

SELECT

  u.traffic\_source AS Traffic\_Source,

  COUNT(DISTINCT oi.user\_id) AS Total\_Customers,

FROM bigquery-public-data.thelook\_ecommerce.users AS u

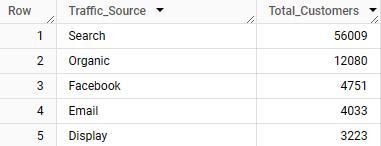
INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items AS oi

ON u.id = oi.user\_id

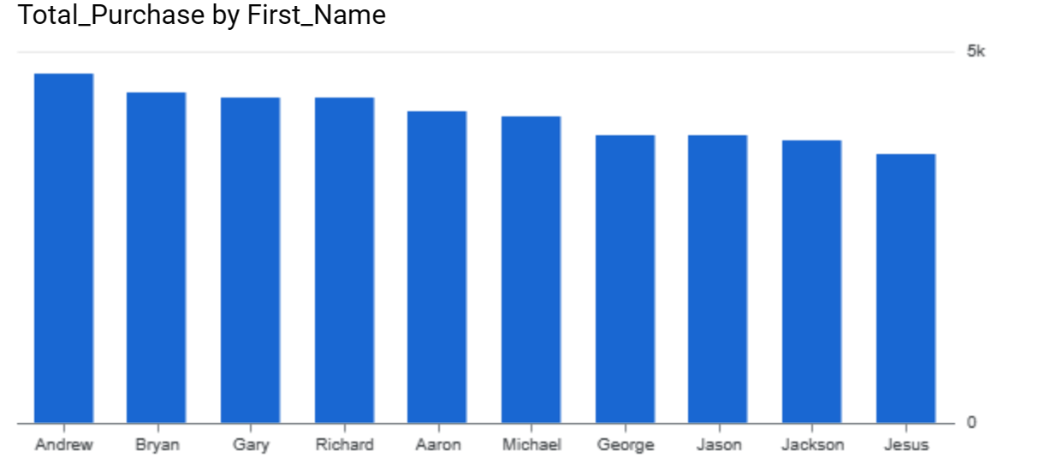
GROUP BY Traffic\_Source

ORDER BY Total\_Customers DESC

* Output:



1. **Siapa saja pelanggan yang melakukan pembelian paling besar?**



1. Pelanggan yang melakukan pembelian paling besar adalah Andrew Wilson yang berada di negara Poland dengan total pembayaran sebesar 4.702,3.
2. Andrew Wilson memiliki total pembayaran paling besar karena melakukan pembelian produk dalam jumlah besar atau membeli barang dengan harga yang tinggi.

*Call to action*: Agar pelanggan menjadi loyal, perusahaan perlu memberikan layanan pelanggan layanan yang efisien, mempertahankan produk yang berkualitas, serta menawarkan program loyalitas. Selain itu, diskon dengan nominal yang lebih tinggi bisa dipertimbangkan untuk diberikan kepada pelanggan yang mencapai minimal pembayaran tertentu.

* Query:

SELECT

  DISTINCT u.id AS user\_id,

  u.first\_name AS First\_Name,

  u.last\_name AS Last\_Name,

  u.country AS Country,

  ROUND(SUM(o.num\_of\_item \* oi.sale\_price),1) AS Total\_Purchase

FROM bigquery-public-data.thelook\_ecommerce.users as u

INNER JOIN bigquery-public-data.thelook\_ecommerce.orders as o

ON u.id = o.user\_id

INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items as oi

ON  o.order\_id = oi.order\_id

WHERE o.status = 'Complete'

GROUP BY  user\_id, First\_Name, Last\_Name, Country

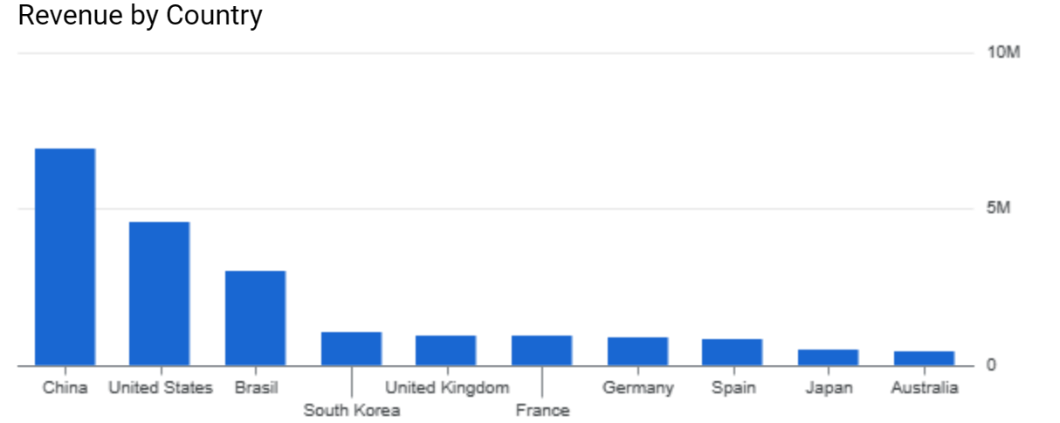
ORDER BY Total\_Purchase DESC

LIMIT 10

* Output:



1. **Negara apa yang memberikan pendapatan paling besar?**



1. China memberikan pendapatan untuk *e-commerce* paling besar dengan total pendapatan sebesar 6.936.191,9, diikuti oleh United States dan Brasil.
2. Hal ini dikarenakan China memiliki populasi yang sangat besar, pertumbuhan pasar dan ekonomi yang pesat, dan infrastruktur teknologi yang canggih sehingga banyak warga China yang melakukan transaksi melalui *e-commerce*.

*Call to action*: Perusahaan dapat meningkatkan pendapatan *e-commerce* dengan melakukan ekspansi global, terus berinovasi pada teknologi, dan melakukan penguatan kemitraan sembari memantau perubahan pasar dan kebutuhan konsumen supaya negara-negara yang menghasilkan pendapatan besar dapat meningkatkan transaksinya terhadap *e-commerce*

* Query:

SELECT

    DISTINCT u.country as Country,

    ROUND(SUM(oi.sale\_price \* o.num\_of\_item),1) AS Revenue

FROM bigquery-public-data.thelook\_ecommerce.users as u

INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items as oi

ON u.id = oi.user\_id

INNER JOIN bigquery-public-data.thelook\_ecommerce.orders as o

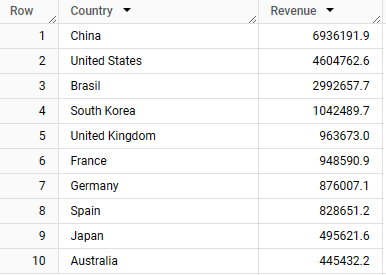
ON oi.order\_id = o.order\_id

GROUP BY Country

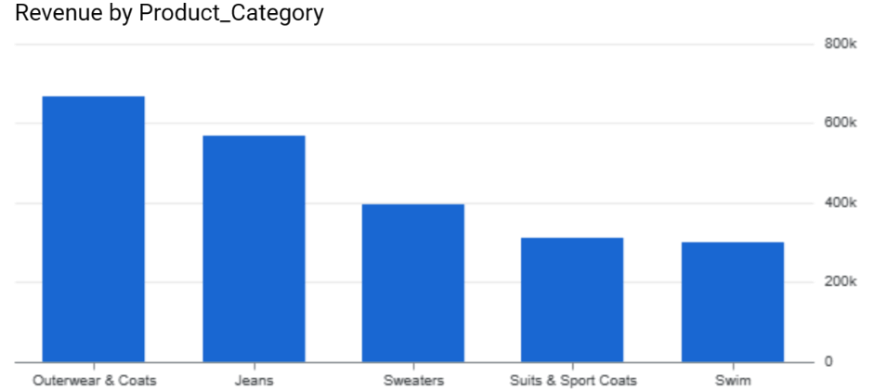
ORDER BY Revenue DESC

LIMIT 10

* Output:



1. **Apa kategori produk yang menghasilkan pendapatan paling besar?**



1. Kategori yang memberikan pendapatan paling besar adalah *Outwear & Coats*, diikuti oleh *Jeans*, *Sweaters*, *Suit* & *Sport Coats*, dan *Swim*.
2. Kategori *Outwear & Coats* kemungkinan besar terkait dengan musim tertentu, seperti musim dingin atau hujan sehingga konsumen membutuhkan penggunaan mantel atau pakaian luar.

*Call to action*: Untuk mempertahankan keuntungan ini, perusahaan dapat meningkatkan stok dan strategi promosi atau diskon ketika musim hujan atau musim dingin akan/sedang berlangsung.

* Query:

SELECT

  category AS Product\_Category,

  SUM(num\_of\_item) AS Quantity,

  ROUND(SUM(sale\_price \* num\_of\_item),1) AS Revenue

FROM bigquery-public-data.thelook\_ecommerce.order\_items AS oi

INNER JOIN bigquery-public-data.thelook\_ecommerce.orders AS o

ON oi.order\_id = o.order\_id

INNER JOIN bigquery-public-data.thelook\_ecommerce.products AS p

ON p.id = oi.product\_id

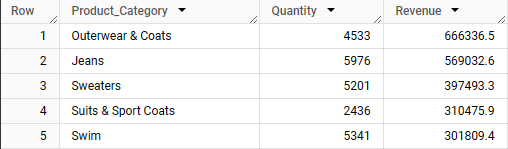
WHERE oi.status = 'Complete'

GROUP BY Category

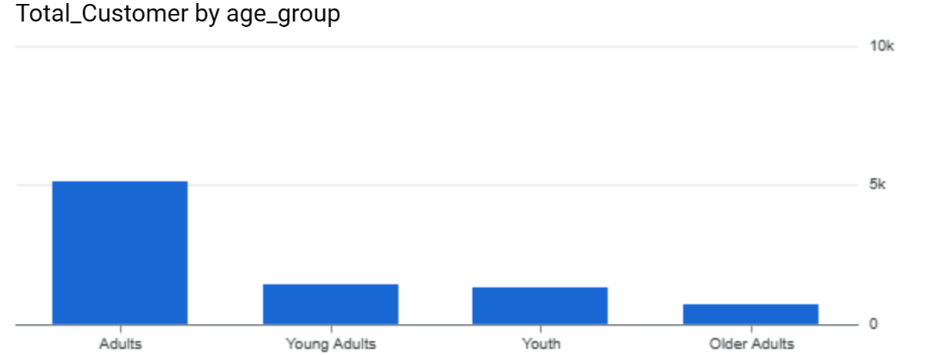
ORDER BY Revenue DESC

LIMIT 5

* Output:



1. **Bagaimana kategori umur pelanggan pada kategori produk yang paling laris?**



1. Pada kategori produk yang paling laris, yaitu *Outerwear & Coats*, pelanggan terbanyak berada pada kelompok umur *Adults* sebanyak 5.120 pelanggan, diikuti kelompok umur *Young Adults* sebanyak 1.433 pelanggan, kemudian kelompok umur *Youth* sebanyak 1.302 pelanggan, dan kelompok umur *Older Adults* sebanyak 345 pelanggan.
2. Pakaian *outerwear & coats* menjadi populer di kalangan umur yang lebih dewasa, hal ini dikarekanakan kebutuhan yang lebih besar untuk pakaian *outer* atau *fashion*. Sementara itu untuk kelompok umur *adults* dan *young adults*, kemungkinan mereka lebih peka terhadap tren mode dan gaya, mereka cenderung membeli produk *outerwear & coats* untuk menjaga penampilan mereka di musim dingin.

*Call to action*: Perusahaan dapat mengarahkan strategi pemasaran dan promosi untuk lebih tepat sasaran, seperti di kelompok umur *adults* dan *young adults* dan memenuhi kebutuhan pelanggan yang dominan pada kategori tersebut.

* Query:

SELECT

  CASE

    WHEN u.age < 12 THEN 'Children'

    WHEN u.age BETWEEN 12 AND 20 THEN 'Youth'

    WHEN u.age BETWEEN 21 AND 30 THEN 'Young Adults'

    WHEN u.age BETWEEN 31 AND 65 THEN 'Adults'

    WHEN u.age > 65 THEN 'Older Adults'

    END AS age\_group,

  COUNT(DISTINCT u.id) AS Total\_Pelanggan

FROM bigquery-public-data.thelook\_ecommerce.users AS u

INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items AS oi

ON u.id = oi.user\_id

INNER JOIN bigquery-public-data.thelook\_ecommerce.inventory\_items AS ii

ON ii.product\_id = oi.product\_id

WHERE ii.product\_category = 'Outerwear & Coats'

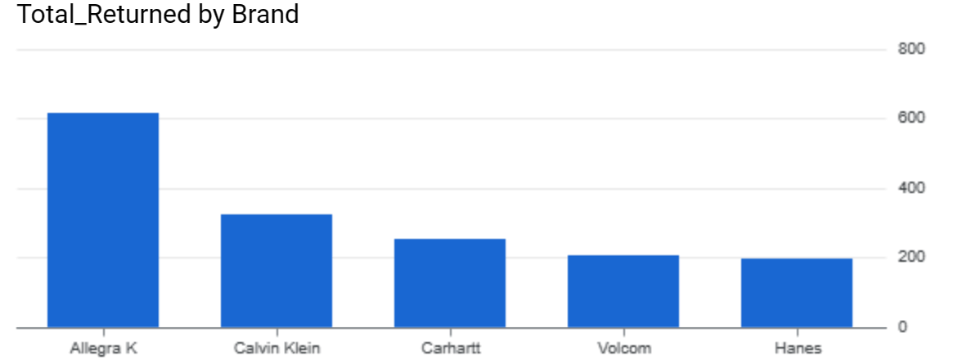
GROUP BY age\_group

ORDER BY Total\_Pelanggan DESC

* Output:



1. ***Brand* apa yang paling sering dikembalikan oleh pelanggan?**



1. Brand yang sering dikembalikan oleh pelanggan adalah *brand clothing*, seperti Allegra K sebanyak 615 unit, diikuti oleh Calvin Klein sebanyak 324 unit, dan Carhartt sebanyak 252 unit.
2. Hal ini dikarenakan pakaian yang tiba tidak sesuai dengan ukuran yang diinginkan oleh pelanggan maupun kualitas produk yang tidak sesuai dengan ekspektasi pelanggan.

*Call to action*: Perusahaan dapat meninjau kembali kualitas produk yang dihasilkan, dan meningkatkan detail deskripsi produk agar barang lebih sesuai dan mengurangi tingkat pengembalian produk.

* Query:

SELECT

  DISTINCT p.brand AS Brand,

  COUNT(o.num\_of\_item) AS Total\_Returned

FROM bigquery-public-data.thelook\_ecommerce.products as p

INNER JOIN bigquery-public-data.thelook\_ecommerce.order\_items as oi

ON p.id = oi.product\_id

INNER JOIN bigquery-public-data.thelook\_ecommerce.orders as o

ON oi.order\_id = o.order\_id

WHERE o.status = 'Returned'

GROUP BY Brand

ORDER BY Total\_Returned DESC

LIMIT 5

* Output:

