**TheLook E-Commerce Data Analysis Using SQL**

**Pendahuluan**

Dalam proyek ini, saya menggunakan SQL untuk menganalisis data e-commerce dari TheLook. TheLook adalah toko pakaian e-commerce yang dikembangkan oleh tim Google Looker. Data e-commerce ini dihosting di Google Big Query dan berisi informasi tentang pelanggan, produk, pesanan, logistik, acara web, dan kampanye pemasaran digital. Konten dataset ini disediakan untuk praktisi industri untuk tujuan penemuan produk, pengujian, dan evaluasi.

Pada saat analisis ini dilakukan, toko e-commerce mencatat data transaksi penjualan dari Agustus 2019 hingga September 2023.

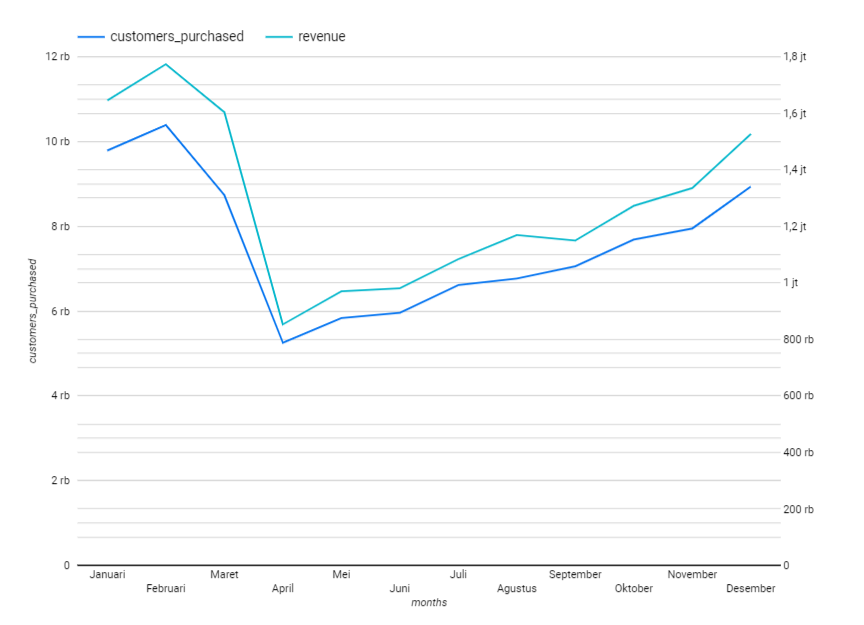
**Tujuan**

Tujuan dari analisis ini adalah untuk memahami dan menjawab beberapa pertanyaan bisnis terkait performa marketplace Look e-commerce, mendapatkan wawasan, dan memberikan beberapa rekomendasi untuk meningkatkan pendapatan.

Kita akan fokus pada pesanan yang tidak dibatalkan atau dikembalikan karena kita akan berurusan dengan metrik yang terkait dengan pendapatan.

**Menjawab Pertanyaan Bisnis**

1. **Bagaimana trend pendapatan pada e-commerce setiap bulannya ? Apakah tinggi atau rendah dibandingkan dengan bulan lalu ?**
2. Bagaimana proyeksi pendapatan tiap bulannya ?
3. Adakah bulan tertentu yang mendomisasi penjualan tiap tahunnya? Mengapa hal tersebut dapat terjadi?



**Insight:**

1. Dari hasil Analisa diperoleh fakta bahwa pendapatan pada perusahaan, mengalami trend naik-turun dari bulan ke bulan.
2. Bulan Februari menjadi puncak keuntungan dan jumlah pembelian customer. Hal ini bisa terjadi dikarenakan perayaan valentine bersamaan dengan suksesnya event yang dilakukan.
3. Bulan Januari dan Desember juga menjadi penyumbang keuntungan dan jumlah pembelian customer yang masif. Hal ini bisa terjadi dikarenakan adanya perayaan natal, tahun baru, dan musim liburan sekolah.
4. Bulan April menjadi bulan yang mencatatkan keuntungan dan jumlah pembelian customer paling rendah dibandingkan dengan bulan lainnya.

**Source code :**

SELECT

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

ROUND(SUM(oi.sale\_price \* o.num\_of\_item), 2) AS revenue,

COUNT(DISTINCT oi.order\_id) AS order\_count,

COUNT(DISTINCT oi.user\_id) AS customers\_purchased

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

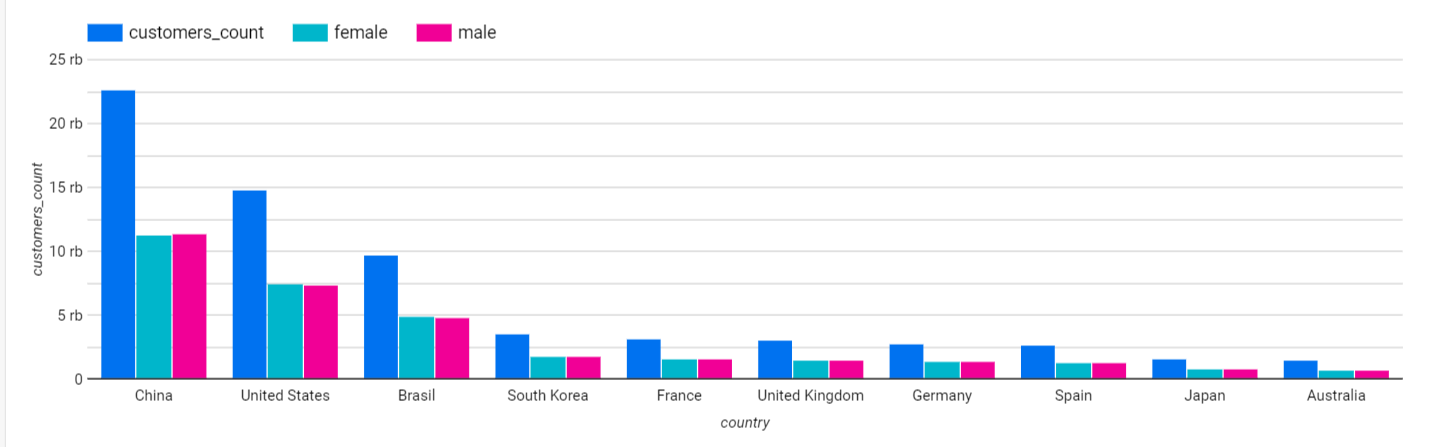
WHERE oi.status NOT IN ('Cancelled', 'Returned')

GROUP BY months

ORDER BY months ASC;

**Rekomendasi :**

1. Dengan demikian perusahaan perlu mempertahankan atau meningkatkan keuntungan dan jumlah pembelian customer pada bulan Januari, Februari, dan Desember dengan mengadakan berbagai event-event yang menarik.
2. Perusahaan juga dapat mengadakan berbagai event lain untuk meningkatkan keuntungan dan jumlah pembelian customer pada bulan-bulan lainnya.
3. **Siapa customer kita ? Dari negara mana, kelompok jenis kelaminnya, dan usia customer ?**
4. Darimana saja pelanggan kita berasal ?
5. Kelompok jenis kelamin mana yang memberikan ?
6. Berapa rentang usia customer kita ?
7. **Customer berdasarkan negara**



**Insight :**

1. Mayoritas customer kita berasal dari Cina, Amerika Serikat, dan Brasil.
2. Proporsi pendapatan yang dihasilkan oleh masing-masing gender hampir sama di setiap negara.

**Source Code :**

WITH customers AS (

  SELECT

    DISTINCT oi.user\_id,

    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,

    u.country AS country

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

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

    ON oi.user\_id = u.id

    WHERE oi.status not in ('Cancelled', 'Returned')

    GROUP BY user\_id, country

    )

SELECT

  country,

  COUNT(DISTINCT user\_id) AS customers\_count,

  COUNT(female) AS female,

  COUNT(male) AS male

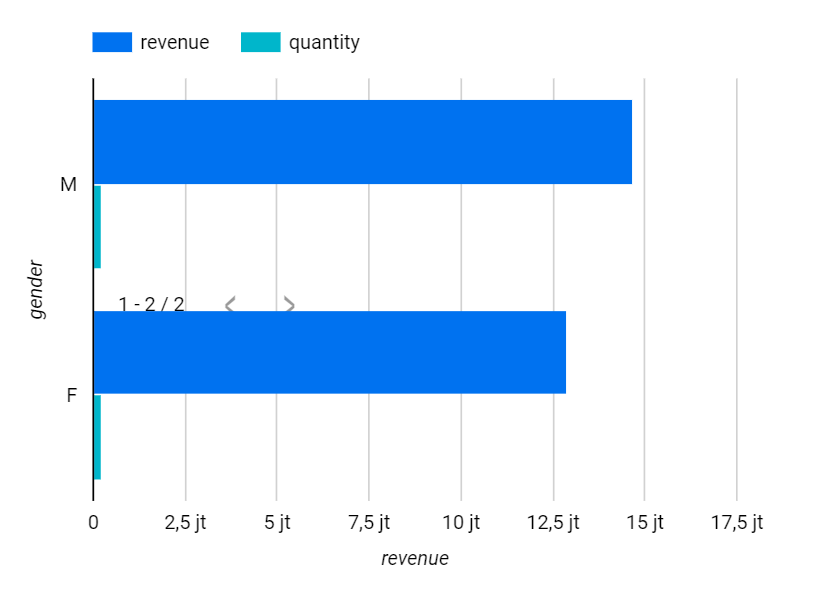
FROM customers

GROUP BY country

ORDER BY customers\_count DESC;

**Rekomendasi :**

1. Perusahaan dapat mengadakan berbagai event khusus pada setiap negara sesuai karakteristik demografinya sebagai cara untuk memperkuat branding dan menarik customer baru. Negara dengan jumlah customer lebih banyak, lebih baik diprioritaskan untuk menjaga loyalitas costumer, namun hal ini tidak berarti mengabaikan negara dengan jumlah customer yang rendah.
2. **Customer berdasarkan jenis kelamin**



**Insight :**

1. Customer laki-laki memberikan lebih banyak pendapatan dan jumlah pembelian dibandingkan dengan customer perempuan. Namun, jumlah pembelian antara laki-laki dan perempuan memiliki jumlah yang hampir sama.

**Rekomendasi :**

1. Basis pelanggan yang terbagi rata memberikan keuntungan untuk dapat melakukan berbagai strategi pemasaran yang fokus kepada laki-laki atau perempuan atau bahkan keduanya, tanpa khawatir nilai resiko yang terlalu besar.

**Source Code :**

SELECT

  o.gender,

  SUM(oi.sale\_price\*o.num\_of\_item) revenue,

  SUM(o.num\_of\_item) quantity

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

LEFT JOIN `bigquery-public-data.thelook\_ecommerce.orders` o

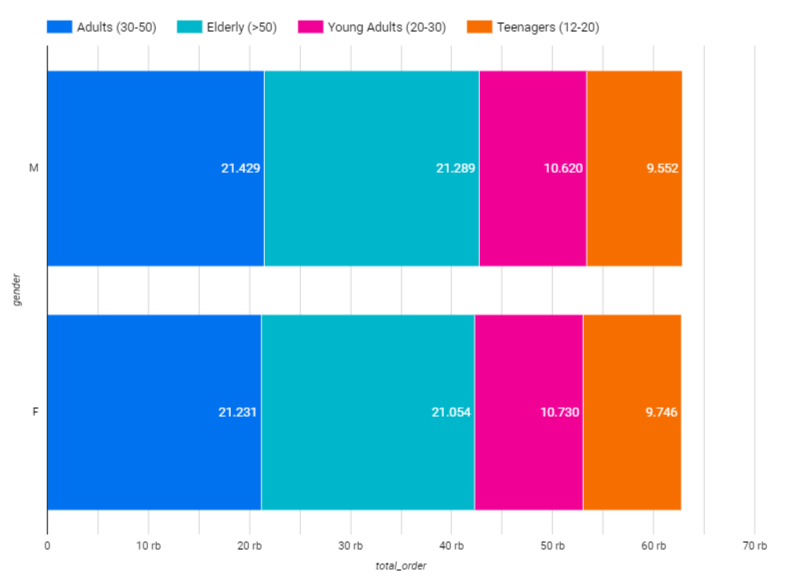
ON oi.order\_id = o.order\_id

WHERE oi.status NOT IN ('Cancelled','Returned')

GROUP BY 1

ORDER BY 2

1. **Customer berdasarkan grup usia**



**Insight :**

1. Mayoritas customer kita adalah lansia (>50) dan orang dewasa (30-50). Sisanya adalah orang dewasa muda (20-30) dan remaja (12-20).
2. Jumlah customer laki-laki dan perempuan menunjukkan jumlah yang hampir sama di setiap grup usia.

**Source Code:**

SELECT

 CASE

  WHEN u.age BETWEEN 12 AND 20 THEN 'Teenagers (12-20)'

  WHEN u.age BETWEEN 20 AND 30 THEN 'Young Adults (20-30)'

  WHEN u.age BETWEEN 30 AND 50 THEN 'Adults (30-50)'

  WHEN u.age > 50 THEN 'Elderly (>50)'

 END AS age\_segment,

 U.gender,

 COUNT(O.num\_of\_item) AS total\_order

 FROM

 bigquery-public-data.thelook\_ecommerce.users U

 LEFT JOIN

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

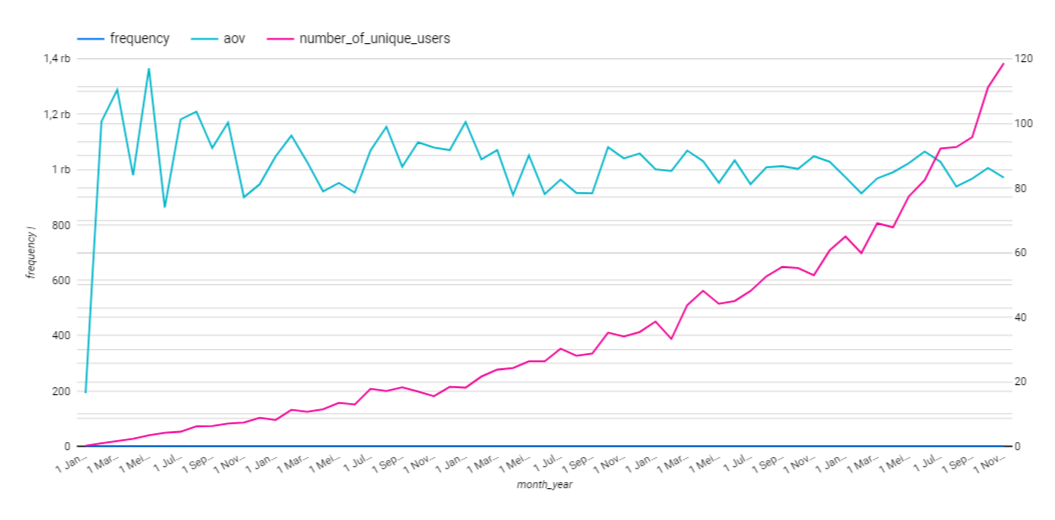
 ON U.id = O.user\_id

 GROUP BY age\_segment, U.gender

 ORDER BY age\_segment, U.gender;

**Rekomendasi :**

1. Tetap mempertahankan dan meningkatkan strategi pemasaran yang dilakukan untuk segmen orang dewasa untuk menjaga loyalitas customer.
2. Mencoba menarik anak-anak muda dengan cara menambahkan produk-produk yang sedang trend di tingkat global atau local dan memberikan berbagai penawaran khusus pada produk-produk tersebut.
3. **Bagaimana trend nilai** **frekuensi, AOV, dan jumlah pengguna pada e-commerce setiap bulannya pada rentang 1 Januari 2019 - 30 November 2023 ?**
4. Bagaimana proyeksi nilai frekuensi, AOV, dan jumlah pengguna e-commerce setiap bulannya pada rentang 1 Januari 2019 – 30 November 2023 ?



**Insight :**

1. Dari Januari 2019 – November 2023, kita dapat melihat jumlah customer yang selalu meningkatkan dari tahun ke tahun menunjukkan perkembangan yang baik, untuk nilai Average Order Value (AOV) menunjukkan nilai yang naik turun (masih dalam kondisi stabil) dengan puncaknya pada 1 Maret 2019, sedangkan untuk frekuensi menunjukkan kondisi yang stagnan.

**Source Code :**

SELECT DATE\_TRUNC(DATE(shipped\_at), MONTH) AS month\_year,

       ROUND(COUNT(DISTINCT order\_id)/COUNT(DISTINCT user\_id), 2) AS frequency,

       ROUND(SUM(sale\_price)/COUNT(DISTINCT order\_id), 2) AS aov,

       COUNT(DISTINCT(user\_id)) AS number\_of\_unique\_users

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

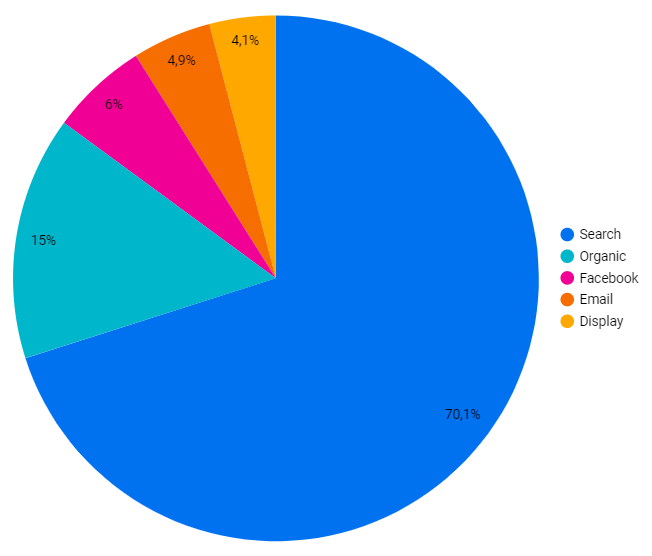
WHERE shipped\_at BETWEEN '2019-01-01' AND '2023-11-30'

      AND status = 'Complete'

GROUP BY 1

ORDER BY 1

1. **Apa media marketing paling baik untuk mempromosikan e-commerce kita ?**



**Insight:**

1. Sebagian besar pelanggan menemukan tentang e-commerce kita melalui pencarian internet (70,1%). Sisanya melalui organic (15%), facebook (6%), email (4,9%), dan display (4,1%).

**Source Code :**

SELECT

  u.traffic\_source,

  COUNT(DISTINCT oi.user\_id) total\_customer

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

LEFT JOIN `bigquery-public-data.thelook\_ecommerce.users` u

ON oi.user\_id = u.id

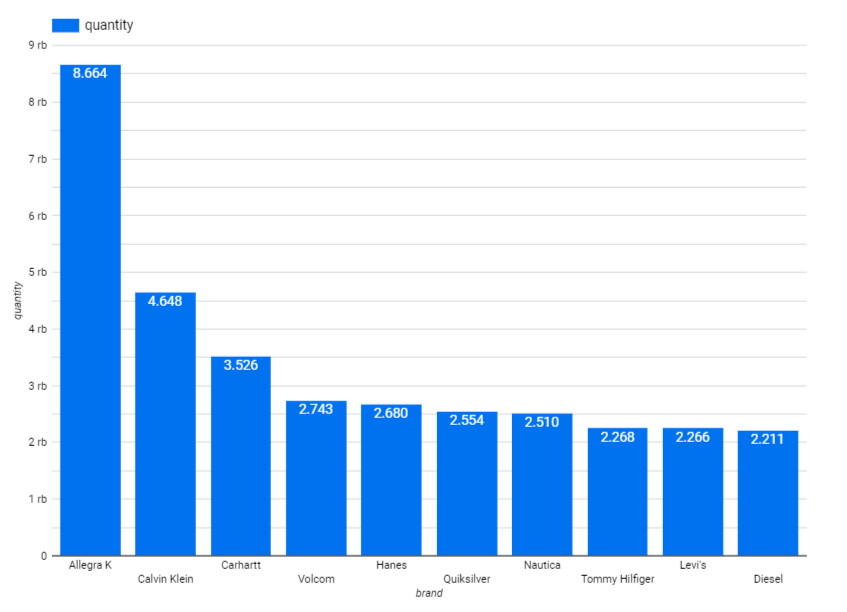
WHERE oi.status NOT IN ('Cancelled','Returned')

GROUP BY 1

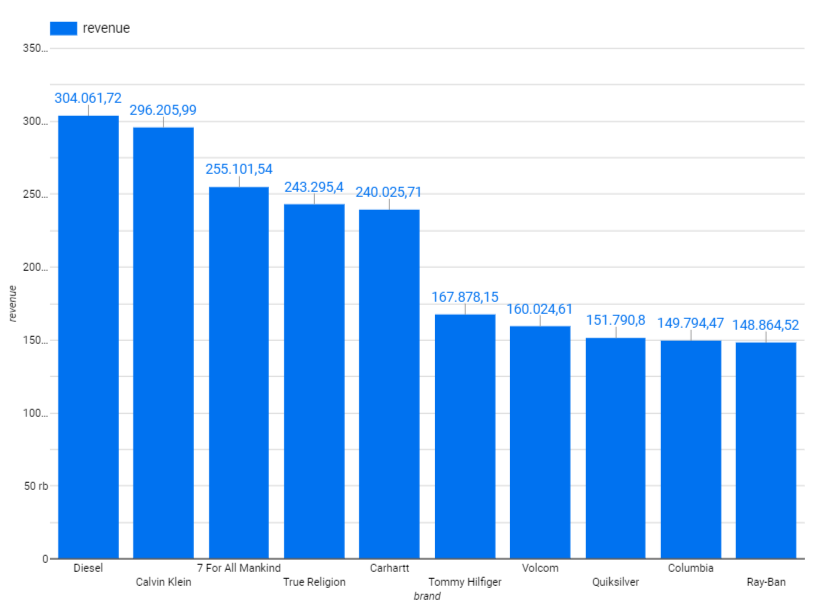
ORDER BY 2 DESC

**Rekomendasi :**

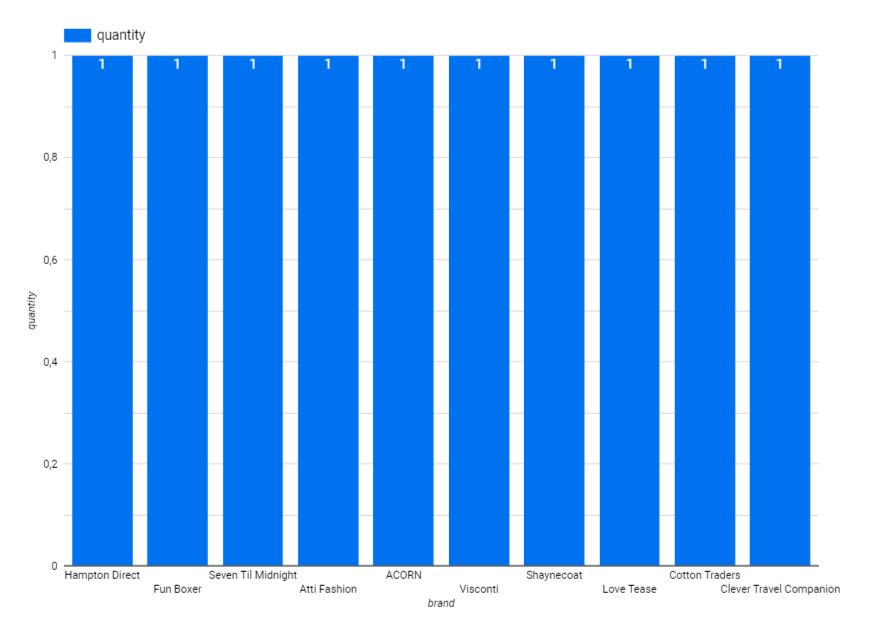
1. Dana marketing yang digunakan untuk iklan di platform pencarian internet lebih baik tidak dikurangi dan akan lebih baik diberikan penambahan.
2. Memperluas jangkauan iklan di platform lain seperti Tiktok, Instagram, dan Twitter dapat dipertimbangkan untuk menjangkau target customer yang lebih luas dan beragam. Hal ini tidak lain dikarenakan pengguna Tiktok, Instagram, dan Twitter yang terus menerus mengalami kenaikan.
3. **Apa kategori produk dan brands yang paling banyak dan paling sedikit memberikan pendapatan dan kuantitas penjualan ?**
4. **Brand Sales**



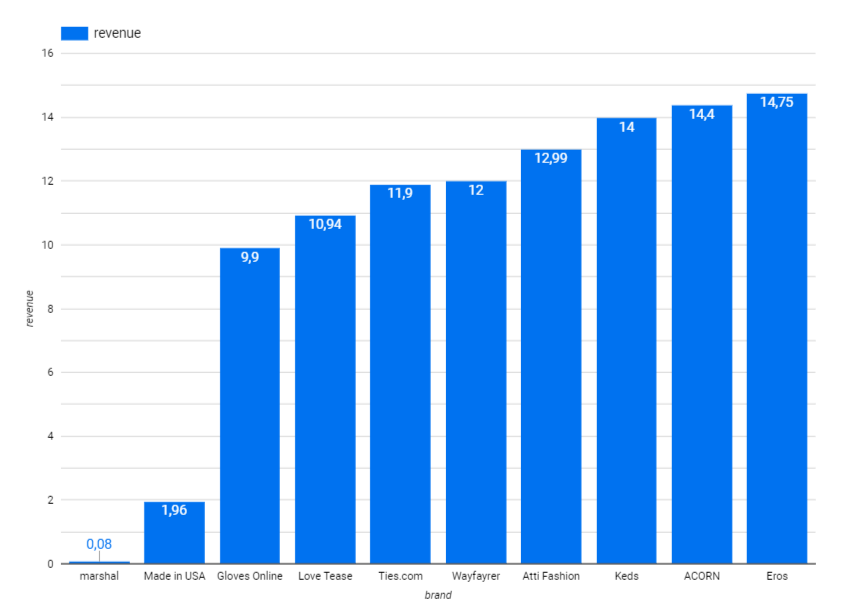
Brand Produk Paling Banyak Jumlah Penjualannya



Brand Produk Paling Banyak Menguntungkan



Brand Produk Paling Sedikit Jumlah Penjualannya



Brand Produk Paling Sedikit Keuntungannya

**Insight :**

1. Brand Diesel memberikan pendapatan paling tinggi di antara brand lain.
2. Brand Allegra K memiliki kuantitas penjualan terbanyak, hampir 2 kali lipat dari jumlah penjualan Calvin Klein.
3. Brand Marshal memberikan pendapatan paling rendah di antara brand lain.
4. Perusahaan kita memiliki banyak brand yang memiliki kuantitas penjualan 1,2, dan 3 produk.

**Source Code :**

SELECT

  p.brand AS brand,

  ROUND(SUM(sale\_price \* num\_of\_item), 2) AS revenue,

  SUM(num\_of\_item) AS quantity

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

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

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

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

ON oi.product\_id = p.id

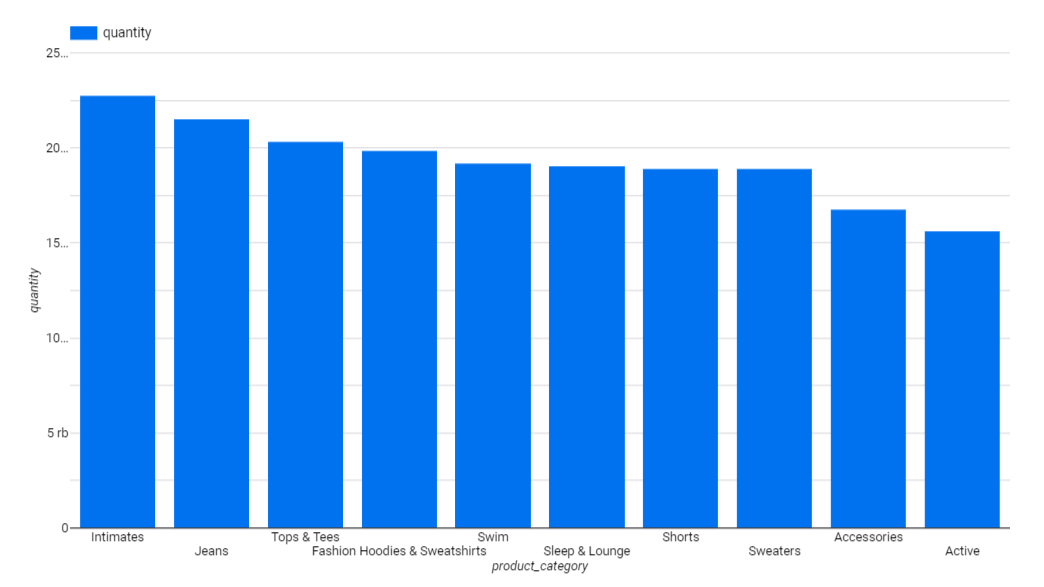
WHERE oi.status NOT IN ('Cancelled', 'Returned')

GROUP BY brand

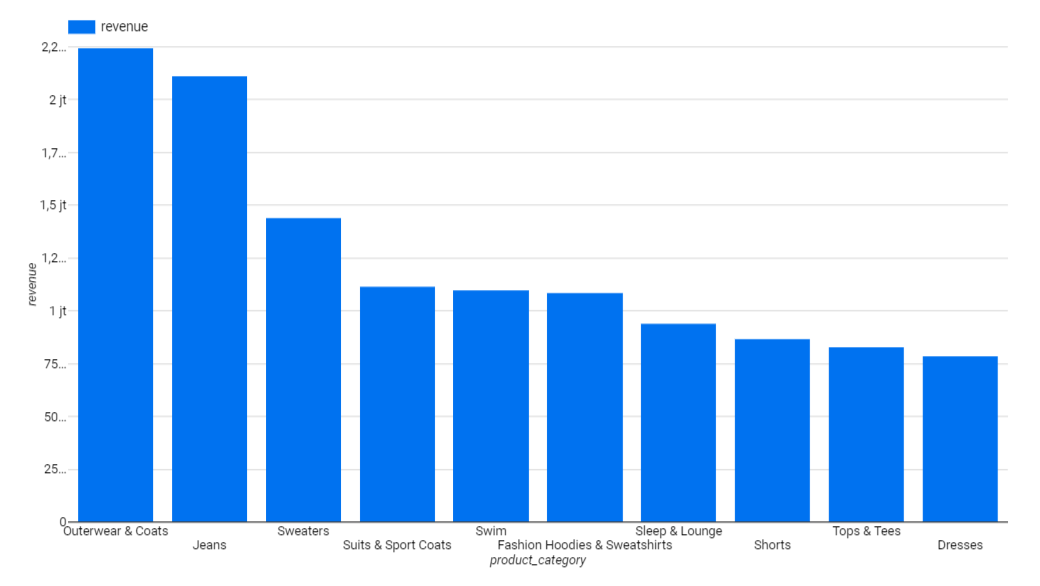
ORDER BY revenue DESC;

**Rekomendasi :**

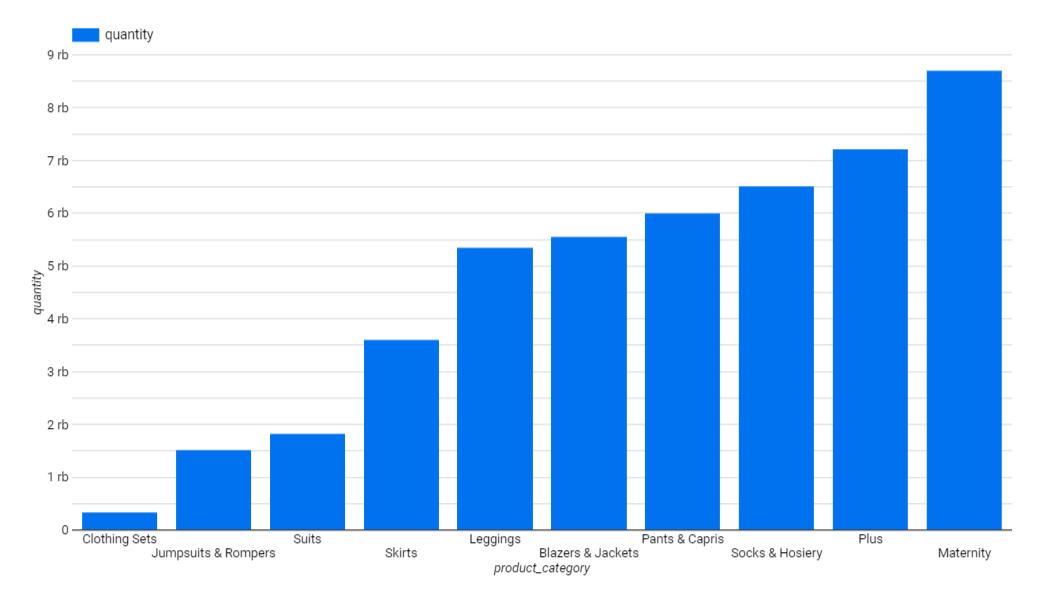
1. Perusahaan dapat melakukan kerja sama kepada brand-brand yang menunjukkan penjualan dan pendapatan yang luar biasa. Dengan melakukan kerja sama, perusahaan dapat mengadakan event-event khusus dengan brand tersebut untuk membuat produk exclusive atau harga khusus.
2. **Product Category Sales**



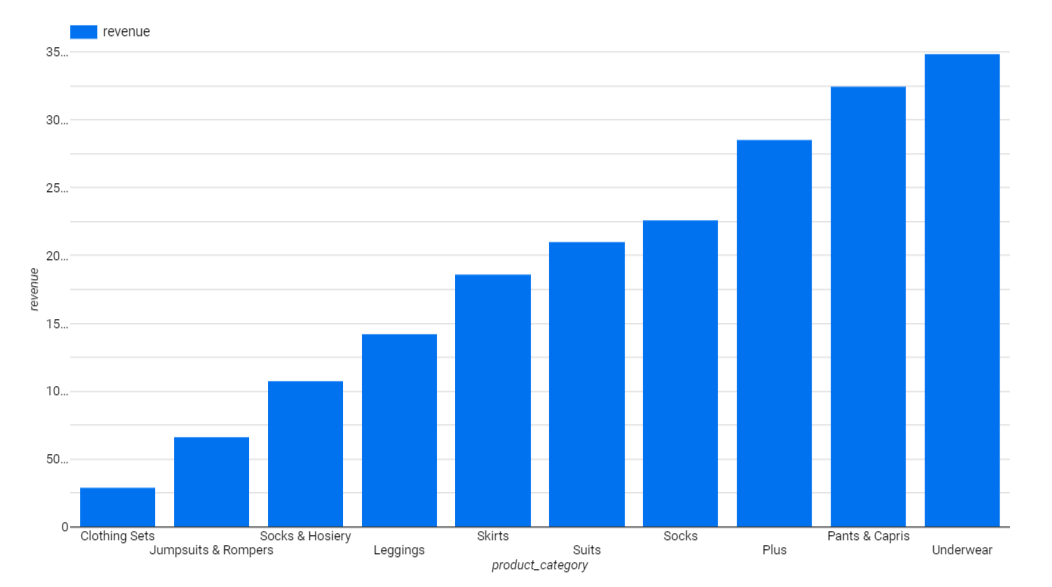
Kategori Produk Paling Banyak Jumlah Penjualannya



Kategori Produk Paling Menguntungkan



Kategori Produk Paling Sedikit Jumlah Penjualannya



Kategori Produk Paling Sedikit Keuntungannya

**Insight :**

1. Jenis produk outerwear dan coats memberikan pendapatan paling tinggi di antara produk lainnya.
2. Jenis produk intimates memiliki kuantitas penjualan terbanyak di antara produk lainnya.
3. Jenis produk clothing sets mencatatkan pendapatan dan kuantitas penjualan paling rendah di antara produk lainnya.

**Source Code :**

SELECT

  category AS product\_category,

  ROUND(SUM(sale\_price \* num\_of\_item), 2) AS revenue,

  SUM(num\_of\_item) AS quantity

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

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

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

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

ON oi.product\_id = p.id

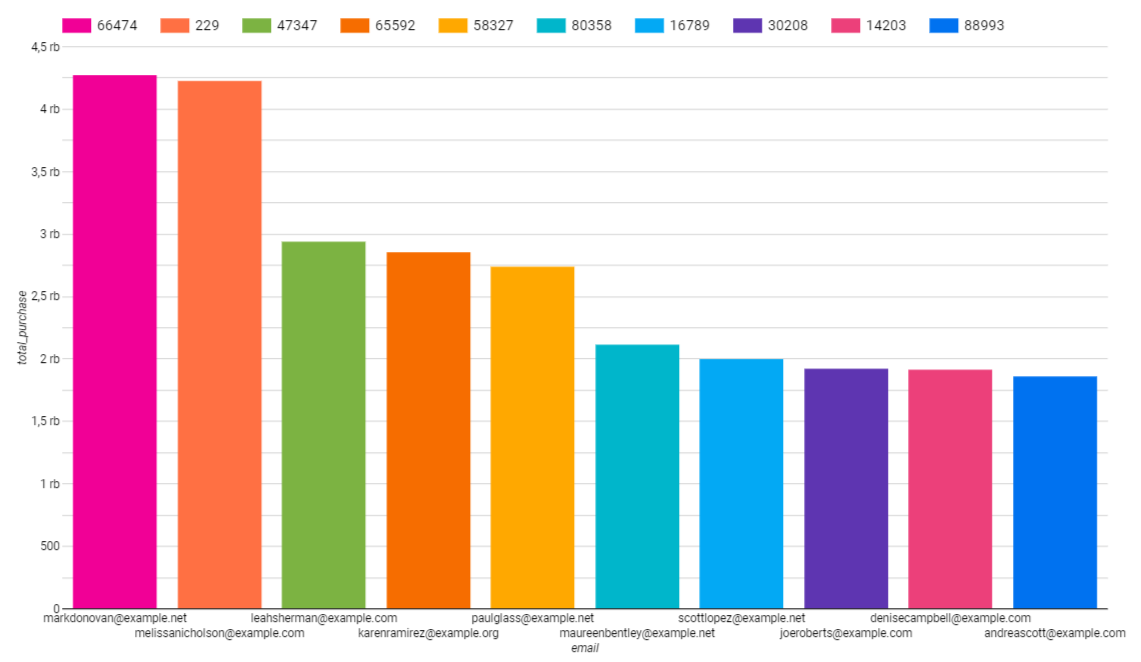
WHERE oi.status NOT IN ('cancelled', 'Returned')

GROUP BY category

ORDER BY revenue DESC;

**Rekomendasi :**

1. Perusahaan dapat melakukan berbagai event khusus yang melibatkan jenis produk dengan kinerja tinggi untuk menghasilkan lebih banyak keuntungan.
2. Dalam hal ini, Perusahaan juga dapat melakukan kerja sama khusus dengan brand ternama pada jenis produk brand yang menunjukkan penjualan dan pendapatan yang luar biasa.
3. Perusahaan juga dapat melakukan event khusus yang berisi jenis produk dengan keuntungan dan jumlah penjualan yang sedikit untuk mengenalkan produk tersebut ke customer.
4. **Mengidentifikasi 10 customer dengan total pembelian terbanyak di negara Australia untuk diskon terbesar event perayaan hari Kemerdekaan Autralia ?**



**Insight :**

1. Pengguna dengan id 66474 menjadi penerima diskon terbesar dengan jumlah pembelian sebanyak 4276,3.
2. Pengguna dengan id 88993 menjadi penerima diskon terbesar terakhir dengan jumlah pembelian sebanyak 1.865,2.
3. Rata-rata jumlah pembelian penerima diskon terbesar adalah 2.689,573.

**Source Code :**

SELECT

  u.id AS customer\_id,

  u.email AS  email,

  ROUND(SUM(oi.sale\_price \* o.num\_of\_item), 2) AS total\_purchase

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.users AS u

ON o.user\_id = u.id

WHERE country = 'Australia'

GROUP BY customer\_id, email

ORDER BY total\_purchase DESC

LIMIT 10;