



# RevoU Case Study - eCommerce Analysis

BY : AJIMAS GALUNG CANDRA



## Background

Data analytics case study is part of mini course data analytics program held by RevoU. In this case study, the participant asked to create their own problem definition and solved it individually based on the dataset given in the course. Here, the dataset used is from BigQuery Public Data about Clothes eCommerce Sales from all around the world.

## Problem Definition

1. Comparison between the product that has been ordered and shipped with the total product ordered from all country.
2. Total product bought by both female and male and clustered it into the product category so the we could summarize which most-bought product category.
3. The most product category that has been shipped to customer clustered by gender
4. Sales clothes trend that was ordered monthly from 2019 -2023
5. The top 10 sales price for product category

## Tools

1. SQL BigQuery: To query and combine all data from 4 table given
2. Microsoft Excel: To do EDA and clean the data
3. Draw.io: To make ER diagram from 4 table
4. Tableau; To visualize and get insight from data

## Data Overview

The data is taken from BigQuery public data. The name of the dataset is “thelook\_ecommerce”. This dataset is a fictitious eCommerce clothing site developed by the Looker. The dataset contains several information about customer, product, and orders. Thelook\_eCommerce contains some table:

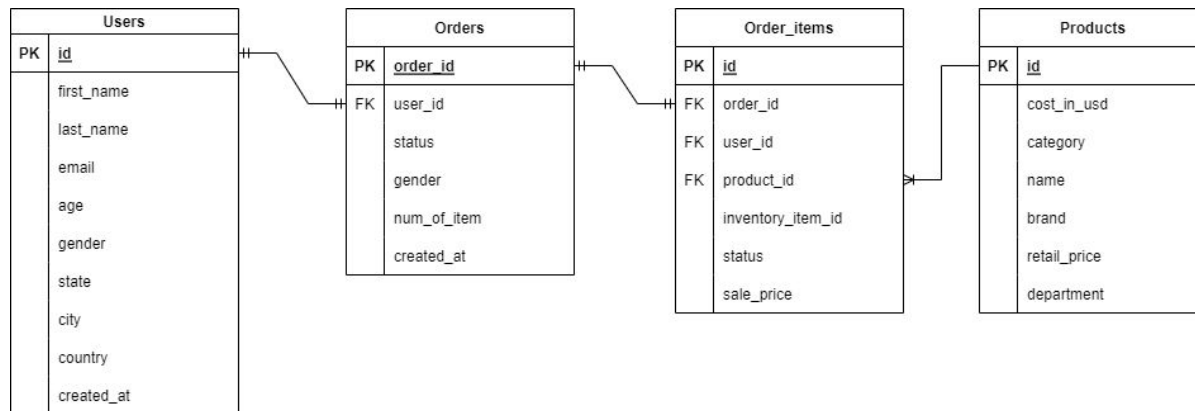
1. **Users Table:** Contains the customer information around the world such as name, age, gender, country, city, email, etc
2. **Orders Table:** Contains the customer that has ordered the product with the order id, the status of ordering, and the timestamp
3. **Order Items Table:** This table give an information about the orders and the product was ordered such as user ID, product ID, order ID, etc.
4. **Product Table:** it gives us an information related with the product in the eCommerce. It include the product ID, product name, product category, the brand, etc.

# Methodology

1. Query all the dataset needed for this project using BigQuery and extract it to CSV file to be analyzed

```
SELECT *  
FROM  
    bigquery-public-data.thelook_ecommerce.users,  
    bigquery-public-data.thelook_ecommerce.orders,  
    bigquery-public-data.thelook_ecommerce.order_items,  
    bigquery-public-data.thelook_ecommerce.products
```

2. Make ER Diagram using draw.io to easily understand relationa schemas between the table



### 3. Set the table to be made from query

timestamp	user_id	order_id	product_id	customer_name	age	gender	city	country	product_name	product_category	product_brand	total_item_order	total_item_shipped	percentage_of_product_shipped	sale_price	rank
29/09/2022	10669	13425	24314	Michael Smith	32	M	Bogatynia	Poland	Nobis Men's Cartel Jacket	Outerwear & Coats	Nobis	1	0	0	750	1
08/08/2022	10669	13425	24314	Michael Smith	32	M	Bogatynia	Poland	Nobis Men's Cartel Jacket	Outerwear & Coats	Nobis	2	0	0	750	1
28/10/2022	10669	13425	24314	Michael Smith	32	M	Bogatynia	Poland	Nobis Men's Cartel Jacket	Outerwear & Coats	Nobis	2	0	0	750	1
01/11/2022	10669	13425	24314	Michael Smith	32	M	Bogatynia	Poland	Nobis Men's Cartel Jacket	Outerwear & Coats	Nobis	1	0	0	750	1
03/02/2021	26971	33647	20368	Willie Morgan	40	M	Bogatynia	Poland	Diesel Men's Jimeneo Jacke Suits & Sport Coats		Diesel	4	0	0	698	2
23/01/2023	19283	24073	21531	Shawn Wilson	58	M	Żagań	Poland	Diesel Men's Safado Slim Str Jeans		Diesel	2	0	0	325	3
15/12/2021	28878	36034	18612	Alexander Wood	33	M	Gorzów Wielkop	Poland	Champion Men's Double Dry Active		Champion	1	1	100	293.33	4
03/10/2020	68222	85119	24418	Gregory Martin	37	M	Złotendorf	Poland	Mountain Hardwear Men's E Outerwear & Coats		Mountain Hardwear	1	0	0	270	5
02/10/2022	43118	53892	20683	Ross Romero	13	M	Bogatynia	Poland	True Religion Men's Bobby Z' Jeans		True Religion	4	4	100	268	6
19/12/2022	43118	53892	20683	Ross Romero	13	M	Bogatynia	Poland	True Religion Men's Bobby Z' Jeans		True Religion	1	0	0	268	6
26/01/2022	92879	115957	4232	Meghan Farmer	46	F	Bogatynia	Poland	True Religion Women's TRBJ Jeans		True Religion	1	0	0	253	7
04/03/2022	57007	71136	8806	Lori Gonzalez	43	F	Gorzów Wielkop	Poland	FactoryExtreme Quilted Chai Outerwear & Coats		FactoryExtreme	1	0	0	238.98	8
19/02/2022	57007	71136	8806	Lori Gonzalez	43	F	Gorzów Wielkop	Poland	FactoryExtreme Quilted Chai Outerwear & Coats		FactoryExtreme	1	0	0	238.98	8
28/12/2022	86683	108176	24444	Angel Leon	57	M	Bogatynia	Poland	Carhartt Men's Big-Tall Flami Outerwear & Coats		Carhartt	1	0	0	234.95	9

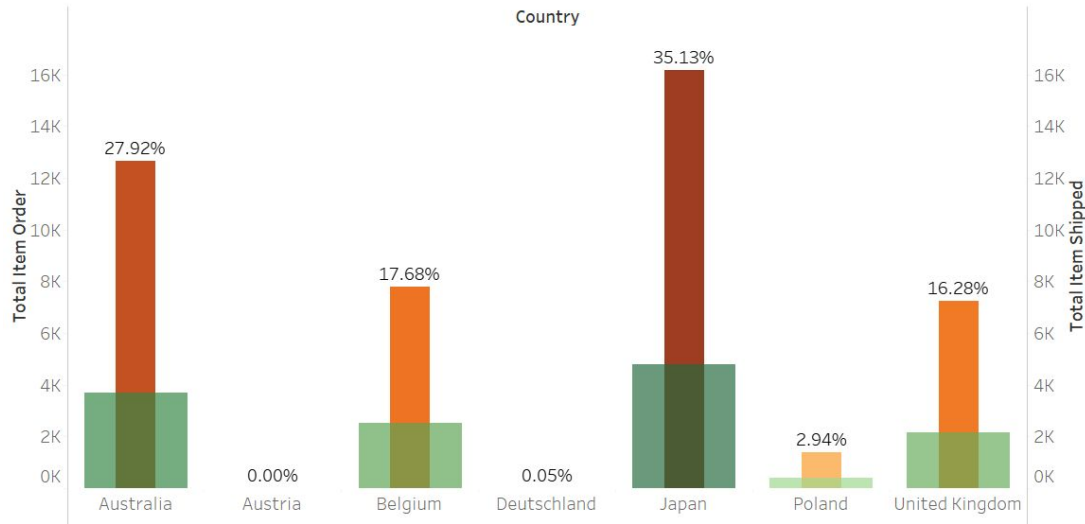
### 4. Write a query that is needed in BigQuery. Here is the link for the query:

<https://console.cloud.google.com/bigquery?sq=1073902271305:a9feab1e40b84252aebea8a552a52161>

### 5. The last one is extract data queried to CSV file and transport it to the tableau to be visualized

# Result

Total Item Shipped Compare with Total Order



Country:

% of Total Percentage Of Product Shipped along Table (Across): **35.13%**

Total Item Order:

**16,164**

Total Item Shipped:

**4,804**

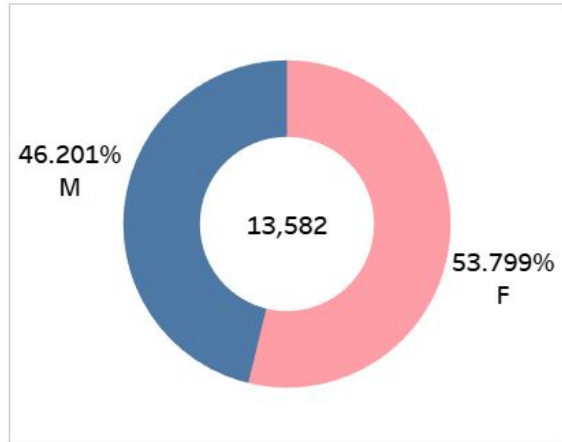
## Descriptive Analysis:

- From the graph above, we can conclude that **Japan** is the-most ordered and product shipped country comparing to other country. Here, percentage of total product shipped respect to the total order is **35.13%**.
- The most-clothes order is only 35.13%. We can take an insight that, based on this data, clothes is not really populer or **less considered**.

Link for Interactive Tableau Dashboard: <https://public.tableau.com/app/profile/ajimas.galung.candra>

# Result

Male and Female Total item Shipped



## Descriptive Analysis:

- This pie chart gives us an information that from total product shipped, which is 13,583 product item, **Woman** take around **54%** of total product shipped compare with **Men** that has **46%** of total product shipped.
- From this chart, actually both men and woman concern about what they use, especially for clothes.

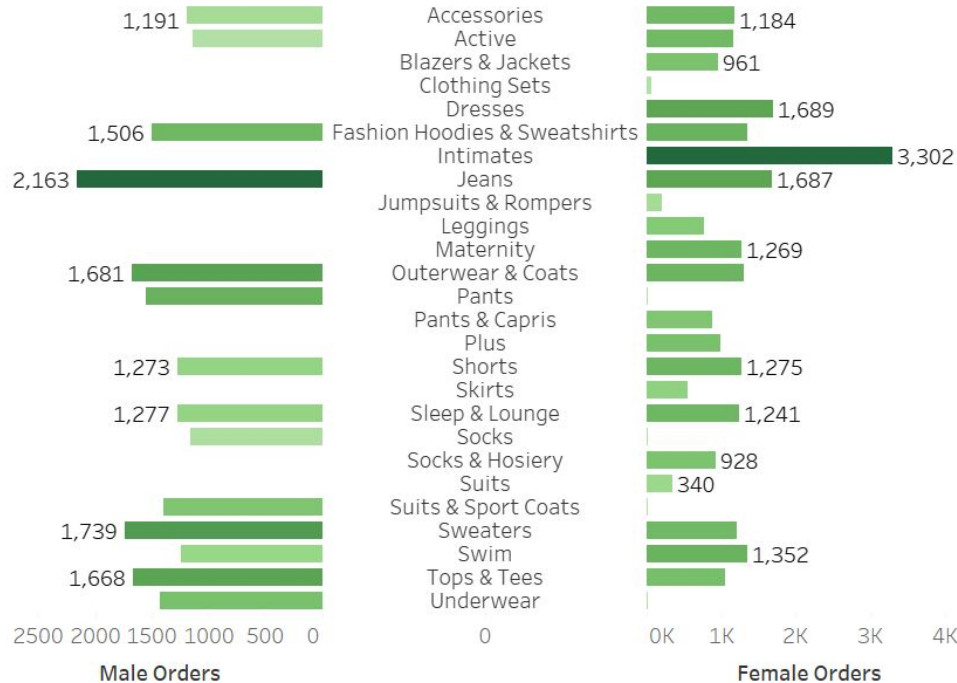
## Diagnostic Analysis:

- Woman slightly more concern about their outfit if we compare with men because woman may think that their appearance impact a lot to their social and woman tends to show their beauty with her new clothes

Link for Interactive Tableau Dashboard: <https://public.tableau.com/app/profile/ajimas.galung.candra>

# Result

Total Order for Each Product Category



## Descriptive Analysis:

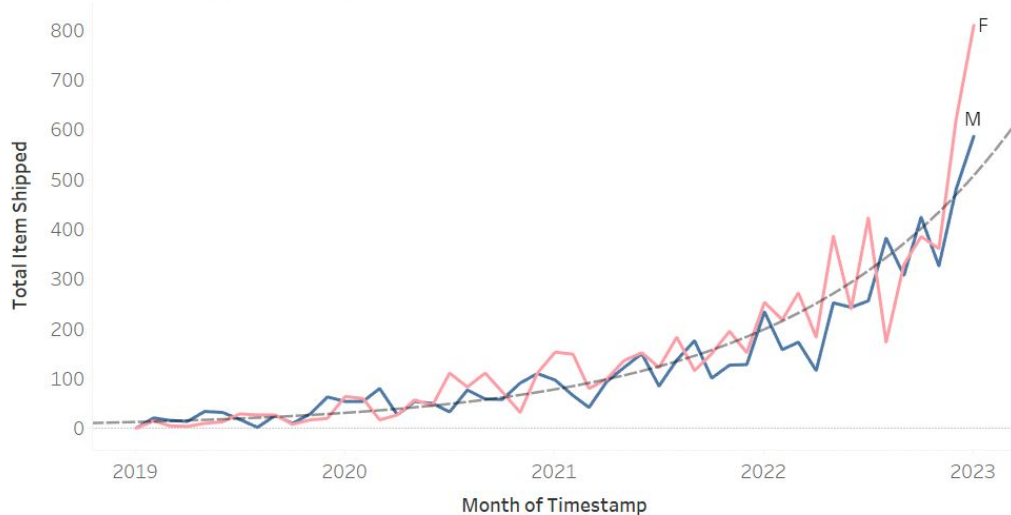
- **Jeans** is the most favorite product category for **Male** and **Intimates** is the most favorite clothes for **Woman**
- For Men, **sweater** and **outwear&coats** is the second and third most clothes category ordered because it is as part of men's style
- For woman, **dresses** and **jeans** are the second and third most clothes category ordered.

Link for Interactive Tableau Dashboard: <https://public.tableau.com/app/profile/ajimas.galung.candra>



# Result

Trend Product Shipped Monthly



## Descriptive Analysis:

- From this graphic, we can conclude that the number of products bought by the customer, both Men and Women, **increase exponentially**

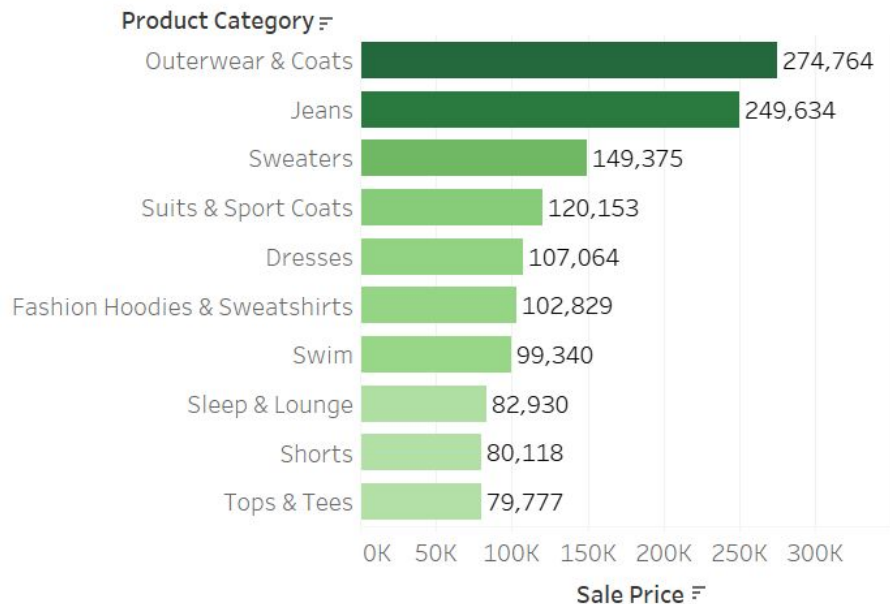
## Diagnostic Analysis:

- Number of adults with age range 20-30 increase and it correlates with the number of clothes that are bought because adults really concern with their looks.
- A lot of new clothes that are eye-catching so it attracts people to buy.

Link for Interactive Tableau Dashboard: <https://public.tableau.com/app/profile/ajimas.galung.candra>

# Result

Top Sale Price by Product Category



## Descriptive Analysis:

- **Outerwear & Coats** the most expensive clothes category with **\$274,764** followed by **Jeans** with **\$249,634** as the second-most expensive clothes category.

## Diagnostic Analysis:

- It is reasonable to be outerwear and jeans as top two most expensive clothes because it will impact your style a lot.

Link for Tableau Public: <https://public.tableau.com/app/profile/ajimas.galung.candra>