

# CAPSTONE PROJECT DATA ANALYTICS

TEAM 2 - SABANG

# TABLE OF CONTENTS

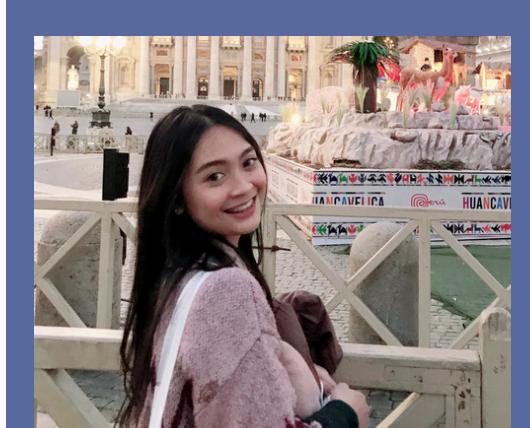
- Business Analysis
- Data Cleaning
- Data Visualization
- Insight & Recomendation



# Our Team Members



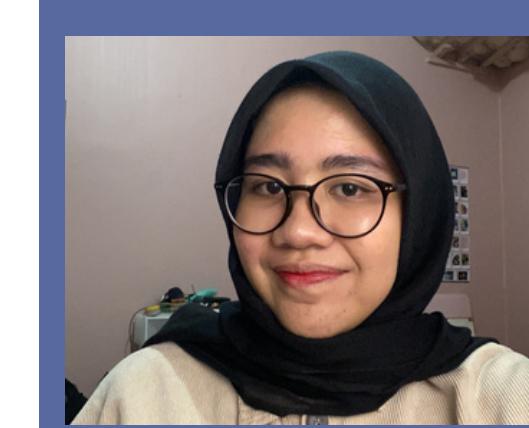
**A'raf Adi Nugraha  
Putra**  
Group Captain



**Ogi Wellen Barus**  
Setting Objective &  
Problem Statement



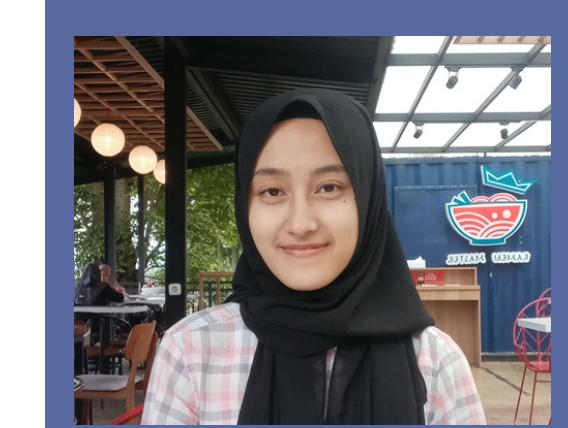
**Valerie Marsha Halim**  
Setting Objective &  
Problem Statement



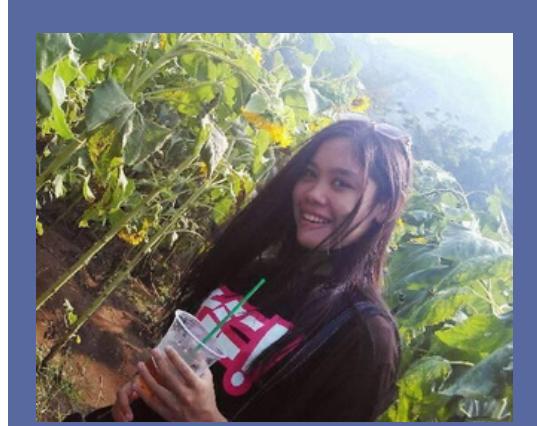
**Fariska Maharani**  
Data Cleaning



**Muhamad Rayhand  
Prasetya**  
Data Cleaning



**Sri Handayani**  
Data Cleaning



**Dyah Ayu Wiranti**  
Data Visualization



**Meysiska Anggraini**  
Data Visualization



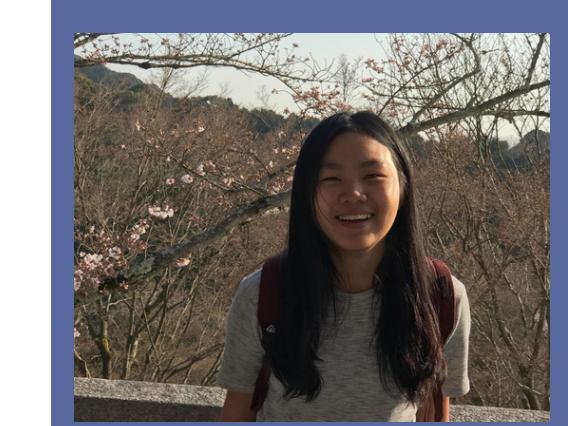
**Muhammad Farhan  
Alvianto**  
Data Visualization



**Paber Yeremia**  
Data Communication



**Safira Karin Hidayah**  
Data Communication



**Wanda Tahir**  
Data Communication



**Business Analysis**

# Company Overview

- The company is a United Kingdom based e-commerce company and registered non-store online retailer.
- Most of the customers are wholesalers both local and international.
- Companies want to find various insights that can be utilized to improve their performance.



## Problem Statement

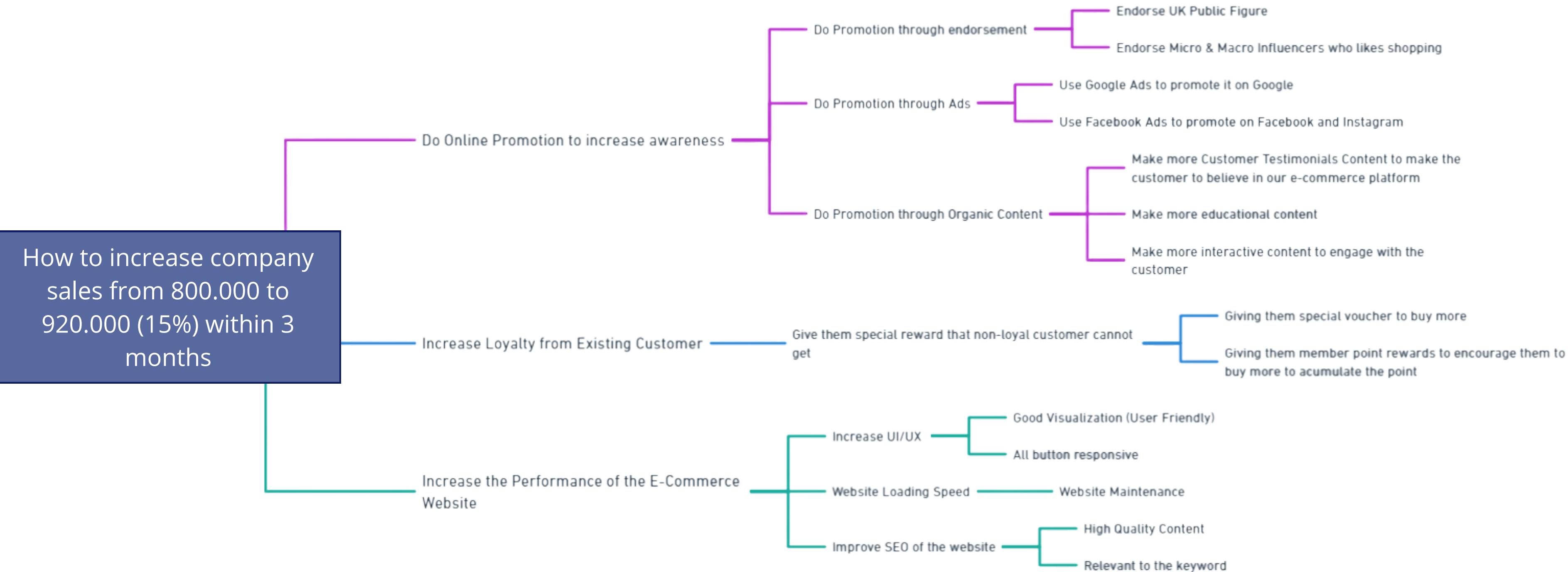
**How to improve company sales**  
**Specific**

**from 800.000 to 920.000 (15%)**  
**Measurable, Realistic**

**within 3 months?**  
**Attainable, Timebound**

# Root Cause Analysis

## Issue Tree



# Metrics

How to increase Sales Performance by 15% in the next 3 months?

Problems	Metrics
Do Promotion through endorsement	<ul style="list-style-type: none"><li>• CAC</li><li>• Engagement Rate</li></ul>
Do Promotion through Ads	<ul style="list-style-type: none"><li>• CAC</li><li>• Engagement Rate</li><li>• Conversion Rate</li></ul>
Do Promotion through Organic Content	<ul style="list-style-type: none"><li>• CAC</li><li>• Engagement Rate</li></ul>
Give them special reward that non-loyal customer cannot get	<ul style="list-style-type: none"><li>• CAC</li><li>• CLTV</li><li>• Average Order Value (AOV)</li></ul>
Increase UI/UX	<ul style="list-style-type: none"><li>• Bounce Rate</li><li>• Session</li></ul>
Website Loading Speed	<ul style="list-style-type: none"><li>• Bounce Rate</li><li>• Session</li></ul>
Improve SEO of the website	<ul style="list-style-type: none"><li>• Bounce Rate</li><li>• Session</li><li>• Number of Profile Visits</li></ul>

# Metrics Detail

## Customer Lifetime Value (CLTV)

To determine the lifetime value and the tendency of customers to generate sales back for us.

With giving them special voucher, giving member point rewards to encourage buy more to accumulate the point

## Customer Acquisition Cost (CAC)

To determine how much it will cost us to get 1 customer, the smaller the CAC the better

## Average Order Value (AOV)

To determine how many customers on average order in 1 transaction.

To measure the results after we give them special vouchers and reward point.

## Conversion Rate

Determine the percentage of effectiveness of the business page/e-commerce platform to attract visitors to make a transaction.

To measure the results Google Ads dan Facebook Ads.

## Engagement rate

To see the success of content on social media can attract customer.

Measuring results after promoting through organic content and endorsements.

## Bounce rate

To calculate the percentage of visitors who arrive, then leave without opening another page or interacting with it.

Measuring E-Commerce Website performance results.

## Session

Measure the number of sessions that occur in a website.

## Number of Profile Visit

To know how many visitors in our business page

# Data Cleaning



# Link Google Spreadsheet

[https://docs.google.com/spreadsheets/d/1MaasUrHC8idawVNcqkqTk0R2vibTUS-QV9GovvlnDEU/edit?usp=drive\\_web&ouid=102196232426997899423](https://docs.google.com/spreadsheets/d/1MaasUrHC8idawVNcqkqTk0R2vibTUS-QV9GovvlnDEU/edit?usp=drive_web&ouid=102196232426997899423)

# Data Cleaning

## Process



Remove non-numeric  
characters (letters) from  
"Invoice No" column by using  
find & replace method

Step 1

Step 2

Trim whitespace on the  
"Description" column by using  
trim space method

Remove negative signs from  
"Quantity" column by using  
find & replace method

Remove empty cells on the  
"Customer ID" column by  
using filter & delete row  
method

Step 3

Step 5

Step 4

Step 6

Extract timestamps from "Invoice  
Date" into new column by using  
split data method

Convert USA, RSA, EIRE  
abbreviation on the "Country"  
column into full country name by  
using find & replace method

Add "Revenue" column by multiplying the "Quantity" and "Unit Price" column

Step 7



Step 8

Add "Category" column and get the data from "Description" by duplicate the data, categorize manually, and store the data by Vlookup

# How do we find & remove the outliers?

**Calculate the first (Q1) and  
third (Q3) quartile by:**

$Q1 = \text{Quartile}(\text{dataset}, 1)$

$Q3 = \text{Quartile}(\text{dataset}, 3)$

Step 1



Step 2

**Calculate Interquartile Range**

**(IQR) by:**

subtracting Q3 and Q1

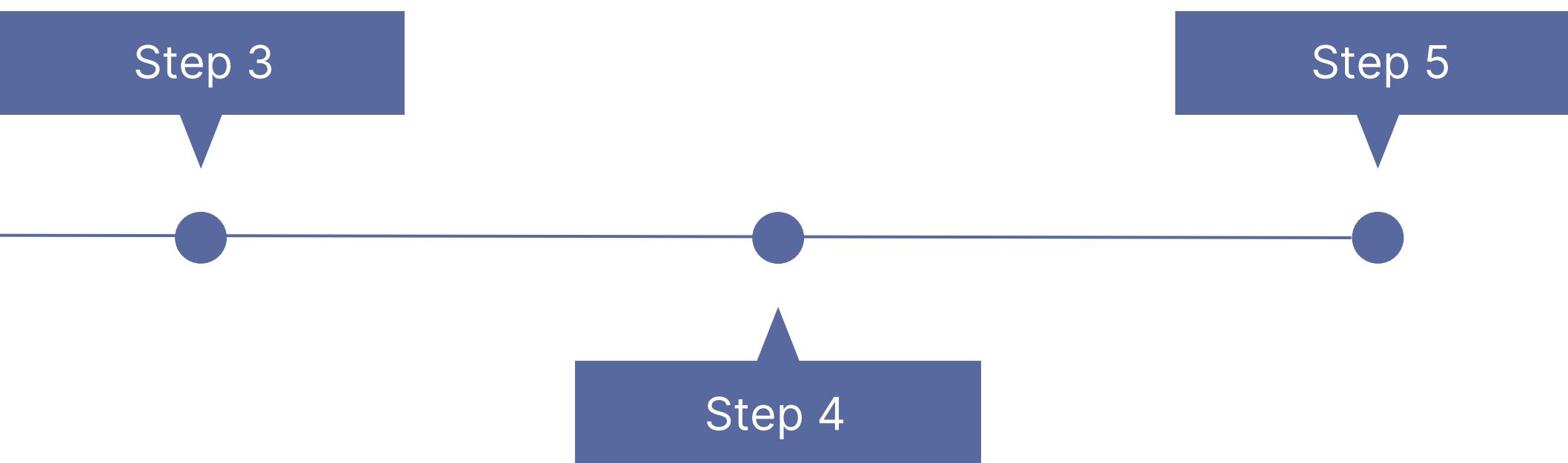
**Calculate lower boundary by:**

$$=Q1-(1.5*IQR)$$

**Calculate upper boundary by:**

$$=Q3+(1.5*IQR)$$

Remove the data outliers



Any data in column "Quantity" and "Unit Price" that is less than the lower boundary or more than the upper boundary is considered as outliers

# Descriptive statistics in finding outliers

Quantity	
Q1	2
Q3	12
IQR	10
Lower	-13
Upper	27

Unit Price	
Q1	1.25
Q3	2.95
IQR	1.7
Lower	-1.3
Upper	5.5

	A	B	C	D	E	F	G	H
1	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
2	C581484	23843	PAPER CRAFT	-80995	12/9/2011 9:27	2.08	16446	United Kingdom
3	C541433	23166	MEDIUM CERA	-74215	1/18/2011 10:17	1.04	12346	United Kingdom
4	556690	23005	printing smudges	-9600	6/14/2011 10:37	0		United Kingdom
5	556691	23005	printing smudges	-9600	6/14/2011 10:37	0		United Kingdom
6	C536757	84347	ROTATING SILV	-9360	12/2/2010 14:23	0.03	15838	United Kingdom
7	556687	23003	Printing smudge:	-9058	6/14/2011 10:36	0		United Kingdom
8	546152	72140F	throw away	-5368	3/9/2011 17:25	0		United Kingdom
9	573596	79323W	Unsaleable, dest	-4830	10/31/2011 15:17	0		United Kingdom
10	566768	16045		-3667	9/14/2011 17:53	0		United Kingdom
11	565304	16259		-3167	9/2/2011 12:18	0		United Kingdom
12	C550456	21108	FAIRY CAKE FL	-3114	4/18/2011 13:08	2.1	15749	United Kingdom
13	560039	20713	wrongly marked.	-3100	7/14/2011 14:27	0		United Kingdom
14	545990	84598	check	-3000	3/8/2011 13:07	0		United Kingdom
15	569466	23270	incorrect stock e	-2880	10/4/2011 11:42	0		United Kingdom
16	560600	18007		-2834	7/19/2011 17:04	0		United Kingdom
17	573593	79323LP	Unsaleable, dest	-2618	10/31/2011 15:16	0		United Kingdom
18	540564	22617	mouldy, thrown a	-2600	1/10/2011 10:36	0		United Kingdom
19	547560	72732	thrown away-car	-2472	3/23/2011 17:28	0		United Kingdom
20	552733	23059	Thrown away-rus	-2376	5/11/2011 10:48	0		United Kingdom
21	573594	79323P	Unsaleable, dest	-2003	10/31/2011 15:16	0		United Kingdom

# 541910

Before Data Cleaning

	A	B	C	D	E	F	G	H	I	J	K
1	Invoice No	Stock Code	Description	Category	Quantity	Invoice Date	Time	Unit Price	Revenue	Customer ID	Country
2	550965	23137	Zinc Wire Sweetheart Letter Tray	Home	1	4/21/2011	16:53	\$3.75	\$3.75	16750	United Kingdom
3	557406	23137	Zinc Wire Sweetheart Letter Tray	Home	1	6/20/2011	11:30	\$3.75	\$3.75	15747	United Kingdom
4	563709	23137	Zinc Wire Sweetheart Letter Tray	Home	1	8/18/2011	15:17	\$3.75	\$3.75	15472	United Kingdom
5	570653	23137	Zinc Wire Sweetheart Letter Tray	Home	1	10/11/2011	13:49	\$3.75	\$3.75	14710	United Kingdom
6	573152	23137	Zinc Wire Sweetheart Letter Tray	Home	1	10/27/2011	20:16	\$3.75	\$3.75	17530	United Kingdom
7	577174	23137	Zinc Wire Sweetheart Letter Tray	Home	1	11/18/2011	10:58	\$3.75	\$3.75	17841	United Kingdom
8	578521	23137	Zinc Wire Sweetheart Letter Tray	Home	1	11/24/2011	14:06	\$3.75	\$3.75	17812	United Kingdom
9	579885	23137	Zinc Wire Sweetheart Letter Tray	Home	1	11/30/2011	17:37	\$3.75	\$3.75	15444	United Kingdom
10	555110	23137	Zinc Wire Sweetheart Letter Tray	Home	2	5/31/2011	14:34	\$3.75	\$7.50	17841	United Kingdom
11	565445	23137	Zinc Wire Sweetheart Letter Tray	Home	2	9/4/2011	14:50	\$3.75	\$7.50	13548	United Kingdom
12	578307	23137	Zinc Wire Sweetheart Letter Tray	Home	2	11/23/2011	15:53	\$3.75	\$7.50	17290	United Kingdom
13	551014	23137	Zinc Wire Sweetheart Letter Tray	Home	4	4/26/2011	11:04	\$3.75	\$15.00	13209	United Kingdom
14	553889	23137	Zinc Wire Sweetheart Letter Tray	Home	4	5/19/2011	16:31	\$3.75	\$15.00	13405	United Kingdom
15	560035	23137	Zinc Wire Sweetheart Letter Tray	Home	4	7/14/2011	13:41	\$3.75	\$15.00	17706	United Kingdom
16	577361	23137	Zinc Wire Sweetheart Letter Tray	Home	4	11/18/2011	16:13	\$3.75	\$15.00	13012	United Kingdom
17	581101	23137	Zinc Wire Sweetheart Letter Tray	Home	5	12/7/2011	11:48	\$1.25	\$6.25	17835	United Kingdom
18	570342	23137	Zinc Wire Sweetheart Letter Tray	Home	6	10/10/2011	12:26	\$3.75	\$22.50	17706	United Kingdom
19	555401	23137	Zinc Wire Sweetheart Letter Tray	Home	7	6/2/2011	17:51	\$3.29	\$23.03	14088	United Kingdom
20	581003	23137	Zinc Wire Sweetheart Letter Tray	Home	8	12/7/2011	8:38	\$1.25	\$10.00	14911	Ireland
21	551351	23137	Zinc Wire Sweetheart Letter Tray	Home	24	4/28/2011	10:20	\$3.29	\$78.96	14298	United Kingdom
22	580695	23143	Zinc Wire Kitchen Organiser	Kitchen	1	12/5/2011	16:03	\$3.95	\$3.95	13230	United Kingdom

# 344869

After Data Cleaning



# Data Visualization

Link Google Data Studio:

<https://datastudio.google.com/u/0/reporting/27c84b5d-fbf0-4bbb-a98b-251bcef70878/page/GfNvC/edit>

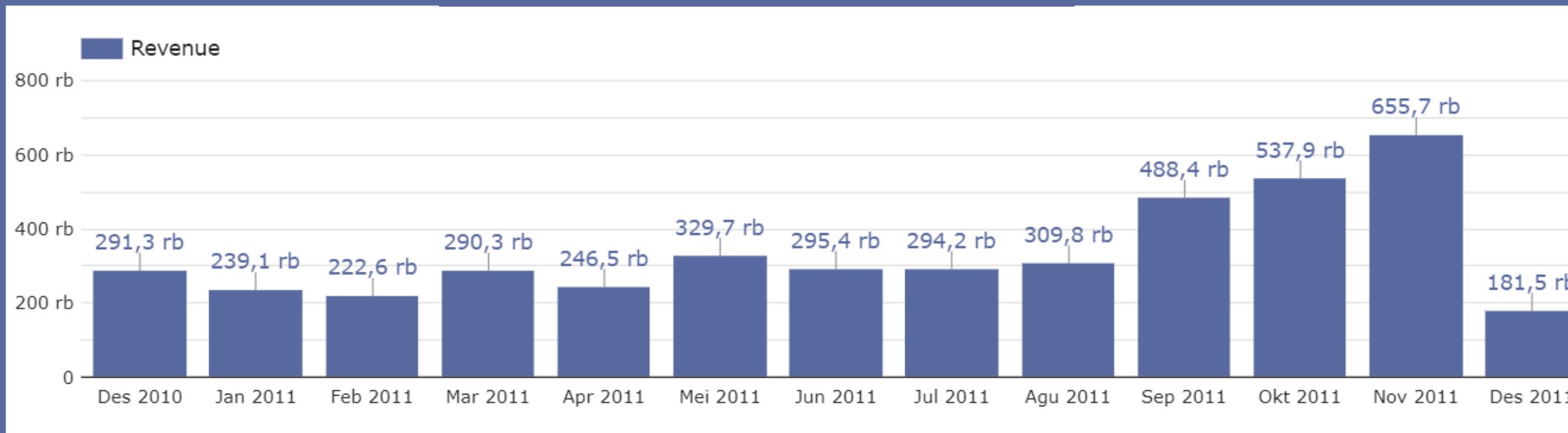
# Insight & Recommendation



## Sales Trend (2010-2011)



## Monthly Revenue (2010 - 2011)

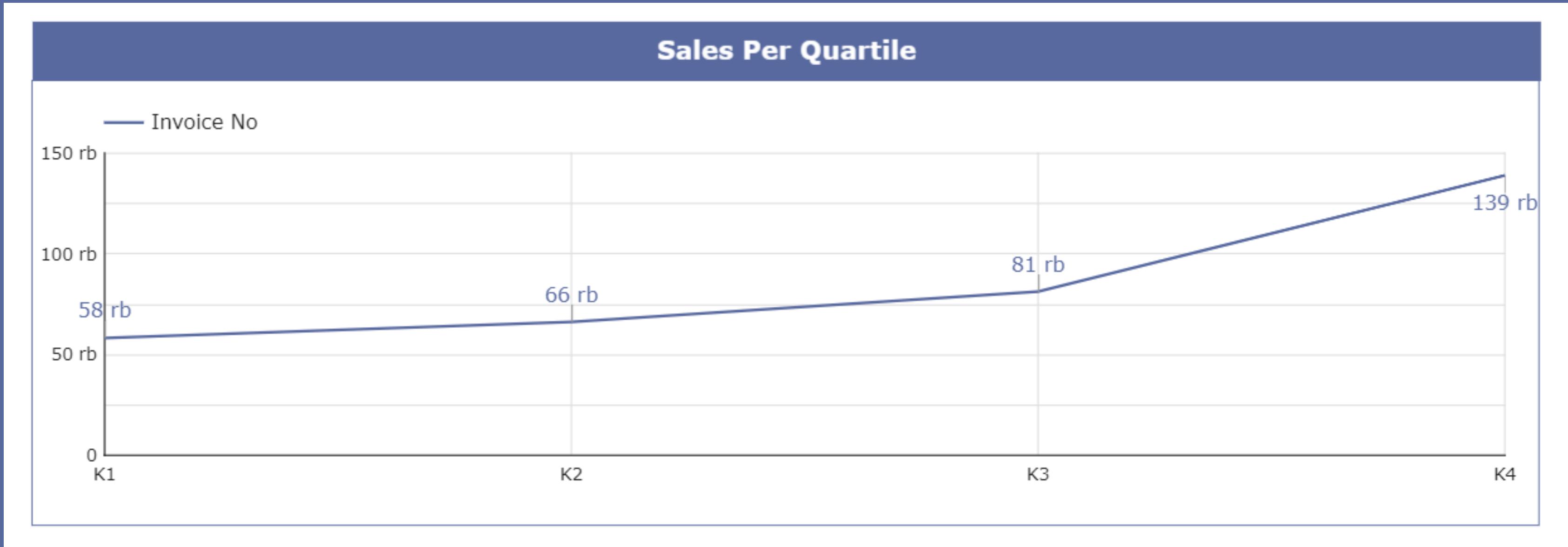


## Insight

- Penjualan dari bulan Juni 2011 - Agustus 2011 mengalami kenaikan dengan stabil namun tidak signifikan, dari 150 rb - 190 rb
- Penjualan bulan September 2011 - November 2011 mulai melonjak dengan sangat signifikan sampai 198% jika dibandingkan dengan bulan agustus 2011 yaitu sebanyak 378 ribu
- Penjualan dari November 2011 - Desember 2011 mengalami penurunan yang sangat signifikan karena data penjualan di Desember 2011 baru sampai tanggal 9 Desember 2011.
- Penjualan mengalami kenaikan selinier dengan revenue yang meningkat juga.

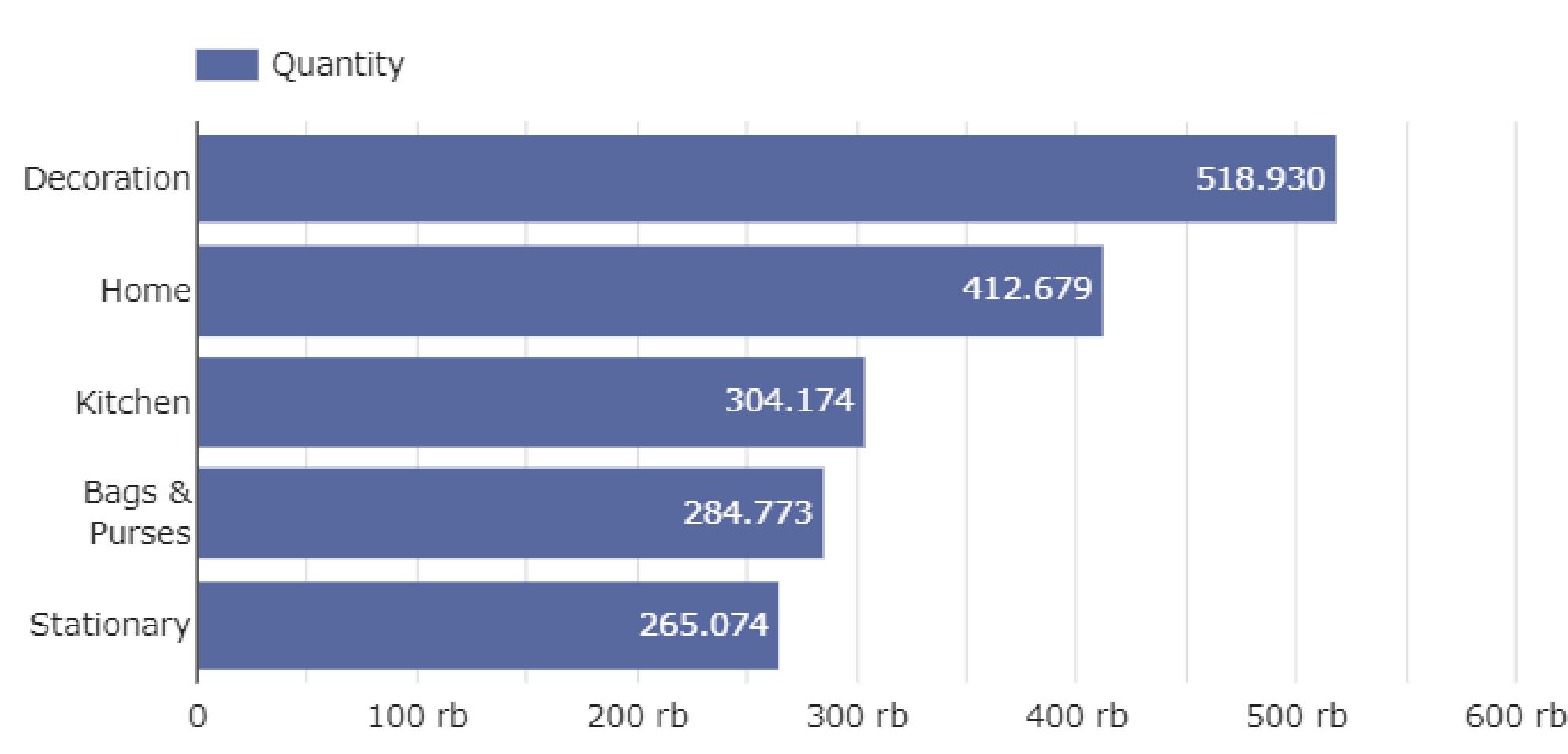
## Recommendation

- E-Commerce harus berusaha untuk membuat penjualan di awal tahun naik secara stabil.
- Dataset yang kurang lengkap di bulan Desember 2011 harus dilengkapi agar bisa dibandingkan secara apple to apple.

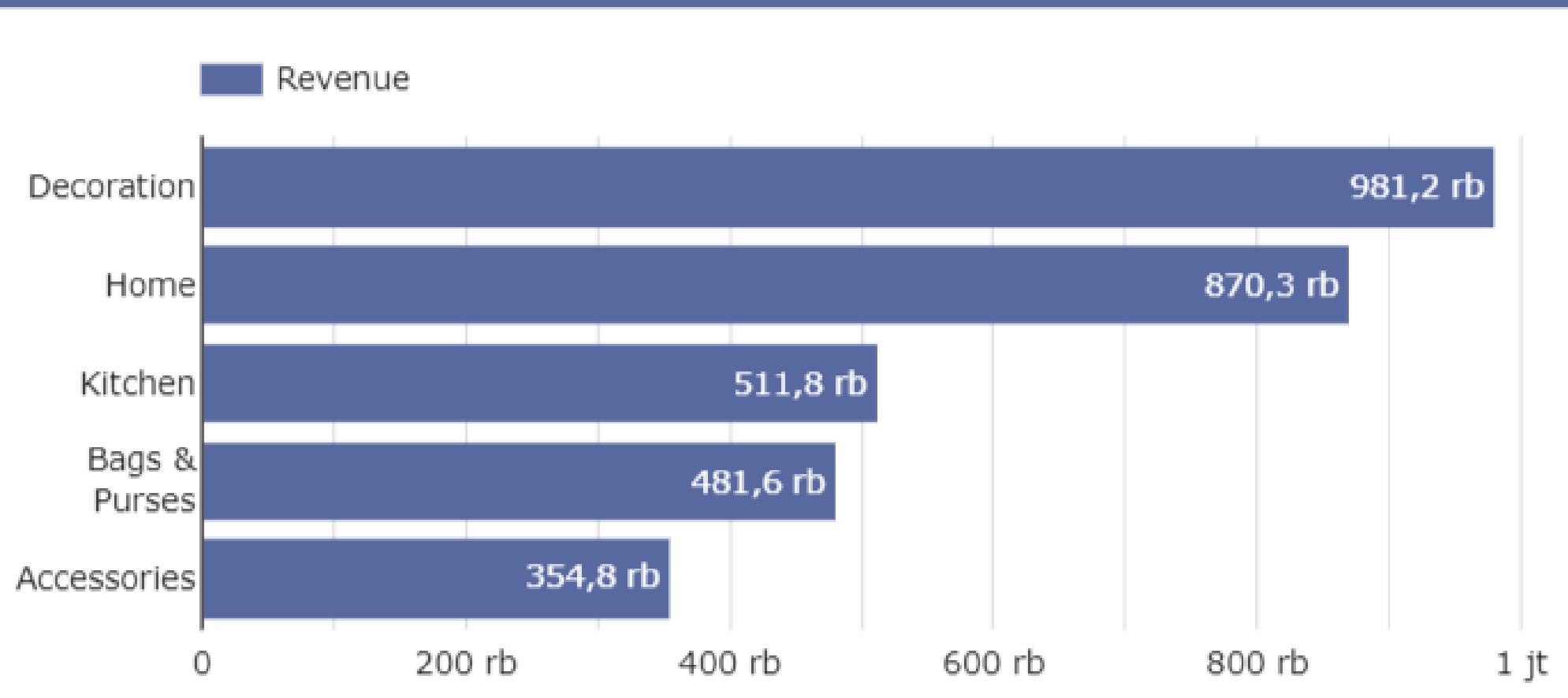


Terjadi kenaikan penjualan per kuartilnya dan kenaikan yang cukup signifikan terjadi pada kuartil 3 ke kuartil 4 meskipun data penjualan dibulan Desember 2011 belum terekap sepenuhnya

## Top 5 Category Based on Sales



## Top 5 Category Based On Revenue



## Insight

- Kategori dari dekorasi menyumbang sales dan revenue tertinggi yaitu 518 pcs dengan revenue 981,2 ribu.
- Revenue dari kategori kitchen jauh **lebih rendah** daripada **decoration** dan **home**, meskipun salesnya cukup tinggi.
- Sales dari kategori **stationary cukup baik**, namun **revenuenya lebih rendah** daripada accessories, kemungkinan karena harganya yang lebih rendah.

## Recommendation

- Memperbanyak **sales quantity** dari dengan cara melakukan lebih banyak promosi ke public figure & influencers, paid ads, dan organic content.
- Mengoptimalkan penjualan untuk kategori yang sales nya masih rendah
- Menawarkan **kategori-kategori sejenis** kepada pelanggan untuk meningkatkan basket size & AOV

## Top 5 Category -All Source

Category	Quantity	Revenue	Buying Power
1. Decoration	518.930	981.186,01	1,89
2. Home	412.679	870.273,26	2,11
3. Kitchen	304.174	511.790,48	1,68
4. Bags & Purses	284.773	481.567,23	1,69
5. Stationary	265.074	351.172,34	1,32

## Insight

**Buying power tertinggi** adalah produk kategori **Home** sebesar **2,11**. Namun, **sales & revenue tertinggi** berasal kategori **Decoration** dengan buying power **1,89**.

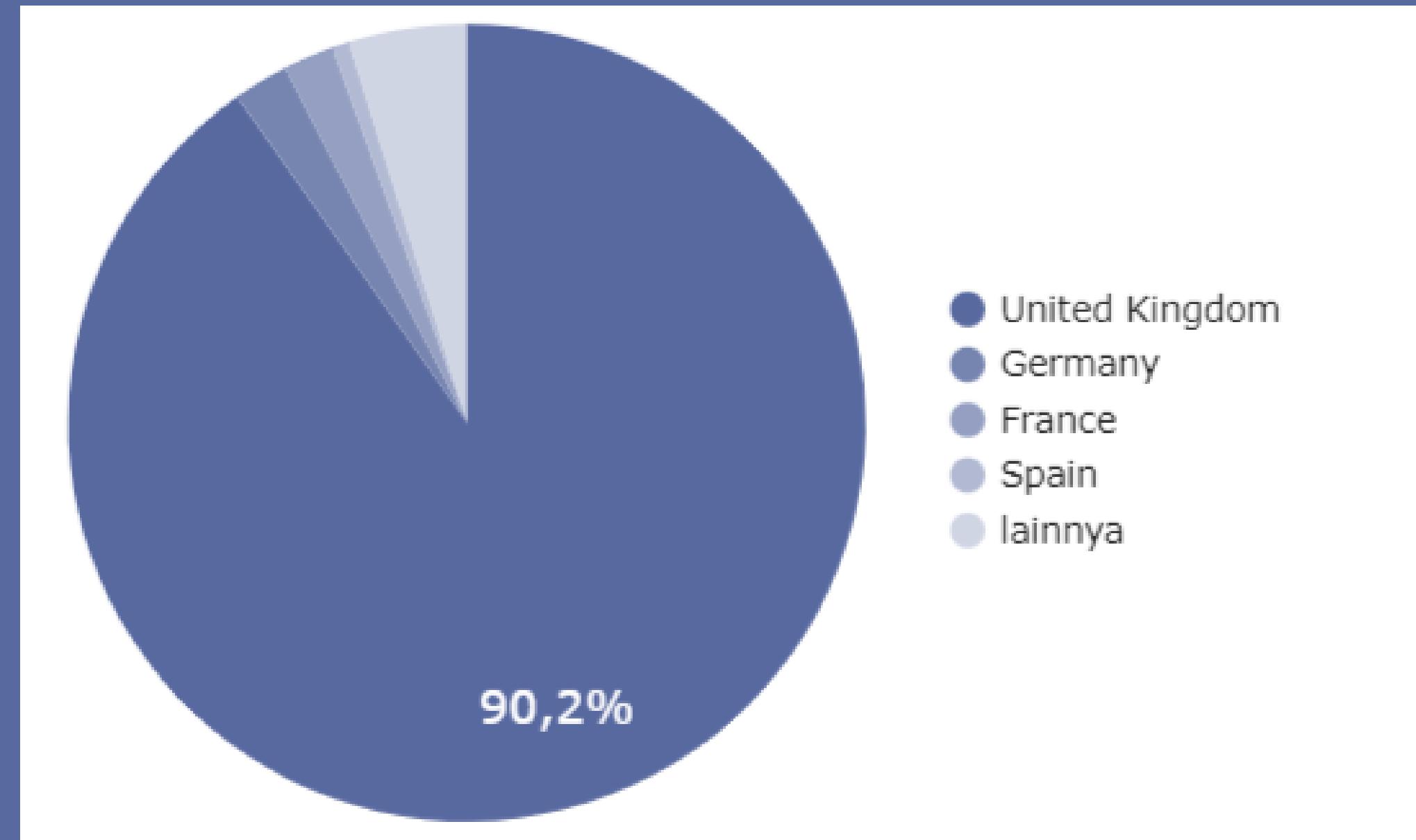
**Buying power terendah** adalah pada kategori **Stationary** sebesar **1,32**.

## Recommendation

Melakukan **adjustment** berdasarkan **buying power** seperti meningkatkan harga dari kategori **Stationary**, **Bags & Purses**, dan **Kitchen** yang memiliki **buying power rendah**.

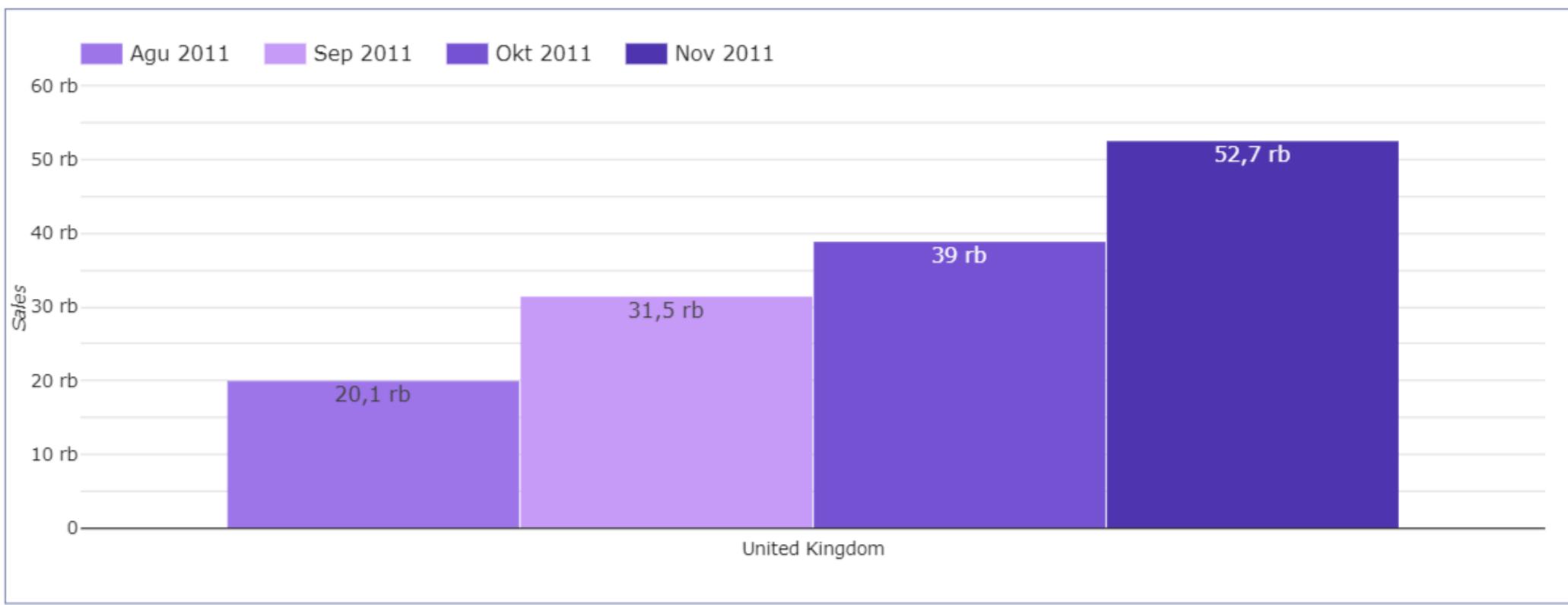
Membuat **bundling package** yang berisi produk yang memiliki **buying power tinggi** dari kategori yang memiliki **sales tertinggi** dan **terendah**.

## Top 5 Customer Based On Country

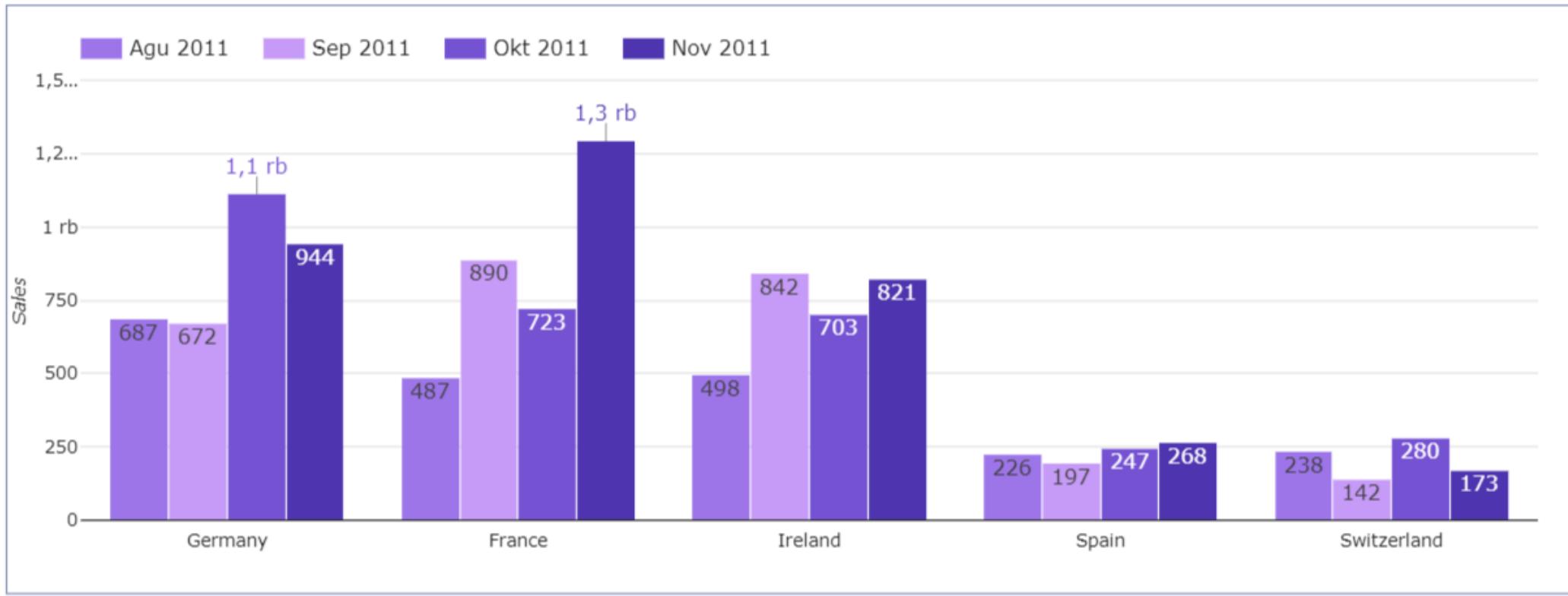


Sales didominasi oleh negara United Kingdom

## Comparison of Sales Based on Country In Aug-Nov 2011 United Kingdom



## Comparison of Sales Based on Country In Aug-Nov 2011 Exclude United Kingdom



## Insight

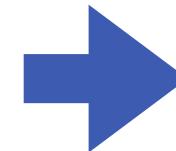
- Sales di negara **United Kingdom** mengalami **kenaikan yang stabil** dari bulan Agustus 2011 sampai November 2011.
- Sales di negara lain (Germany, France, Ireland, Spain, dan Switzerland) setiap bulan nya **tidak stabil** (naik turun) seperti United Kingdom.

## Recommendation

- **Memaksimalkan penjualan produk** yang banyak dibutuhkan oleh masyarakat di negara masing-masing (selain UK).
- **Mengalokasikan budget** dan menjalankan kegiatan **pemasaran** secara gencar di negara **selain UK**
- **Memberikan voucher** gratis atau pemotongan ongkir dengan minimal pembelian tertentu agar konsumen dari negara lain lebih tertarik untuk melakukan pembelian

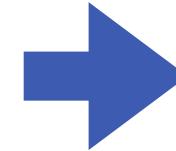
# Recommendation Summary

Kenaikan sales signifikan pada bulan Agustus - November 2011.



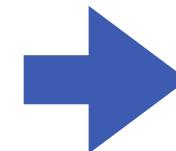
Manfaatkan lonjakan kenaikan sales pada akhir tahun untuk bulan-bulan selanjutnya.

Tren sales dan revenue bulanan.



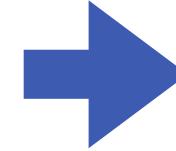
- Memperbanyak kuantitas dari kategori yang penjualannya paling populer.
- Mengoptimalkan pemasaran bagi kategori yang sales quantity-nya rendah.

Buying power tertinggi adalah produk kategori Home. Namun, sales & revenue tertinggi berasal kategori Decoration.



Melakukan adjustment berdasarkan buying power.

Sales didominasi oleh negara UK sebanyak 90,2%.



Mengevaluasi strategi bisnis untuk meningkatkan penjualan di negara lain selain United Kingdom.

# THANK YOU!

---

#SalamHangatDariTeam2-Sabang