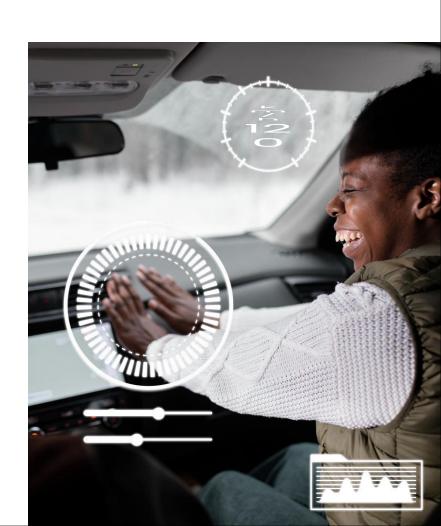
Final Project Data Analyst

Full Stack Data Scientist 30



LETS!

Introduction



"Without learning, there is no change. Without change, there is death."



Hai Everyone!,

I'm, Ferdyansyah Permana Putra

Data Analyst/Scientist Enthusiast

I am a data analyst/scientist enthusiast currently enrolled in a bootcamp program at **DIBIMBING.ID**. I have a strong passion for data, analytics, computation, and visualization to support valuable decision-making.

To support data-driven decisions, I am proficient in various analytical techniques and tools such as Power BI, Tableau, Looker Studio, Excel, SPSS/Minitab, Python, and SQL. I enjoy exploring patterns and insights from data and understanding how data can be optimized to drive business growth. With a strong learning spirit and high curiosity, I am always enthusiastic about facing new challenges in the data world.



@fer.dyput_



linkedin.com/in/ferdypput

Education and Experiences

Education

Full Stack Data Science Batch 30
Bootcamp – Dibimbing.id
26 Nov 2024 - Present

02

Business Statistics – Sepuluh Nopember Institute of Technology 20 Sep 2020 – 29 Sep 2024



Experiences

Data Analyst - PT.
Siantar Top, Tbk.
23 Jan 2025 - Present

02



Data Processing and Visualization - PT. Telekomunikasi Indonesia, Tbk. 16 Feb – 7 Jul 2024

03

Assistant Lecturer of Statistics and Visualization – Sepuluh Nopember Institute of Technology 11 Mar – 26 Des 2024



ON

Overview Project



What have I been working on so far?





Feature Engineering Task

I have been performing data cleaning, data manipulation, and data visualization on various datasets such as Telco Churn, Superstore, and Spotify/Netflix using Python.



SQL Task and Analysis

Understanding and performing SQL analysis on datasets such as DS Salaries, dvdrental, and retail using Window Functions



EDA Analysis

Conducting Exploratory Data Analysis (EDA) on an E-Commerce dataset using Python.

Web Scrapping

Performing web scraping analysis on the Tiket.com website.



Dashboard & Customer Segementation Analysis

Creating dashboards using Power BI with a customer segmentation approach based on RFM.

Dashboard Titles:

- Bee Cycles Market: Customer
 Segmentation and Behavioral Insights
 Dashboard
- Superstore Dashboard: Customer Segmentation Analysis

ON

Final Project (Study Case)



Project Workflow Systematics

01	Study Case Dataset Background (Data Understanding)	04	Case Project Approach Method
02	Business Objective (Problem and Goals)	05	Analysis and Results
03	Feature Engineering (Data Preprocessing)	06	Insights and Recommendations

Link Project:

1. Drive:

https://drive.google.com/drive/fold ers/1raWx31NgAz9-jn ZG51HiQEp1fmtbzj?usp=sharing

2. Google Colab:

https://colab.research.google.com/ drive/1gnLA3gxwt8Tg4uPr9aejeTh P hUBRYFp?usp=sharing



01

Study Case Dataset Background (Data Understanding)



Dataset Project -



I used data obtained from the <u>Kaggle</u> website related to Customer Shopping data in Istanbul. The dataset contains a total of 99,457 rows.



The dataset covers a date interval from the year 2021 to 2023.



The dataset contains the following columns: Invoice_No, Customerld, Gender, Age, Category, Quantity, Price, Payment Method, Invoice_date, and Shopping Mall (10 columns).

DATA UNDERSTANDING

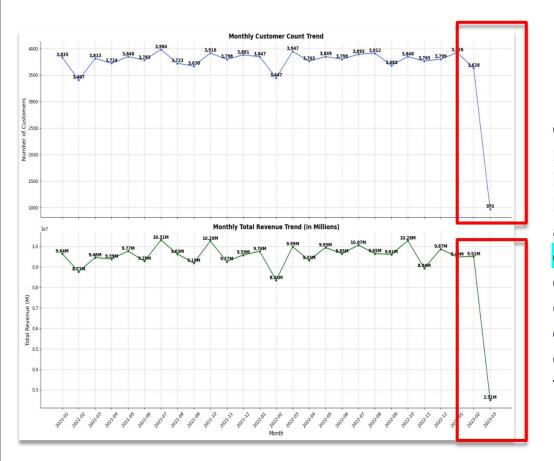
- · invoice_no: Invoice number. A unique combination of the letter 'I' and 6 digits for each transaction.
- · customer_id: Customer ID. A unique combination of the letter 'C' and 6 digits for each customer.
- · gender: Customer's gender.
- · age: Customer's age (positive number).
- · category: Product category purchased.
- · quantity: Number of items purchased per transaction.
- · price: Unit price of the product in Turkish Lira (TL).
- payment_method: Payment method used (cash, credit card, or debit card).
- · invoice_date: Date when the transaction/invoice was created.
- · **shopping_mall**: Name of the shopping mall where the transaction took place.



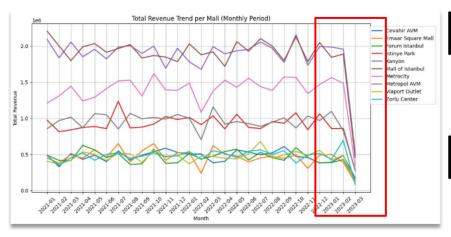
02

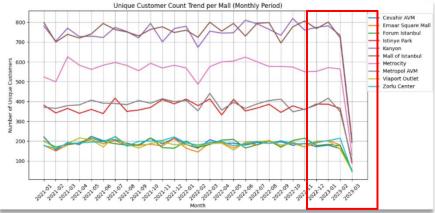
Business Objective (Problem and Goals)





There has been a significant decline in total revenue and the number of transactions since the beginning of 2023. Previously, the business revenue was stable at around 10 million, but it suddenly dropped to just 2.51 million by the end of 2023. In addition, the number of customers, which previously reached over 3,000 per month, has drastically decreased to only 970 customers by the end of 2023.





GENERAL PATTERN

- 1. The trend of customer count and total revenue shows a consistent pattern with several notable fluctuations.
- Large malls remain dominant in attracting customers and generating high revenue.
- 3. There was a sharp decline across all malls in the year 2023.

MALL PERFORMANCES

- Mall of Istanbul & Kanyon: Highest number of customers and revenue (>2 million/month).
- 2. Metrocity: Stable performance with revenue around 1.5 million/month.
- Istinye Park & Metrovol AVM: Mid-range performance with occasional spikes during specific periods.
- 4. Viaport Outlet, Emaar Square Mall, Cevahir AVM, Forum Istanbul, Zorlu Center:
- a. Lower customer count and revenue (rarely exceeding 600K/month).
- b. Sharp fluctuations during popular shopping periods (major discounts/holidays).

WHY?

- 1. There has been a shift in shopping trends, which may be influenced by specific events/promotions, external accessibility factors, shopping experience, and competition.
- 2. Large malls continue to excel in attracting customers and generating revenue. Meanwhile, Smaller malls face challenges in maintaining performance amid changing market conditions.

Both of these issues—the decline in revenue and number of customers since early 2023, as well as the existing performance gap—may be attributed to changes in customer behavior.

Is it possible that competition from similar businesses to Istanbul Store is causing this? Or is there indeed a shift in customer preferences? Or perhaps customers are less interested in this product?

We Need....



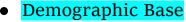
Customer Segmentation



Metode Clustering (Mirip RFM)







Looking at the data, RFM cannot be used because the Frequency is unique for all entries. Therefore, we cannot assess the effectiveness of the RFM segmentation function.

In the case of Istanbul Store, this is used to understand customer shopping patterns (what kind of customer segmentation), anticipate further declines, and find more effective sales strategies.



The main goal is to identify customer segments based on their demographics, develop retention and win-back campaign strategies, and understand purchasing patterns to optimize product offerings.

This analysis is expected to enhance customer experience and brand engagement through more personalized communication, as well as optimize product and promotional strategies by maximizing the sales of flagship products and increasing product appeal



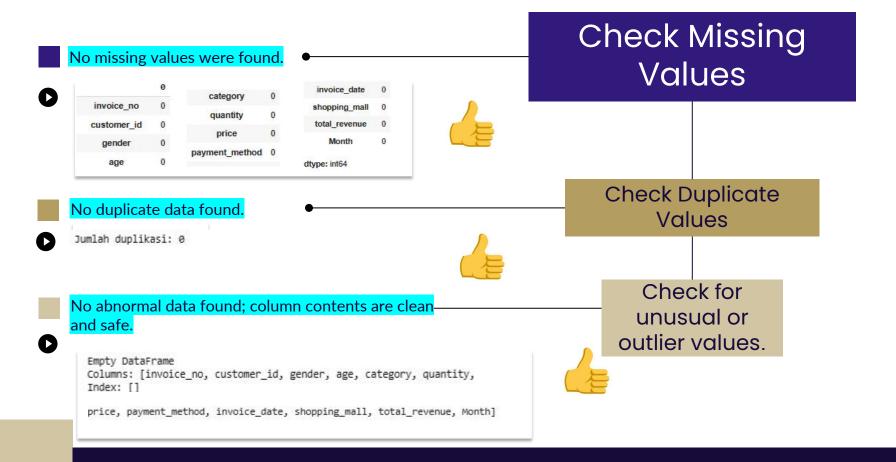
03

Feature Engineering

(Data Cleaning dan Data Manipulation)



Data Cleaning - Data cleaning process to prepare the data for analysis



Data Manipulation -

The process of transforming/creating certain variables to prepare for analysis.

Change of Date Types

shopping_mall	invoice_date	payment_method	price	quantity	category	age	gender	customer_id	invoice_no	
Kanyon	08/05/2022	Credit Card	1500.40	5	Clothing	28	Female	C241288	1138884	0
Forum Istanbul	12/12/2021	Debit Card	1800.51	3	Shoes	21	Male	C111565	1317333	1
Metrocity	11/09/2021	Cash	300.08	1	Clothing	20	Male	C266599	1127801	2
Metropol AVM	05/16/2021	Credit Card	3000.85	5	Shoes	66	Female	C988172	1173702	3
Kanyon	10/24/2021	Cash	60.60	4	Books	53	Female	C189076	1337046	4

D

	invoice_no	customer_id	gender	age	category	quantity	price	payment_method	invoice_date	shopping_mall
0	1138884	C241288	Female	28	Clothing	5	1500.40	Credit Card	2022-08-05	Kanyon
1	1317333	C111565	Male	21	Shoes	3	1800.51	Debit Card	2021-12-12	Forum Istanbul
2	1127801	C266599	Male	20	Clothing	1	300.08	Cash	2021-11-09	Metrocity
3	1173702	C988172	Female	66	Shoes	5	3000.85	Credit Card	2021-05-16	Metropol AVM
4	1337046	C189076	Female	53	Books	4	60.60	Casi	2021-10-24	Kanyon
										sx e

Create of Revenue

Categories

Create of Total Revenue

shopping_mall	total_revenue	Month	Kategori_Umur	Kategori_Pelanggan_Revenue
Kanyon	7502.00	2022- 08	Dewasa	Rendah Revenue
Forum Is <mark>t</mark> anbul	5401.53	2021- 12	Dewasa	Rendah Revenue
Metrocity	300.08	2021- 11	Dewasa	Rendah Revenue
Metropol AVM	15004.25	2021- 05	Tua	Sedang Revenue
Kanyon	242.40	2021- 10	Tua	Rendah Revenue

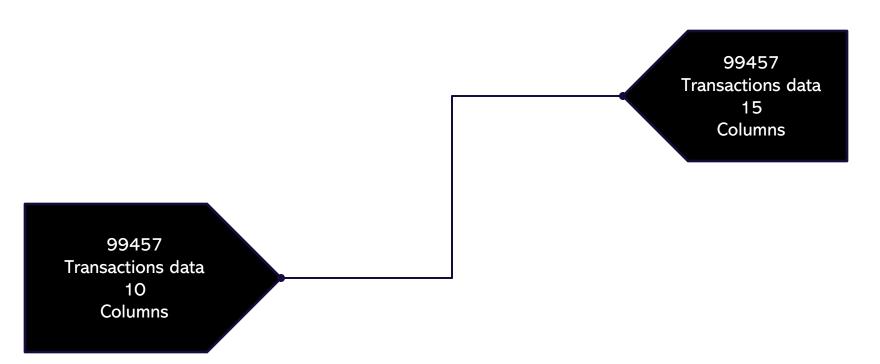
Create of Age Categories

	1mvo1ce_no	customer_1d	Beuger	age	category	quantity	price	payment_method	invoice_date	shopping_mail	total_revenue	Monto	Kategori_Umu
0	1138884	C241288	Female	28	Clothing	5	1500.40	Credit Card	2022-08-05	Kanyon	7502.00	2022-00	Dewas
1	1317333	C111565	Male	21	Shoes	3	1800.51	Debit Card	2021-12-12	Forum Istanbul	5401.53	2021-12	Dewas
2	1127801	C266599	Male	20	Clothing	1	300.08	Cash	2021-11-09	Metrocity	300.08	2021-1	Dewas
3	1173702	C988172	Female	66	Shoes	5	3000.85	Credit Card	2021-05-16	Metropol AVM	15004.25	2021-00	Tu
4	1337046	C189076	Female	53	Books	4	60.60	Cash	2021-10-24	Kanyon	242.40	2021-10	Tu
												-	
99452	1219422	C441542	Female	45	Souvenir	5	58.65	Credit Card	2022-09-21	Kanyon	293.25	2022-00	Dewas
99453	1325143	C569580	Male	27	Food & Beverage	2	10.46	Cash	2021-09-22	Forum Istanbul	20.92	2021-00	Dewas
99454	1824010	C103292	Male	63	Food & Beverage	2	10.46	Debit Card	2021-03-28	Metrocity	20.92	2021-00	Tu
99455	1702984	C800631	Male	56	Technology	4	4200.00	Cash	2021-03-16	Istinye Park	16800.00	2021-00	Tu
99456	1232867	C273973	Female	36	Souvenir	3	35.19	Credit Card	2022-10-15	Mall of Istanbul	105.57	2022-10	Dewas

Create of District/Area

Distrik	Kategori_Pelanggan_Revenue	Kategori_Umur	Month	total_revenue
Levent Beşiktaş	Rendah Revenue	Dewasa	2022- 08	7502.00
Bayrampaşa	Rendah Revenue	Dewasa	2021- 12	5401.53
Levent Beşiktaş	Rendah Revenue	Dewasa	2021- 11	300.08
Ataşehii	Sedang Revenue	Tua	2021- 05	15004.25
Levent Beşiktaş	Rendah Revenue	Tua	2021- 10	242.40

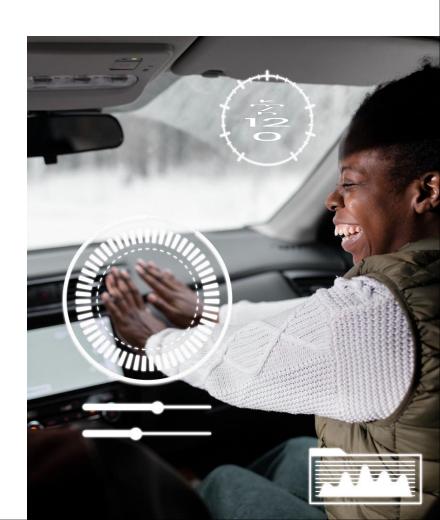
Feature Engineering Result



04

Case Project Approach method

(Customer Segmentation of Demography)



Customer Segementation

(Demographic Base)



Age

Categorizing age into three groups: Young, Adult, and Senior to analyze customer purchase preferences based on age groups



Revenue

Categorizing revenue into three groups: low, medium, and high to analyze contribution preferences toward a product. Additionally, this helps in developing pricing strategies and loyalty programs



Payment Method

Categorizing payment types into three categories: cash, credit, and debit to analyze customer preferences for payment methods, which can be used by the team for targeted promotions or other strategies.

Scoring of Customer Segmentation



Age

< 20 Years</p>

Ordinal: 1

20 - 50 Years

Ordinal: 2

> 50 Years

Ordinal: 3



Revenue

< 10000 Turkish Lira</p>

Ordinal: 1

■ 10000 - 20000 Turkish Lira

Ordinal: 2

> 20000 Turkish Lira

Ordinal: 3





Payment Method

Cash

Ordinal: 1

Debit Card

Ordinal: 2

Credit Card

Ordinal: 3

Segmentation Customer Category

	Total Score	Definition	Recommendations
High-Value Customers	6-9 *with certain criteria	Customers with high spending, usually more mature (adults, teenagers), using modern payment methods (credit/debit card/cash). They are ideal targets for premium or exclusive loyalty programs.	 ✓ Loyalty & VIP Membership with cashback or reward points. ✓ Exclusivity: Special discounts or early-bird access to premium products. ✓ Partnerships with banks: O% installment promos or credit card cashback.
Growing Customers	4-8 *with certain criteria	Customers with medium spending, often from young to adult age groups, flexible in payment methods. Some have high spending with young age and flexible payment. They have potential to become high-value customers if given proper incentives.	 ✓ Incentive-based promotions: Progressive discounts based on increased purchases. ✓ Bundle Deals: Product packages to increase transaction value. ✓ Personalized Offers: Product recommendations based on shopping patterns.
Casual Spenders	4-7 *with certain criteria	Customers with low spending, generally young to adult, often using debit or credit cards. They may only purchase during promotions or urgent needs.	 ✓ Flash Sales & Seasonal Promotions to attract purchases during specific events. ✓ Buy More, Save More: Discounts for larger quantity purchases. ✓ Payment Method Promotions: Discounts/cashback when using certain payment methods.
Low- Engagement Customers	3-5 *with certain criteria	Customers with very low spending, usually paying cash and from less active shopping age groups (very young or very old). They may not be main targets but can be increased with more attractive offers.	 ✓ Customer Reactivation: Special discounts for next purchase. ✓ Referral Program: Incentives for inviting friends to shop. ✓ Interactive Communication: Educational content or attractive promotions via social media.

*with certain criteria: I have categorized/classified it based on the criteria according to its definition



ISTANBUL STORE: CUSTOMER SHOPPING DASHBOARD

05 + 06

Analysis Insights and and

nu and

Results Recommen

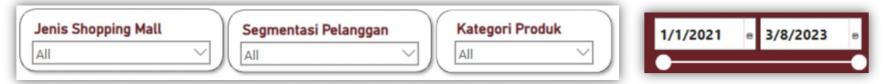
dations

(Dashboard Visualization Results from the Project)

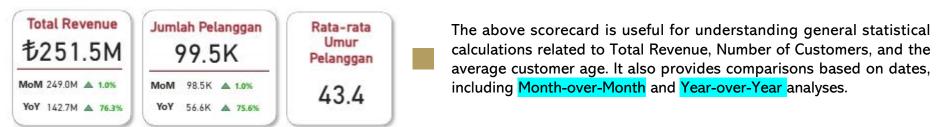


Filters

Filters to customize dashboard visualizations according to user needs



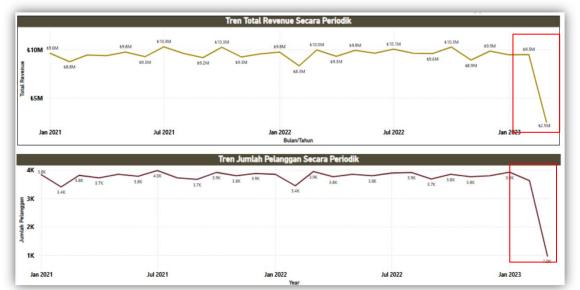
Scorecard



In general, what is the condition of total revenue, the number of customers, and the average customer age?

- The total revenue amounts to approximately **251.5 million** Turkish Lira.
- The total number of customers is around 99,500 during the period from 2021 to 2023.
- The average customer age is about 43 years old.

How are the sales trends and customer counts behaving over time (periodically)?



Insights:



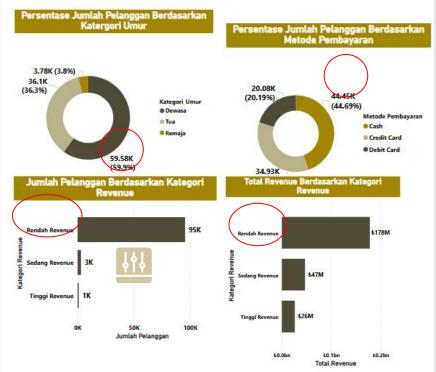
Sales and customer counts dropped sharply in 2023:

- Revenue fell from 9.5M to 2.5M
- Customers dropped from 3.8K to 1K
- Indicates significant external or internal factors

From 2021-2022, fluctuations were stable, but the 2023 decline shows issues with customer retention

- Analysis of the Causes of the Drastic Decline (Check for seasonal factors, economic crisis, or business policies) + Customer Surveys to understand preferences
- Customer Retention Strategy (Loyalty programs for existing customers, optimizing exclusive discounts/promotions to attract customers)
- Marketing Strategy based on Digital Advertising and Collaboration with Influencers or other effective marketing methods
- Product or Business Model Diversification (Adjusting Products/Services according to customer demand)

What are the customer characteristics based on age, payment method, and their contribution to total revenue, and how can business strategies improve the transaction value per customer?



Insights:



The majority of customers are **adults and seniors**, likely with higher purchasing power. However, they prefer cash or debit over credit cards, indicating a **tendency toward direct payment** and possibly a more conservative financial culture



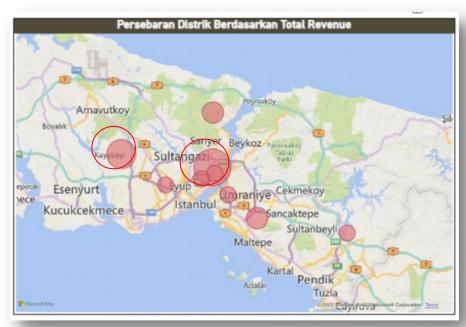
Most customers fall into the low-revenue category and contribute the most to total revenue, indicating that the business relies heavily on high transaction volume from small spenders. Therefore, revenue growth strategies should focus on increasing the value of each transaction through up-selling, cross-selling, or encouraging larger purchases per customer.



- 1. Encourage Card Usage (through promotions or cashback)
- 2. Increase Purchase per Customer (via up-selling, cross-selling, or product bundling)
- **3.** Optimize High-Value Segments (create loyalty programs or exclusive offers to attract customers with higher purchasing power)

Revenue Distribution

Which District Contributes the Most to Revenue?



Insight:

The most prominent districts in terms of total revenue contribution are **Levent and Besiktas**, generating around **87 million** Turkish Lira. The second highest is **Basaksehir**, with approximately **50 million** Turkish Lira.



D Pelanggan	Jenis Kelamin	Umur	Produk	Jenis Mall	Kategori Revenu
C100795	Female	44	Books	Mall of Istanbul	Rendah Revenue
C100919	Male	69	Books	Metrocity	Rendah Revenue
C101035	Male	18	Books	Metrocity	Rendah Revenue
C101055	Male	46	Books	Kanyon	Rendah Revenue
C101133	Female	46	Books	Metrocity	Rendah Revenue
C101238	Female	22	Books	Zorlu Center	Rendah Revenue
C101347	Male	37	Books	Kanyon	Rendah Revenue
C101482	Male	23	Books	Metropol AVM	Rendah Revenue
C101531	Female	62	Books	Forum Istanbul	Rendah Revenue
C101612	Female	37	Books	Kanyon	Rendah Revenue
C101709	Male	68	Books	Kanyon	Rendah Revenue
C101726	Male	59	Books	Kanyon	Rendah Revenue
C101833	Male	69	Books	Istinye Park	Rendah Revenue
C102024	Female	46	Books	Emaar Square Mall	Rendah Revenue
Total		4319128			

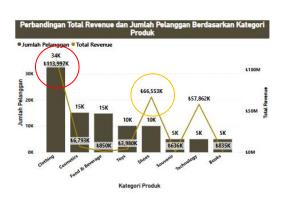
- 1. Strengthen Levent & Besiktas (Maximize promotions and customer retention strategies)
- Boost Basaksehir (Increase market penetration through promotions and product expansion)
- **3. Develop Other Districts** (Analyze expansion opportunities in potential areas)





What is the **contribution** of each customer segment and product category to total revenue, and what strategies can be implemented to increase transaction value?





Insights:



Customer Contribution:

The majority of revenue comes from Casual Spenders (53K customers, \$98M) and Low-Engagement Customers (43K customers, \$80M), while Growing Customers (\$48M) and High-Value Customers (\$26M) contribute less and are fewer in number. This indicates that revenue heavily relies on low-value, less loyal customers.



Kontribusi Produk:

Clothing & Accessories generate the **highest revenue** (\$113.9M) with the largest customer base (34K), indicating strong popularity. Shoes and Bags have high revenue (\$66.5M and \$57.8M) despite fewer customers, suggesting higher average prices. Food & Beverages attract a fair number of customers (15K) but contribute lower revenue (\$6.7M), reflecting lower spend per customer.



- 1. Increase Loyalty (Create Reward/Discount Programs)
- Promote High-Value Products (Focus on Shoes and use bundling strategies for Food & Beverages)
- 3. Optimize High-Value Customers (Offer Exclusive Deals)

Full Dashboard Table Filter ISTANBUL STORE: CUSTOMER SHOPPING DASHBOARD Total Revenue Scorecard Segmentasi Pelanggan Kategori Produk ₺251.5M 99.5K Pelanggan 43.4 Line Chart Pie Chark **Bar Chart Bubble Map** Combo Chart Chart Copyrights @2024 by Ferdyansyah Permana Putra

Conclusions Overall can we get



Increase Loyalty & Retention + Conduct Customer Preference Surveys

Create reward programs, discounts, and membership plans to encourage repeat transactions.

At the same time, conduct customer or market surveys to understand customer preferences and behaviors.



Promote high-value products

Focus on promoting Shoes and bundling Food & Beverages.



Optimize Potential Segments & Regions

Maximize marketing efforts in Levent, Besiktas, and Basksehit.



Market Expansion & Growth

Analyze districts and customer segments with low contributions to identify growth opportunities.



Thanks!









Do you have any questions? ferdyansyahpputra@gmail.com