Report

Brazilian E-Commerce Sales

& Inventory Optimization

(ABC/XYZ Analysis)

Date Analyst: Andrii Isachenko

2025

1. Business Problem

The goal of the project is to optimize inventory management of the Olist marketplace.

Key business questions:

- Which product categories generate 80% of revenue?

- Which products have stable demand, and which have chaotic demand?

- How to prevent the warehouse from being overstocked with dead stock and a shortage of sales leaders?

2. Data Acquisition

1. **Source:** [Brazilian E-Commerce Public Dataset by Olist](#) on [Kaggle.](#)

2. **Download:** The following relational tables were downloaded in CSV format:

olist_order_items_dataset.csv - Main table with prices and ID

olist_products_dataset.csv - Directory of products and categories

product_category_name_translation.csv - Vocabulary: Portuguese -> English

olist_orders_dataset.csv - Order dates

3. **Import:** The data was imported into Google Sheets / MS Excel for further processing.

3. Data Preparation and Consolidation

This stage included the creation of a single **Master_Data** register:

1. **Merging:** Using the XLOOKUP function, data from different tables was merged into one using the keys order_id and product_id.

2. **Localization:** Portuguese category names were replaced with English ones using a translation table.

3. **Cleaning:**

◦ Converting the Order_Date column to a date format for correct grouping by month.

◦ Data types processed: Price (Number), Order_Date (Date), Category (Text).

4. Metrics Calculation and ABC Analysis

For each category, aggregated indicators were calculated:

1. **Revenue:** Sum of price for all units sold.

2. **Ranking:** Categories sorted by decreasing revenue.

3. **Cumulative Share:** Cumulative percentage calculation to determine group boundaries:

◦ **Group A:** Top 80% of revenue.

◦ **Group B:** Next 15%.

◦ **Group C:** Remaining 5%.

ABC Analysis Matrix

Category	Revenue	Cumulative Revenue	Cumulative %	ABC
housewares	923k	923k	10,55%	A
health_beauty	771k	1 694k	19,35%	A
sports_leisure	665k	2 360k	26,95%	A
electronics	623k	2 983k	34,07%	A
garden_tools	557k	3 539k	40,43%	A
computers_accessories	458k	3 998k	45,66%	A
furniture_decor	330k	4 327k	49,43%	A
baby	321k	4 649k	53,10%	A
cool_stuff	318k	4 967k	56,74%	A
watches_gifts	305k	5 272k	60,22%	A
perfumery	271k	5 543k	63,32%	A
Other	260k	5 803k	66,28%	A
toys	258k	6 060k	69,23%	A
auto	256k	6 317k	72,15%	A
stationery	240k	6 556k	74,89%	A
home_appliances	220k	6 777k	77,41%	A
telephony	198k	6 974k	79,66%	A
bed_bath_table	168k	7 142k	81,58%	B
home_appliances_2	156k	7 298k	83,36%	B
office_furniture	122k	7 420k	84,76%	B
fashion_bags_accessories	118k	7 538k	86,10%	B
computers	106k	7 644k	87,31%	B
pet_shop	89k	7 733k	88,33%	B
musical_instruments	84k	7 818k	89,30%	B
small_appliances	80k	7 897k	90,21%	B
consoles_games	74k	7 972k	91,06%	B
construction_tools_construction	68k	8 040k	91,84%	B
luggage_accessories	65k	8 105k	92,58%	B
food	59k	8 164k	93,25%	B
art	54k	8 218k	93,87%	B
signaling_and_security	52k	8 270k	94,46%	B
christmas_supplies	49k	8 319k	95,02%	C
agro_industry_and_commerce	35k	8 354k	95,42%	C
home_construction	35k	8 388k	95,81%	C
furniture_living_room	33k	8 421k	96,19%	C
fixed_telephony	30k	8 451k	96,53%	C

home_confort	28k	8 479k	96,85%	C
air_conditioning	25k	8 503k	97,13%	C
small_appliances_home_oven_and_coffee	24k	8 527k	97,40%	C
kitchen_dining_laundry_garden_furniture	24k	8 550k	97,67%	C
construction_tools_safety	22k	8 573k	97,92%	C
audio	22k	8 594k	98,17%	C
books_general_interest	21k	8 616k	98,41%	C
construction_tools_lights	18k	8 634k	98,62%	C
industry_commerce_and_business	18k	8 652k	98,83%	C
market_place	12k	8 664k	98,96%	C
fashion_shoes	10k	8 674k	99,08%	C
costruction_tools_garden	10k	8 684k	99,19%	C
books_technical	9k	8 693k	99,30%	C
drinks	7k	8 701k	99,38%	C
furniture_bedroom	7k	8 708k	99,47%	C
food_drink	7k	8 715k	99,55%	C
fashion_male_clothing	6k	8 721k	99,62%	C
costruction_tools_tools	5k	8 727k	99,68%	C
fashion_underwear_beach	5k	8 731k	99,73%	C
music	4k	8 736k	99,78%	C
tablets_printing_image	3k	8 739k	99,82%	C
cine_photo	3k	8 742k	99,85%	C
furniture_mattress_and_upholstery	2k	8 744k	99,88%	C
party_supplies	2k	8 746k	99,90%	C
dvds_blu_ray	2k	8 747k	99,91%	C
books_imported	2k	8 749k	99,93%	C
fashio_female_clothing	1k	8 750k	99,95%	C
fashion_sport	1k	8 751k	99,96%	C
arts_and_craftmanship	922	8 752k	99,97%	C
la_cuisine	719	8 753k	99,98%	C
diapers_and_hygiene	599	8 753k	99,98%	C
cds_dvds_musicals	400	8 754k	99,99%	C
flowers	297	8 754k	99,99%	C
home_comfort_2	275	8 754k	100,00%	C
security_and_services	183	8 755k	100,00%	C
fashion_childrens_clothes	160	8 755k	100,00%	C
Overall result	**8 754 696,06**			

Explanation of the methodology

-       **Group A** (80% of revenue): The most valuable categories. These are the "locomotives" of the business. They require maximum attention and strict control of availability.

-       **Group B** (15% of revenue): Medium importance. Categories that provide stability but are not the main drivers.

- **Group C** (5% of revenue): The least profitable categories. They make up the "tail" of the assortment. They need optimization to avoid wasting resources on their maintenance.

Key findings

High concentration: Only 17 categories (out of over 70) generate 80% of the company's total revenue. This indicates a high dependence of the result on a narrow range of products.

Leaders: Top 3 categories (housewares, health_beauty, sports_leisure) generate over 25% of revenue. Any disruptions in their operations will have a critical impact on profits.

Tail efficiency: Over 40 categories are in Group C. They generate less than 5% of revenue in total. This is a reason to reconsider the feasibility of keeping them in stock.

Recommendations

For group A: Ensure 100% availability in stock (Service Level 98-99%). Conduct regular analysis of competitors' prices.

For Group B: Review purchases once a month. Look for opportunities to move them to Group A through marketing.

For group C: Consider withdrawing from the assortment or switching to the "Drop-shipping" model (selling directly from the supplier's warehouse) so as not to freeze funds.

5. XYZ-analysis: Estimation of stability (Demand Volatility)

The coefficient of variation (CV) was used to determine the unpredictability of demand:

1. A summary table of sales by month was formed.

2. The average value (AVERAGE) and standard deviation (STDEV.P) were calculated.

3. Formula CV: CV = STDEV.P / AVERAGE

4. **Custom segmentation logic:** Implemented formula to handle anomalies (e.g., single sales з CV=0):

```
=IFERROR(IFS(CV=0; "Z"; CV>0.8; "Z"; CV>0.4; "Y"; TRUE; "X"); "Z").
```

XYZ Analysis Matrix

Category_EN	Average	STDEV	CV	XYZ
diapers_and_hygiene	3,80	1,30	0,34	X
la_cuisine	1,20	0,45	0,37	X
air_conditioning	6,43	4,27	0,66	Y
audio	7,90	4,17	0,53	Y
auto	86,67	52,58	0,61	Y
baby	65,71	34,97	0,53	Y
bed_bath_table	238,10	117,18	0,49	Y
books_general_interest	12,19	7,87	0,65	Y
books_imported	2,27	1,49	0,66	Y
cds_dvds_musicals	2,00	1,15	0,58	Y
christmas_supplies	4,91	3,18	0,65	Y
computers	9,18	6,19	0,67	Y
computers_accessories	170,81	112,57	0,66	Y
consoles_games	25,10	15,22	0,61	Y
construction_tools_safety	8,38	5,90	0,70	Y
cool_stuff	79,76	38,93	0,49	Y
costruction_tools_garden	5,00	2,94	0,59	Y
costruction_tools_tools	3,83	2,17	0,57	Y
drinks	9,73	6,82	0,70	Y
dvds_blu_ray	2,20	1,32	0,60	Y
electronics	62,55	48,84	0,78	Y
fashio_female_clothing	2,09	1,22	0,58	Y
fashion_bags_accessories	44,00	20,43	0,46	Y
fashion_male_clothing	4,47	2,64	0,59	Y
fashion_shoes	5,35	3,18	0,60	Y
fashion_sport	1,36	0,92	0,68	Y
fashion_underwear_beach	3,72	2,08	0,56	Y
fixed_telephony	5,10	3,24	0,64	Y
flowers	1,80	0,84	0,46	Y
food_drink	7,31	3,88	0,53	Y
furniture_bedroom	2,59	1,66	0,64	Y
furniture_decor	169,09	82,34	0,49	Y
furniture_living_room	11,55	4,82	0,42	Y
furniture_mattress_and_upholstery	2,71	1,80	0,66	Y
garden_tools	94,00	54,65	0,58	Y
health_beauty	201,57	114,75	0,57	Y
home_appliances	18,47	13,10	0,71	Y
home_appliances_2	5,95	4,03	0,68	Y
home_comfort_2	1,60	1,07	0,67	Y
home_confort	10,42	4,99	0,48	Y
housewares	149,95	81,70	0,54	Y
kitchen_dining_laundry_garden_furniture	6,62	5,24	0,79	Y
luggage_accessories	25,20	13,83	0,55	Y
market_place	6,80	3,14	0,46	Y

music	1,75	0,75	0,43	Y
musical_instruments	16,05	9,11	0,57	Y
office_furniture	36,90	22,94	0,62	Y
Other	34,81	20,79	0,60	Y
perfumery	74,52	36,16	0,49	Y
pet_shop	39,19	22,33	0,57	Y
small_appliances	14,80	7,39	0,50	Y
small_appliances_home_oven_and_coffee	4,44	3,47	0,78	Y
sports_leisure	183,86	92,40	0,50	Y
stationery	54,65	41,97	0,77	Y
tablets_printing_image	3,07	2,20	0,72	Y
telephony	97,14	46,42	0,48	Y
toys	86,67	50,76	0,59	Y
watches_gifts	132,29	87,59	0,66	Y
agro_industry_and_commerce	5,28	4,24	0,80	Z
art	5,53	5,01	0,91	Z
arts_and_craftmanship	3,50	3,79	1,08	Z
books_technical	7,41	6,64	0,90	Z
cine_photo	2,33	2,53	1,09	Z
construction_tools_construction	22,44	22,48	1,00	Z
construction_tools_lights	11,25	11,27	1,00	Z
fashion_childrens_clothes	1,00	0,00	0,00	Z
food	10,86	11,62	1,07	Z
home_construction	15,11	13,83	0,91	Z
industry_commerce_and_business	7,38	7,02	0,95	Z
party_supplies	2,11	1,76	0,84	Z
security_and_services	1,00	0,00	0,00	Z
signaling_and_security	4,82	4,32	0,90	Z

Explanation of methodology

Coefficient of variation (CV): This is the head indicator. The lower the price, the more stable sales.

- **X** (CV ≤ 0.4): Висока стабільність. Коливання попиту менше 40% від середнього. Попит майже «плаский», закупівлі можна автоматизувати.

- **Y** (0.4 < CV ≤ 0.8): Moderate stability. Kolivannya is present (up to 80%), but there is a distinct pattern (seasonality). Safety stock required.

- **Z** (CV > 0.8): Low stability. Chaotic sales. The volume exceeds 80% of the average level. The forecast carries a high risk of death.

For marketplaces (like Olist): limits (40% and 80%) are optimal, as the assortment is huge and demand is often "noisy".

Key Findings

Based on the data provided, the following important points can be highlighted:

Processing feature CV = 0:

Although mathematically zero variation indicates perfect stability, within the scope of this project, a value of 0 is interpreted as a single sale or insufficient data for forecasting.

Therefore, such items are forcibly assigned to class Z to avoid erroneous planning of large purchases.

After analyzing the data, the following can be highlighted:

Lack of stability (Class X): There is almost no "pure" class X among the categories (only diapers_and_hygiene with CV 0.34 and la_cuisine with CV 0.37).

This indicates that Olist's business operates in conditions of high demand volatility.

Class Y Dominance: Most key categories, such as health_beauty (0.57), housewares (0.54), and furniture_decor (0.49), fall into the Y zone.

This is "stable volatility" — demand is always there, but its volume fluctuates within 50% of the average.

High-risk areas **(Class Z**): The categories agro_industry (0.80), art (0.91), and construction_tools_lights (1.00) have critical volatility.

It is impossible to build an accurate forecast for them based on past periods.

Operational recommendations

What to do with categories based on their XYZ status:

For X (Stable): You can switch to automatic replenishment of stocks. The risk of "shortage" is minimal.

For Y (Floating): It is necessary to maintain a Safety Stock. For most of your categories (CV ≈ 0.5-0.6), the recommended stock is 20-25% of expected monthly sales.

For Z (Unpredictable): The optimal strategy is to "make to order" or purchase in very small batches.

Don't try to guess the volume a quarter in advance - you will either freeze money or be left without the product.

6. Creating the final matrix (Total_Matrix)

Combining the results into a matrix (AX, AY, AZ ... CZ). A strategy is developed for each segment:

　　　• **Segment AY:** Maintaining safety stock (Safety Stock).

　　　• **Segment AZ:** Purchasing model "in small batches".

　　　• **Segment CZ:** Withdrawal from the assortment or working according to the dropshipping scheme.

Total Matrix	Category
AY	17
BY	10
BZ	4
CX	2
CY	29
CZ	10
Total	72

Total_Matrix

Category	ABC_Class	XYZ_Class	Final_Segment
cool_stuff	A	Y	AY
pet_shop	B	Y	BY
furniture_decor	A	Y	AY
perfumery	A	Y	AY
garden_tools	A	Y	AY
housewares	A	Y	AY
telephony	A	Y	AY
health_beauty	A	Y	AY
books_technical	C	Z	CZ
fashion_bags_accessories	B	Y	BY
bed_bath_table	B	Y	BY
sports_leisure	A	Y	AY
consoles_games	B	Y	BY
office_furniture	B	Y	BY
luggage_accessories	B	Y	BY
food	B	Z	BZ
agro_industry_and_commerce	C	Z	CZ
electronics	A	Y	AY
computers_accessories	A	Y	AY
construction_tools_construction	B	Z	BZ
audio	C	Y	CY
baby	A	Y	AY
construction_tools_lights	C	Z	CZ
toys	A	Y	AY
stationery	A	Y	AY

industry_commerce_and_business	C	Z	CZ
watches_gifts	A	Y	AY
auto	A	Y	AY
Other	A	Y	AY
home_appliances	A	Y	AY
kitchen_dining_laundry_garden_furniture	C	Y	CY
air_conditioning	C	Y	CY
home_confort	C	Y	CY
fixed_telephony	C	Y	CY
small_appliances_home_oven_and_coffee	C	Y	CY
diapers_and_hygiene	C	X	CX
signaling_and_security	B	Z	BZ
musical_instruments	B	Y	BY
small_appliances	B	Y	BY
costruction_tools_garden	C	Y	CY
art	B	Z	BZ
home_construction	C	Z	CZ
books_general_interest	C	Y	CY
party_supplies	C	Z	CZ
construction_tools_safety	C	Y	CY
cine_photo	C	Z	CZ
fashion_underwear_beach	C	Y	CY
fashion_male_clothing	C	Y	CY
food_drink	C	Y	CY
drinks	C	Y	CY
furniture_living_room	C	Y	CY
market_place	C	Y	CY
music	C	Y	CY
fashion_shoes	C	Y	CY
flowers	C	Y	CY
home_appliances_2	B	Y	BY
fashio_female_clothing	C	Y	CY
computers	B	Y	BY
books_imported	C	Y	CY
christmas_supplies	C	Y	CY
furniture_bedroom	C	Y	CY
home_comfort_2	C	Y	CY
dvds_blu_ray	C	Y	CY
cds_dvds_musicals	C	Y	CY
arts_and_craftmanship	C	Z	CZ
furniture_mattress_and_upholstery	C	Y	CY
tablets_printing_image	C	Y	CY
costruction_tools_tools	C	Y	CY
fashion_sport	C	Y	CY
la_cuisine	C	X	CX
security_and_services	C	Z	CZ
fashion_childrens_clothes	C	Z	CZ

Strategy Matrix:

Group AX / AY (Stars): High income, stable or forecasted demand. This is the "foundation" of profit.

    Strategy: Maximum availability in stock, priority in logistics.

Group AZ (Risk Leaders): Brings a lot of money, but demand is chaotic.

    Strategy: Increased control of balances to avoid losing sales during peak times.

Group BX / BY (Middle class): Stable second-tier income.

    Strategy: Optimize procurement to reduce storage costs.

CX / CY Group (Mass Tail): Little money, but the demand is clear.

    Strategy: Automation. Spend a minimum of manager time on these positions.

Group CZ (Exit Candidates): Low income and complete unpredictability.

    Strategy: Reduction of assortment or work purely "to order".

Strategic conclusions:

Stable Core (AY): Your top categories—health_beauty, computers_accessories, watches_gifts, auto — are all AY. This is ideal. They generate the core revenue, and while demand fluctuates, it is manageable.

Unique Finds (CX): The diapers_and_hygiene and la_cuisine categories are in CX. They don't bring in much money, but they work like clockwork (very stable). These are good magnet products for regular customers.

Clearing Zone (CZ): A large list of categories in the CZ segment (arts_and_craftmanship, security_and_services, agro_industry). These categories "eat up" the attention of managers, do not bring significant profits and constantly create problems with forecasting. They should be checked for profitability.

Dangerous Leaders (AZ): The housewares and perfumery categories have AZ status.

These are your "money bags" that can suddenly run out of stock or, conversely, become stagnant. They require an individual approach to purchasing.

Recommendations

For AY/AX segments: Ensure 98% service level (availability). Any absence of these products is a direct loss of large profits.
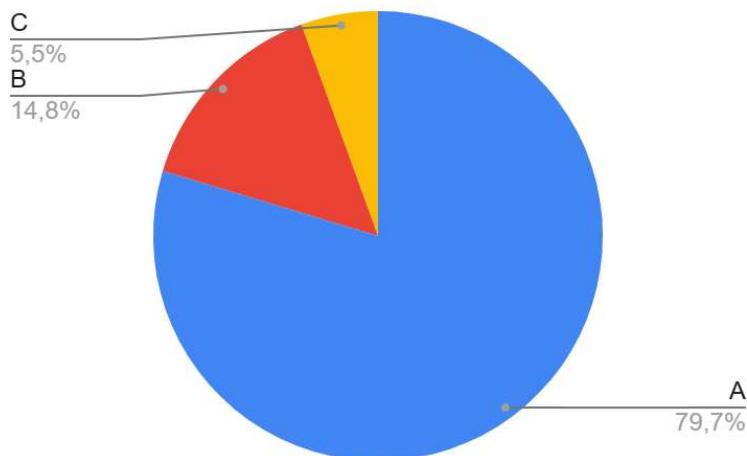
For BY/CY segments: Try to negotiate with suppliers for smaller but more frequent deliveries so as not to freeze working capital.

For CZ segments: Conduct an audit. If these categories are not "strategically important" (e.g., for traffic generation), it is better to remove them from the assortment.
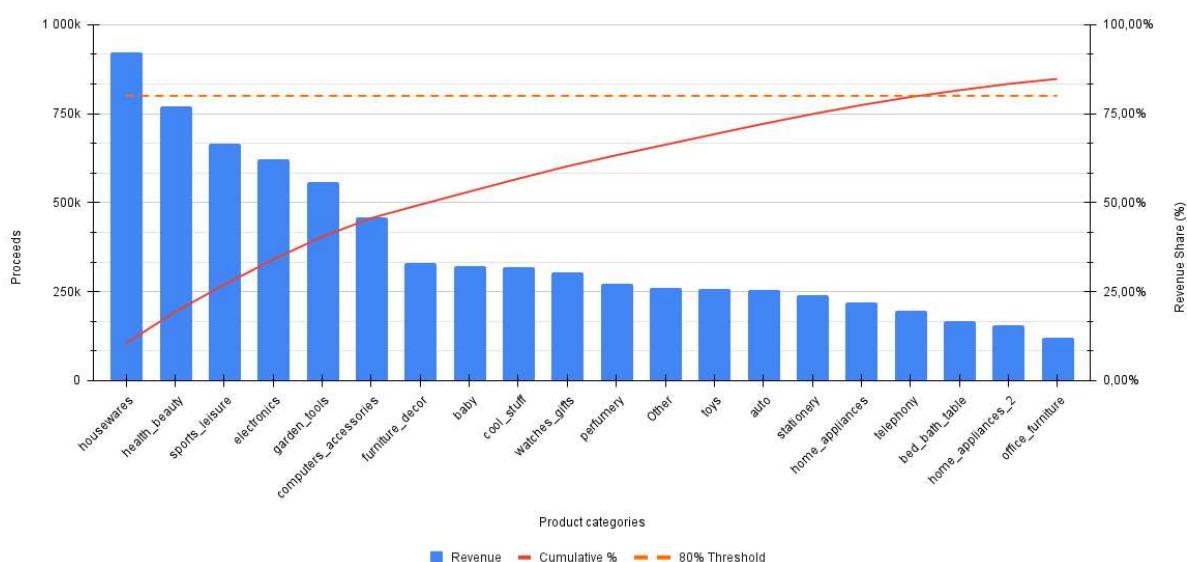
7. Visualization and Dashboard

An interactive dashboard has been created:
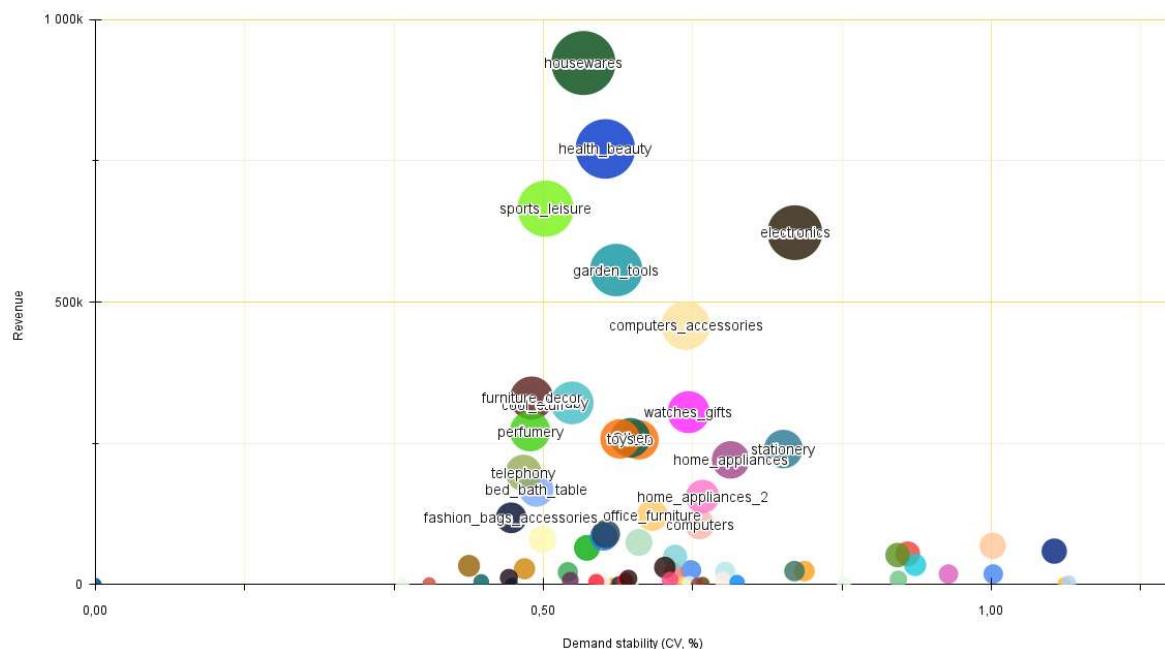
- **Revenue structure by class**

-
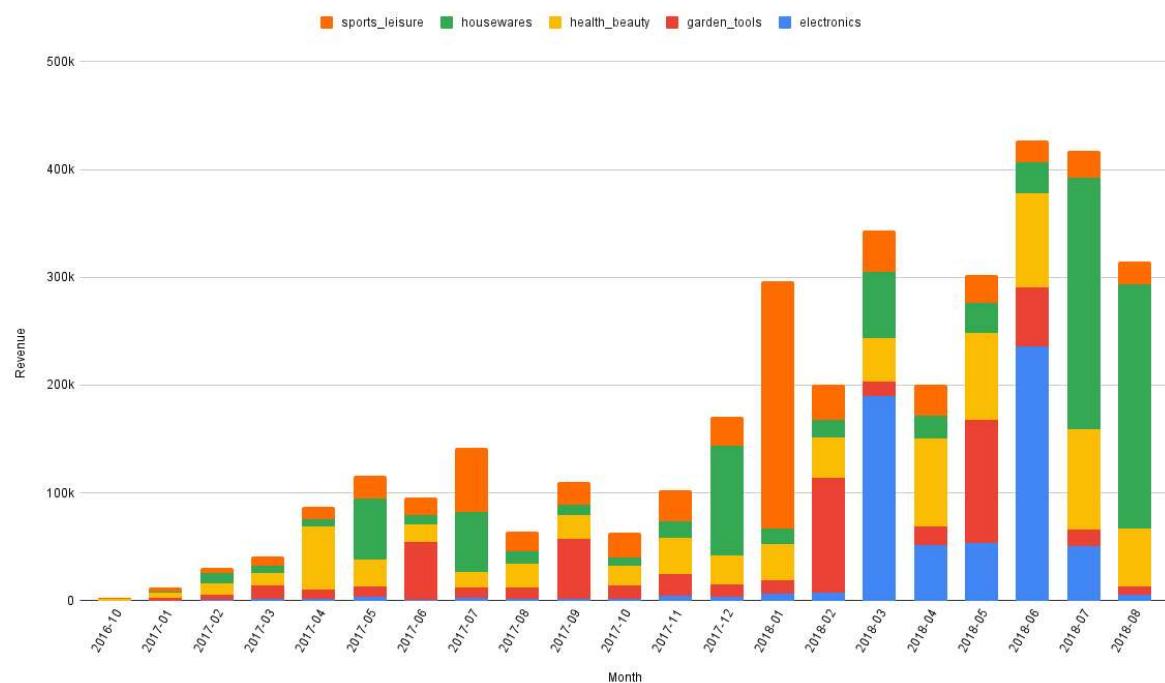
- **Pareto Chart:** Visualization of revenue concentration.

- **ABC-XYZ analysis matrix:** Distribution of categories by matrix segments.

- **Monthly Dynamics:** Seasonality graph for peak load forecasting.

8 Conclusions and recommendations

• **Concentration:** It was found that 15% of categories bring 80% of the money (classic Pareto rule).

• **Seasonality:** Peaks were found in May-July and November.

• **Optimization:** It is proposed to reduce inventories in the CZ group, which will free up up to 15% of warehouse space.

9. Pivot Tables Used (Pivot Tables Structure)

To transform raw data into Master_Data, 3 key pivot tables were built:

1. **Sales by Category (for ABC):**
   ◦ Rows: Category_EN
   ◦ Value: SUM of price
   ◦ Purpose: Identifying leaders by sales.

2. **Monthly Demand Matrix (for XYZ):**
   ◦ Rows: Category_EN
   ◦ Columns: Order_Date (grouped by months)
   ◦ Value: COUNTA of order_id
   ◦ Purpose: Calculating demand stability (CV).

3. **The seasonality chart of the leading categories**
   ◦ Rows: Monts,
   ◦ Columns: Category_EN
   ◦ Value: Price
   ◦ Filters: Category_EN
   ◦ Purpose: Building a seasonality chart of leading categories

10 Detailed conclusions and business recommendations

Analysis of 50,499 orders was conducted. The product portfolio was divided into segments based on their contribution to revenue (ABC) and demand stability (XYZ).

Based on the analysis, a strategic assortment management plan was formed:

Strategic conclusions:

1. **Pareto effect:** 80% of revenue is provided by only 15% of categories. Any error in the availability of group A products leads to critical profit losses.

2. **Stability imbalance:** Most of the top categories (health_beauty, housewares) belong to class Y, which requires mandatory calculation of safety stock (Safety Stock) due to seasonal fluctuations.

3. **Inefficient assets:** A significant list of categories in the CZ segment (low revenue + chaotic demand) was identified that create a load on logistics without economic benefit.

Recommendations (Action Plan):

• **For segments AX / AY (Stars):** Implement automatic replenishment of stocks. Target service level (Service Level) — 98%.

• **For segment AZ (High Risk Leaders):** Switch to the "Just-in-Time" or small-lot procurement model to avoid freezing large amounts of goods that suddenly stopped selling.

• **For segments CX / CY (Stable Tail):** Optimize management costs. Automate orders with a low frequency of checks.

• **For segment CZ (Candidates for Exclusion):**

    1. Conduct an audit of margin.

    2. If the margin is low — remove from the assortment (De-listing).

    3. If the category is image-related — switch to the model of direct deliveries from the manufacturer (Dropshipping).