Technical analysis of pricing push-ups strategy

BY IGNAS JANCIUKAS

1.1 Dataset description

The dataset contains aggregated metrics for the pricing push-up feature at the sub-category level. Descriptions of columns include detailed explanations of what each represents, data types, aggregation types, and relationship rules. Constraints and specific business rules are omitted as they do not exist and are not relevant for this analysis. NULL values are not expected in any column; if found, handling will be explained during quality testing. Current price of the feature for the sellers is 2 Euros for 3 days increased exposure of an item.

To ensure accurate analysis, the following assumptions are made:

- All items in the dataset are unique
- Each item belongs to exactly one sub-category
- Each sub-category name is unique across the dataset

1.2 Granularity of the dataset

Each row represents a unique sub-category (category_3) with aggregated metrics across all items within that sub-category. The dataset provides category level summaries rather than individual transaction or item details.

Missing Dimensions:

- Individual seller information
- Individual item details
- Temporal data (dates, time periods)
- Geographic coverage
- User demographics (gender, age, nationality)

This coarse granularity limits analysis to category level patterns and prevents deeper investigation of seller behavior, temporal trends, item level performance, or demographic segmentation.

1.3 Column descriptions

Category_2: The column represents the name of the main category that items fall under, where different related sub-categories are grouped into one category. Many sub-categories can have one main category, so this forms a hierarchical relationship in the table. Data type is a string.

Category_3: The column represents the name of the sub-category which belongs to the main category. Every sub-category must belong to exactly one main category. Data type is a string.

Number_of_listings: The column shows the number of total items listed on the platform per sub-category. It represents an aggregated sum of items. Data type is an integer.

Avg_listing_price_eur: The column represents the average price of a listed item per sub-category. It is an aggregated average price of items in Euros. Data type is float.

Revenue_from_push_ups: The column represents the total revenue of the pricing push-up feature per sub-category. It is an aggregated sum of push-up payments from sellers in Euros. Data type is float.

Sample data:

	category_2	category_3	number_of_listings	avg_listing_price_eur	revenue_from_push_ups
1	BOOKS_AND_SCHOOL	LITERATURE_FOR_PARENTS	10,813	7.260281	228
2	MEN_ACCESSORIES	SCARVES_SHAWLS	1,380	23.003616	52
3	COSMETICS_AND_BEAUTY_PRODUCTS	WOM_COS_BOD_COSMETICS	19,031	10.380822	906
4	MEN_FOOTWEAR	MEN_FOO_TRAINERS	73,941	149.24828	9,700
5	MEN_CLOTHING	MEN_TROUSERS	37,116	28.793058	1,860
6	TOYS_AND_GAMES	DECORATION_PARTIES	787	8.776811	28
7	BAGS_BACKPACKS	TOTE_BAGS	8,233	12.31953	440
8	ACCESSORIES_JEWELLERY	SOCKS_TIGHTS_STOCKINGS	8,741	6.5553517	270
9	CLOTHING_FOR_BOYS	TWINS_BOYS	3,816	9.121208	138

1.4 Dataset quality and exploratory analysis

1.4.1 Dataset quality

To start the analysis, the data must be examined, and its quality must be tested to see if it's fit to use and meets the standards:

For testing the quality Python Pandas library for data frames is used with a combination of SQL validation checks to verify the data integrity. For visualization Matplotlib and Seaborn libraries are used.

Data completeness check:

a) Data set consists of 5 columns and 238 entries (rows)

```
RangeIndex: 238 entries, 0 to 237
Data columns (total 5 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   category_2             237 non-null    object
 1   category_3             230 non-null    object
 2   number_of_listings     238 non-null    int64
 3   avg_listing_price_eur  238 non-null    float64
 4   revenue_from_push_ups  223 non-null    float64
dtypes: float64(2), int64(1), object(2)
memory usage: 9.4+ KB
```

b) Check of continuous columns. Table shows count of records, percentage of missing values, min. value, Q1, mean, median, Q3. max. value and standard deviation.

	Feature	Count	% Miss	Card.	Min	Q1	Mean	Median	Q3	Max	Std. Dev.
0	number_of_listings	238	0.00%	232	1.0	1669.25	58586.80	6328.50	38330.75	1844512.00	156578.00
1	avg_listing_price_eur	238	0.00%	236	3.5	11.84	34.10	21.29	44.38	329.73	40.74
2	revenue_from_push_ups	223	6.30%	187	2.0	125.00	4008.71	498.00	3215.00	79424.00	10131.36

c) Check of categorical columns. Table shows count of records, percentage of missing values, cardinality(uniqueness), most frequent title and its frequency, second most frequent title and its frequency.

	Feature	Count	% Miss	Card.	Mode	Mode Freq	Mode %	2nd Mode	2nd Mode Freq	2nd Mode %
0	category_2	237	0.42%	35	GIRLS_CLOTHING	18	7.59%	WOMENS	17	7.17%
1	category_3	230	3.36%	228	MEN_FOO_TRAINERS	2	0.87%	MEN_FOO_TRAINERS	2	0.87%

d) Check for duplicates using .duplicated() method by checking all columns. 2 rows were identified as complete duplicates rather than legitimate separate records, matching across all columns, so the dataset will be cleaned, and the duplicate entries will be removed.

	category_2	category_3	number_of_listings	avg_listing_price_eur	revenue_from_push_ups
155	MEN_FOOTWEAR	MEN_FOO_TRAINERS	73941	149.248275	9700.0
234	WOMENS	SKIRTS	214910	13.239253	11830.0

e) Handling of NULL Values after removing duplicates. Tables above show they exist, and these values will be handled as follows:

	Count	Miss %
category_2	1	0.42%
category_3	8	3.39%
number_of_listings	0	0.00%
avg_listing_price_eur	0	0.00%
revenue_from_push_ups	15	6.36%

Revenue_from_push_ups NULL values will be replaced with 0, representing no revenue. Using the mean value would be inappropriate as it would artificially inflate revenue metrics and misrepresent actual push-up feature usage.

```
Unknown categories revenue: 6214.0
Unknown categories listings: 115385
Percentage of total revenue: 0.71%
Percentage of total listings: 0.85%
```

Rows with NULL values in *category_2* and *category_3* will be deleted from the dataset. Although these rows contain data that could be relevant for analysis, their impact on the dataset is minimal (0.71% of total revenue and 0.85% of total listings). The analysis is based on category-level data, therefore unknown categories would not provide sufficient analytical value. The row with a NULL value in category_2 also has NULL value on category_3 and the remaining 7 rows have NULL values only in category_3. Additionally, since category_3 serves as the unique identifier for each row, labeling these as 'unknown' would violate the uniqueness constraint and compromise data integrity.

f) Check for negative values in numeric columns. No negative values were found, so no further handling is needed.

	Column	Negative
0	number_of_listings	False
1	avg_listing_price_eur	False
2	revenue_from_push_ups	False

g) The dataset represents data without time series information. Checking for date values is not possible.

Data Quality Summary

- Original dataset: 238 rows, 5 columns, 35 unique main categories (category_2)
- Rows removed: 10 (4.20%)
- Final dataset: 228 rows, 5 columns, 28 unique main categories (category_2)
- Data retention: 99.29% of revenue, 99.15% of listings
    - Impact: The 7 removed main categories represented <1% of total data
- Quality issues found and resolved:
    - Duplicates: 2 found and removed
    - NULL revenue values: 15 replaced with 0
    - Missing category identifiers: 8 rows removed
    - Negative values: None found

1.4.2 Exploratory analysis

After data inconsistencies are handled, exploratory analysis can proceed. As shown in the chart below, the 28 unique main categories are distributed differently and vary in the number of subcategories they have.

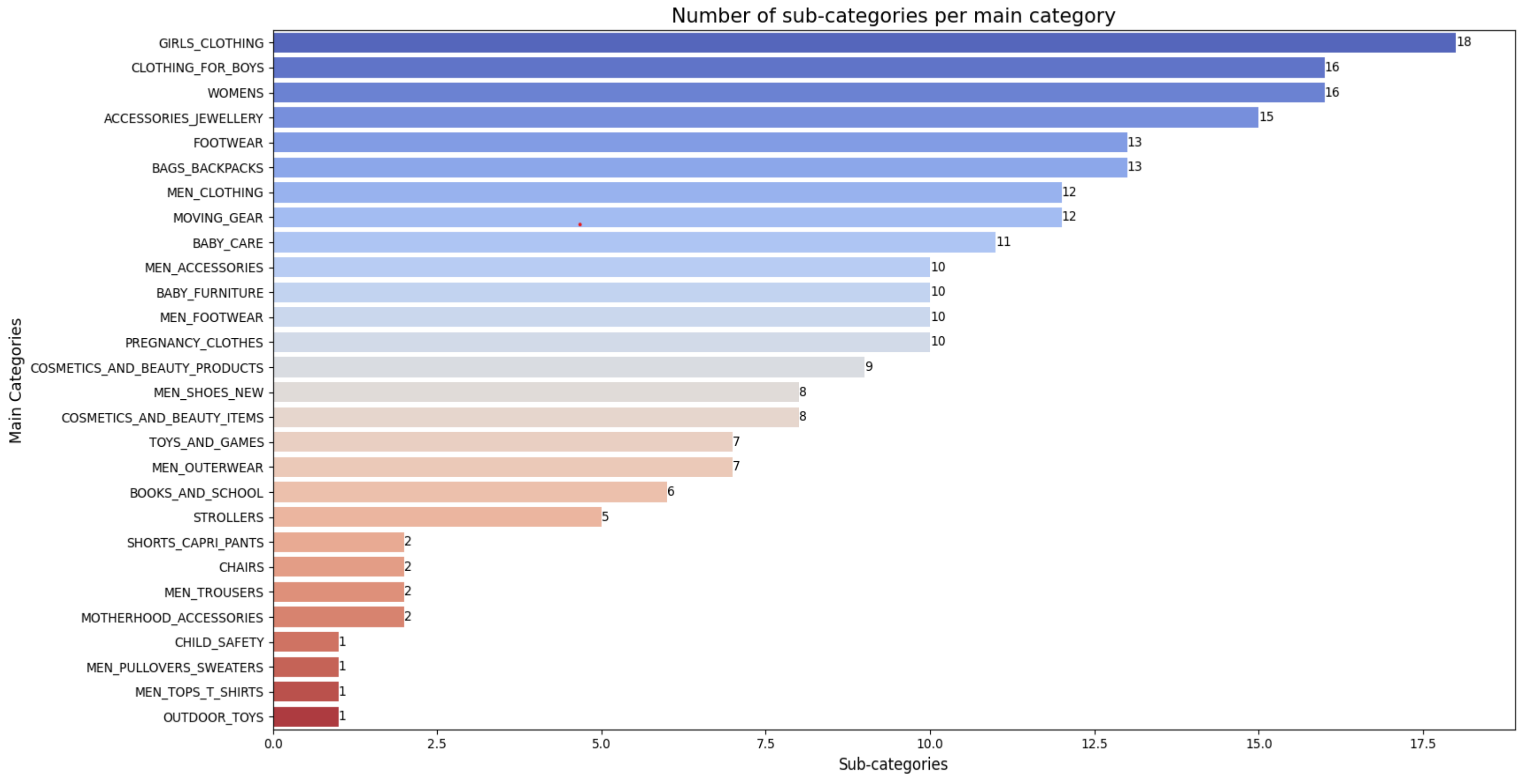
High diversity categories:

- GIRLS_CLOTHING leads with 18 sub-categories
- CLOTHING_FOR_BOYS and WOMENS each have 16 sub-categories
- ACCESSORIES_JEWELLERY contains 15 sub-categories

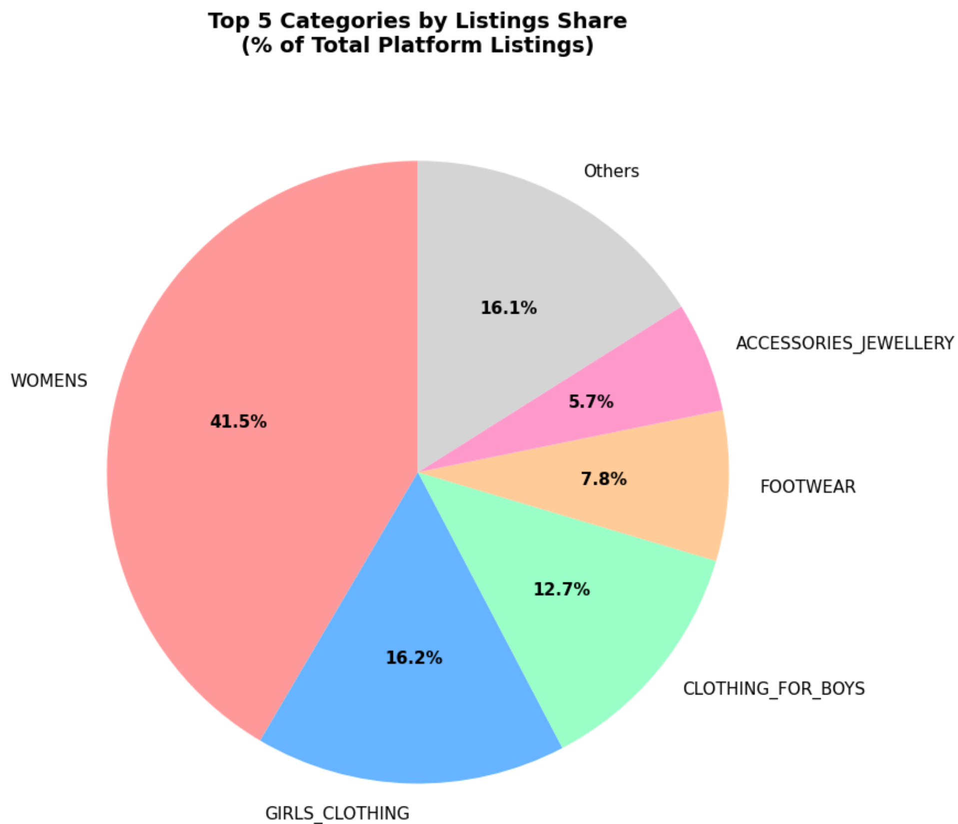
Low diversity categories:

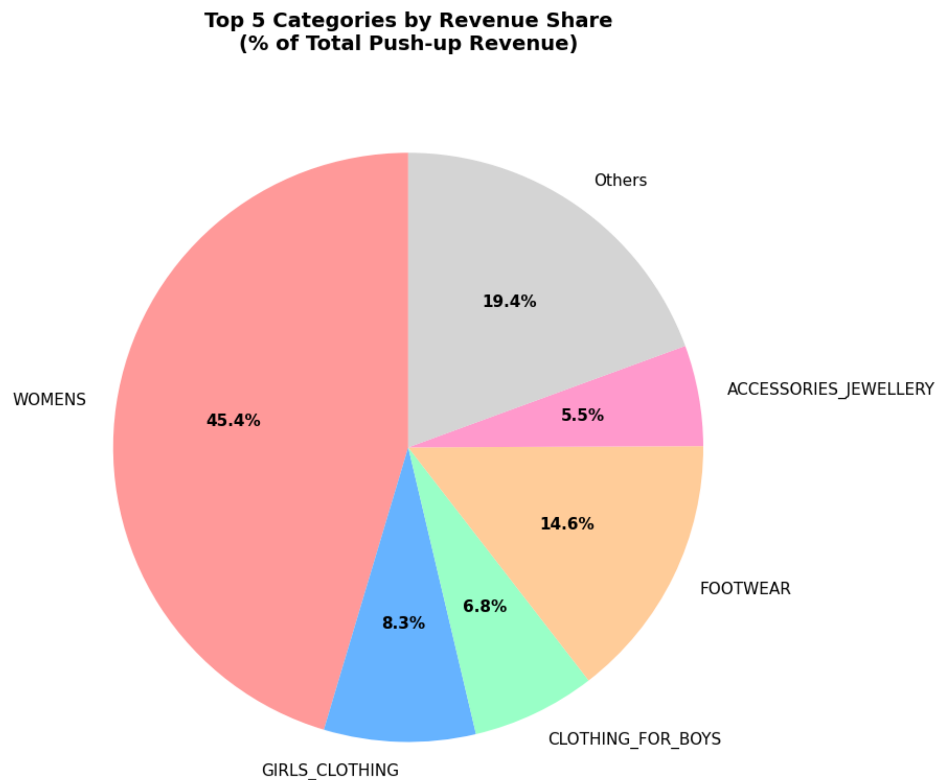
- OUTDOOR_TOYS, MEN_TOPS_T_SHIRTS, MEN_PULLOVERS_SWEATERS, CHILD_SAFETY: 1 sub-category each
- CHAIRS, MEN_TROUSERS, MOTHERHOOD_ACCESSORIES, SHORTS_CAPRI_PANTS: 2 sub-categories each

This distribution reveals a clear pattern: Women's and children's categories average 15 sub-categories compared to only 4 for men's categories, demonstrating the platform's primary focus on women and children's products. This nearly 4 times difference in category diversity suggests significantly greater product variety and market attention in women's and children's segments.



To examine whether category dominance correlates with product diversity, the charts below show the top 5 categories by share of listings and revenue.





The data reveals significant market concentration: the top 5 categories represent ~ 84% of all listings and ~81% of push-up revenue. This pattern confirms a strong correlation between product diversity (as shown previously) and market dominance.

Notably, performance varies within the top 5. While GIRLS_CLOTHING accounts for 16.2% of listings, it generates only 8.3% of revenue, suggesting lower push-up adoption in this category. Conversely, FOOTWEAR (7.8% of listings) generates a disproportionate 14.6% of revenue, indicating higher push-up effectiveness.

Findings show that women's and children's categories dominate both product diversity and platform activity, particularly in push-up feature usage. This dominance is statistically confirmed by strong right sided skewness, indicating that only several categories drive most outcomes.

```
---------------------------
Revenue skewness: 5.06
Listings skewness: 7.28
Price skewness: 3.97
---------------------------
```

- Revenue concentration: A small number of categories generate most of the earnings, highlighting dependence on top performers.
- Listings imbalance: Most categories have fewer listings, while a handful dominate with high numbers of listings.
- Price disparity: Most items are priced in a lower range, but a few categories with rather expensive items stretch the distribution.

1.5 Analysis Limitations

The dataset limits the depth of analysis possible, and the important factors below will not be included in the actual analysis:

1. What are the habits of the sellers using push-ups:

    a) Do they have more active listings, or is it diversified?
    b) Retention rate: do those using push-ups continue to do so when successfully selling an item, or not?
    c) Drop off rate: Do they stop using when sale is unsuccessful and vice versa?
    d) The number of unique sellers using push-ups versus repetitive usage of same sellers from total users.
    e) Sellers adoption rate: percentage of sellers using push-ups versus total sellers.
    f) Time to adoption: from the moment the seller signs up on the platform (or the feature is introduced) until sellers make their first use of the feature.

2. Ratio of successful sales after using push-ups.

3. Feature reach ratio: number of reached audience members when using the feature versus the reach without the feature.

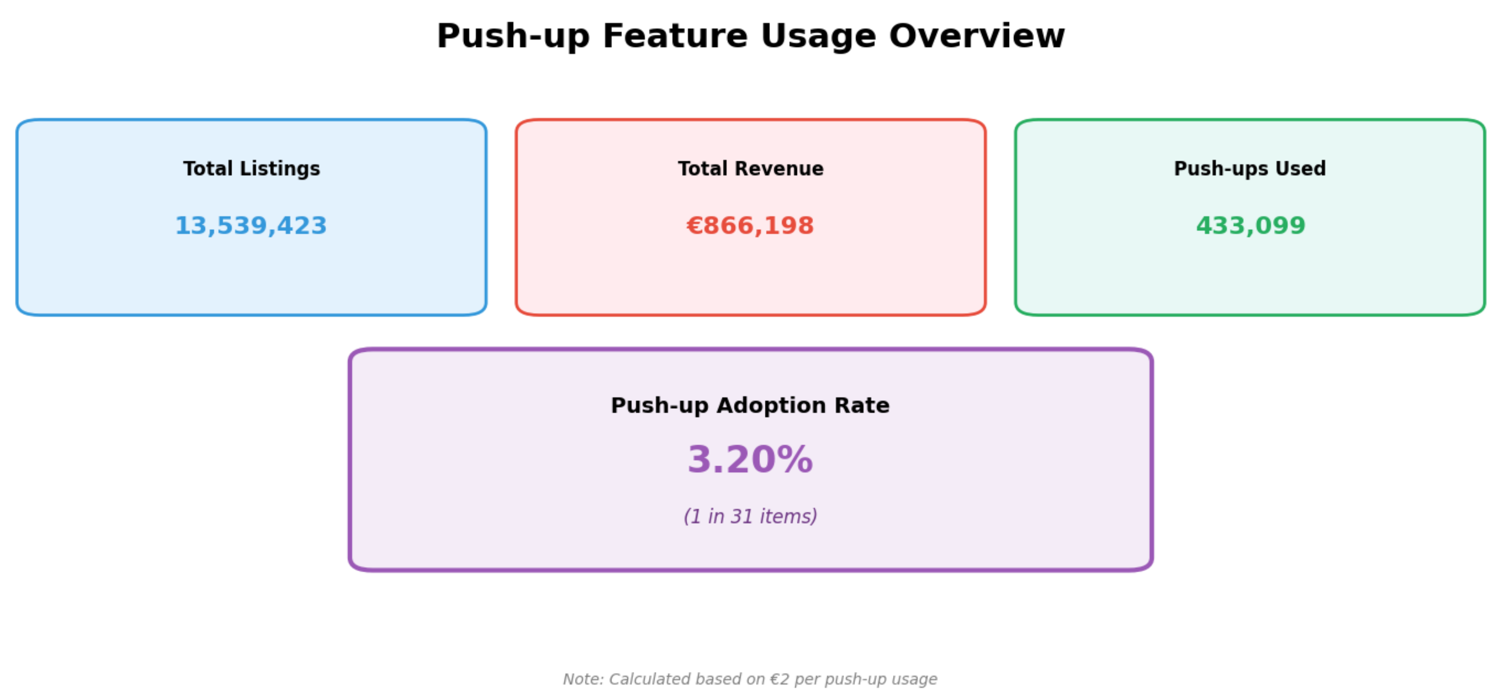
Considering the limitations of the dataset and these constraints, the analysis will focus on available category-level metrics to identify relative performance patterns and revenue distribution across product categories for the pricing push-up feature. While temporal trends and seller level insights are not possible, the dataset enables meaningful cross category comparisons and identification of high performing product segments based on listing volumes, average pricing, and revenue generation.

2. Analysis of push-up feature usage

To assess sellers interest in the push-up feature, a usage ratio metric should be defined. This metric measures the proportion of push-up feature usage relative to the total dataset and can be further applied at the category level to evaluate performance, highlighting which categories show the highest and lowest adoption.

The calculation is performed by dividing total revenue from push-ups by the push-up price (€2), yielding the number of times the feature was used. This value is then divided by the total number of listings, producing the usage ratio that indicates how frequently the feature is applied relative to overall listings.

This is the main metric that can represent the overall trend of the push-up feature and its frequency from the data available.
The parameters required for calculations are represented in the table below and the result is presented.



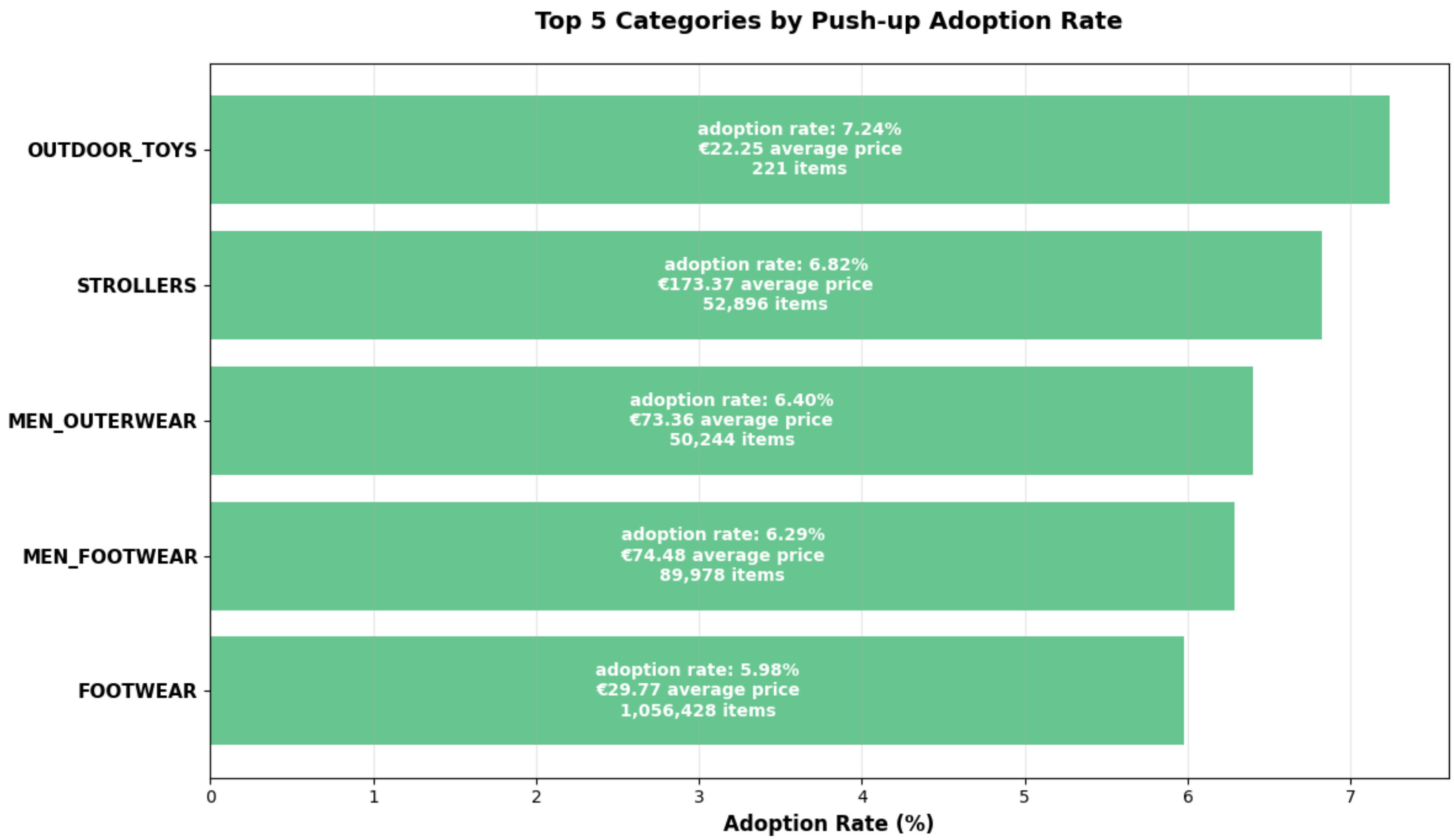
Summary of the results:

- Total listings: 13,539,423
- Total revenue: 866,198 Euros
- Push-ups used: 433,099 times
- Push-up Adoption Rate: 3.20%.

Based on these metrics, the overall adoption rate is relatively low. Out of 100 items listed, approximately only 3 use the push-up feature, meaning around 97% of total listings never utilize the feature. This low adoption may indicate that the perceived value of the feature is insufficient, awareness among sellers is underdeveloped, or the price (€2 per use) may be too high relative to lower priced items.

Despite the low adoption rate, engaged users find value in the feature, generating €866,198 in total revenue. This presents a significant growth opportunity through strategies such as targeted seller education, price optimization testing, and category specific promotion campaigns focused on segments that demonstrate higher engagement.

Based on this metric, the 5 top performing main categories are indicated in the chart below:

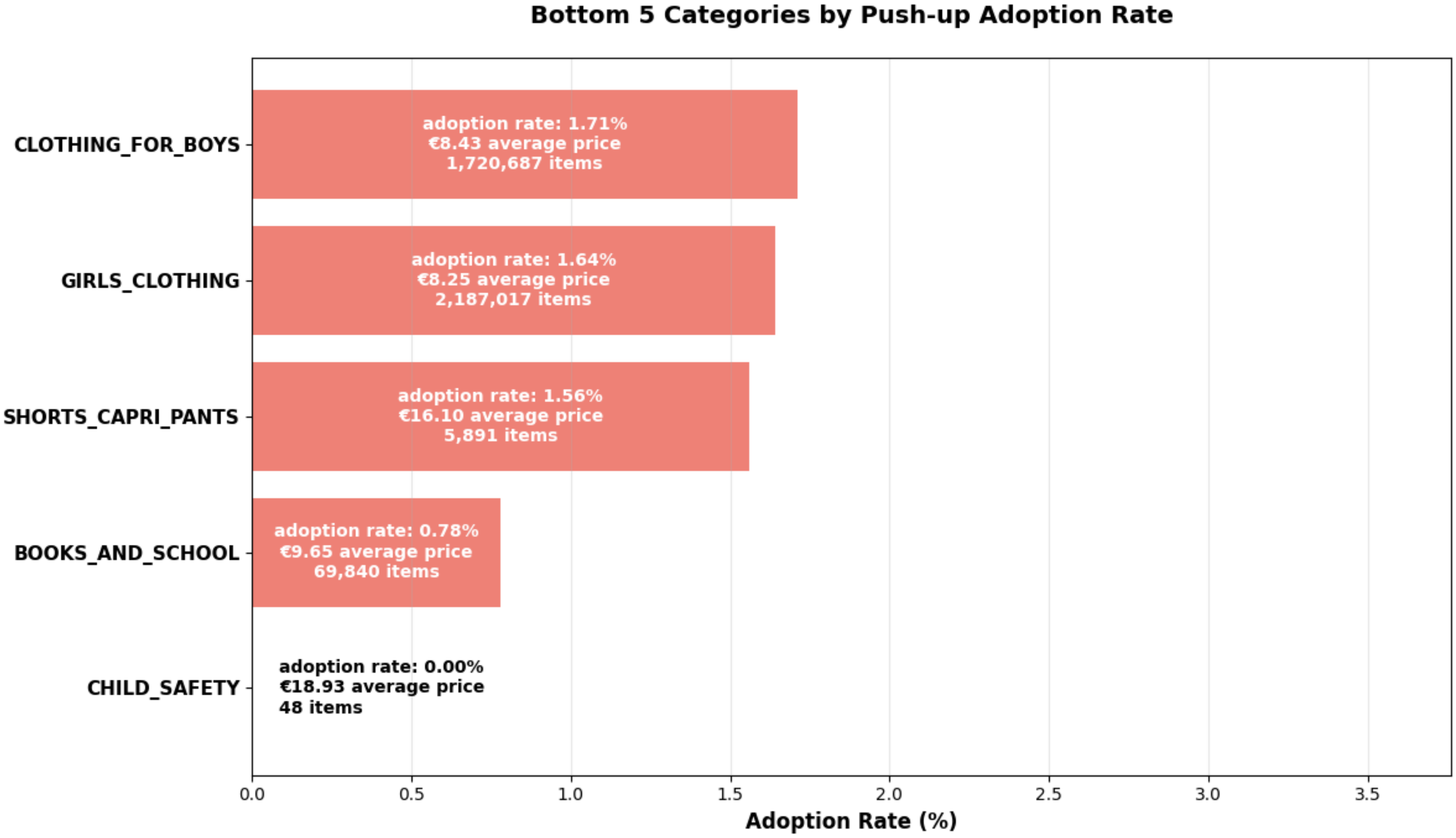


OUTDOOR_TOYS: While appearing as the top performer by adoption rate, this category has an extremely low number of listings (221), suggesting this may be an anomaly rather than a true top performer. The small sample size limits the reliability of this metric for this category.

STROLLERS, MEN_OUTERWEAR, and MEN_FOOTWEAR: share similar characteristics: moderate listing volumes, higher average item prices (€70-€180), and adoption rates ranging between 6-7%. These categories demonstrate consistent, reliable engagement with the push-up feature across meaningful sample sizes.
FOOTWEAR: As indicated in the exploratory analysis, this category showed early signs of overperformance, which this metric confirms. With an average price of ~€30 and over 1 million total listings, it achieves the 5th highest adoption rate among main categories, demonstrating strong push-up engagement at scale.

The most meaningful top performers are FOOTWEAR (combining scale with adoption) and the mid-tier categories (STROLLERS, MEN_OUTERWEAR, MEN_FOOTWEAR), which balance higher price points with strong adoption rates. Notably, WOMENS, the largest category by volume (41% of total listings), has a 3.50% adoption rate, slightly above the overall platform rate (3.20%), presenting the greatest opportunity for growth through targeted strategies.

On the other hand, the bottom 5 performers are indicated in the chart below:



CHILD_SAFETY: While appearing as the worst performer with a 0.00% adoption rate, this category cannot be considered suitable for analysis as the number of listings (48) is statistically irrelevant for the dataset.

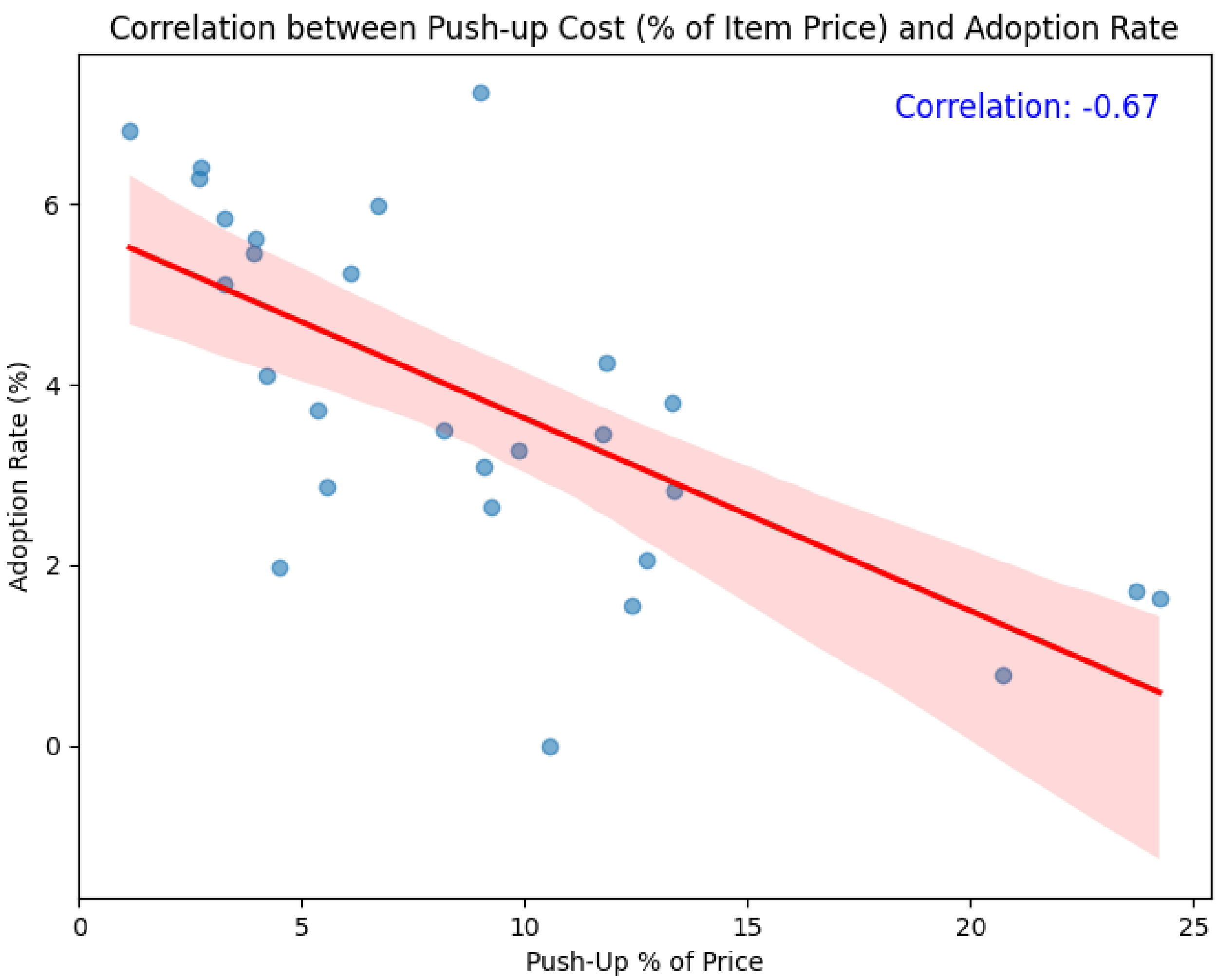
BOOKS_AND_SCHOOL has a meaningful sample size, yet its adoption rate is approximately 4 times lower than the overall platform rate (3.20%), making it a significant underperformer. The lower average price (~€10) may contribute to reduced push-up usage, as the €2 feature cost represents a higher percentage of item value.

SHORTS_CAPRI_PANTS: With a relatively small sample and an adoption rate 2 times below the platform average, this category offers limited analytical value. The combination of low average price (~€16) and small volume do not provide meaningful insights into strategic decisions.

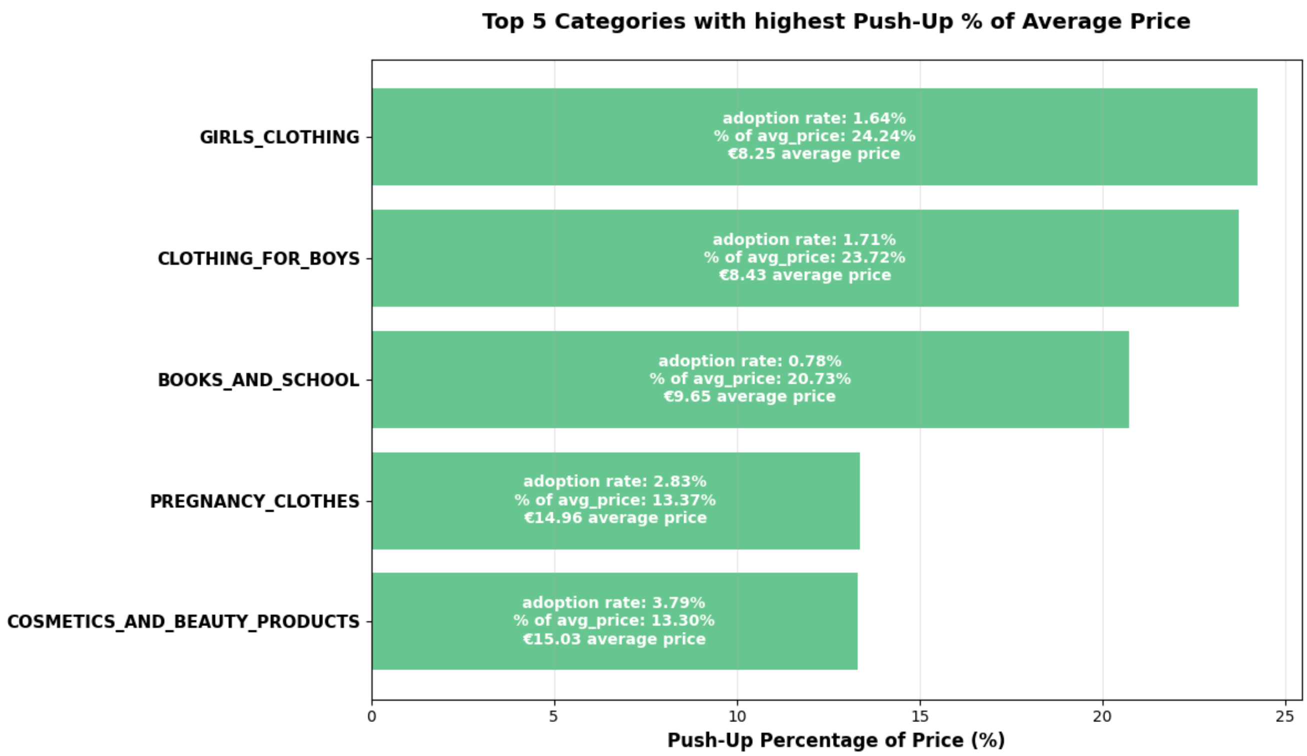
GIRLS_CLOTHING and CLOTHING_FOR_BOYS: These two categories are the most concerning underperformers identified in this analysis. Despite significant listing volumes (16.2% and 12.7% of total listings respectively), both demonstrate adoption rates well below the platform average. With similar average prices (~€8-9), the €2 push-up cost represents higher part of item value, likely creating a significant barrier to adoption. These categories belong to the women's and children's segments that dominate the platform (identified in exploratory analysis), yet they severely underutilize the push-up feature. This represents both a major performance gap and a substantial growth opportunity requiring deeper investigation into seller behavior, price sensitivity, and feature awareness within these segments.

Analysis of top and bottom performers reveals that higher-priced items demonstrate greater push-up adoption compared to lower priced items. Large categories such as CLOTHING_FOR_BOYS and GIRLS_CLOTHING, which have lower average prices, fall among the bottom performers. To validate whether lower average prices correspond to lower adoption rates, introducing a secondary metric: push-up cost as percentage of average item price (€2 ÷ average price × 100). This metric quantifies the relative cost burden sellers face and helps confirm the relationship between pricing levels and feature adoption.

The graph below shows that these two metrics exhibit a moderate to strong negative correlation of -0.67, which indicates that as the push-up percentage of price increases, adoption rates tend to decrease. This reflects price sensitivity, as sellers are less likely to adopt the feature when their relative cost rises.

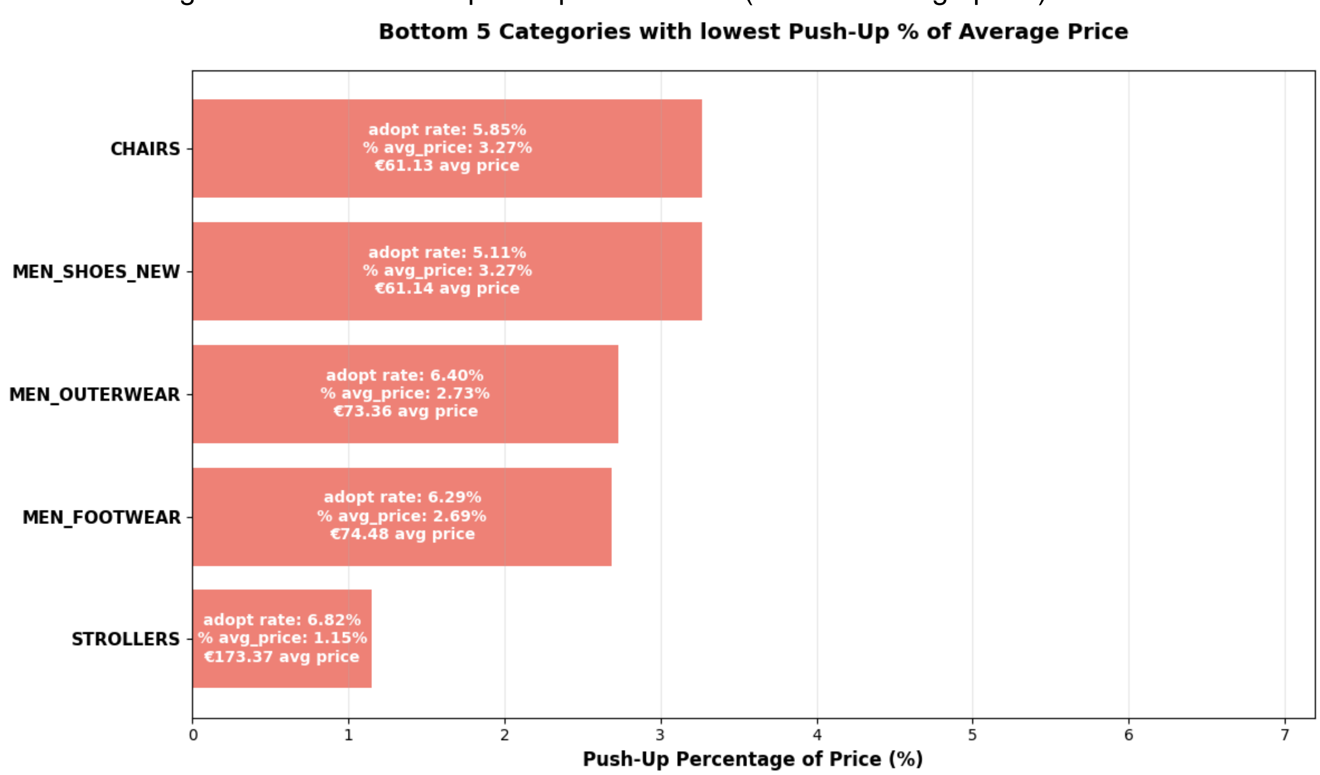


To further illustrate this relationship, the chart below highlights the top five categories with the highest push-up cost burden (as % of item price):



The results align with earlier findings: the same three categories (GIRLS_CLOTHING, CLOTHING_FOR_BOYS, and BOOKS_AND_SCHOOL) previously identified as the weakest performers also face the highest relative cost burden. A clear threshold emerges, when the push-up cost exceeds 20% of average item price, adoption rates drop significantly, and these categories consistently underperform.

The bottom 5 categories with the lowest push-up cost burden (as % of average price) are shown below:



These results also validate the earlier adoption rate analysis. Notably, the 3 bottom categories in this chart (STROLLERS, MEN_FOOTWEAR, MEN_OUTERWEAR), which show the lowest push-up cost burden, are the same categories identified as top performers in adoption rate. This perfect alignment strongly supports the correlation between relative cost and adoption behavior.

This pattern demonstrates that the flat €2 pricing model disproportionately impacts lower priced categories, creating a significant barrier to adoption in high volume segments like children's clothing. Categories most affected include GIRLS_CLOTHING (~25% of price), CLOTHING_FOR_BOYS (~24% of price), and BOOKS_AND_SCHOOL (~21% of price), all showing adoption rates below 2%.

While some individual sub-categories show push-up costs exceeding 50% of average price, these represent statistical anomalies rather than meaningful patterns. For example, the GIRLS_CLOTHING / FOR_BABIES sub-category has an average listing price of €3.72 and generated €274 in push-up revenue (137 uses) yet it demonstrates an extremely low adoption rate of 0.35%. This apparent contradiction likely stems from:

1. A small number of higher-priced items within the sub-category received push-ups, skewing the revenue data while most low-priced items did not use the feature.
2. Items listed for extended periods may have received push-ups as a final attempt to generate sales, rather than as part of regular selling strategy.
3. With only 137 uses across ~39,000 listings, individual seller decisions create disproportionate impact on metrics.

These edge cases reinforce rather than contradict the overall finding: when push-up costs represent an excessively high percentage of item value, adoption remains minimal regardless of isolated usage instances. The category_2 level analysis remains the most reliable for identifying actionable patterns.

3. Price change implications

Current analysis shows strong negative correlation (r = -0.67) between push-up cost as % of item price and adoption rate, providing clear predictions for price changes.

Price Decrease (e.g., €2 to €1):

Positive Impacts:

- Adoption increases significantly, especially in low price categories
- GIRLS_CLOTHING: burden drops from 25% to 12.5%, adoption could double (1.64% to 3-4%)
- Platform-wide adoption: 3.20% to 5-6% (+60-90%)
-

Negative Impacts:

- Revenue per use cuts in half
- Generating an additional 433K uses would only yield €433K instead of €866K (immediate -50% loss on existing volume)
- Visibility competition increases (more sellers using feature = less individual impact)

In conclusion, volume increase may not offset revenue loss. Additionally, there is a risk of devaluing the feature and losing seller trust if increased usage does not translate to sales, leaving limited room for future price increases.

Price Increase (e.g., €2 to €3):

Positive Impacts:

- Revenue per use increases 50%
- Premium positioning, better value perception

Negative Impacts:

- Adoption collapses in price sensitive categories (GIRLS_CLOTHING burden: 25% to 37.5%)
- Platform-wide adoption: 3.20% to 1.5-2.0% (-40-55%)
- Total uses drop: 433K to 200K-250K

In conclusion, volume loss exceeds per use gain. Significant risk of alienating existing users, particularly in the platform's dominant women's and children's segments.

To measure the success of a newly implemented pricing strategy (e.g., decrease in price of the feature), an A/B test with statistical validation would be most appropriate.

Test Design:

1. Sample Selection: Randomly assign sellers into two groups:
   - Control Group: Current €2 pricing
   - Test Group: New pricing (e.g., €1 for items <€10)

Focus the test on price sensitive categories such as GIRLS_CLOTHING and CLOTHING_FOR_BOYS, where the current cost burden is highest (20-25% of item value) and potential for improvement is greatest.

2. Test Duration: Run the test for 4-6 weeks. This period is sufficient to capture seller behavior patterns while allowing for timely decision making. Longer periods risk competitive disadvantages and delay optimization.

3. Success Metrics: Define clear targets before the test:
   - Primary: Adoption rate increases by at least 25% (e.g., from 1.64% to 2.05% in GIRLS_CLOTHING)
   - Secondary: Total revenue remains stable or increases (minimum -10%)

4. Statistical Method: Two-proportion z-test:
   - Compares adoption rates between control and test groups
   - Determines if the difference is statistically significant ($p < 0.05$) or due to random chance

5. Analysis Process:
   - After 4-6 weeks, compare test group vs control group:
     - Adoption rate (primary metric)
     - Total revenue generated
     - Number of push-up uses
     - Revenue per seller

- Conduct statistical test to confirm significance
- Perform segment analysis to identify differential impacts

6. Decision Criteria:

If test group shows statistically significant improvement ($p < 0.05$) AND meets revenue targets, implement broadly. If test group underperforms or results are not significant, reject change or modify approach.

This approach ensures data-driven decision making while minimizing risk through controlled experimentation.

4. Conclusions and alternative strategies

1. **Low platform-wide adoption (3.20%):** Only 1 in 31 listings uses push-ups, indicating significant untapped potential across the platform.
2. **Price sensitivity barrier**: Strong negative correlation (r = -0.67) between push-up cost as percentage of item price and adoption rate demonstrates that relative cost is the primary adoption barrier.
3. **Critical threshold identified**: When push-up costs exceed 20% of item value, adoption rates drop below 3%, creating a clear price sensitivity threshold.
4. **High volume underperformers**: GIRLS_CLOTHING and CLOTHING_FOR_BOYS (28.9% of total listings combined) demonstrate adoption rates 40-50% below platform average, representing the largest growth opportunity.
5. **Market concentration**: Top 5 categories account for 84% of listings and 81% of revenue, with women's and children's segments dominating platform activity.

The push-up feature shows strong potential but suffers from a pricing model misaligned with platform demographics. The flat €2 pricing creates an insurmountable barrier for the platform's dominant segments (women's and children's low-priced items), while leaving revenue opportunities untapped in premium categories.

Recommendations for alternative Strategies:

While both price increase and decrease present tradeoffs, alternative strategies could better optimize adoption and revenue:

1. Tiered pricing strategy: Implement pricing based on percentage of item value with a maximum cap to protect high value item sellers (e.g., 5-7% of item price, capped at €4). This maintains affordability for low priced categories while preventing excessive costs for premium items. To avoid manipulation, the price could not be increased when the item is “bumped” or sold for a higher price than listed.

2. Bundle Packages: Offer volume discounts for sellers purchasing multiple push-ups simultaneously (e.g., 5 push-ups for €8 instead of €10), encouraging higher usage frequency and increasing lifetime value per seller. These approaches balance accessibility, revenue optimization, and seller retention more effectively than a single fixed price change.

5. Notes

All ideas, analytical insights, metrics, and methodology in this analysis were developed by me. Data processing and calculations were performed using Python's Pandas library, with visualizations created using matplotlib and seaborn. I did use AI for grammar checking and proofreading, rephrasing and structuring written insights for better readability, enhancing visualization design and style (pie charts, adoption rate dashboard).

I have created a repository in my GitHub with full script, if needed it is available at public repository: https://github.com/Ignas-Jan/Pricing_push_ups, after the further selection process I will make the repository private or delete it.