

a data-driven approach to

Product Portfolio Management

in e-commerce

Marc Puls | Code Analytics | 23.05.2025

AGENDA

Introduction | p. 3-5

Data Insights & Key Findings | p. 6 - 9

Product Portfolio Management (PPM) | p. 10 - 16

PPM Applied | p. 17 - 25

Clustering | p. 26 - 30

With More Time | p. 31

Summary & Outlook | p. 32

INTRODUCTION



INTRODUCTION

Superstore, e-commerce business based in US, selling products used in offices

Three main categories: furniture, office supplies and technology

Time period: 2014 - 2017

Dataset:

1. 9.994 order lines, 5.009 orders
2. 1.894 products (sales, profit, discount, units sold, etc.)
3. 793 customers (names, addresses, etc.)

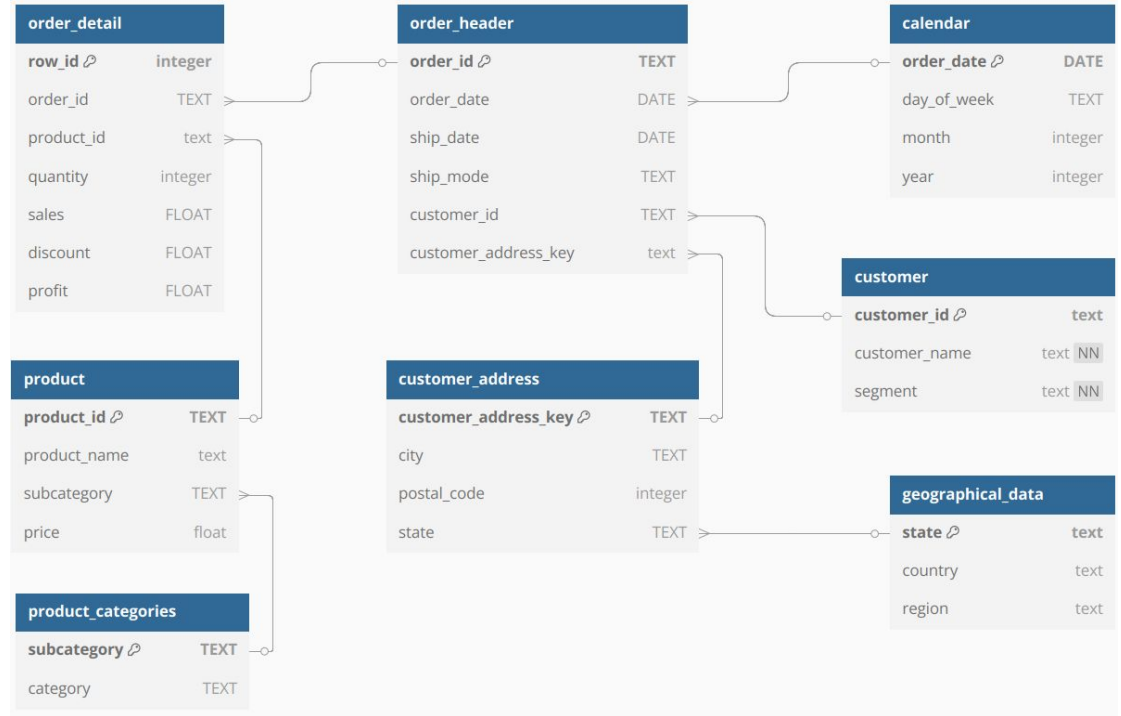
INTRODUCTION

Created:

Entity - Relationship - Diagram

Why?

More efficient
Data Storage



DATA INSIGHTS & KEY FINDINGS

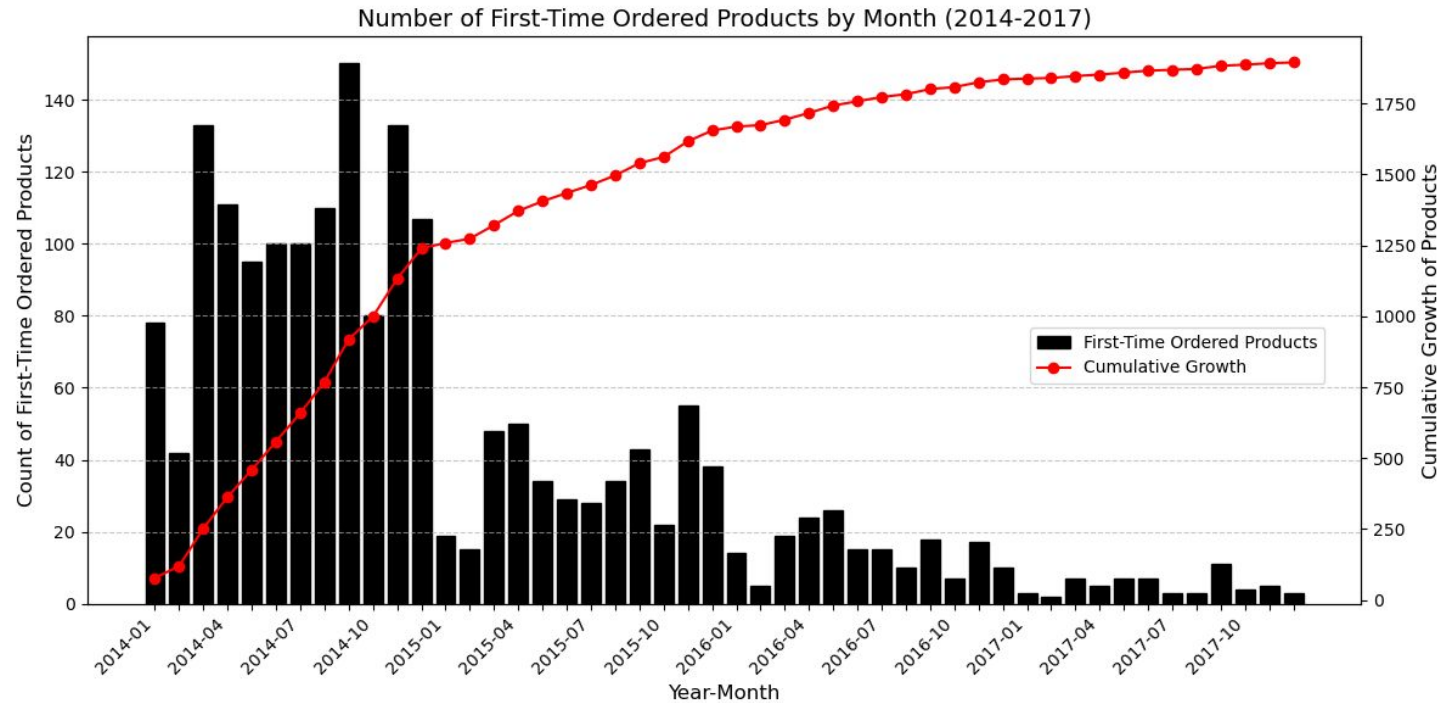
A tree in need of pruning.



FIRST DATA INSIGHTS

- 2.4 times more products than customers!
- 3.42 orders / day (not a lot of traffic)
- each product **only** ordered 5.3 times [=avg, min: 1, max: 15]
- only 20 units sold of every product [=avg, min: 1, max: 67]
- 52% of order lines with discount

KEY FINDINGS & FOCUS SELECTION



KEY FINDINGS & FOCUS SELECTION

- Not a lot of activity for such a wide product portfolio
- Aged product portfolio (products launched per year and %-share)
1.239 (2014, 65.4%), 415 (2015, 21.9%), 180 (2016, 9.5%), 60 (2017, 3.2%)

Main focus: Product Portfolio Management

Side focus: Impact of Discounts on Profitability (Pricing Strategy)

PRODUCT PORTFOLIO MANAGEMENT (PPM)

Setting the goal.



PRODUCT PORTFOLIO MANAGEMENT (PPM)

Domain knowledge

Definitions

Visualization

Product portfolio management (PPM)

PPM is the strategic process of overseeing and coordinating your company's entire range of products. It involves evaluating and optimizing your product line to ensure each product aligns with your overall business objectives.

By doing so, you can maximize revenue, minimize risks, and ensure long-term growth.

PPM allows you to make informed decisions about which products to develop, enhance, or retire.

This ensures your portfolio remains competitive and in line with market demands.

<https://www.simon-kucher.com/en/consulting/commercial-strategy-pricing-consulting/customer-product-market-strategy/product-portfolio-management>

PRODUCT PORTFOLIO MANAGEMENT (PPM)

Domain knowledge

Definitions

Visualization

Product lifecycle management (PLM)

In industry, PLM is the process of managing the entire lifecycle of a product from its inception through the engineering, design, and manufacture, as well as the service and disposal of manufactured products.[1][2]

PLM integrates people, data, processes, and business systems and provides a product information backbone for companies and their extended enterprises.[3]

https://en.wikipedia.org/wiki/Product_lifecycle

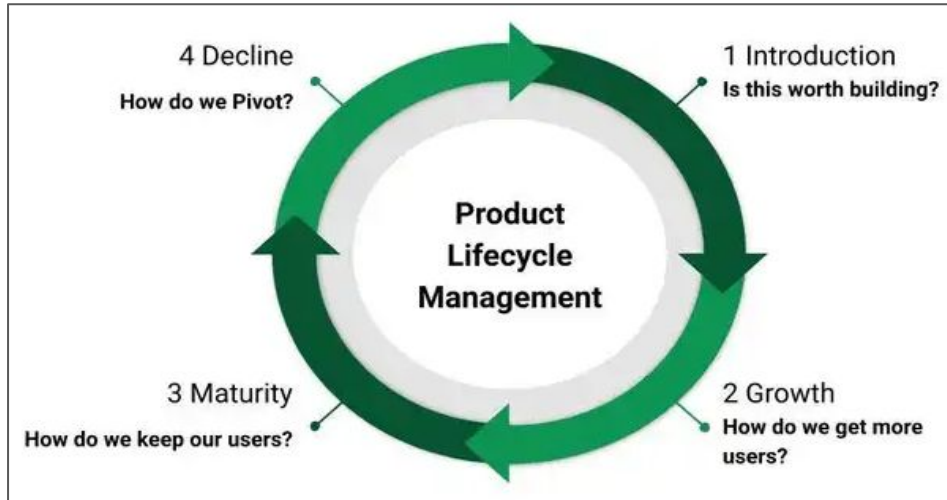
PRODUCT PORTFOLIO MANAGEMENT (PPM)

Domain knowledge

Definitions

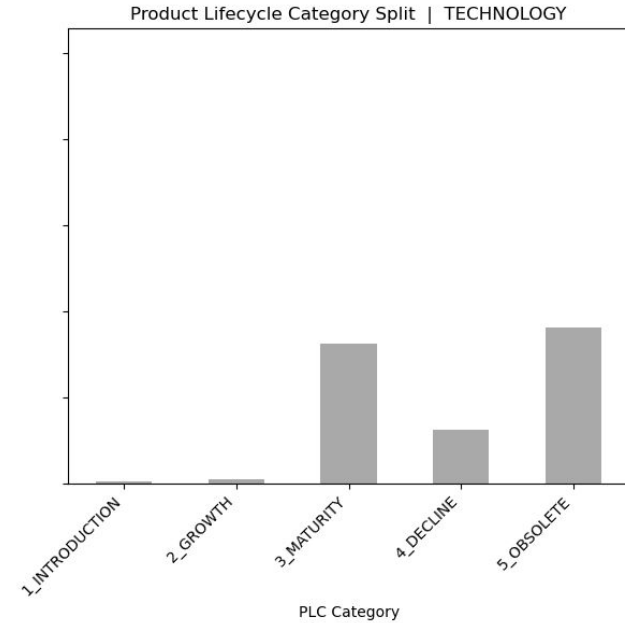
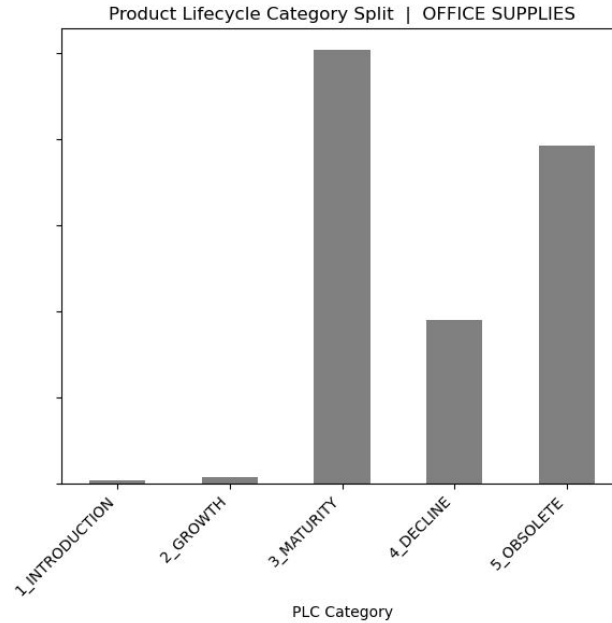
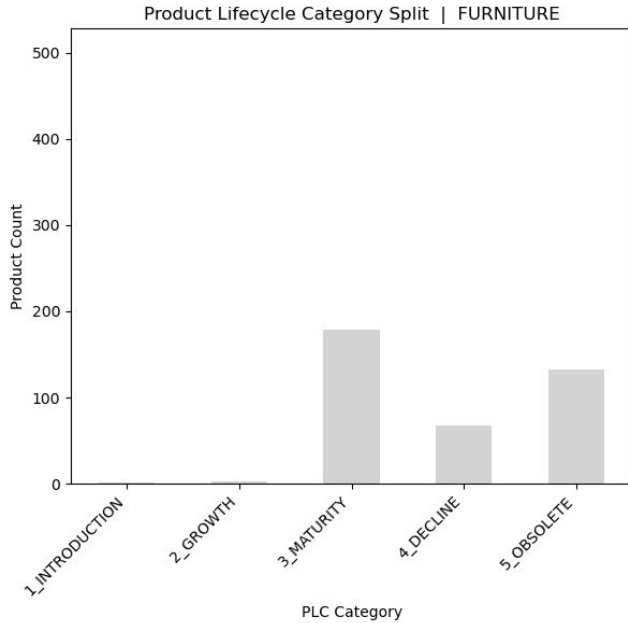
Visualization

Product lifecycle stages

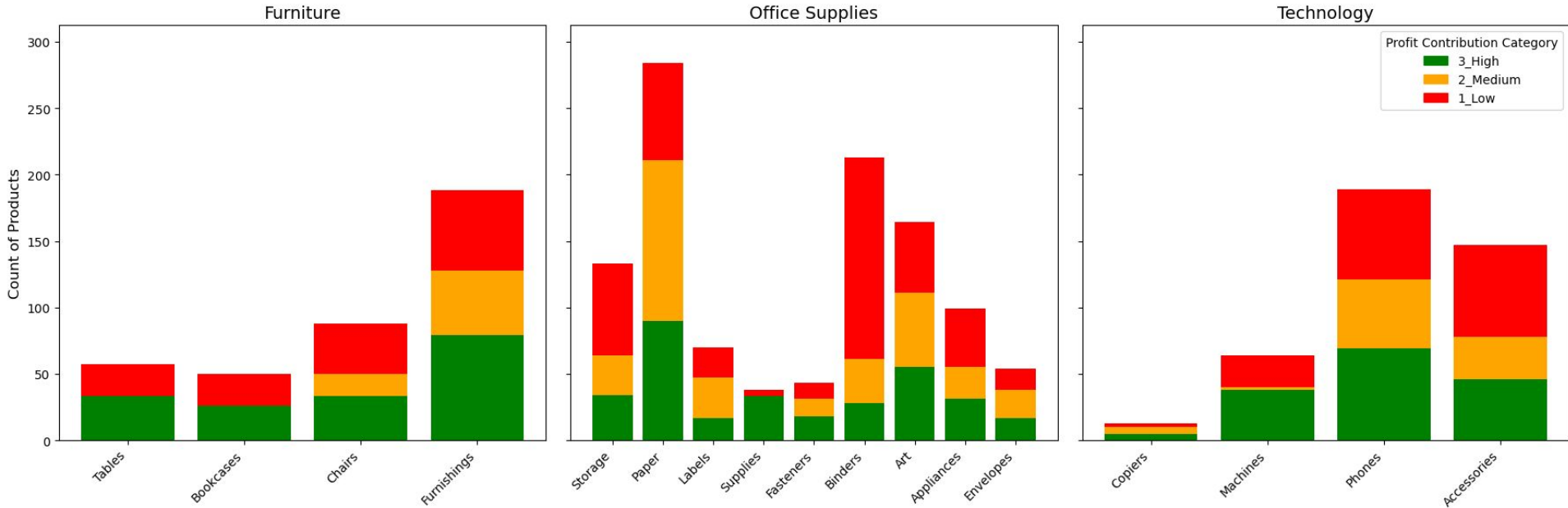


<https://www.dezkr.com/blog/product-lifecycle-management-meaning/>

PPM | PRODUCT LIFECYCLE



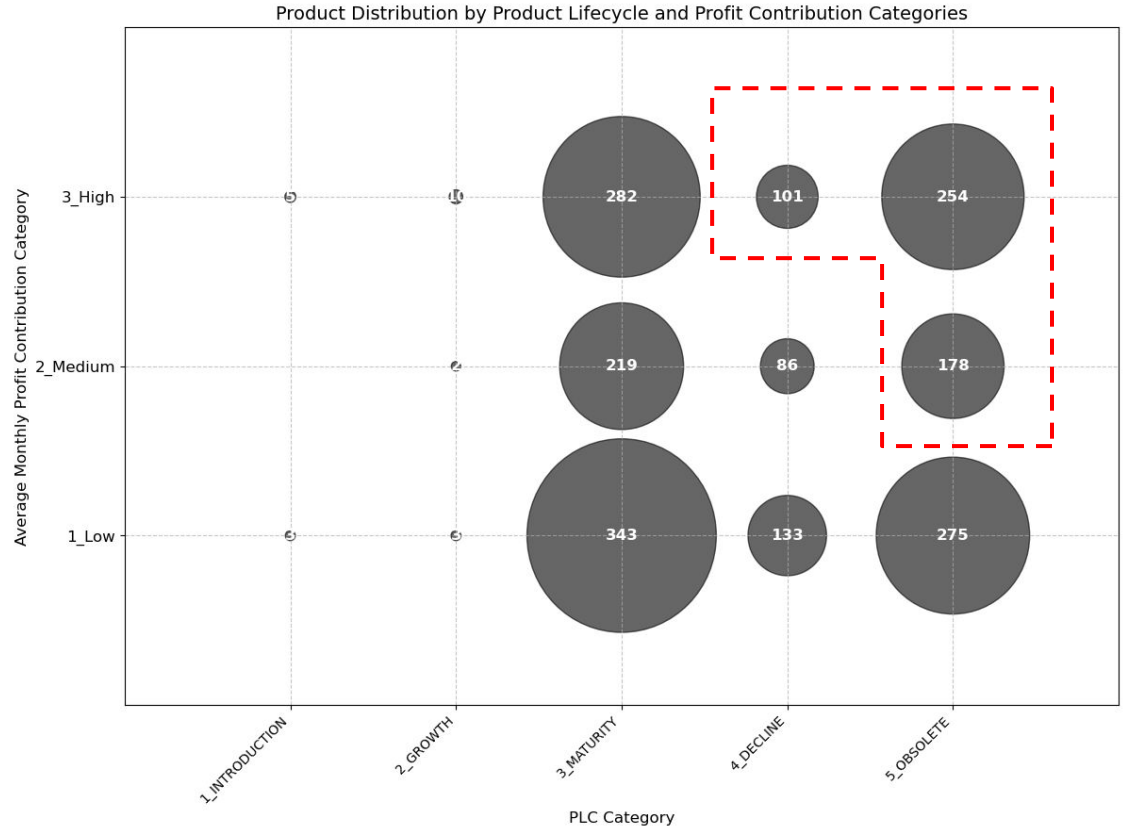
PPM | AVERAGE PROFIT CONTRIBUTION



PPM | COMBINED

Focus:

1. High-profit | Obsolete
2. High-profit | Decline
3. Medium-profit | Obsolete



PPM APPLIED

Low hanging blossoms, first



PPM APPLIED | ONE-TIME CLEAN-UP

Before Product Lifecycle Categorization, general clean-up required:

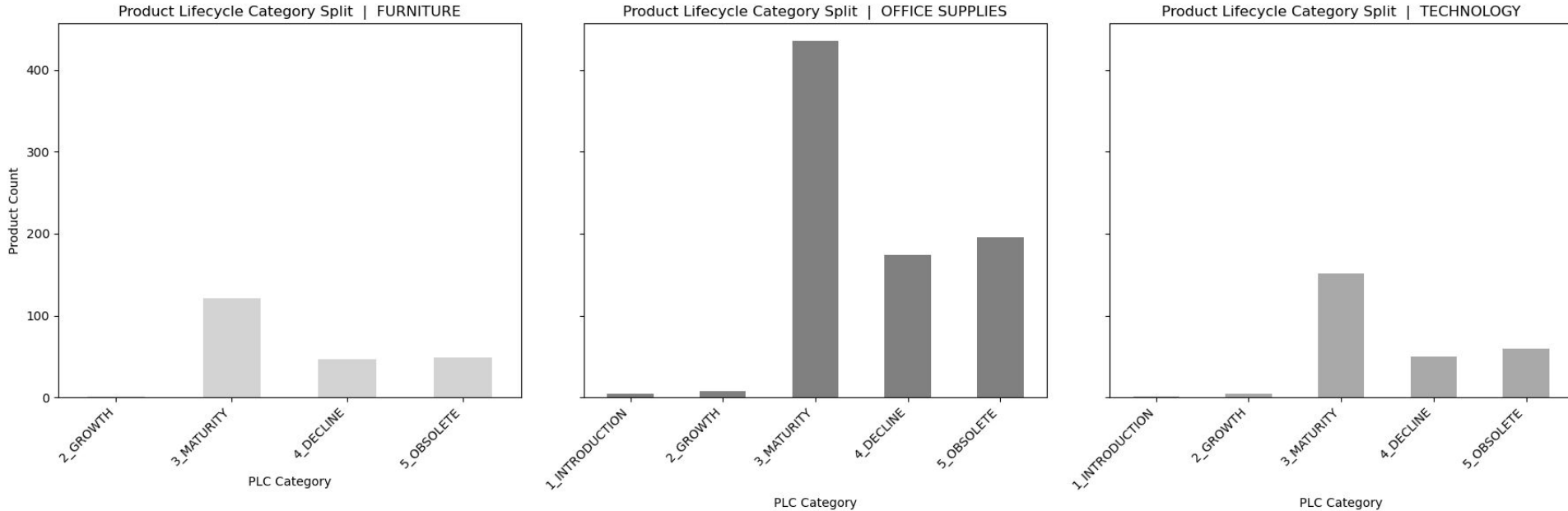
1. Cancel products with no sale in 2017

⇒ - 347 products | -18%

2. Cancel products with negative total profit contribution

⇒ - 247 products | ~-13%

PPM APPLIED | ONE-TIME-EFFORT



⇒ Portion of 4_Decline, 5_Obsolete significantly reduced

PPM APPLIED | ONE-TIME-EFFORT

- Focus on: **Binders** subcategory
- Observation: Multiple brands for same or similar products
- Use multiple selection criteria to cancel entire brands and/or products
- Reason: Reduce costs associated with maintaining wider portfolio

PPM APPLIED | ONE-TIME-EFFORT

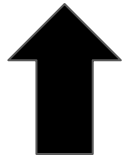
Subcategory:

Binders

Metric:

Product Count

Brand	Keep	Remove	Total Count	Count Change
n.a.	38	22	60	-36.67%
Acco	1	20	21	-95.24%
Avery	0	33	33	-100.00%
Cardinal	0	10	10	-100.00%
Fellowes	3	4	7	-57.14%
GBC	26	20	46	-43.48%
Ibico	7	5	12	-41.67%
Wilson Jones	17	7	24	-29.17%
Total Count	92	121	213	-56.81%



PPM APPLIED | ONE-TIME-EFFORT

Subcategory:

Binders

Metric:

Total Profit

Brand	Keep	Remove	Total Profit	Profit Change
n.a.	\$2,527.13	-\$236.38	\$2,290.75	10.32%
Acco	\$55.54	\$299.78	\$355.32	-84.37%
Avery	\$0.00	\$577.99	\$577.99	-100.00%
Cardinal	\$0.00	\$136.81	\$136.81	-100.00%
Fellowes	\$10,343.05	\$707.75	\$11,050.80	-6.40%
GBC	\$10,070.38	-\$2,431.83	\$7,638.54	31.84%
Ibico	\$7,173.09	\$204.10	\$7,377.19	-2.77%
Wilson Jones	\$836.87	-\$42.51	\$794.35	5.35%
Total Profit	\$31,006.05	-\$784.29	\$30,221.76	2.60%



PPM APPLIED | MAINTENANCE

Use Tableau **Dashboard**

Advantages:

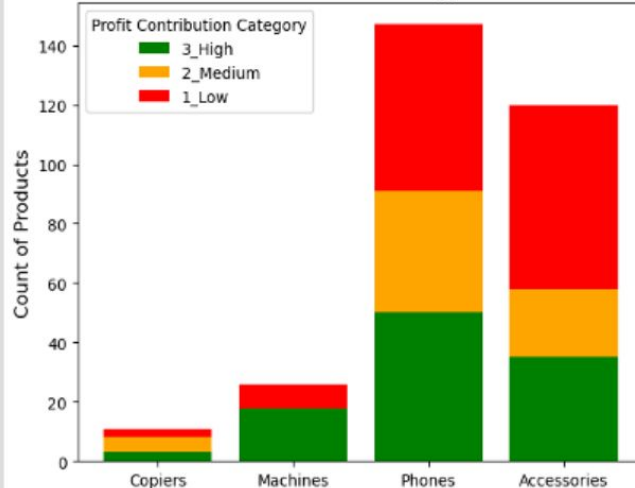
- Standardized approach of how to review data
- User does not spent time on data preparation aggregation
- Background Data refreshed regularly

PPM APPLIED | MAINTENANCE

Product portfolio review

SUPERSTORE

Product Count & Profit Contribution by Subcategory
Focus : Technology



List output for category that is selected in graph on top. Limited to 20 items

- [1] List Obsolete Products
 - [2] List Decline Products > -75%
 - [3] List Decline Products > -50%
 - [4] List Low-Profit-Contributors - FU
 - [5] List Low-Profit-Contributors - OF
 - [6] List Low-Profit-Contributors - TE
 - [7] TBD
 - [8] TBD
- [BUTTON – BACK TO MAIN]

Offer Drill-down on right
0. Start: All products
1. Level: Category
2. Level: Subcategories

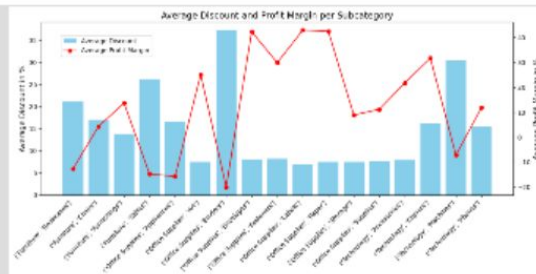
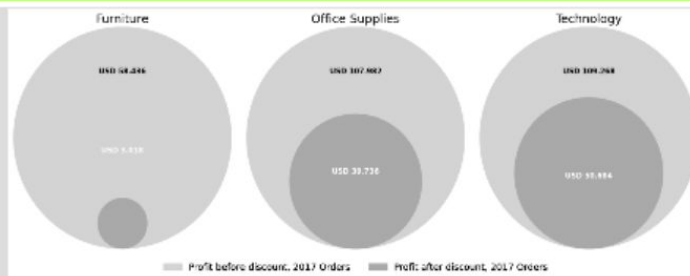
Offer possibility to toggle
between Combination or to
only focus on PLC or
Profit_Contribution

Predefined reports available
here.

PPM APPLIED | MAINTENANCE

Pricing & Discount Strategy Review

SUPERSTORE



List output for category that is selected in graph on top. Limited to 20 items

Start with 3 Circles (Actual vs. Potential profit by Category)

Offer Drill-down on the right (go down to Subcategories and potentially TOP-5 products from subcategory)

Predefined reports available here.

- [1] List Products Avg Disc > 75%
- [2] List Products Avg Disc > 50%
- [3] List Products Avg Disc > 25%
- [4] Buy-Sell-Ratio by Subcategory
- [5] Top-selling products w/o discount
- [6] TBD
- [7] TBD
- [8] TBD

[BUTTON - BACK TO MAIN]

CLUSTERING

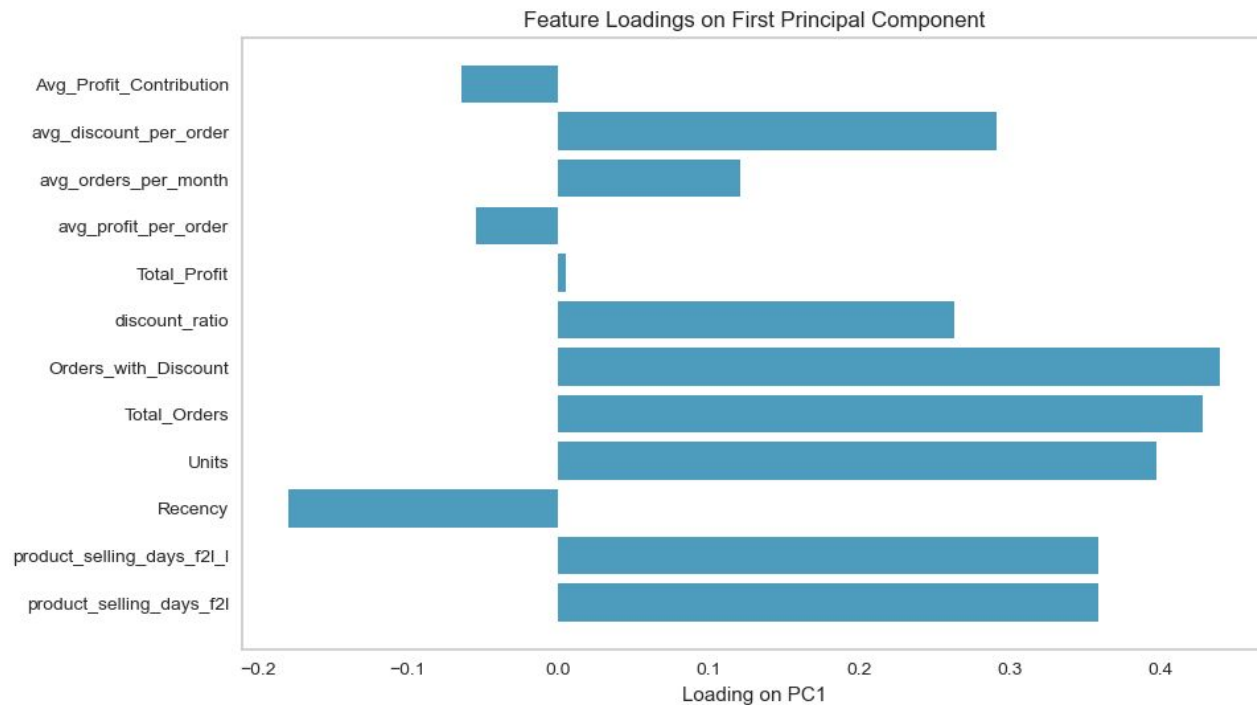
Bringing order into chaos.



CLUSTERING | STEPS

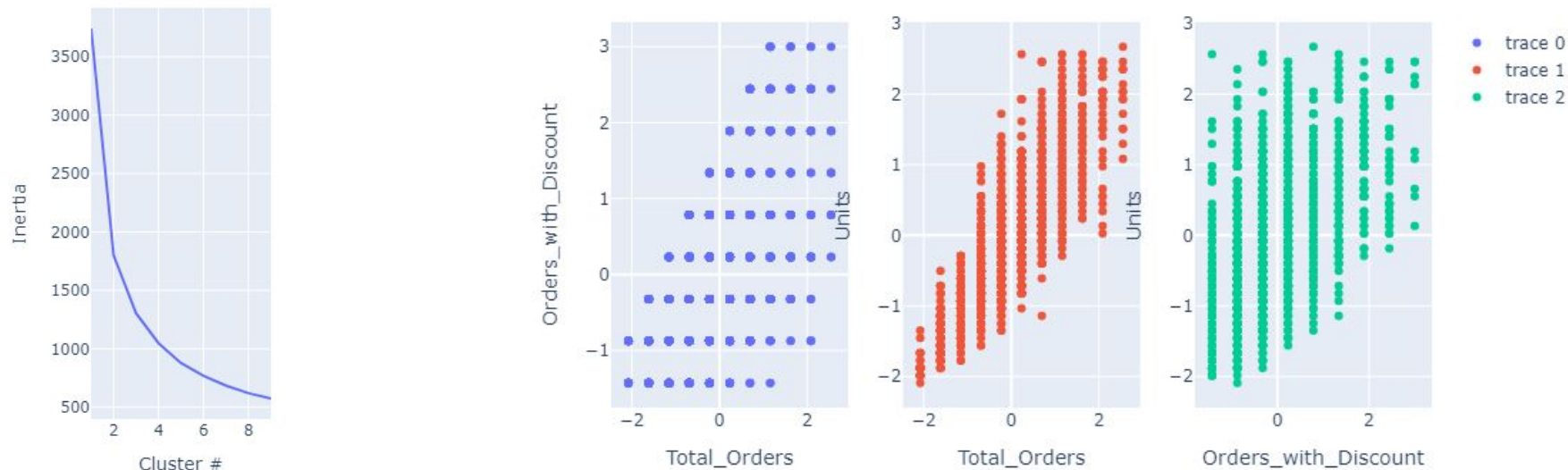
1. Comparison between raw & cleaned product portfolio as starting point
2. Engineered 17 numerical features
3. Principal Component Analysis (PCA) \Rightarrow Identification of best features
4. Outlier treatment ($1.5 \times \text{IQR}$) \Rightarrow removed 52 out of 1.300 products (-4%)
5. Scaling
6. Trial of four different Machine Learning (ML) Clustering models

CLUSTERING | PCA



CLUSTERING | ELBOW METHOD & FEATURES

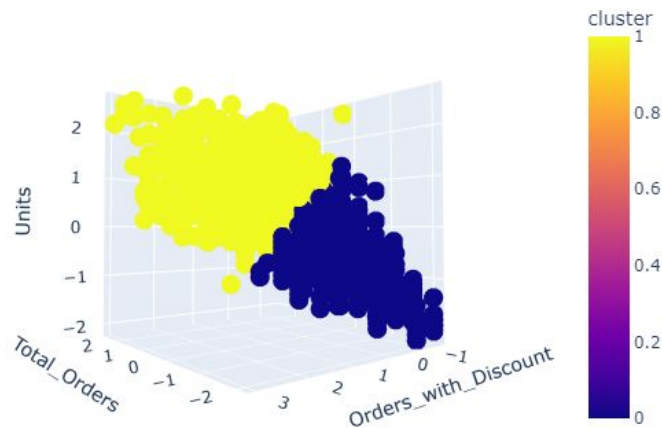
Elbow method



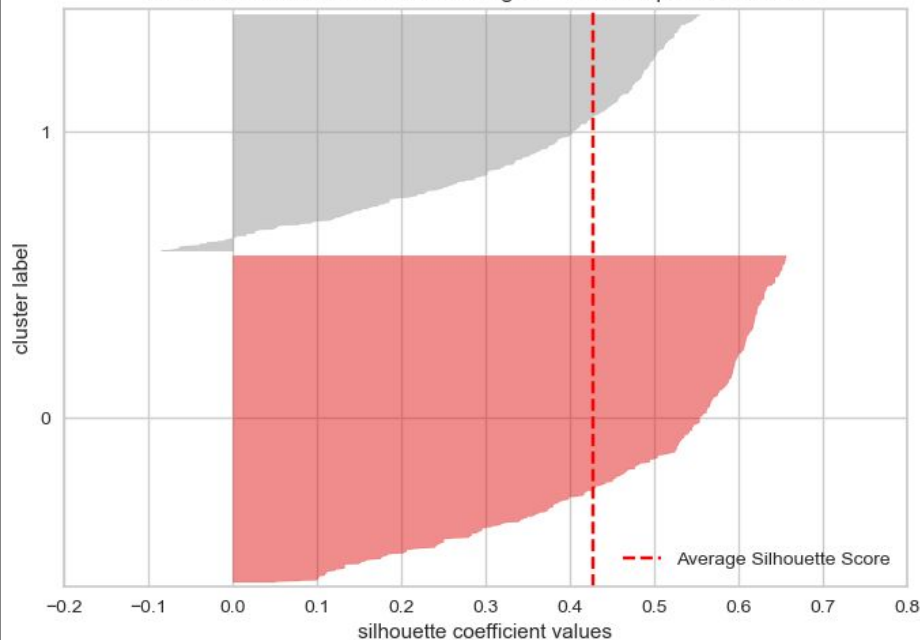
CLUSTERING | MODEL EVALUATION

KMeans | 2 Clusters

Silhouette Average: 0.427 | Davies-Bouldin Index: 0.890



Silhouette Plot of KMeans Clustering for 1248 Samples in 2 Centers



WITH MORE TIME



Spiderweb-plot for subcategories to compare multiple categories

Include selling and purchasing price as features in PCA

Deep dive into centroid placement and how to integrate in clustering routine

Play clustering categorization back to main dataset

Three level-reduction bar chart

Compare average order size by year

SUMMARY & KEY LEARNINGS

- Define **business strategy first**, strategy for **PPM follows**
- ML did not generate new insights for product categorisation so far
- Principal Lessons Learned:
 - Value of combination of domain knowledge with data analysis
 - Importance of Dashboard Conceptualization
 - Intricacies of ML Clustering

THANK YOU



*Which
tree
is
it
going
to
be?*



RESOURCES

<https://inventoro.com/the-e-commerce-owners-guide-to-a-strong-product-portfolio/>

[https://www.planview.com/resources/articles/**product-portfolio-analysis**-make-better-portfolio-decisions/](https://www.planview.com/resources/articles/product-portfolio-analysis-make-better-portfolio-decisions/)

[https://codesignal.com/learn/courses/k-means-clustering-decoded/lessons/mastering-k-means-clustering-selection-of-clusters-and-**centroid-initialization**](https://codesignal.com/learn/courses/k-means-clustering-decoded/lessons/mastering-k-means-clustering-selection-of-clusters-and-centroid-initialization)

[https://medium.com/@tarammullin/**dbscan-parameter-estimation**-ff8330e3a3bd](https://medium.com/@tarammullin/dbscan-parameter-estimation-ff8330e3a3bd)

Gemini → Tree Image Generation, Pruning strategies for cherry trees

BACK-UP

Product Lifecycle Category Definitions

Phase	Conditions
1_Introduction	days_since_first_sale <= 60
2_Growth	60 < days_since_first_sale <= 120
3_Mature	days_since_last_sale > 150 ... AND ... total order period / 2 if total orders in period 1 > period 2
4_Decline	days_since_last_sale > 150 ... AND ... total order period / 2 if total orders in period 1 > period 2
5_Obsolete	days_since_last_sale > 150

Average Profit Contribution Category Definitions

1. Compute Average Profit Contribution per Product for each Subcategory (total profit of subcategory divided by total quantity / 48 Months)
2. Assign Categories:

If Average Profit Contribution per Product above 1.5 times Average of Subcategory \Rightarrow 'High'

If Average Profit Contribution per Product ≤ 1.5 and ≥ 0.5 times Average of Subcategory \Rightarrow 'Medium'

If Average Profit Contribution per Product below 0.5 times Average of Subcategory \Rightarrow 'Low'

Cost of maintaining a wide product portfolio

Setup cost of article (ERP system, website, etc.)

Maintain relationship with supplier (orders, shipments, payments)

Stocking cost, carrying cost of inventory \Rightarrow Bound Capital

Fewer suppliers

Bargaining power with remaining suppliers \uparrow

Number received shipments \downarrow

Radar plot

Used for:

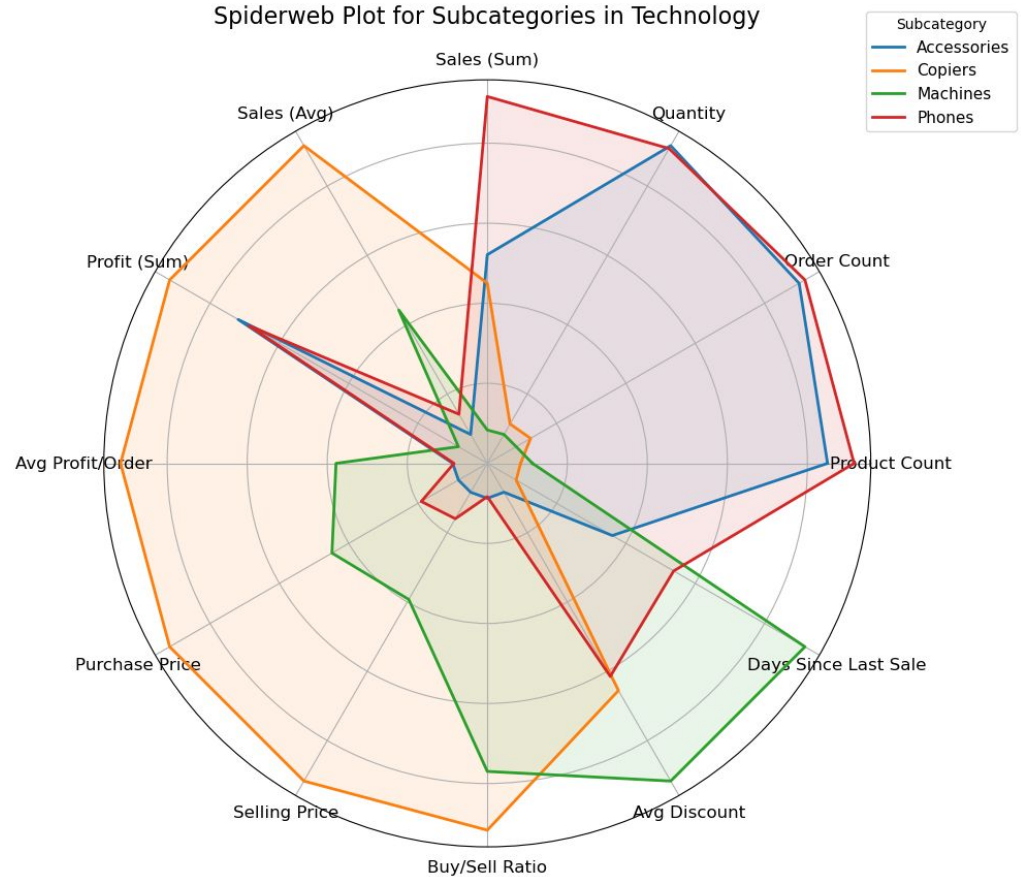
Subcategory Comparison

Selection:

Technology

Advantage:

Compare 12+ features at a glance



CLUSTERING | Methods & Metrics Explained

- **Elbow method:** measures sum of squares within, desired state: low inertia & low number of clusters
- **Silhouette Coefficient:** measures intra-cluster distance and distance to nearest cluster; from -1 to +1, if negative, then associated with wrong cluster; the higher the better
- **Davies-Bouldin Index:** measures compactness of clusters and separation between clusters; <1 means clusters well-separated and compact