Algorithmic Price Optimization with Machine Learning at Allegro

Marcin Lebiedź

Data Science Manager, Allegro

Jabir Kangarli

Product Analyst, Allegro

Event organized by: Academic Partners Foundation

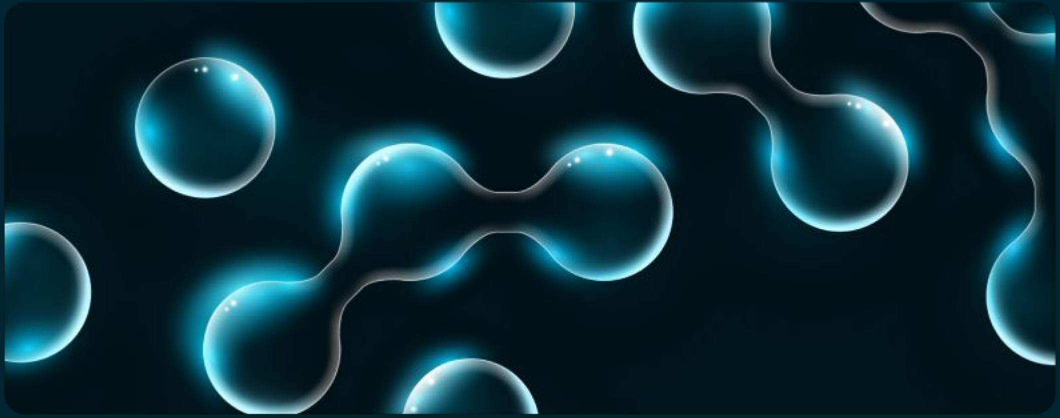
DATA SCIENCE SUMMIT

MACHINE LEARNING EDITION

A lecture selected by a Program Council consisting of recognized leaders in the IT and Data Science field.

Warsaw,
05.06.2025 - 06.06.2025

OFFICIAL LECTURE OF THE DATA SCIENCE SUMMIT MACHINE LEARNING EDITION



allegro

Introduction to Price Subsidy Mechanics

What is ML Pricing and how does the price subsidy work?

Example for Allegro Ceny program

Wełna mineralna szklana Eurowool 100 mm

4,87 ★★★★★ 90 ocen i 12 recenzji

Stan: Nowy

WEŁNA ATESTOWANA MINERALNA MATA SZKLANA EUROWOOL 0,039 10CM 10,8M2

176,00 zł SMART

(16,30 zł/m²)

zapłać później z pay sprawdź

51 osób kupiło

Liczba sztuk

1 z 311 sztuk

DODAJ DO KOSZYKA

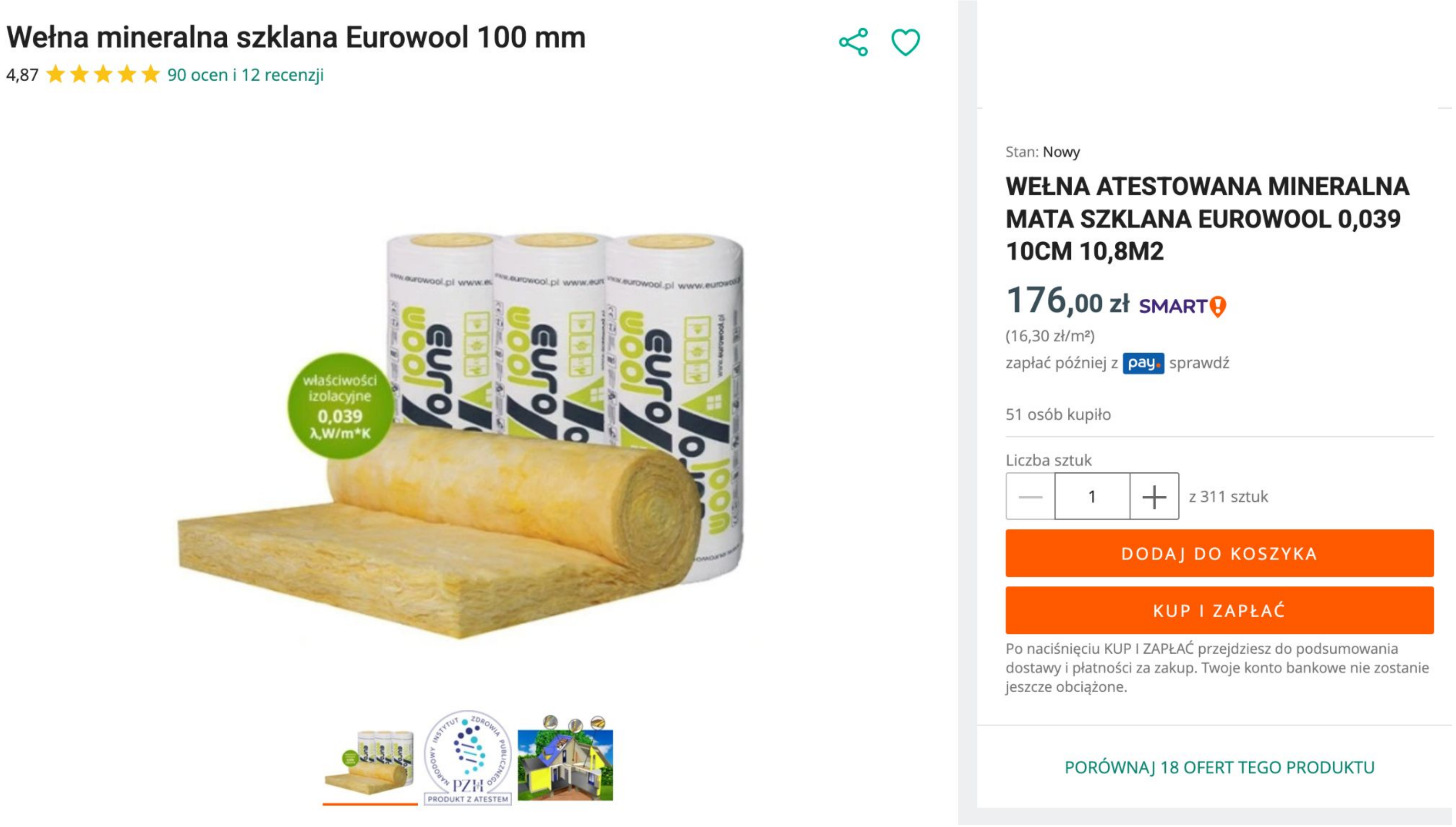
KUP I ZAPŁAĆ

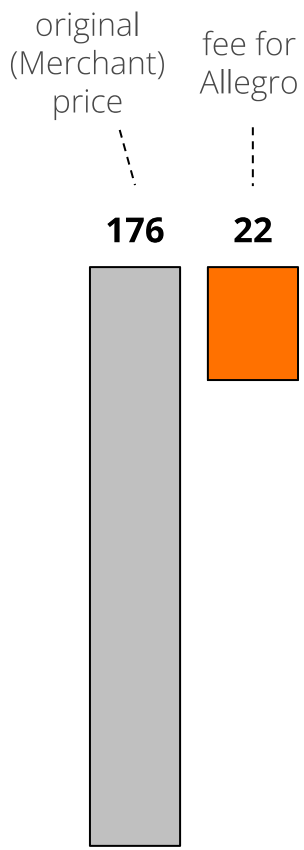
Po naciśnięciu KUP I ZAPŁAĆ przejdziesz do podsumowania dostawy i płatności za zakup. Twoje konto bankowe nie zostanie jeszcze obciążone.

PORÓWNAJ 18 OFERT TEGO PRODUKTU

What is ML Pricing and how does the price subsidy work?

Example for Allegro Ceny program

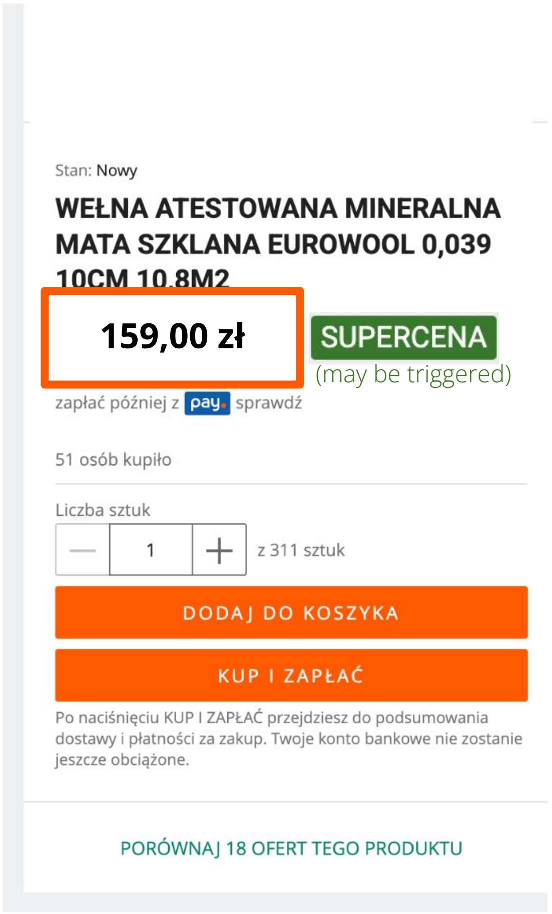


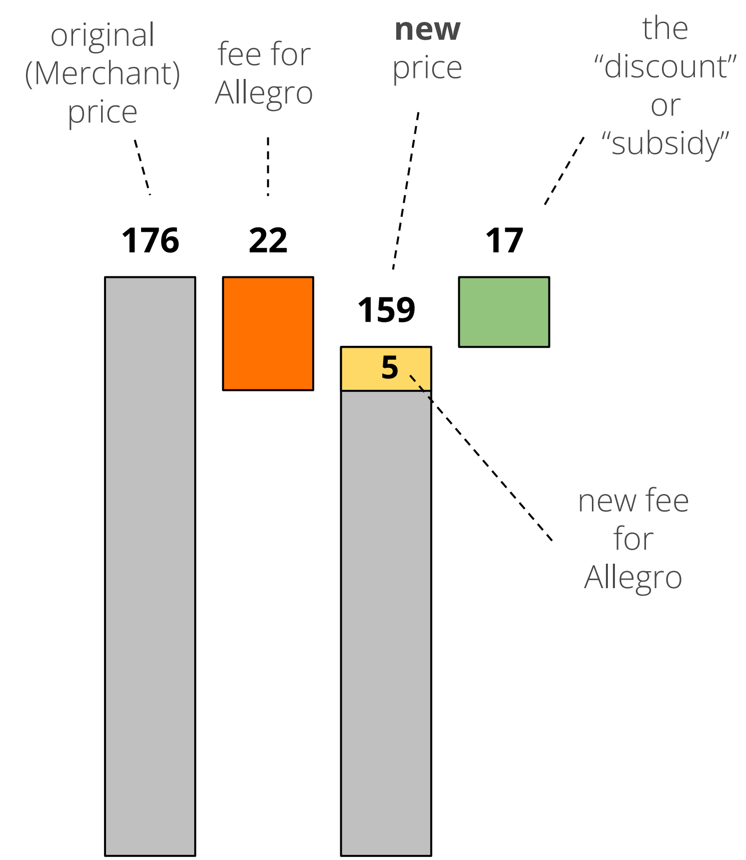


allegro

What is ML Pricing and how does the price subsidy work?

Example for Allegro Ceny program





allegro

Solution plan and challenges

Modeling & Optimization

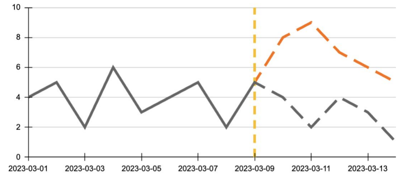
We ideated two major components:

- **MODELING** – predict expected demand given the offer's features, including price
- **OPTIMIZATION** – optimize discount levels within limited budget

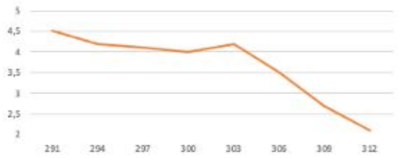
Product	Offer	…	Price
123	455	…	105
123	777	…	47
…		…	
789	990	…	28

MODELING

For each offer prepare:



Next day demand forecast model.
Demand = f(past data, future price)



Price elasticity curve:
Price (hence discount) vs demand

OPTIMIZATION

Offer	EP* (2%)	EP* (4%)	…	EP* (20%)
455	-2	0	…	1
777	50		…	
…			…	
990	30		…	28

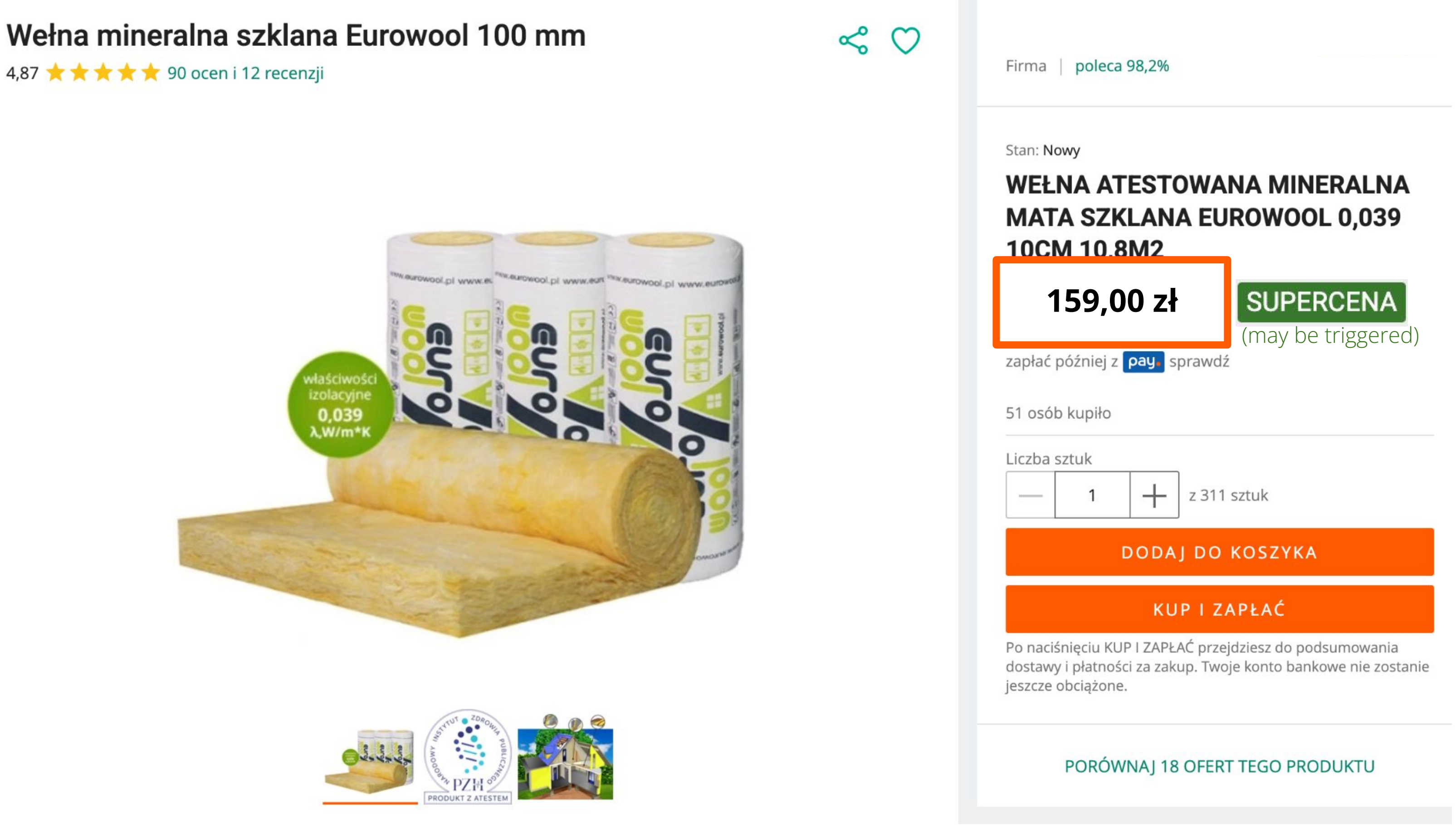
EP(n%) = estimated profit from applying an n% discount.

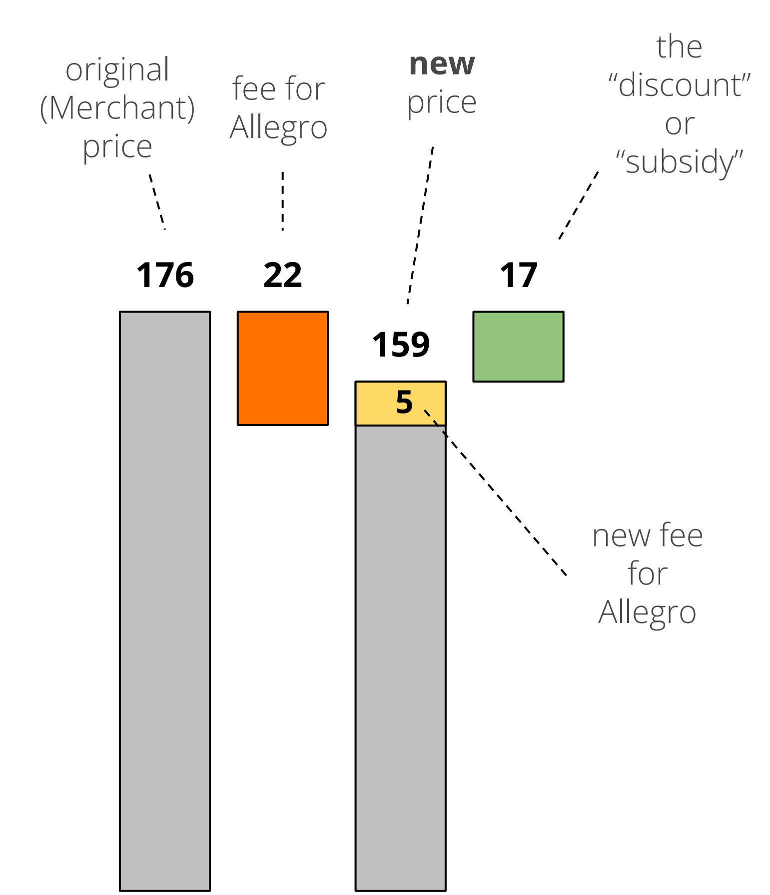
Solve Knapsack problem:
maximize expected GMV/profit/…
with constraints:
budget, one offer per product etc.

Challenge 1 – it is easy to discount non-profitably

Extra care required not to burn money

10% price discount → 77% fee lost





allegro

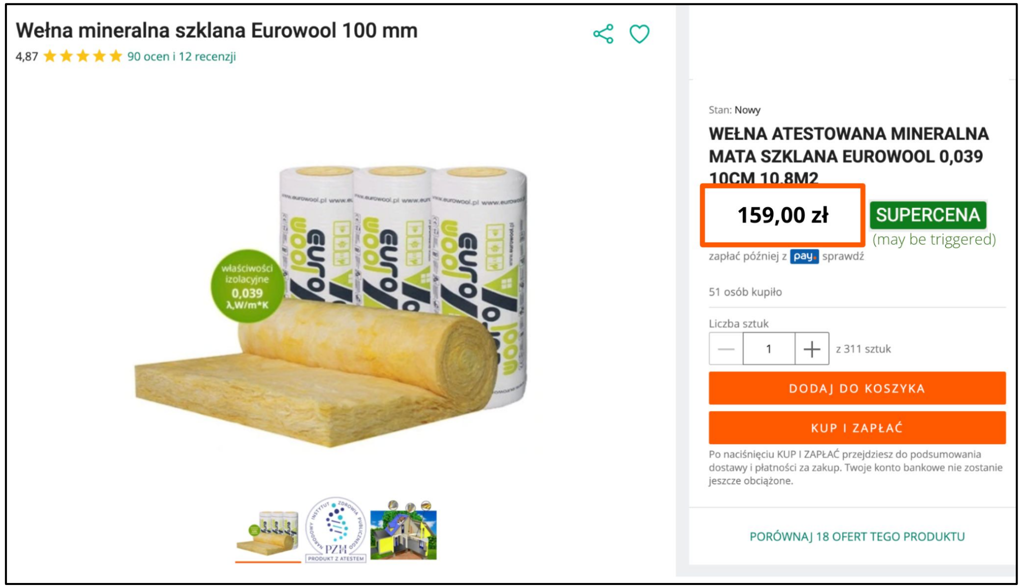
Challenge 2 – Allegro is a marketplace, offers are not independent

- Demand on offers is non-i.i.d. – “sales boost” on discounted offer may come from cannibalizing other offers.
- Hence we added price-rank related features for offer-level modeling, and established that the discounts will be selected based on the total impact on all offers of the same product…
- … but this created another challenge – of defining “the same product”…
- We rely on Allegro’s Productization (a separate project on its own).
- Product perspective brings in yet another issue: there’s an abundance of variants – multipacks, sets etc, which cannibalize each other only partially, depending on the customer needs.
- You can see how moving from *the data* to *the dataset ready for model fitting* seems to be a struggle to select the *least imperfect solution*… :)

offer 2

offer 3

discounted offer 1



offer 4

offer 5

Challenge 3 – Evaluation metrics analysis paralysis

Assessing the performance of demand forecasting models:

- absolute error metrics bias towards products with high demand (e.g. products that sell in thousands daily, as opposed to singular sales)
- relative error metrics (MAPE etc.) bias towards low demand products, zero-sales weighing the most
- we really mostly care about business performance – profitably generating non-cannibalized sales...
- we never found a “winner in all metrics”, at some plateau point moved towards experimentation and measuring business impact.

Lesson learned: real world experimentation can get you unstuck and help deliver value much faster than over analyzing.

allegro

Challenge 4 – Optimizer impact on metrics

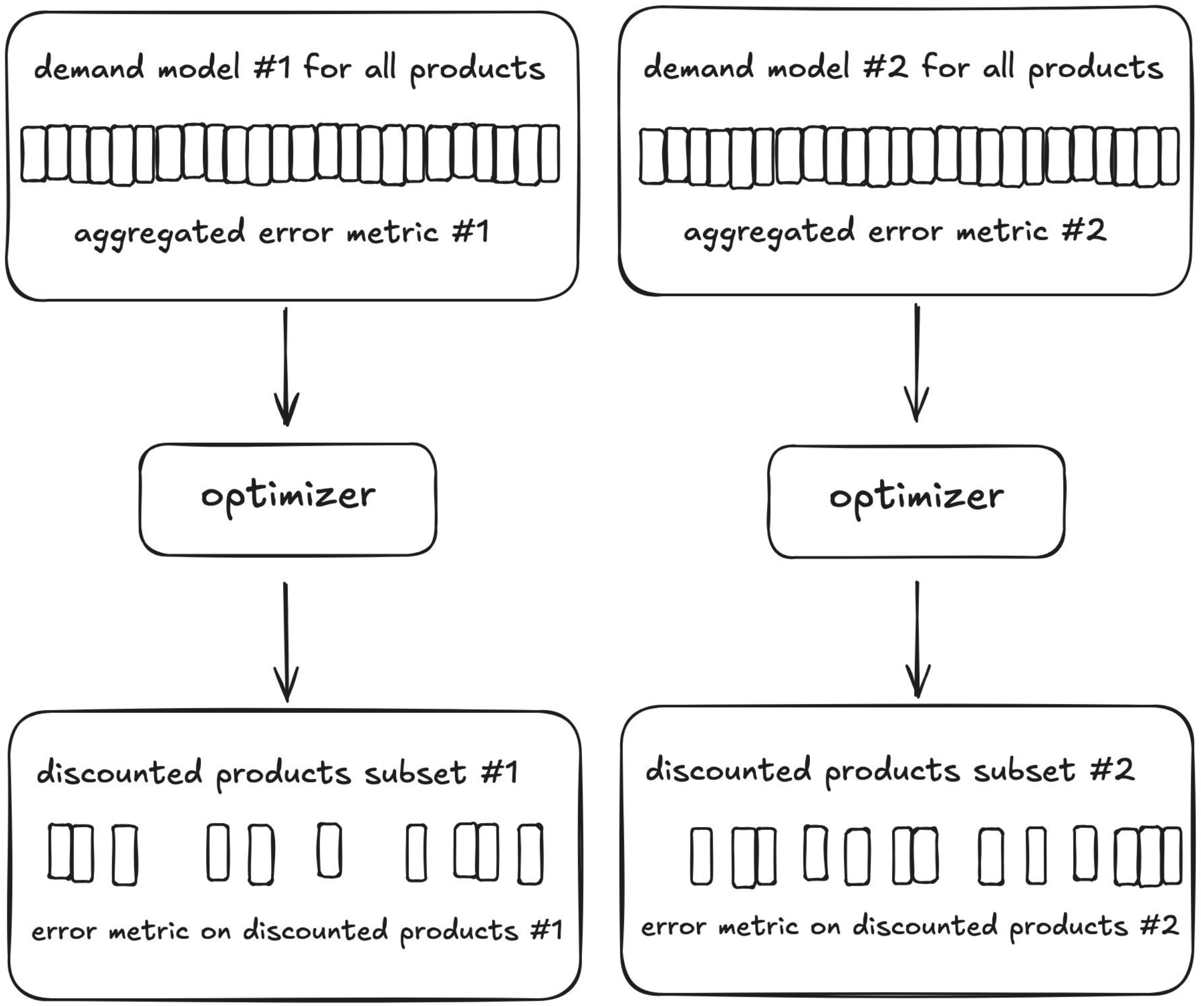
Discount Optimizer on top of model predictions makes the model selection even more difficult.

Optimizer acts as a filter selecting the “best discount candidates” according to model predictions.

It’s <u>not impossible</u> that you’ll have Model #1 <u>better</u> than Model #2 in every aggregated error metric, but <u>worse</u> on the subset of products finally selected for discount 🤷‍♂️

You can’t estimate the final performance while looking at the model metrics alone and not moving to the last step in the downstream task.

Lesson learned: Model metrics are helpful, but not conclusive. Downstream task performance is conclusive.



Challenge 5 – estimating causal interventional effect

By default, most observational-data based regression ML models estimate:

$$P(Y|X = x)$$

... but here we're interested in (using Pearl's *do-calculus* notation):

$$P(Y|\mathrm{do}(X = x))$$

... and these two **are not the same!**

Toy example. Consider:

$$P(\text{water boils} \mid \text{thermometer shows } 100^{\circ}\text{C})$$

vs

$$P(\text{water boils} \mid \text{we manually set thermometer to } 100^{\circ}\text{C})$$

allegro

Our approach

Keeping "**good**" and removing "**bad**" (biasing) features ("*controls*") based on causality theory.

Main point: more features will generally help you with observational prediction, but <u>may harm</u> interventional effect estimation. Select them carefully!

Sociological Methods & Research
1–34

DOI: 10.1177/00491241221099552
journals.sagepub.com/home/smr
SAGE

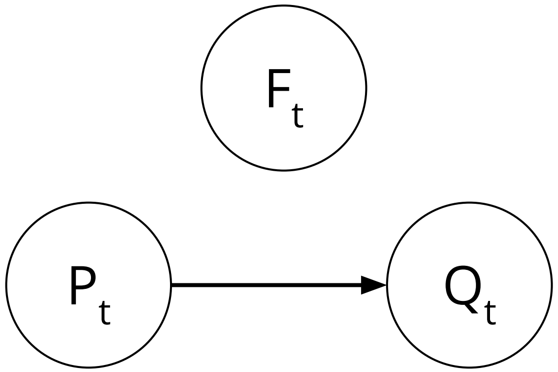
A Crash Course in Good and Bad Controls

Carlos Cinelli[1], **Andrew Forney**[2], **and Judea Pearl**[3]

allegro

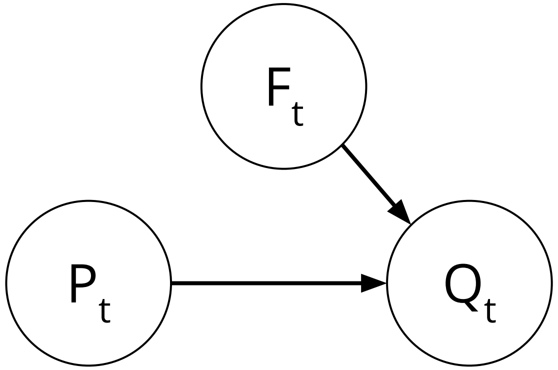
Example – day of the week

P_t – Price at time t
Q_t – Demand at time t
F_t – Day of the week



Example – day of the week

P_t – Price at time t
Q_t – Demand at time t
F_t – Day of the week

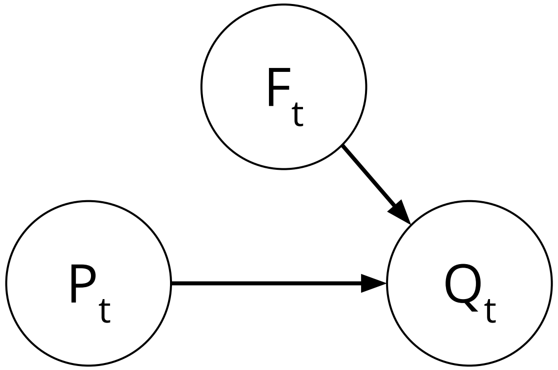


Assume:

- day of the week influences the demand
- the merchants don't change prices according to weekday.

Example – day of the week

P_t – Price at time t
Q_t – Demand at time t
F_t – Day of the week



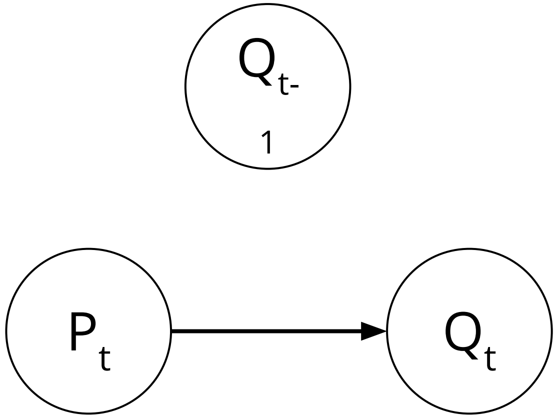
Assume:

- day of the week influences the demand
- the merchants don't change prices according to weekday.

It's a **good control** ("model 8" in the paper).
Does not introduce bias in estimation of P→Q, but (possibly) explains some variations of Q making this estimation more "sample-efficient".

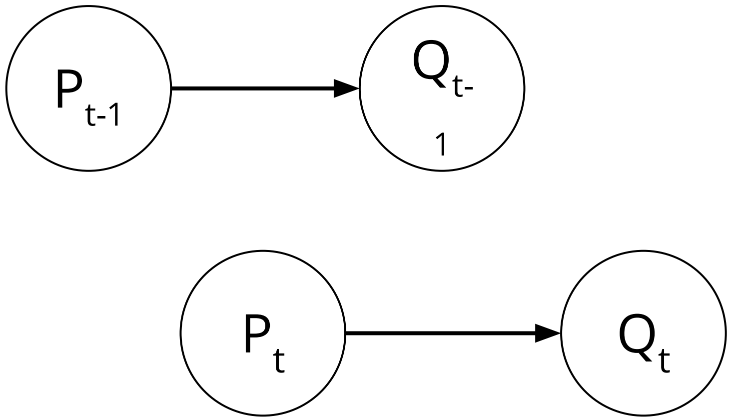
Example 2 – lagged demand

P_t – Price at time t
Q_t – Demand at time t



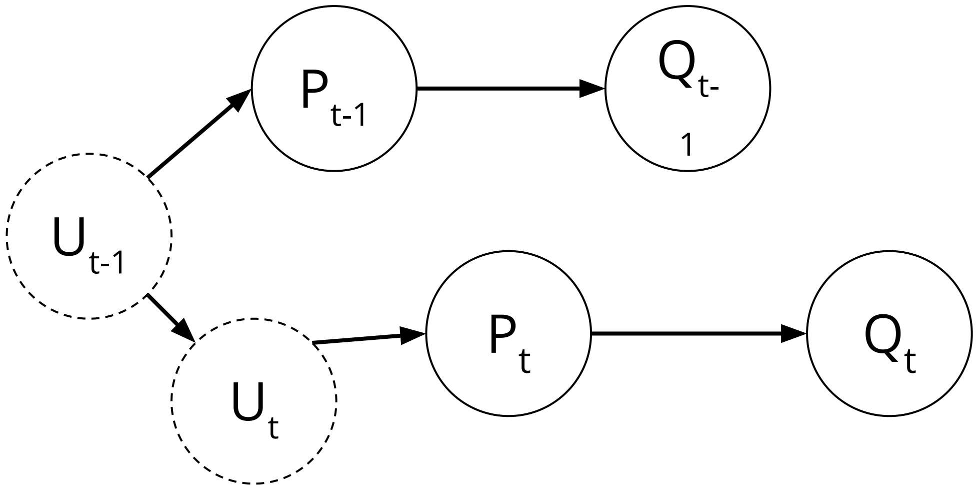
Example 2 – lagged demand

P_t – Price at time t
Q_t – Demand at time t



Example 2 – lagged demand

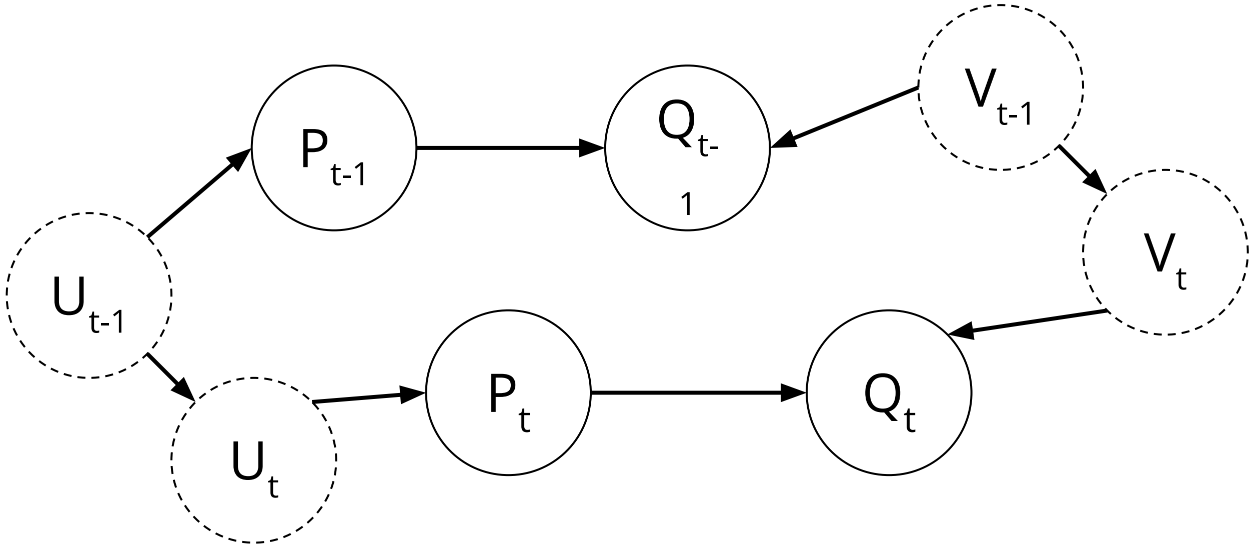
P_t – Price at time t
Q_t – Demand at time t
U_t – Unobserved confounder – hidden merchant state



allegro

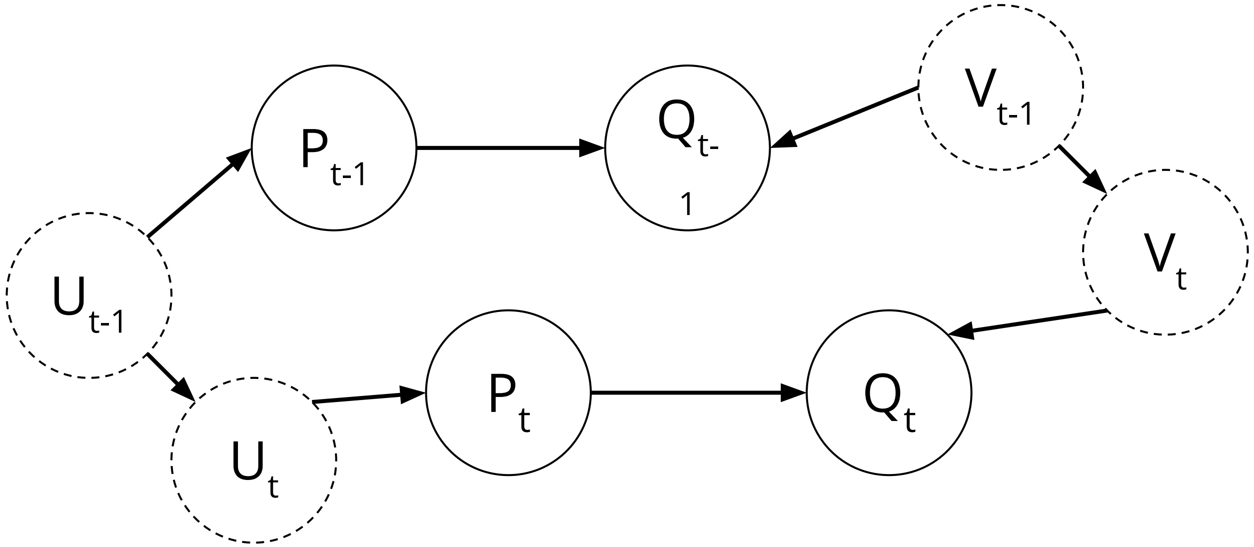
Bad control example – lagged demand

P_t – Price at time t
Q_t – Demand at time t
U_t – Unobserved confounder – hidden merchant state
V_t – Unobserved confounder – hidden market state



Bad control example – lagged demand

P_t – Price at time t
Q_t – Demand at time t
U_t – Unobserved confounder – hidden merchant state
V_t – Unobserved confounder – hidden market state



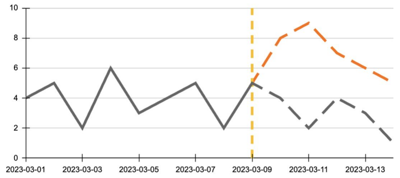
Bad control, ("model 7" in the paper), a.k.a M-bias.
Controlling for Q_{t-1} biases estimation of $P_t \rightarrow Q_t$.
Caveat – we could add P_{t-1} to close the path again, but that's strongly correlated with P_t and generates even more issues

allegro

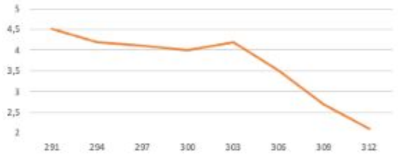
Final setup – wrap-up

MODELING

For each offer prepare:



Next day demand forecast model.
Demand = f(past data, future price)



Price elasticity curve:
Price (hence discount) vs demand

Tweedie regression (internal upgraded implementation based on sklearn)
Features relate to:

- price, price-rank relative to other offers
- offer and product quality
- delivery and other badges visible by customer
- merchant quality
- calendar context

OPTIMIZATION

Offer	EP* (2%)	EP* (4%)	…	EP* (20%)
455	-2	0	…	1
777	50		…	
…			…	
990	30		…	28

EP(n%) = estimated profit from applying an n% discount.

Solve Knapsack problem:
maximize expected GMV/profit/...
with constraints:
budget, one offer per product etc.

Linear programming-based solver

scikit learn

Google Big Query

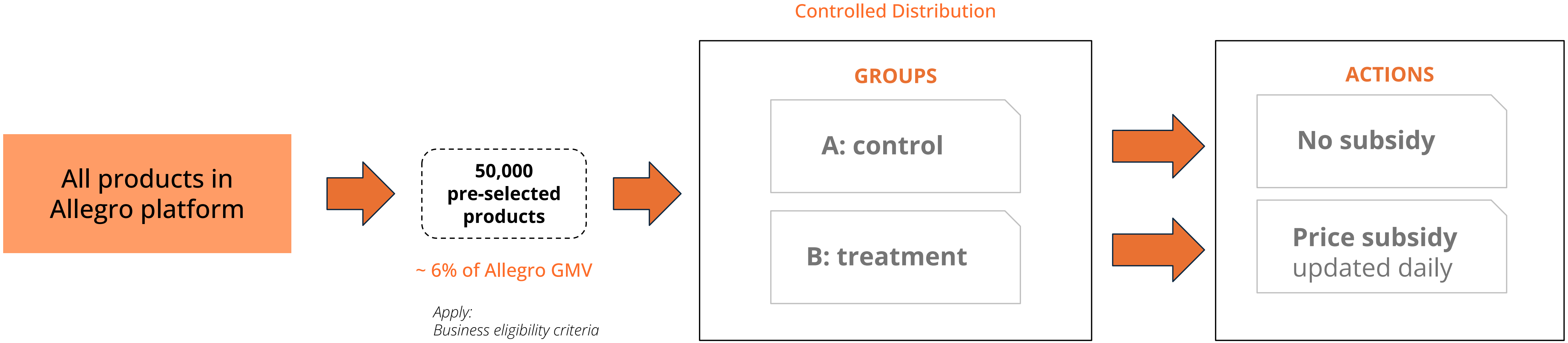
Apache Airflow

allegro

Experimentation

Experimental setup

Product level A/B testing

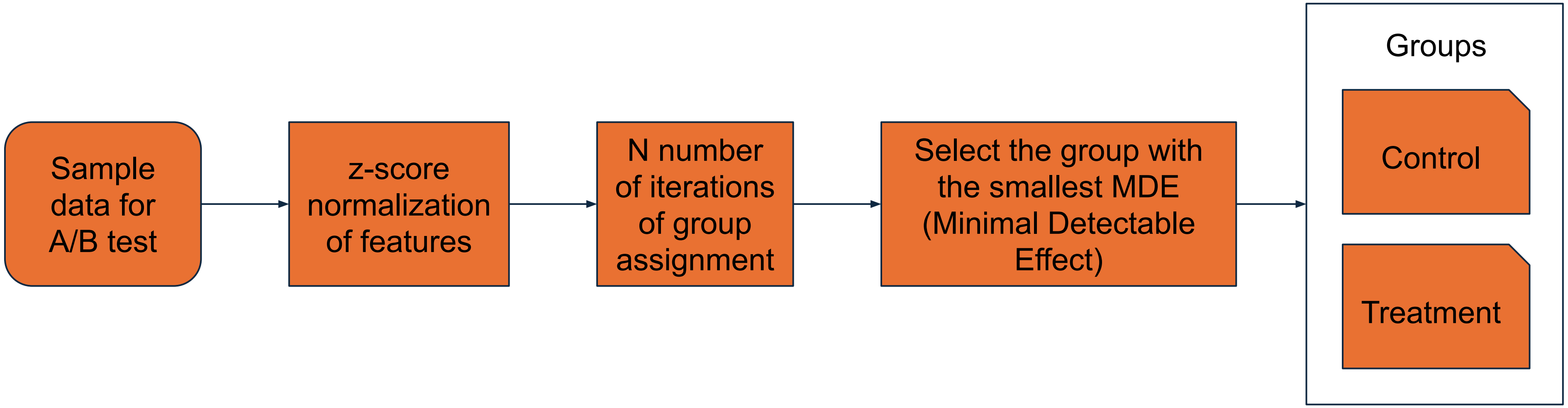


Controlled Distribution

Part of internal A/B testing package developed at Allegro

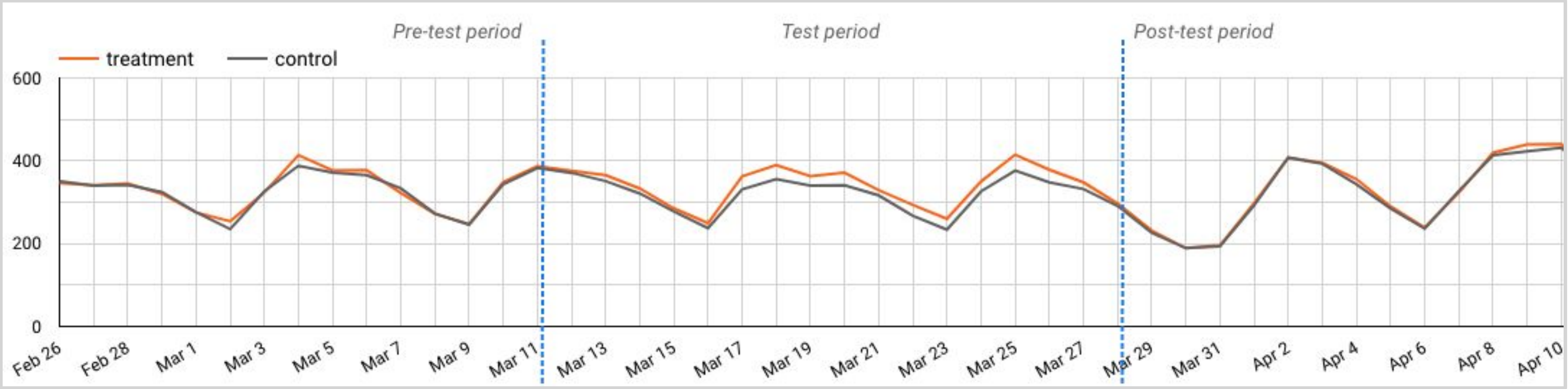
Key features

- Our recommended approach especially for limited sample size, so that random distributional differences don't skew the results
- Smart group assignment: Distributes products into groups while maintaining balanced metric distributions
  - Sampling-based approach to minimize distributional differences of normalized group metric means with Euclidean distance
- Input based slice aware splitting: Handles splits across multiple slicing dimensions (i.e category)
- Using Monte Carlo sampling, control and treatment group attributes like GMV, profit, and category are nearly identical



A/B test Overview
Product level A/B testing

Average GMV per group



- **Consistent growth** in GMV translating into **significant yearly impact** on the whole Allegro level.
- Generated GMV is profitable - incremental sales pay off the subsidy investments.
- The Evaluation was based on bootstrapping due to several advantages:
    - No distributional assumptions
    - Particularly useful when you are dealing with a limited sample size
    - Confidence intervals are derived from the N iterations of resampled data
    - Each bootstrap iteration captures around 63–65% of unique data points, but with enough iterations, you can eventually cover the entire dataset

Results

ML Pricing Optimization in numbers

50k	6%	8 m PLN	+10%	~100 m PLN
Individual price elasticity models	allegro.pl GMV portfolio covered with models	subsidy budget for 2025	incr. GMV on treated products	incremental yearly GMV

Thank you!

We wouldn't be here without the amazing team of contributors:

Data Scientists:

Paweł Olszewski

Krystian Korzec

Monika Sołtysek

Marzena Wojtkowiak

Klaudia Walewska-Łubian

Marta Szczęsna

Product Analysts:

Jabir Kangarli

Agata Redmerska

Mateusz Kasprowicz

Product Managers:

Karolina Cuch

Maciej Karbowski

Data Engineers:

Dawid Szweda

Jakub Wiercimak

Jakub Krajniak

Alicja Nowak

Team Managers:

Marcin Lebiedź

Luboš Květ

Wojciech Smolak

allegro

Thank you for watching!

Remember to leave your questions and rate the presentation in the section below.

linkedin.com/in/jabirkangarli
linkedin.com/in/marcin-lebiedz

Event organized by: Academic Partners Foundation

DATA SCIENCE SUMMIT

MACHINE LEARNING EDITION

A lecture selected by a Program Council consisting of recognized leaders in the IT and Data Science field.

Warsaw,
05.06.2025 - 06.06.2025

OFFICIAL LECTURE OF THE DATA SCIENCE SUMMIT MACHINE LEARNING EDITION

ACADEMIC PARTNERS

allegro