

A thick black L-shaped frame is positioned on the left and bottom edges of the slide, framing the central text.

# SUPERMARKET PRODUCTS: A COMPARISON

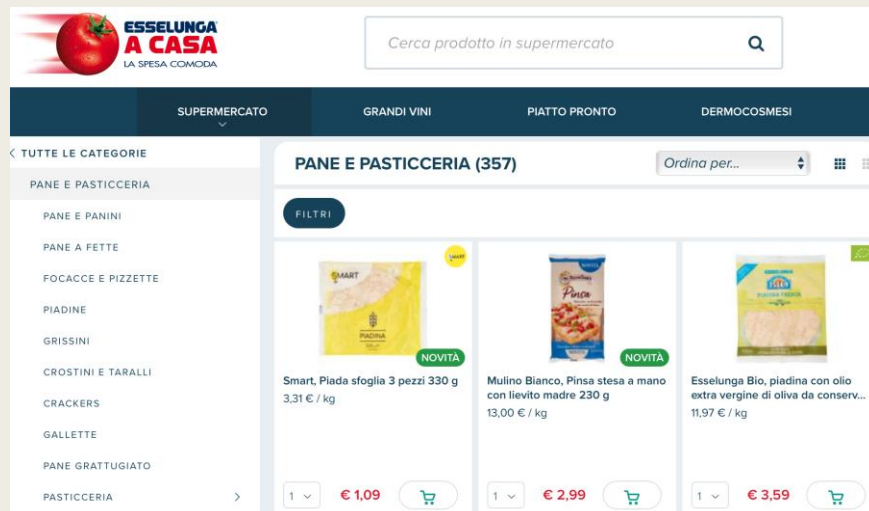
between Carrefour, Esselunga and Conad

# Main objective

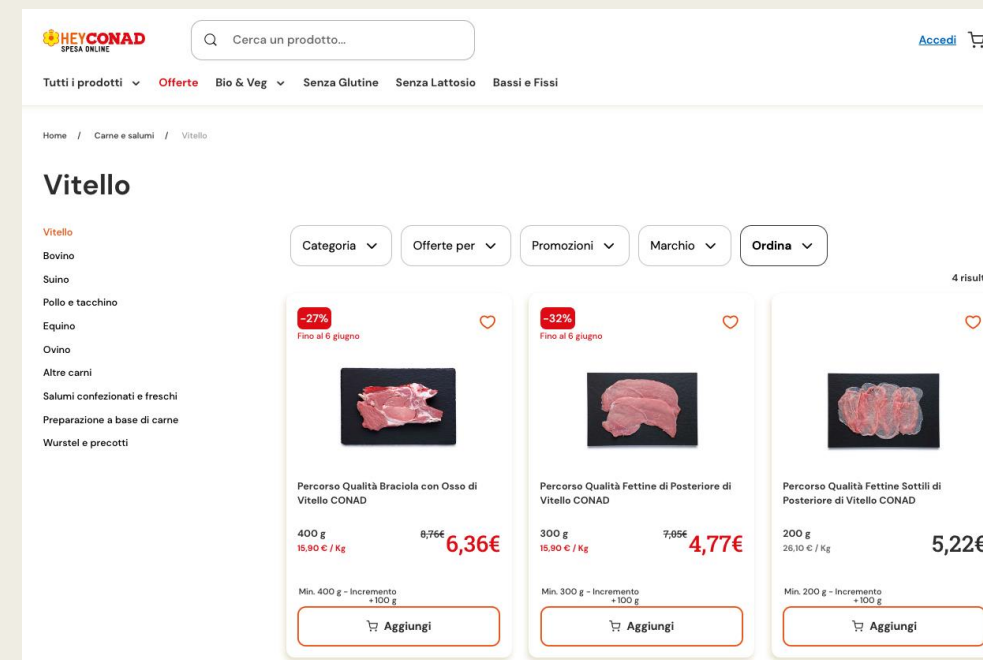
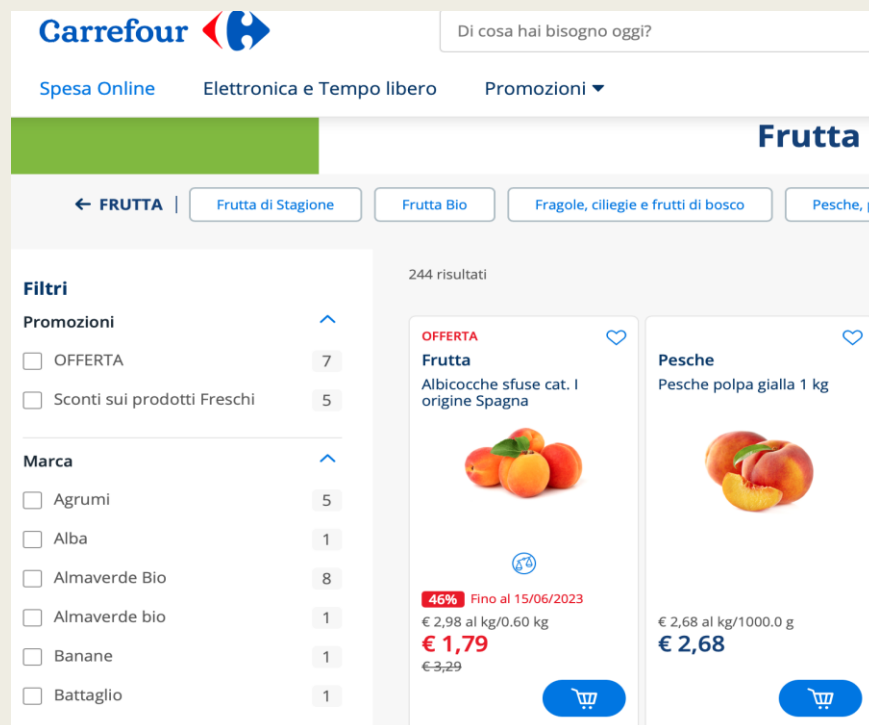
- Our main goal was to compare prices of the same products across 3 supermarkets in Milan: Esselunga, Conad and Carrefour.
- The technique we employed to carry out this task is web scraping.
- For each supermarket, we obtained all product names, their prices and the categories they belong to.



	Product	Price	Category
0	Esselunga Bio, cipolle dorate 500 g	Euro 3,16 / kg	SPESA BIO
1	Esselunga Bio, datterini biologici 300 g	Euro 8,27 / kg	SPESA BIO
2	Esselunga Bio, Insalatina 100 g	Euro 14,80 / kg	SPESA BIO
3	Esselunga Bio, 6 uova fresche biologiche	Euro 0,49 / pz	SPESA BIO
4	Esselunga Bio, prezzemolo tritato biologico su...	Euro 23,00 / kg	SPESA BIO



From....



...to

	Product	Price	Category
0	Esselunga Bio, cipolle dorate 500 g	Euro 3,16 / kg	SPESA BIO
1	Esselunga Bio, datterini biologici 300 g	Euro 8,27 / kg	SPESA BIO
2	Esselunga Bio, Insalatina 100 g	Euro 14,80 / kg	SPESA BIO
3	Esselunga Bio, 6 uova fresche biologiche	Euro 0,49 / pz	SPESA BIO
4	Esselunga Bio, prezzemolo tritato biologico su...	Euro 23,00 / kg	SPESA BIO

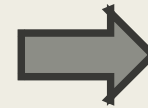
	Product	Price	Category	Subcategory
0	Arance Valencia Italia cal. 7 1,5 kg Percorso ...	2,65 € / Kg	Frutta e verdura	Frutta fresca
1	Pompelmi Rossi Cal. 4/5 Cat. I	2,48 € / Kg	Frutta e verdura	Frutta fresca
2	Limone Costa d'Amalfi I.G.P. Primofiore Italia...	5,56 € / Kg	Frutta e verdura	Frutta fresca
3	Limoni Verdello Italia 0,500 kg Verso Natura C...	3,96 € / Kg	Frutta e verdura	Frutta fresca
4	Percorso Qualità Limoni Italia Cal. 4 Cat. I R...	3,96 € / Kg	Frutta e verdura	Frutta fresca

	Product	unit Price	Category
0	Carrefour Correttore a nastro maxi 5x12 mm	€ 1,89 al\mpz/1.0 pz	articoli per la casa
1	Carrefour Confezione di 4 Penne a Sfera Blu Pe...	€ 1,49 al\mpz/1.0 pz	articoli per la casa
2	Carrefour Forbici punte arrotondate 17cm	€ 2,99 al\mpz/1.0 pz	articoli per la casa
3	Carrefour Confezione di 4 Penne a Sfera Nero P...	€ 1,49 al\mpz/1.0 pz	articoli per la casa
4	Carrefour Quaderno A4 punto metallico con cope...	€ 1,59 al\mpz/1.0 pz	articoli per la casa

# DATA PREPROCESSING

- We standardised the prices across the three supermarkets
- extracted quantities when present

	Category	Product	Price
0	SPESA BIO	Esselunga Bio, cipolle dorate 500 g	Euro 3,16 / kg
1	SPESA BIO	Esselunga Bio, datterini biologici 300 g	Euro 8,27 / kg
2	SPESA BIO	Esselunga Bio, Insalatina 100 g	Euro 14,80 / kg
3	SPESA BIO	Esselunga Bio, 6 uova fresche biologiche	Euro 0,49 / pz
4	SPESA BIO	Esselunga Bio, prezzemolo tritato biologico su...	Euro 23,00 / kg



	Category	Product	Price	Quantity
0	spesa bio	esselunga bio, cipolle dorate 500 g	3.16 € / kg	0.5 kg
1	spesa bio	esselunga bio, datterini biologici 300 g	8.27 € / kg	0.3 kg
2	spesa bio	esselunga bio, insalatina 100 g	14.80 € / kg	0.1 kg
3	spesa bio	esselunga bio, 6 uova fresche biologiche	0.49 € / pz	6.0 pz
4	spesa bio	esselunga bio, prezzemolo tritato biologico su...	23.00 € / kg	0.05 kg

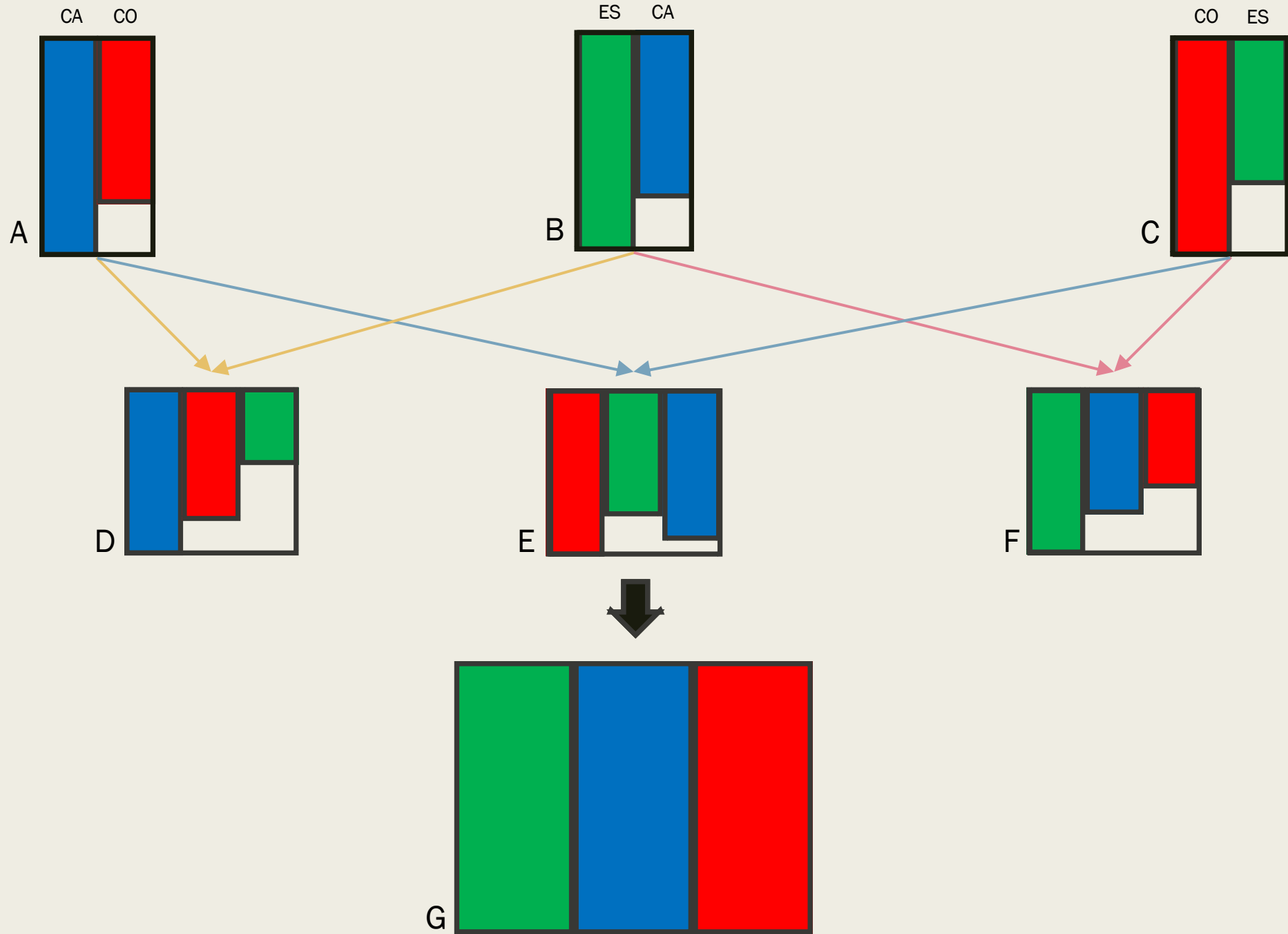
# DATA INTEGRATION

from three different sources

- Firstly we performed pairwise comparisons. For each supermarket, we looked for matching products from one other supermarket, provided that the products shared the same weight
- In order for there to be a match, we imposed the condition that two product strings had to have a cosine similarity of at least 70%
- We created a Category column, where we saved the category names of the identified products

	Carrefour	Conad	Carrefour price	Conad price	Carrefour category	Conad category	Quantity
961	carrefour bio latte parzialmente scremato 1 l	Latte Parzialmente Scremato Biologico 1 l Vers...	1.35 € / l	1.79 € / L	prodotti biologici	Formaggi, latte e uova	1.0 l
966	carrefour bio latte parzialmente scremato micr...	Più Gusto Latte UHT Microfiltrato Parzialmente...	1.75 € / l	1.29 € / L	prodotti biologici	Formaggi, latte e uova	1.0 l
967	carrefour bio crescita 100 g	Crescenza 100 g Conad	17.50 € / kg	11.50 € / Kg	prodotti biologici	Formaggi, latte e uova	0.1 kg

- We decided to keep products even in the case they presented no matches





# DATA INTEGRATION

- We merged the three datasets, obtaining a final one with all products from all supermarkets with their matches from other supermarkets, if any.
- Duplicates were dropped

	Esselunga	Carrefour	Conad	Esselunga price	Carrefour price	Conad price	Esselunga category	Carrefour category	Conad category	Quantity
957	esselunga bio, latte microfiltrato biologico p...	carrefour bio latte parzialmente scremato micr...	Più Gusto Latte UHT Microfiltrato Parzialmente...	1.79 € / l	1.75 € / l	1.29 € / L	spesa bio	prodotti biologici	Formaggi, latte e uova	1.0 l
958	esselunga bio, crescita biologica 100 g	carrefour bio crescita 100 g	Crescenza 100 g Conad	18.90 € / kg	17.50 € / kg	11.50 € / Kg	spesa bio	prodotti biologici	Formaggi, latte e uova	0.1 kg
959	sterzing vipiteno, bio yogurt bianco 150 g	sterzing vipiteno bio yogurt da latte fieno bi...	Sterzing Vipiteno bio yogurt da latte fieno bi...	7.21 € / kg	7.67 € / kg	6.60 € / Kg	spesa bio	prodotti biologici	Formaggi, latte e uova	0.15 kg
967	esselunga bio, fagiolini biologici surgelati 4...	carrefour bio fagiolini finissimi surgelati bi...	Fagiolini Finissimi Biologici Surgelati 450 g ...	4.65 € / kg	4.87 € / kg	4.64 € / Kg	spesa bio	prodotti biologici	Surgelati e gelati	0.45 kg
974	fiorentini, bio gallette di grano saraceno 100 g	carrefour bio gallette di grano saraceno 100 g	Gallette di Grano Saraceno Biologiche 100 g Ve...	27.50 € / kg	20.90 € / kg	19.90 € / Kg	spesa bio	prodotti biologici	Panetteria e snack salati	0.1 kg



# DATA QUALITY

- We proceeded by checking random matches, to see if the ones we had obtained actually linked together the same products
- For each supermarket, we chose a category and sampled 50 products.
- 87% of the matches showed real correspondences of data
- Most discrepancies were due to flavour differences:

eg. milk, pro 35g high protein cacao 350 g,  
milk pro high protein 35g vaniglia 350 g,  
milk pro high protein 35g vaniglia 350 g

# QUERIES

- We exported the integrated dataset into SQL and made some queries.
- Find the most convenient store to buy Noberasco dried fruit

Esselunga	Carrefour	Conad	Esselunga price	Carrefour price	Conad price	Esselunga category	Carrefour category	Conad category	Quantity	Price um	Most Convenient
noberasco bio uva sultanina biologiche 250 g	carrefour uva sultanina 250 g	Uva Sultanina 250 g Conad	11.96	6.32	6.2	spesa bio	frutta e verdura	Frutta e verdura	0.25 kg	€/ kg	6.2
esselunga mandorle pelate 40 g	mandorle pelate 40 g	Noberasco € 0.99 Mandorle Pelate 40 g	24.75	24.75	24.75	frutta e verdura	frutta e verdura	Frutta e verdura	0.04 kg	€/ kg	24.75
esselunga nocciole sgusciate 40 g	nocciole sgusciate 40 g	Noberasco € 0.99 Nocciole Sgusciate 40 g	24.75	24.75	24.75	frutta e verdura	frutta e verdura	Frutta e verdura	0.04 kg	€/ kg	24.75
noberasco bio uva sultanina biologiche 250 g	carrefour bio uva sultanina biologica 250 g	Uva Sultanina 250 g Conad	11.96	10.72	6.2	spesa bio	prodotti biologici	Frutta e verdura	0.25 kg	€/ kg	6.2
noberasco bio uva sultanina biologiche 250 g	carrefour uva sultanina 250 g	Uva Sultanina 250 g Conad	11.96	6.32	6.2	frutta e verdura	frutta e verdura	Frutta e verdura	0.25 kg	€/ kg	6.2

- Suppose that you need apples, pears and bananas. Where best to buy them to spend less?

min carrefour	min conad	min esselunga
8.17	6.63	6.97



THANK YOU FOR YOUR  
ATTENTION

