Business Case #3

Instacart Market Basket Analysis

Teaching staff: Fernando Bação, João Fonseca and David Silva

NOVA Information Management School – Business Cases with Data Science (Spring 2020/2021)

General context

- Understanding the purchasing patterns/behaviors of consumers is a **key task for any retail company**;

- Identifying the **relationships between the different types of products** (e.g., complementary, substitute and convenience goods), provides the company a wholistic view about its customers and the portfolio of products it has to offer.

General **context**

- **Market basket analysis** is a data mining technique that is used to inform a retailer about relationships across different products.

- This information can be used to optimize the store's layout, develop cross-promotional programs and develop different marketing strategies for its customers.

Business situation

- **Instacart** is an american company that provides a grocery delivery and pick-up service via a website or mobile app in the United States and Canada.

- Within their platform, users can select items from a wide portfolio of products, while allowing you to better manage your shopping by speaking with the personal shopper assigned to you. Your groceries are delivered to you the same day they are ordered.

Business situation

- Jane Doe, district manager at Instacart, has been trying to use this transactional data to understand more about the business, but the lack of qualified in-house data scientists hindered her ability to take full advantage of this data.

- So, your team was hired to address the questions they have been looking for.

Business situation

- You should aim to provide an overview of Instacart's business as complete as possible. Jane is particularly interested in the topics below:

  1) What are the main **types of consumer behavior** in the business?

  2) Which types of products should have an extended amount of product offerings?

  3) Which types of products can be seen as **substitutes**?

  4) Which items are **complementary**?

Business situation

- Remember: Jane is looking to find **as much useful information as possible** out of this dataset.

- Convey your findings as completely as you can, while **maintaining your non-technical audience engaged**.

- **Use visualizations whenever possible**, and **set up a context for the information** you're passing (i.e., in one or two sentences, why is this information you're providing any useful?)

Metadata

Departments

PK	department_id
	department

Products

PK	product_id
FK	department_id
	product_name

Order Products

FK	order_id
FK	product_id
	add_to_cart_order
	reordered

Orders

PK	order_id
FK	user_id
	order_number
	order_dow
	order_hour_of_day
	days_since_prior_order