DATA IMMERSION PORTFOLIO

INSTACART ONLINE GROCERIES

MARY SASSAQUI

DATABASE:

- Programs: Anaconda/Jupiter, Python, Matplot/Seaborn;

- Data Souce: Instacart (open-source), Dictionary, Customer dataset (ficticious).

HANDS ON DATA!

- Learn Python programming language and expand through Stackflow;

- Clean, Wrangle, Label, Create new columns, Merge datasets;

- Write a final report in Excel.

The Objective

PERFORM INITIAL DATA AND EXPLORATORY ANALYSIS

The Business

The sales are performing well but in order to improve the business, some information must be uncovered:

- Customer's behaviour and profile;
- Purchases patterns;
- Segmentation;
- Obtain Insights and make Recommendations.

Instacart is a online grocerie store

Operates through an app

The Search

What are we are looking for?

Target

Email-marketing
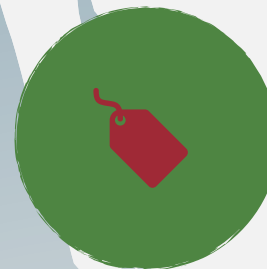- Right people
- Right time

Habits

Most Popular:

- Departments
- Products
- Days/Hours

Profile

Customer's
Demographics

Categorize

- Loyalty
- Frequency
- Price range
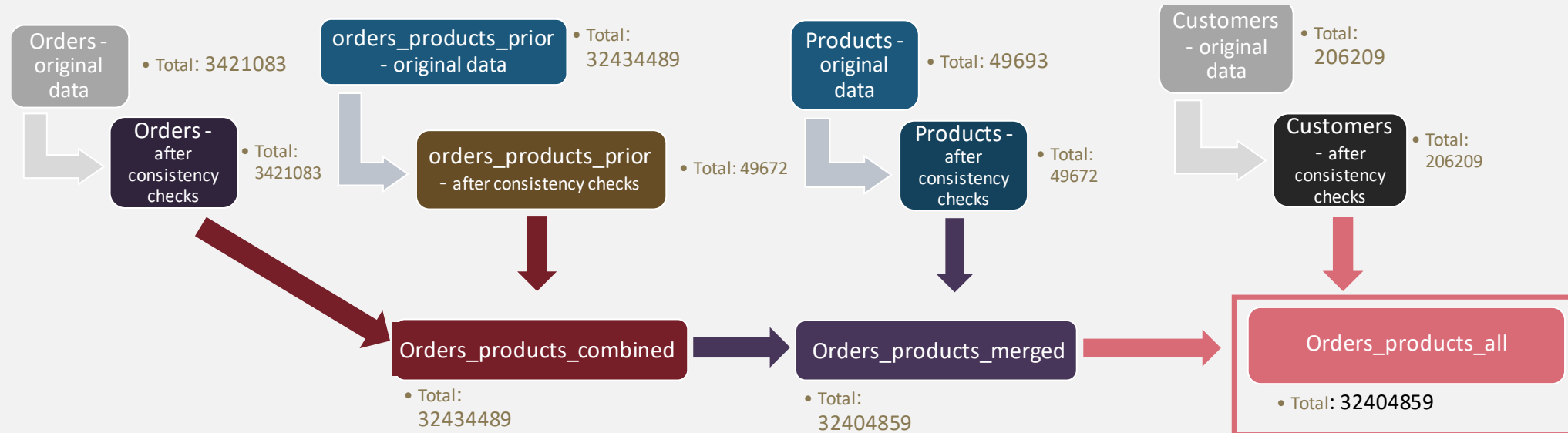- Income
- Regions
- Marital Status
- Age Groups

Visualizations

Show the findings
to Marketing and
Sales team

The Data: Population Flow

- ✓ Clean orders and products datasets;
- ✓ Transponsder departments dataset;
- ✓ Clean customers dataset;
- ✓ Merge;
- ✓ Labeling, creating new columns, data aggregation;
- ✓ All done to obtain insights and business solutions.

instacart   5

The Challenges

Several appeared during the project

PC Crash

Due to the huge amount of data my computer crashed.

I decided to buy a new ssd card and extended memory.

Stackflow

I wanted to explore more information from the dataset.

I created new labels, categories and different forms of visualizing data.

Stackflow was the key to figure out more codes in Python.

Tic toc!!

Due to the problems and my curiosity/interest on data, I spent a lot of time on this project.

It was in total **45 days.** A lot longer than intended.

instacart

THE FINDINGS

Visualizations and Recommendations

↑ **AVG Ticket**

Saturday and Sunday
3:00 – 5:00
20:00 – 21:00

Busiest Hour & Day

Saturday and Sunday
9:00 – 10:00
15:00 – 16:00

Popular Departments

1. Produce
2. Dairy and eggs
3. Snacks
4. Beverages
5. Frozen

Nº of orders per Region

1. South   **31%***
2. West   **28%**
3. Midwest   **24%**
4. Northeast   **18%**

The Habits

Customer's purchase patterns

* % of orders

instacart

The Profile

Many hours in Stackflow for the rewarding results below!

Loyalty

All loyal customers are frequent customers

- Have > 40 orders
- Retun in < 10 days

- Only married people purchase high range priced products

Marital Status

- Living w. parents	5%*
- Single	16%
- Married	70%
- Div/Widowed	9%

- Living w. parents < 21 y/o
- Div/widowed > 51 y/o

Price Range

- Only Meat & Seafood and Pantry Departments have high range produts (> $15)

- 67% of the orders in all regions are Mid-range priced (> $5 and < $10)

- **Non-frequent and low income** customers don't order high range priced products.

Income

- High > 100K	46%*
- Mid > 50K and < 100K	43%
- Low < 50K	11%

- High inc. is majority > 41 y/o
- Mid inc. is majority < 41 y/o

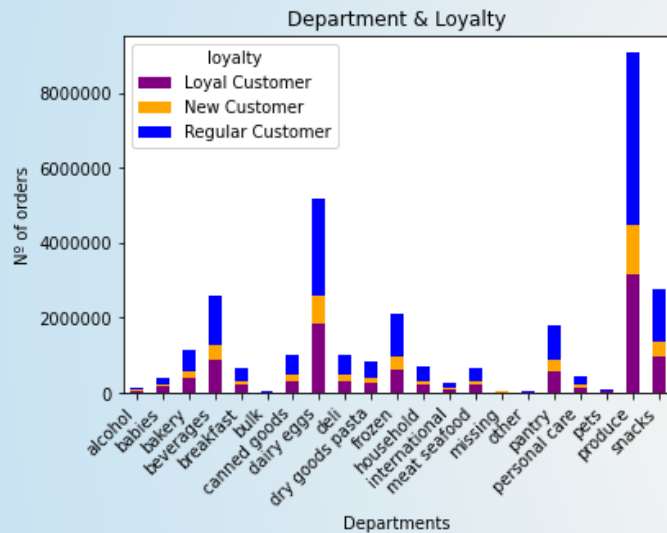## Age Groups

|	
> 60	32%*
51 - 60	16%
41 - 50	16%
31 - 40	16%
18 – 30	20%

* % of orders

instacart

9

Loyalty

- In **Babies and Bulky** departments, loyal customers are responsible for 41% of all the orders, 6% more than what they order in other departments.
- Send promotions for regular customers, to increase the number of orders and frequency.

Marketing

- Email marketing could be sent in the afternoon, targeting clients for the next day.
- Most of the clients purchase in the morning and lunch time.
- Focusing also in promotions for the weekend, the busiest days.

Focus
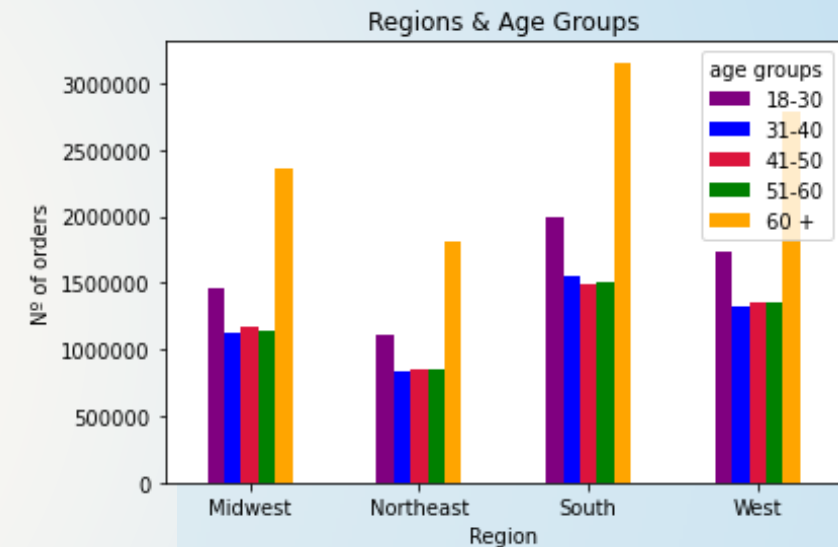
- Regions South and West
- Age > 60+ y/o
- Popular Departments
- Married people
- High range priced products in Seafood &   Meat and Pantry

The overview

On this Project, caught my attention:

Priorities
Management

I tend to go deeper in the Dataset, setting aside the priorities.
Focus was key to finish the Project.

Data
Huge and Complex

Cleaning and Analysing was a great exercise.

I wish I could access real customer's dataset.

Time
Management

I wish I had more time to explore more and discover more codes in Python.

THANK YOU

Mary Sassaqui 👤

maryasti@gmail.com ✉

LinkedIn 🔗

Github 🔗

Tableau 🔗