COMMERCE INSIGHTS

# Background

Decision makers in a Brazil e-commerce need a dashboard that provides a concise and comprehensive overview of KPIs and key metrics. This dashboard describes general company performance for C-levels such as CEO, CMO, and CFO, as well as more detailed insights such as sales performance, seller performance for the Sales department, promo campaign performance and customer profiles for the Marketing team, as well as the results of AB testing experiments of the old and new landing page for the Business Development (Bisdev) team.

# Objectives

## Vision

Providing insightful data visualizations for everyone to get ahead with what has been done and what they can do in the future.

## Goals

* 1. Sales: to get a glimpse of overall sales performance such as total monetary, product category contribution, sellers performance, and customer’s sentiment in their review.
  2. Marketing: promo campaign effectiveness (acceptance rate), customer segmentation by age group, income group, educational level, and marital status. Also an insight on RFM segmentation by each monetary size.
  3. Bisdev: to provide them the visualization of their experiment’s performance and to conclude if the experiment significantly gives a better result.

## Initiative

To provide each team with product data as a dashboard which will provide all of their work performance in a glimpse.

## Features:

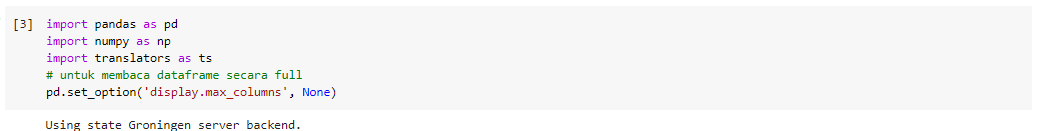
* 1. C-levels dashboard with prime key metrics from all departments; sales, marketing.
  2. Sales: in one glimpse, the dashboard is divided into 3 parts for product sales performance, seller performance and contribution by state, and the third is customer review and the sentiment they give.
  3. Marketing: acceptance rate for each promo campaign, customer segmentation proportion to give a slight information of who is the majority of the customer, and also RFM (Recency, Frequency, Monetary) segmentation to give the information of customer profiling and what to do with them in the future.
  4. Bisdev: to give information on which group, whether control or treatment group, whose conversion is significantly better than the other, to give a slight info on daily conversion and if the p-value is smaller than 0.05, with accumulative power daily to see if the sampling is strong enough to represent the population.

# Overview and Data Processing

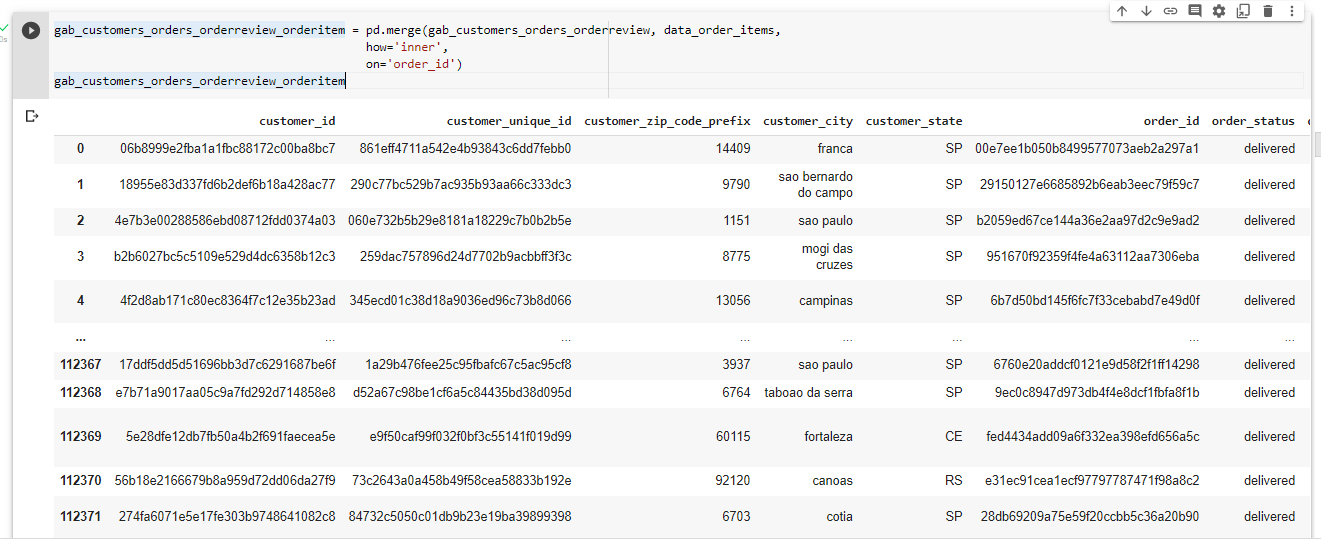
## Data Wrangling

### Pre-processing

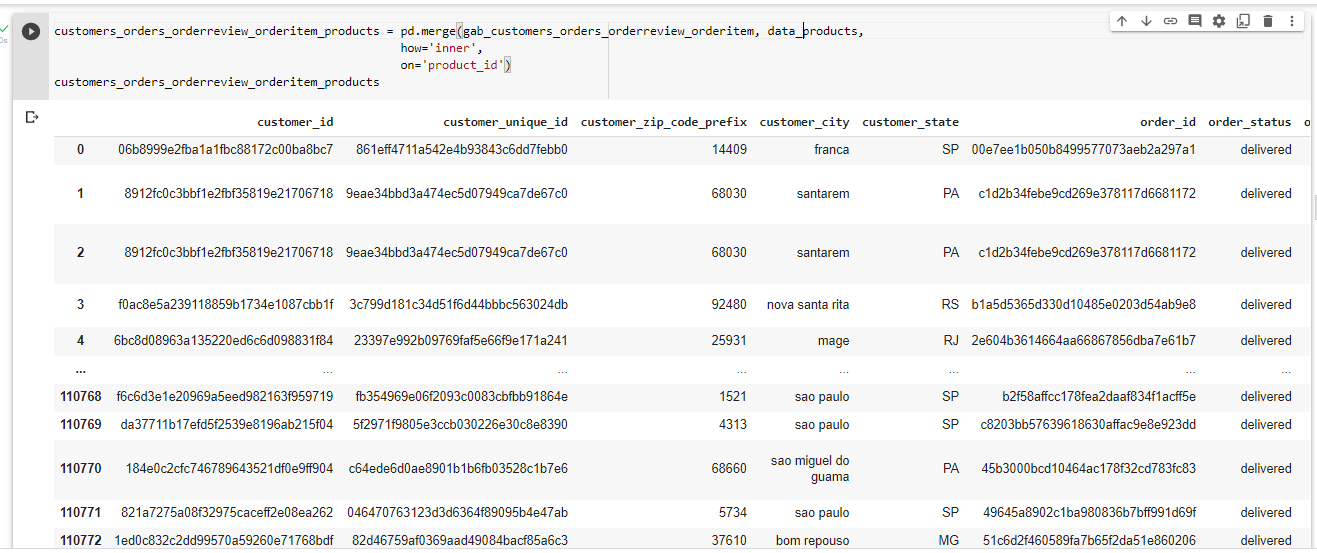
We compile several datasets from orders, seller, customer review, and product, all in Jupyter Notebook using Python. The process flow captured in the following:



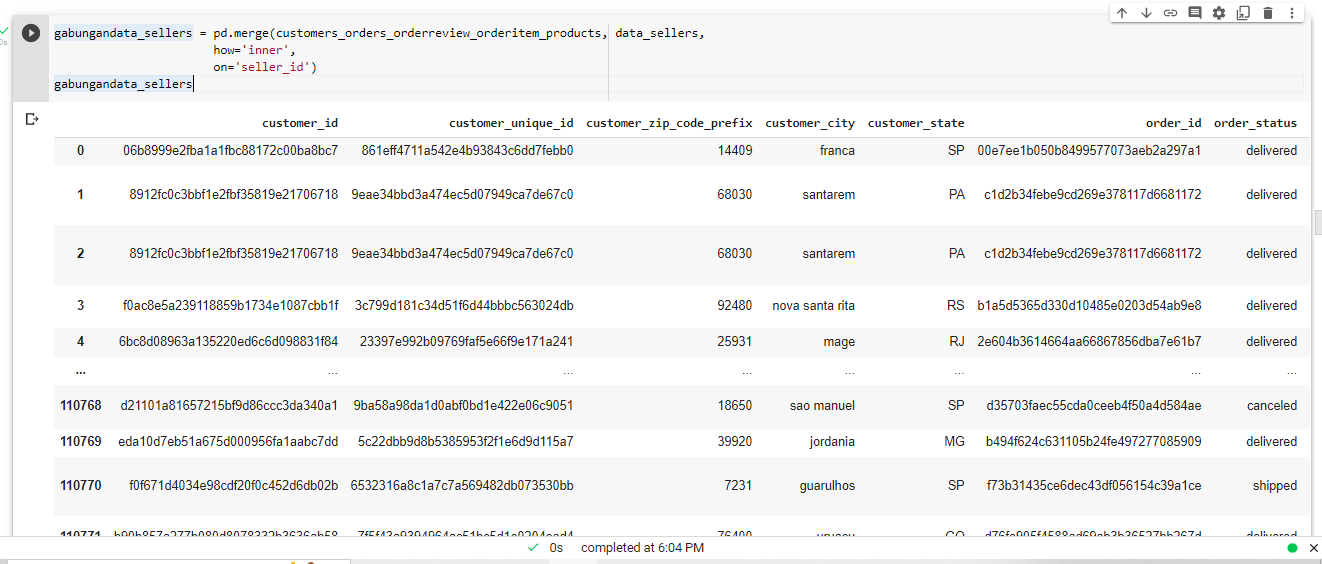
(import libraries)



(joining all the clean data of orders)



(merging with data products)

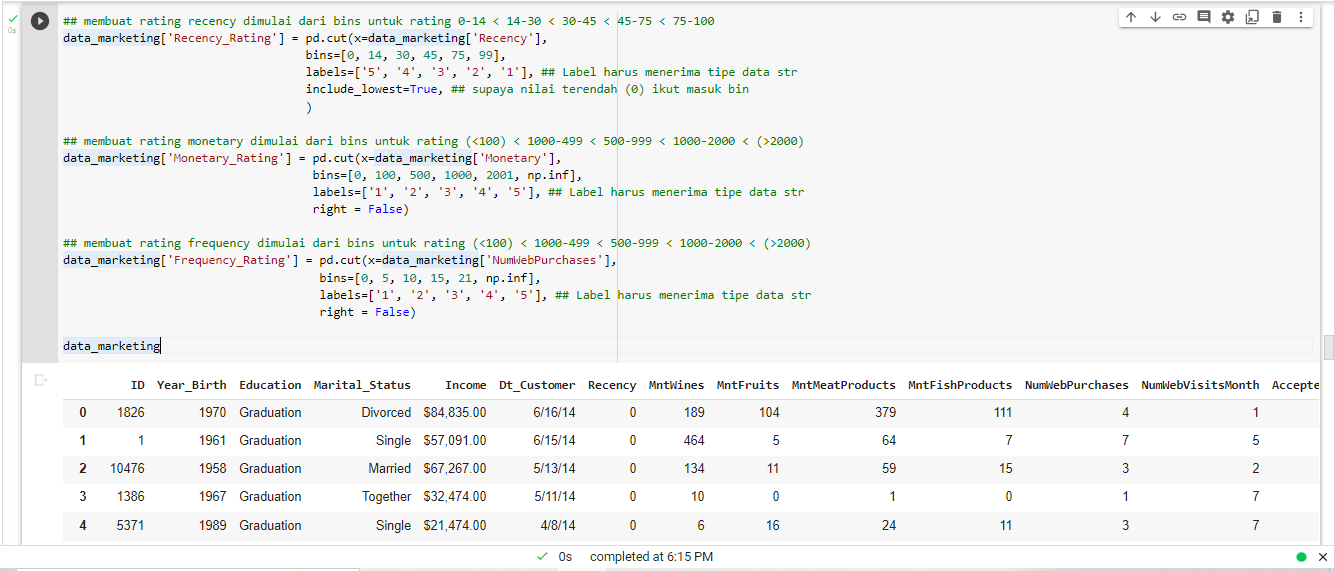


(joining with sellers data)



(getting datetime data)

### Marketing dataset

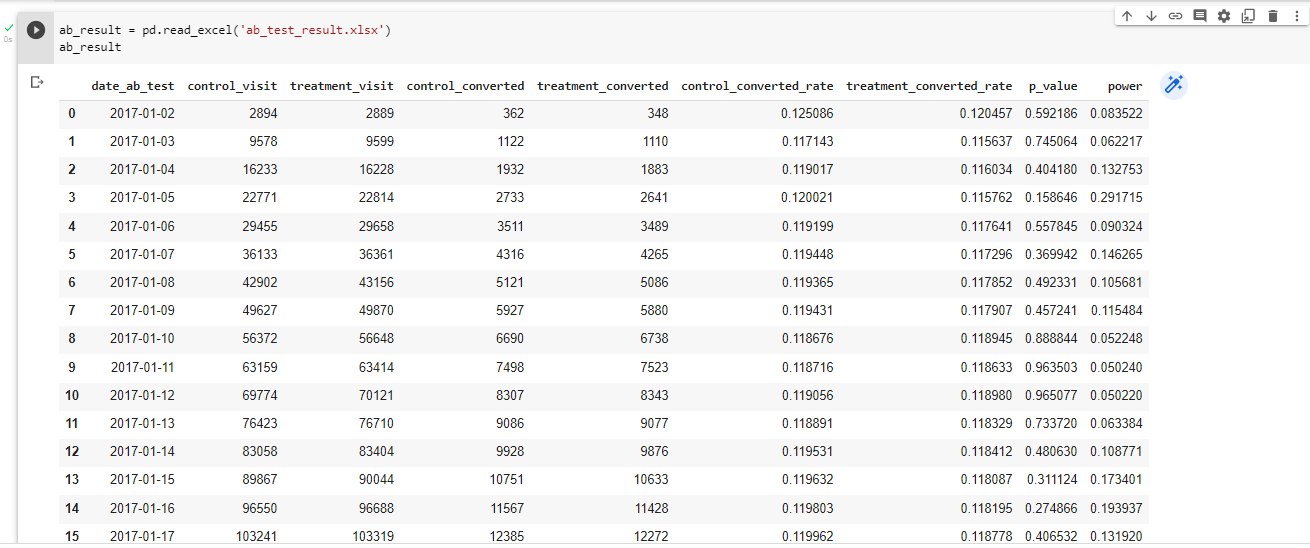


(creating RFM segmentation)



(joining sales & marketing data)

### AB Testing dataset



(joining all values)

## Findings on the Data

* 1. Sales data: Even though we find that more than 58% of **comment reviews** are missing values (n/a), we need to know if the product review gives a positive or negative sentiment to what they sell. We then translate the comment review from Portuguese to English.

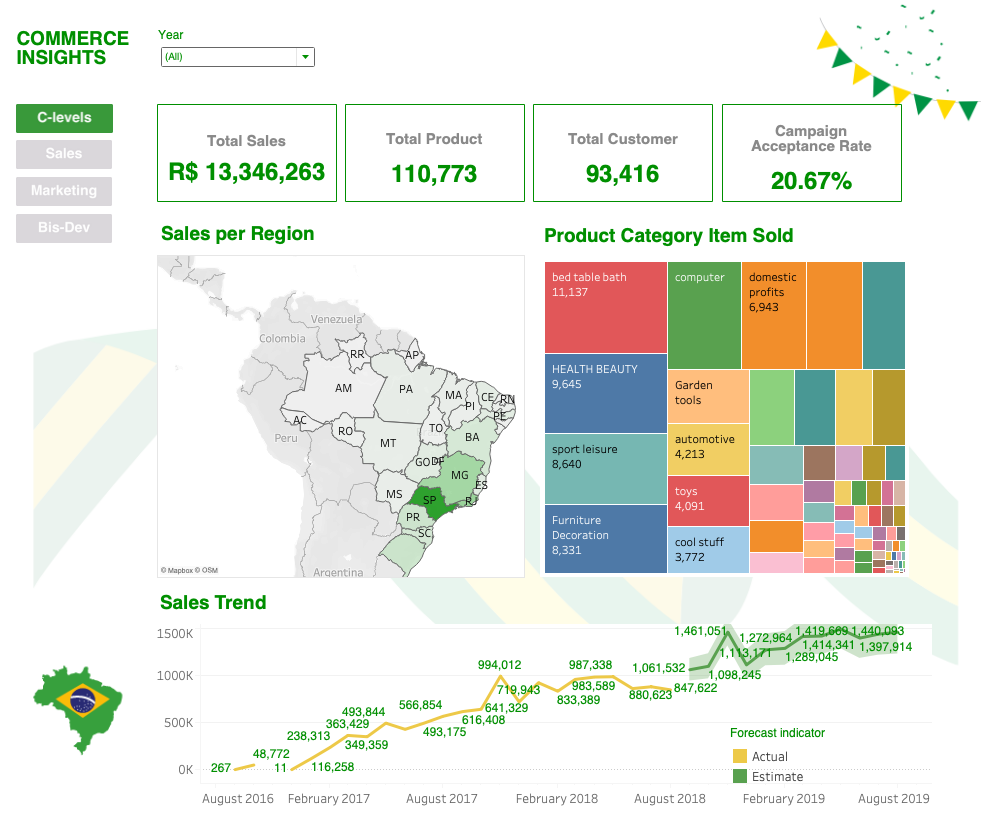
On **sales forecast**, we use Tableau Analytics forecast feature and use weekly frequency on sales monetary so that the Tableau can get the seasonality within the timeframe. When we experiment with Monthly frequency (for simplicity reason), Tableau did not get the seasonality and it shows only the median value of the data and the forecast shows one line which was the median value only for several months after the last sales value.

* 1. Marketing dataset:
  2. AB Testing dataset:

# Minimum Viable Product (MVP) and Analysis.

With the datasets all set, here are the MVP result:

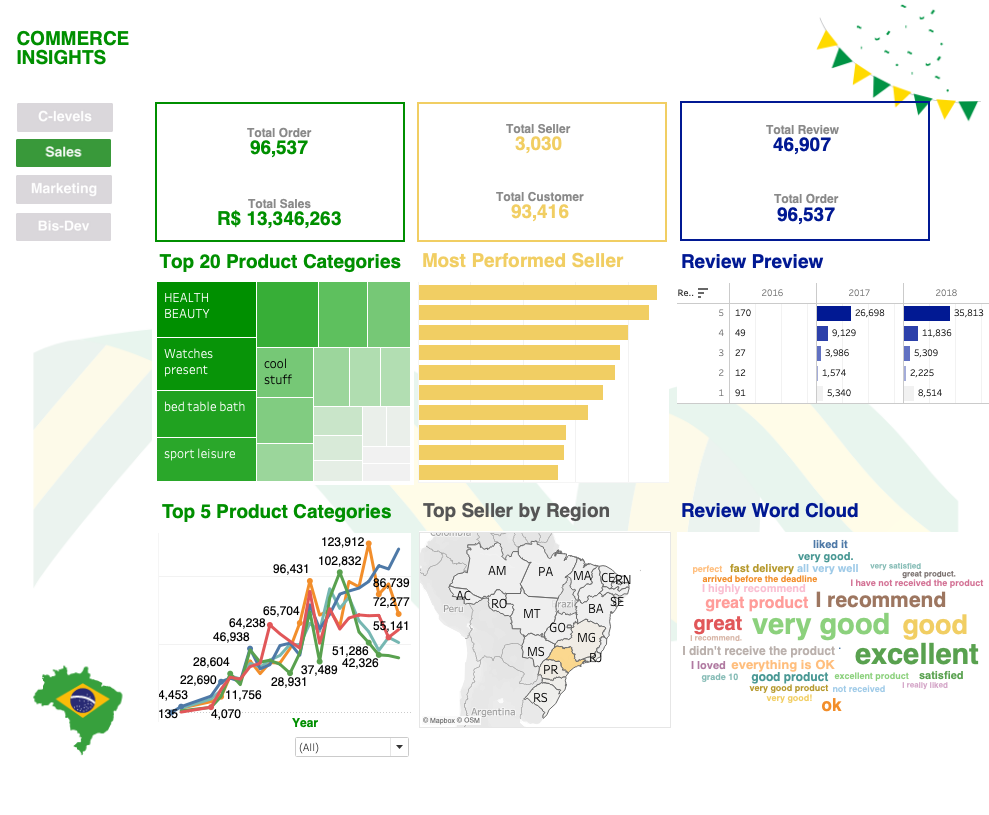
## C-levels



### C-levels analysis

Based on the data shown, CEO, CMO, and CFO can see the total sales for each year, total product, total customer, and how their campaign performs overall. They can see where their business is thriving, it’s in Sao Paulo, and the second one is Rio de Janeiro. They can also see how many items are sold based on each category from the heatmap to the right. This can give them insights into which category to get more support in terms of promo campaign, or seller support. To the bottom of the dashboard, they can see the sales trend with its forecast to put into consideration on their next KPI target of sales monetary.

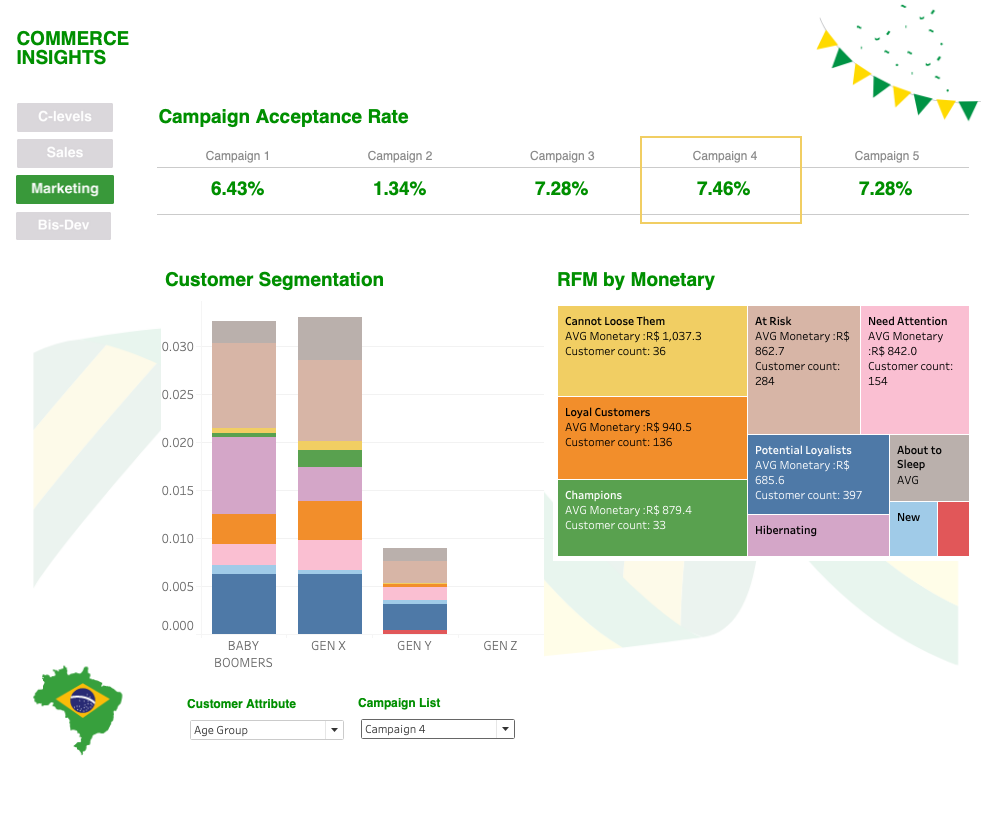
## Sales



### Sales analysis

The dashboard is divided into three parts. The first part (green) is showing the Sales performance and which product category contributes to Sales (heatmap) and how the sales trend behaves for the Top 5 sales contributors. With trendline as the visual (per quarter), the Sales team can quickly decide if they want to continue their current strategy, or pivoting into pushing the sales of other categories.

## Marketing



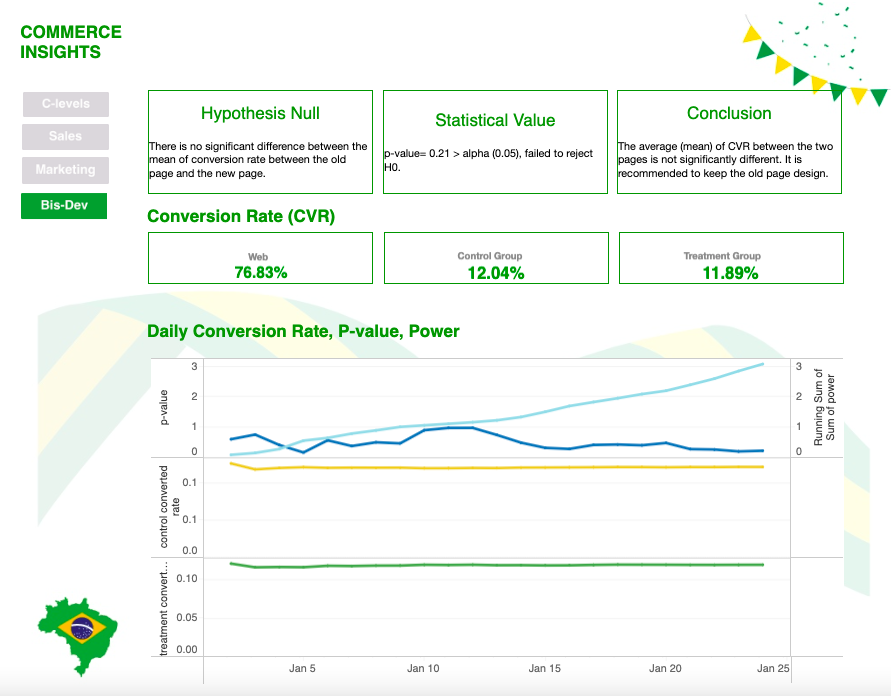
### Marketing analysis

On marketing, Campaign Acceptance Rate, the team can see which promo campaign with the best result, also which one with the lowest acceptance rate. We need more context on the promo visual design, the copywriting, the product category support, and also if the monetary can be compared with one another.

On Customer Segmentation, the team can see who are their most valuable customers based on their age group, income group, marital statuses, and educational levels. For example, when we choose the income group, all of the customers who get the promo came from the High Income group. Maybe they can try to push the promo campaign to the lower income group even though the monetary might be not high. The purpose is to see if we can maintain the customer retency for all income group.

On RFM by Monetary, compared the customer count to the left (Segments), we can see that the majority of our customers by age group is in At Risk RFM category. They might spend a little amount of money, but we shall see if they are actually the most loyal customer. We can initiate a strategy for them so they engage with us more. The Cannot Lose Them RFM category are the high spenders. With the small proportion on the Segment barchart, we should find a strategy to maintain this group of customers so we can see them purchase more items in the future.

## AB Testing



### AB Testing analysis

The Hypothesis Null says: There is no significant difference statistically between the conversion rate on the old page and the new page. Since this is a two-tailed ab-testing, we write them as: H0 = H1. The data sampling uses xx.xxx users compared to all users xxx.xxx. And compared to the web Acceptance Rate (76.83%), the average Control Group Acceptance Rate is 12.04%, and the Treatment Group gets 11.89%. The way we see it, it looks like the Treatment Group failed to drive customers’ conversion compared to the Treatment Group with 0.15% difference. But statistically, the p-value result is 0.21 which is higher than the accepted 0.05, so we failed to reject H0.

—-----------------

Group B - PLBI 2

Andina Septia

Rochyati

Important Links:

Dashboard: ​​<https://public.tableau.com/app/profile/dinnah/viz/CommerceInsightsB/Bisdev#1>

Github: <https://github.com/dinnah88/PLBI_2_B.git>