



# Sales Data Analysis for e-commerce

---

Anastasios Vitsios  
30.3.2021

# Table of Contents

---

Executive Summary .....	3
Introduction .....	4
Methodology .....	5
Initial View – Sales for the 2 first years .....	6
Sales in Top 6 States and Top 15 Products Categories .....	7
Sales between Year_1 and Year_2 - Spike in November 2017 .....	8
Spike in November 2017 – Sales vs MRKT qualified leads .....	9
Decline in Sales after May 2018 .....	10
Sellers for period May - July 2018 .....	11
Top 5 Product_categories - Growth in 2018 .....	12
Decline in Product Categories (top 15 product categories) .....	13
Watches_gifts - Increase and drop in 2018 .....	14
MARKETING - Origins for MRKT Qualified Leads per Month (data for period June 2017 - May 2018) .....	15
Population in States vs Sales .....	16
Conclusion .....	17
Suggestions .....	18
Appendix .....	19

# Executive Summary

---

In this analysis we will examine the sales and marketing performance of the company Olist.

Olist has a web platform from where any Brazilian company can sell their products in Brazil. They started their activity in September 2017 and the data that we possess are for 24 months (until August 2018).

In this analysis we will see that there was an overall increase in sales in these 2 year with some spikes and some drops. The big spike was in November 2017 and it was not related to the marketing of Olist. After this month there was an increase in marketing qualified leads which affected the sales positively.

After May 2018 there was a decline in sales in almost all States in Brazil and in the majority of product categories. Even though the sellers (companies which sell products through Olist) were increased after May 2018, the increase in sales was not possible. We don't have marketing data for the months after May 2018 to compare but it is true that in 2018 there was a 9% decline in GDP per capita which may affect the sales performance of Olist.

We will also realize that there are many States in Brazil where the sales are too low compared to the population, so we can expect increase in sales in those States with the proper strategic plan.

# Introduction

---

**Olist** is the largest department store in Brazilian marketplaces. Olist has a web platform where small businesses from all over Brazil are able to sell their products through the Olist Store and ship them directly to the customers using Olist logistics partners.

The data that we have is from kaggle.com. They are for 24 months, from September 2016 to August 2018. The first 12 months consider as the Year 1 (Sep 2016 - Aug 2017) and the rest of months as the Year 2 (Sep 2017 to Aug 2018)

**In total, in those 24 months Olist had:**

- 99,441 orders
- 112,649 items sold, from
- 71 product categories, in
- 29 States, from
- 2383 sellers and
- 13,591,498 BR total sales

The analysis was made in files olist\_all.xlsx, olist\_groups1.xlsx, mrkt\_leads.xlsx, .IPYNB

# Methodology

---

I used 7 files from Kaggle.com which are: olist\_custome\_dataset.csv, olist\_order\_items\_dataset.csv, olist\_orders\_dataset.csv, olist\_products\_dataset.csv, olist\_sellers\_dataset.csv, product\_category\_name\_translation.csv

For the purposes of this analysis I downloaded 3 more tables from the web with the state population, state names/abbreviations and GDP per capita.

The analysis was made:  
in the excel files olist\_all.xlsx, olist\_groups1.xlsx, mrkt\_leads.xlsx,  
in python file Olist\_groupby.IPYNB and  
in Power BI file olist.pbix

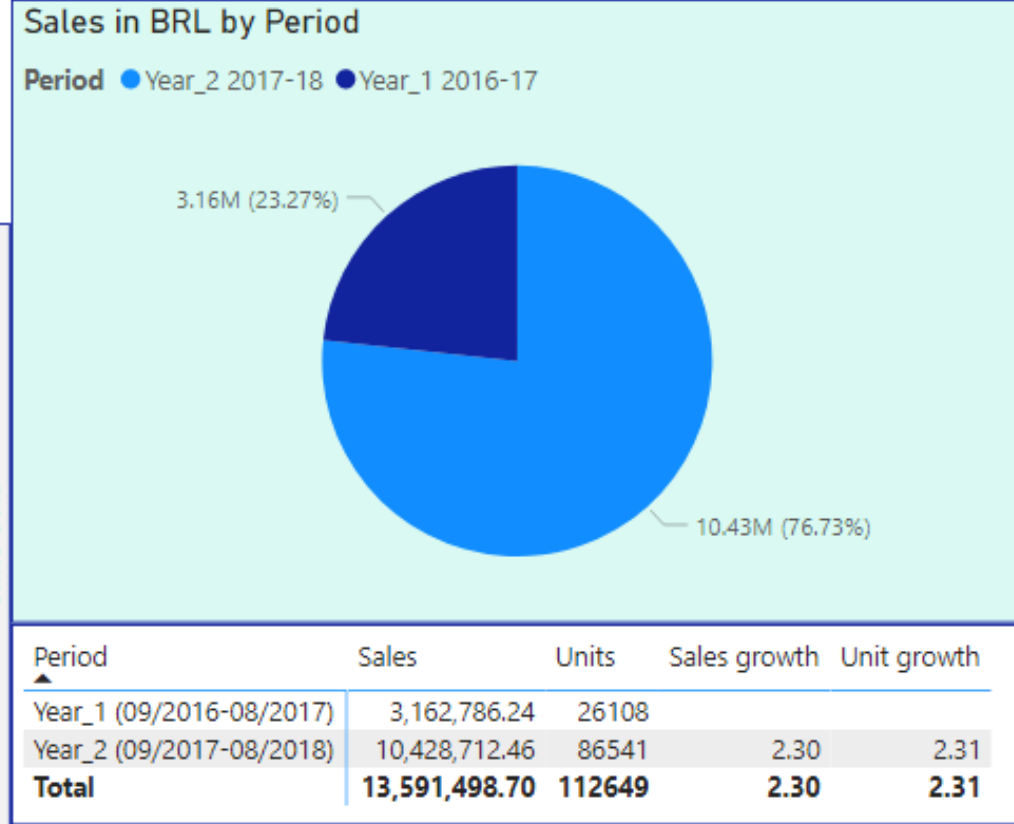
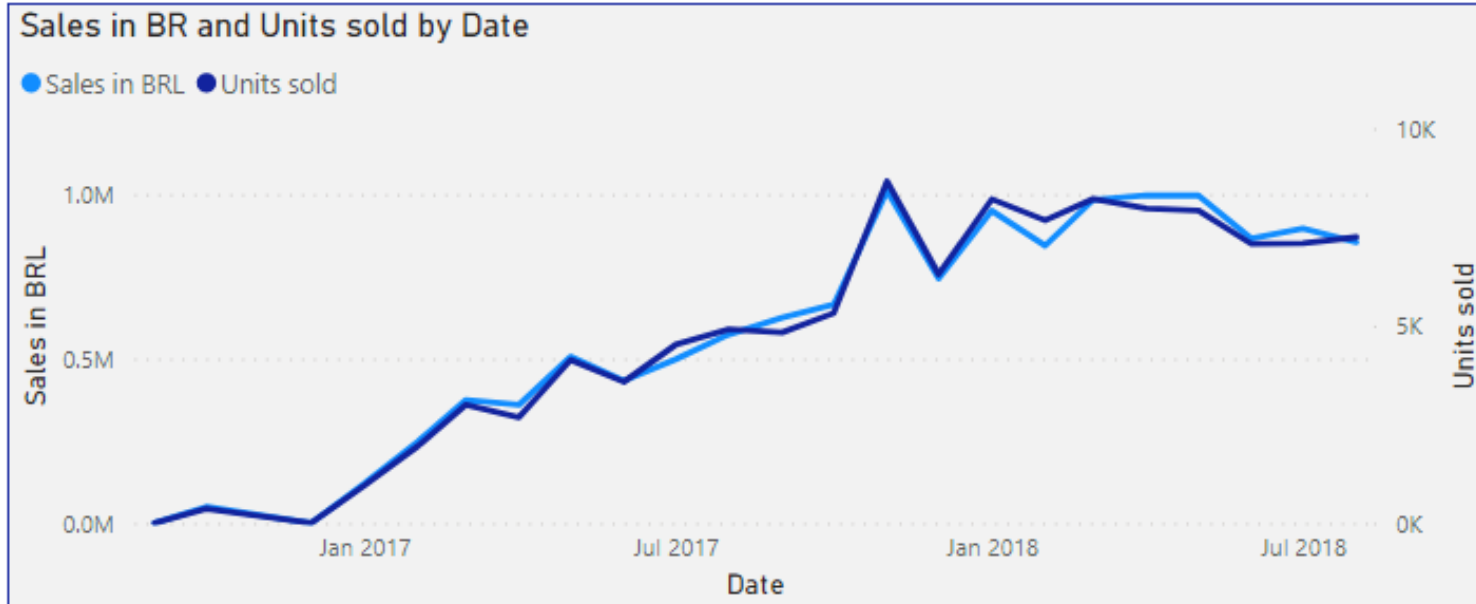
I created tables in excel by joining the tables from Kaggle.com, with the appropriate format in order to be analysed in Power BI.

I grouped by year, month, product category, seller some tables with python in order to aggregate the sales for further analysis.

The first month for each period is September. So, the first year (from now on Year\_1) starts in September 2016 and ends in August 2017. The second year (from now on Year\_2) starts in September 2017 and ends in August 2018.

# Initial View – Sales for the 2 first years

The total revenue from  
Sep 2016 to Aug 2018 in Brazilian real (BRL)  **13.59M**  
Sales in BRL



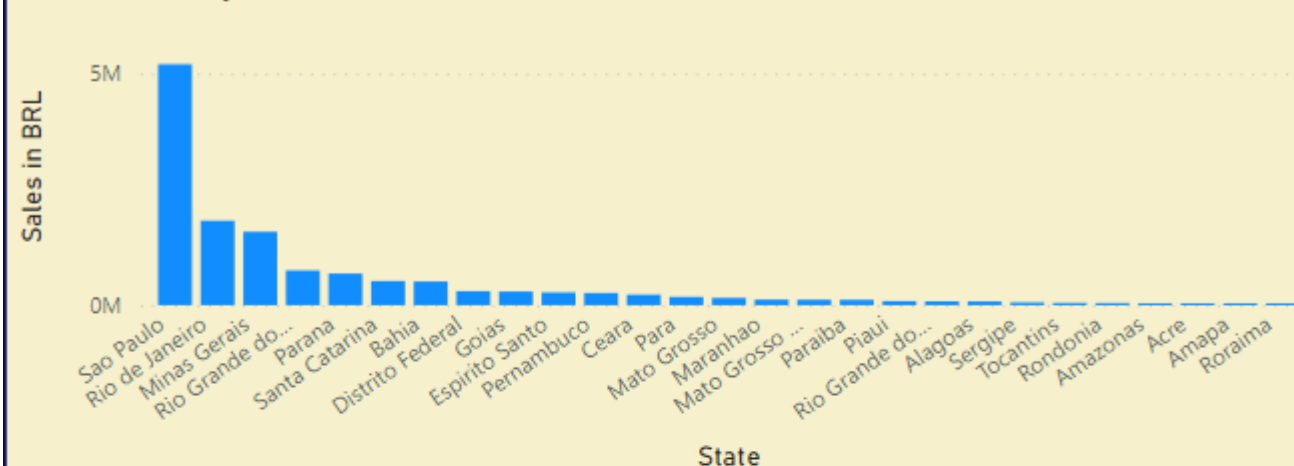
Olist started to sell on September 1st 2016. The data that we have is until August 31st 2018. I split the 24 months sales period in 2 annual periods. The 1st is from 9/1/2016 to 8/31/2017 (named "Year\_1") and the 2nd is from The 1st is from 9/1/2017 to 8/31/2018 (named "Year\_2") .

In Year\_2 there is an increase or 230% in revenue and 231% in units and as we see in the graph "Sales in BR and Units sold" the growth between revenue and units is identically.

- The 76.73% of total sales became in Year\_2.
- There is a spike in November 2017 and a decline in sales after May 2018 which we will examine in the following slides

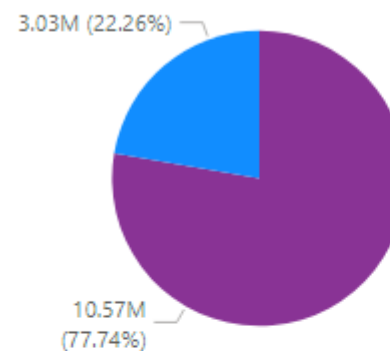
# Sales in Top 6 States and Top 15 Products Categories

Sales in BRL by State



Sales by State\_groups 27 States

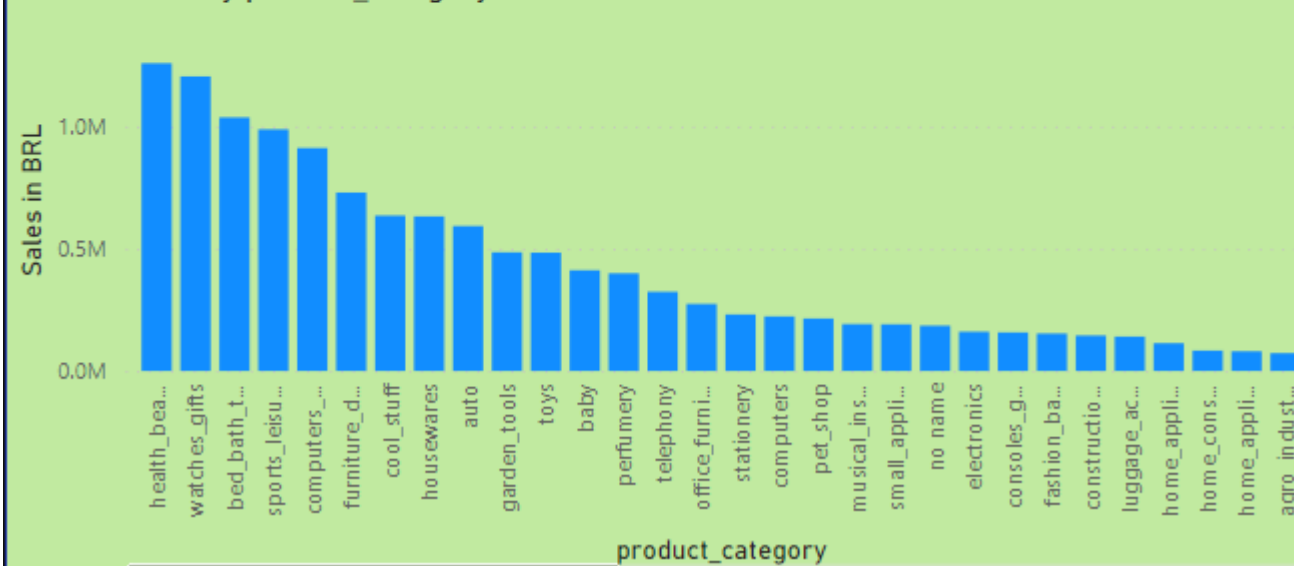
State\_gr... ● Top 6 States ● Other 21 States



Top 6 States - Sales

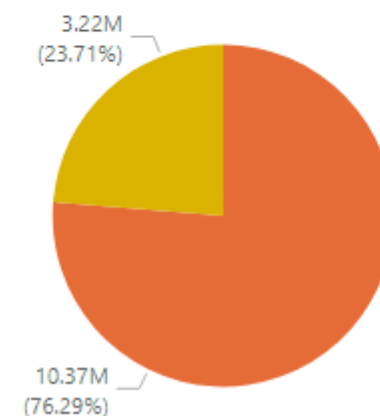
State	Sales in BRL
Sao Paulo	5,202,810.05
Rio de Janeiro	1,824,092.67
Minas Gerais	1,585,308.03
Rio Grande do Sul	750,304.02
Parana	683,083.76
Santa Catarina	520,553.34
<b>Total</b>	<b>10,566,151.87</b>

Sales in BRL by product\_category



Sales by product\_groups 71 Categories

product... ● Top 15 Categories ● Other 56 Cat...



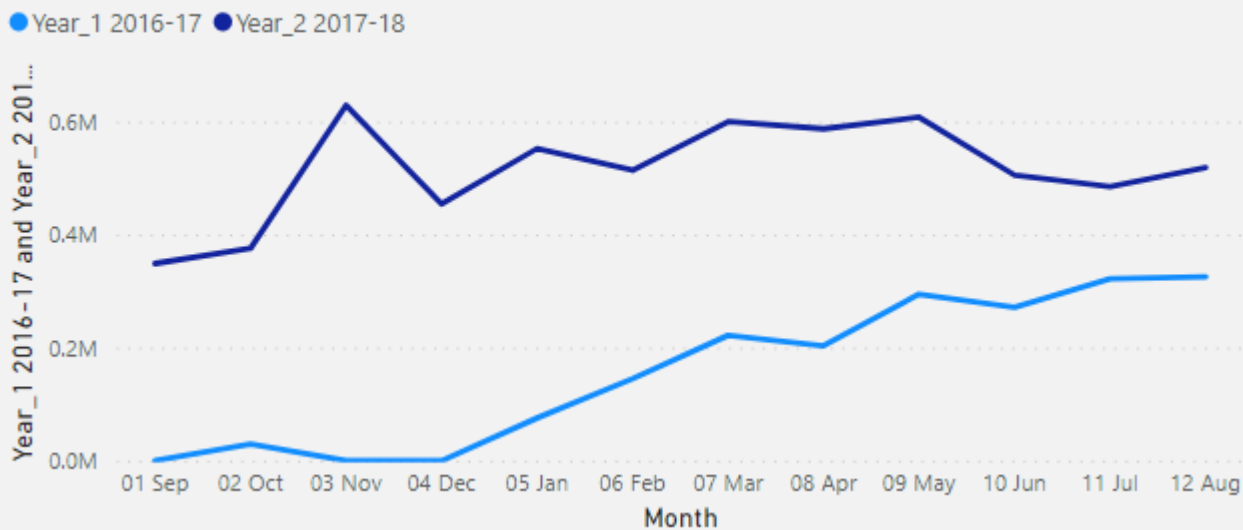
Top 15 Product Categories - Sales

product_category	Sales in BRL
health_beauty	1,258,681.34
watches_gifts	1,205,005.68
bed_bath_table	1,036,988.68
sports_leisure	988,048.97
computers_accessories	911,954.32
furniture_decor	729,762.49
cool_stuff	635,290.85
housewares	632,248.66
auto	592,720.11
garden_tools	485,256.46
toys	483,946.60
baby	411,764.89
perfumery	399,124.87
telephony	323,667.53
office_furniture	273,960.70
<b>Total</b>	<b>10,368,422.15</b>

The sales became in 27 States in Brazil and in 71 different product categories. Here we will analyse the sales performance in the **top 6 States** and in the **top 15 product categories** where Olist has **77.74%** and **76.29%** of total sales respectively.

# Sales between Year\_1 and Year\_2 - Spike in November 2017

Year\_1 2016-17 and Year\_2 2017-18 by Month

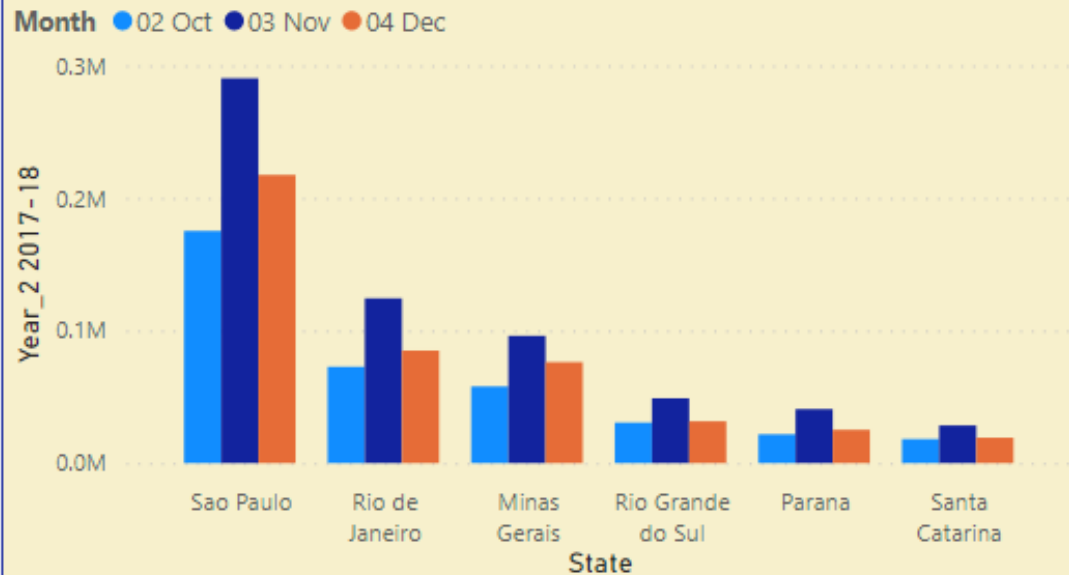


In graph "Year\_1 2016-17 and Year\_2 2017-18 by Month" we see the sales in the annual periods.

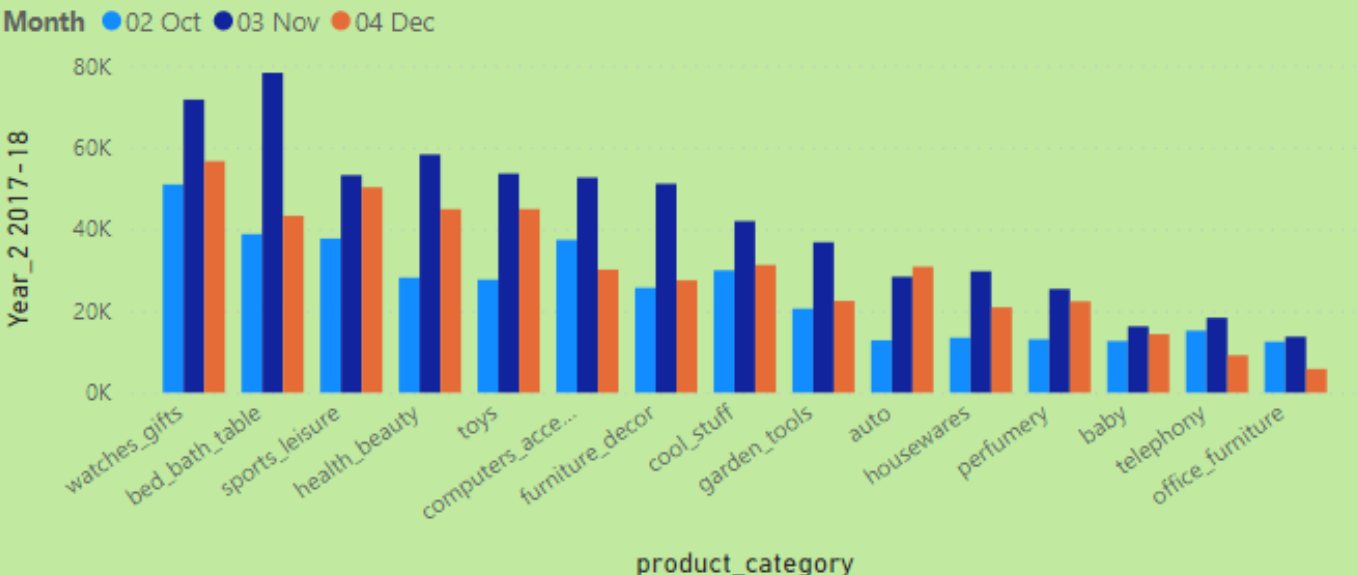
We notice a spike in November 2017 when there is not something similar in November 2016. In graphs "Year\_2 2017-18 by State and Month" and "Year\_2 2017-18 by product\_category and Month" we see that there is an increase in sales from October to November 2017 in all States and in all Product categories. So the spike is not related to some specific States or products.

Also, we notice a decline after May 2018, which we will examine later

Year\_2 2017-18 by State and Month



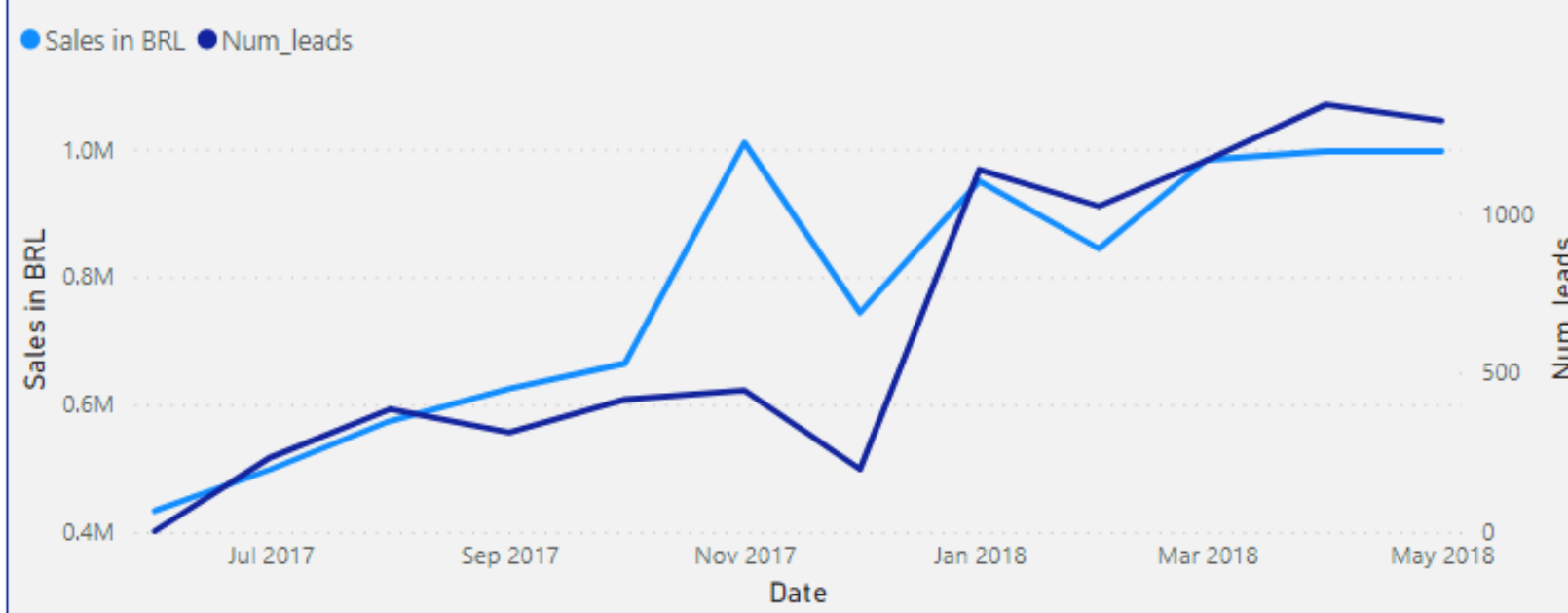
Year\_2 2017-18 by product\_category and Month





# Spike in November 2017 – Sales vs MRKT qualified leads

Sales & Marketing leads for Jun 2017 - May 2018



Sales & Marketing leads for Jun 2017 - May 2018

Year	Month	Day	Sales in BRL	Num_leads	Units sold
2017	June	1	433,038.60	3	3583
2017	July	1	498,031.48	234	4519
2017	August	1	573,971.68	385	4910
2017	September	1	624,401.69	312	4831
2017	October	1	664,219.43	415	5322
2017	November	1	1,010,271.37	444	8665
2017	December	1	743,914.17	197	6308
2018	January	1	950,030.36	1136	8208
2018	February	1	844,178.71	1021	7672
2018	March	1	983,213.44	1165	8217
2018	April	1	996,647.75	1339	7975
2018	May	1	996,517.68	1289	7925
Total			9,318,436.36	7940	78135

The data for marketing qualified leads are for only 12 months, from June to 2017 to May 2018

The increase from October 2017 to November 2017 is:

52% in revenue

62% in units sold

7% in marketing qualified leads

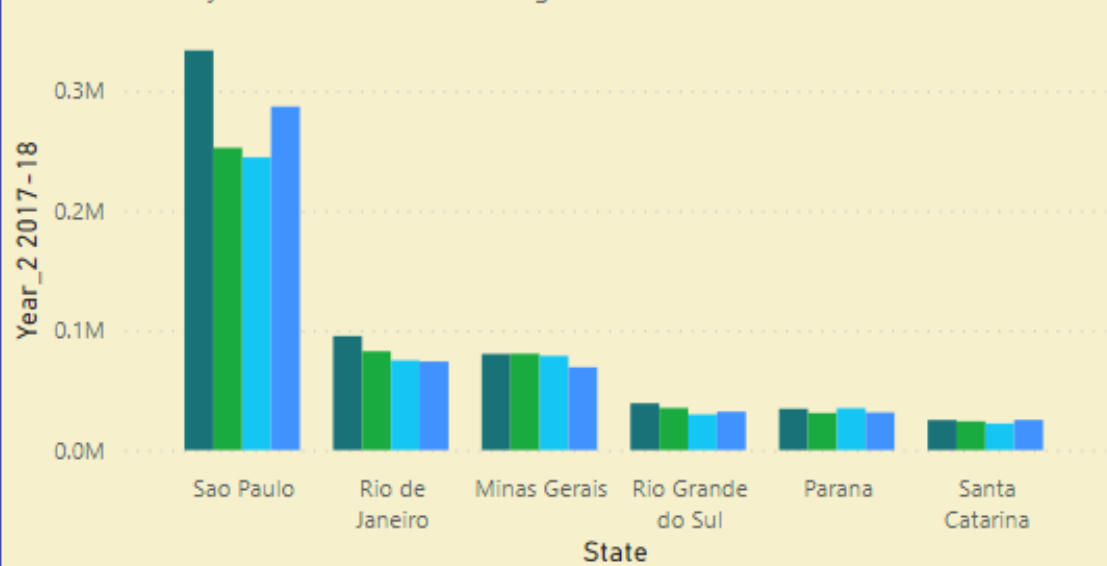
From this we can assume that either the same number of people bought more items on average or more qualified leads were converted to customers (or of course both). This could have happened because of a discount campaign (more units were sold than the revenue which was created) or another marketing campaign.

Another assumption is that the big event of Gran Prix which took place in Brazil in November 2017, could affect positively the growth in sales if it was combined with a marketing campaign.

# Decline in Sales after May 2018

Year\_2 2017-18 by State and Month

Month ● 09 May ● 10 Jun ● 11 Jul ● 12 Aug



Year\_2 2017-18 by product\_category and Month

Month ● 09 May ● 10 Jun ● 11 Jul ● 12 Aug



The decline in sales after May 2018 happened in almost all States in almost for all product categories as we see in the 2 graphs above. Here we can consider as the main reason the 9% drop in GDP per capita in 2018, which could be a reason for this drop (see table below "GDP per capita 2017 & 2018")

GDP per capita 2017 & 2018					
Year	Sum of GDP per capita BR \$	GDP Growth	Period	Sum of Sales	Sales growth
2017	56,076.25	0.14	Year_1 (09/2016-08/2017)	3,162,786.24	
2018	50,855.65	-0.09	Year_2 (09/2017-08/2018)	10,428,712.46	2.30
	106,931.90			13,591,498.70	

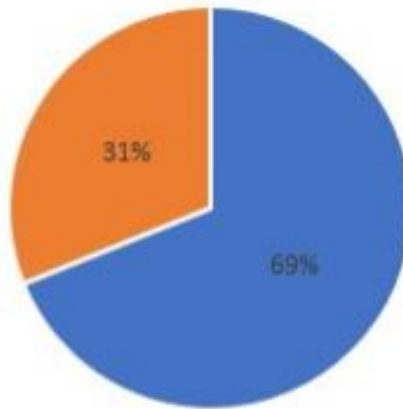
The GDP per capita in 2018 was 9% less than 2017 but in 2018 Olist had increase in sales of 230%. The high increase in sales is normal since is the 2nd year that the company is active but by the fact the GDP per capita is less, we understand that Olist increased their proportion in the Brazilian market versus their competitors.

# Sellers for period May - July 2018

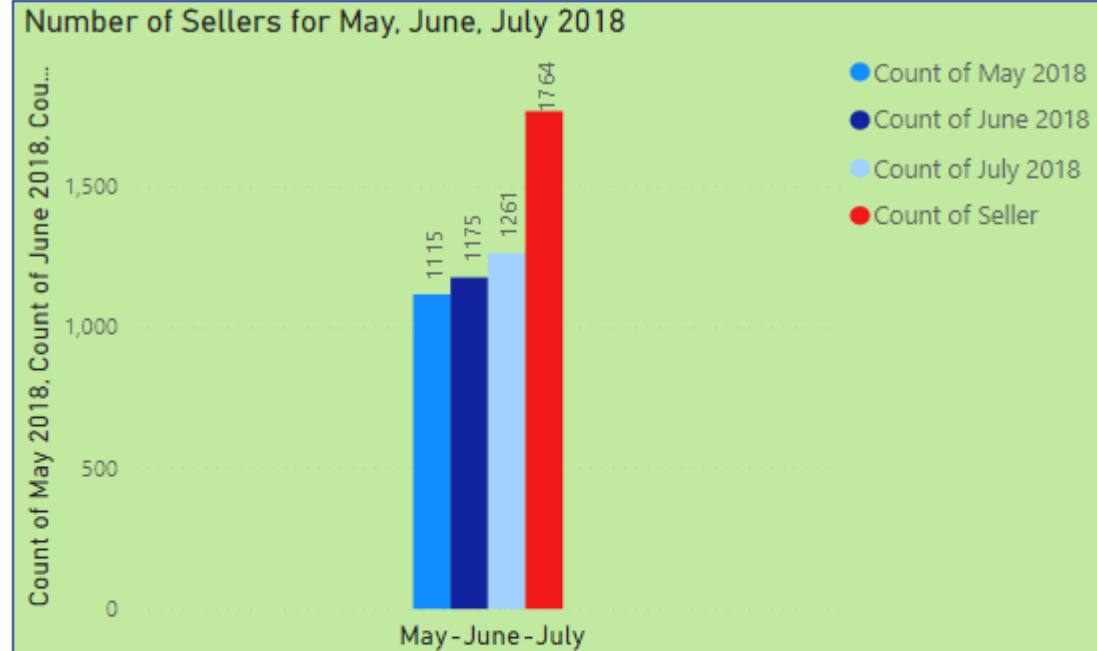
	Sellers	Proportion
Drop after May	768	68.9%
No drop after May	347	31.1%
Only in June & July	649	
Total Sellers	1764	

Sellers May to July 2018

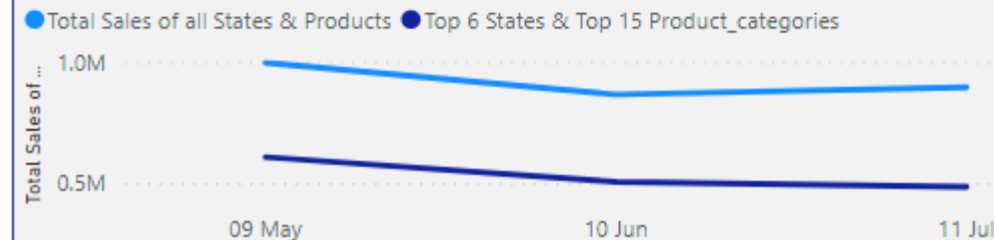
■ Drop after May ■ No drop after May



Graph from file olist\_all.xlsx sheet=Sellers-drop



Sales in Year\_2 for May to July 2018



We see that the 69% (768) of all 1097 Sellers (who sold in May 2018) had drop in sales from May 2018 to July 2018.

Although in June and in July 2018 the sellers increased 58% (649 more) from May 2018, there was a 10% decline in sales from May to July for all States and Product\_categories.

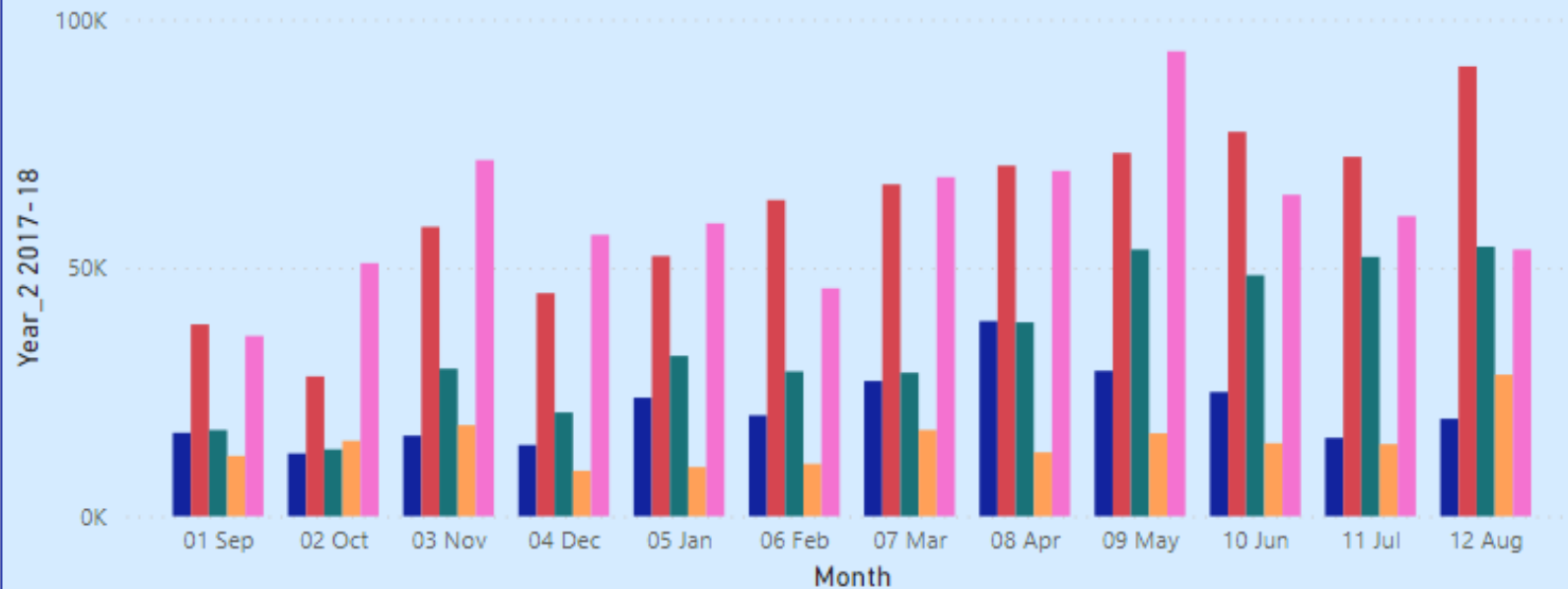
So, the drop is not related to some specific sellers.

However, in July 2018 there was a 3.5% raise in total, but there was a 4% decline in the top 6 States and in top 15 Product categories as we see in graph "Sales in Year\_2 for May to July 2018".

# Top 5 Product\_categories - Growth in 2018

Top 5 Product\_categories in Year\_2

product\_category ● baby ● health\_beauty ● housewares ● telephony ● watches\_gifts



Growth by Product\_category for Year\_2

product_category	Year_1 2016-17	Year_2 2017-18	Growth Sales %
watches_gifts	155,321.72	730368	370.23
baby	55,944.98	260083	364.89
health_beauty	174,684.07	736515	321.63
telephony	42,616.24	179607	321.45
housewares	109,226.01	419163	283.76
computers_accessories	162,874.17	553872	240.06
auto	100,962.49	334004	230.82
sports_leisure	181,765.56	598992	229.54
bed_bath_table	216,017.31	676963	213.38
toys	99,513.72	288145	189.55
furniture_decor	154,364.98	446407	189.19
garden_tools	104,926.16	263550	151.18
office_furniture	63,952.48	155803	143.62
perfumery	99,575.03	220387	121.33
cool_stuff	167,001.34	315585	88.97
<b>Total</b>	<b>1,888,746.26</b>	<b>6179444</b>	<b>227.17</b>

In the graph "Top 5 Product\_categories in Year\_2" we see the performance by month of each 5 product categories which had the highest growth in Year\_2.

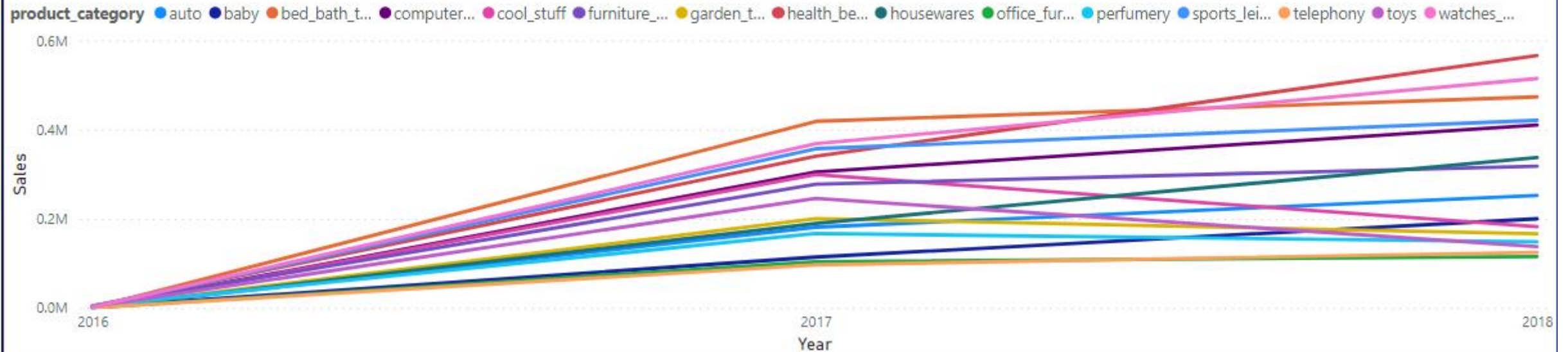
The products "watches\_gifts" had the highest growth in whole 2nd Year (370%).

Only the watches\_gifts and baby had a decline after May with the first to have the highest drop.

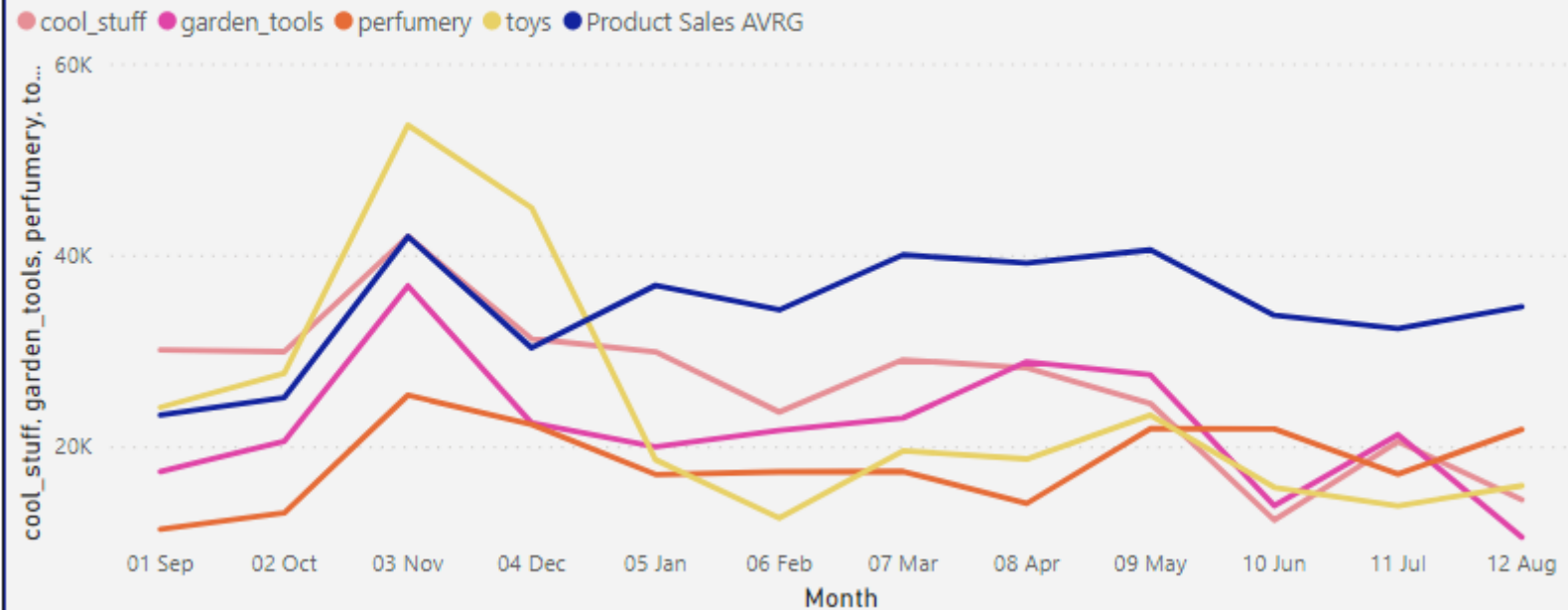
In the next page we see the sales growth of watches\_gifts by State, in order to examine deeper if the drop was related to some specific States

# Decline in Product Categories (top 15 product categories)

Sales by Year and product\_category



Product\_categories with biggest decline in Year\_2 & Product\_categories sales average



In graph "Sales by Year and product\_category" we see the performance of all top 15 product categories. The worst performance in 2018 belongs to product categories **cool\_stuff**, **graden\_tools**, **perfumery** and **toys**. But because this graph is too noisy, we see separately these 4 categories in the next graph.

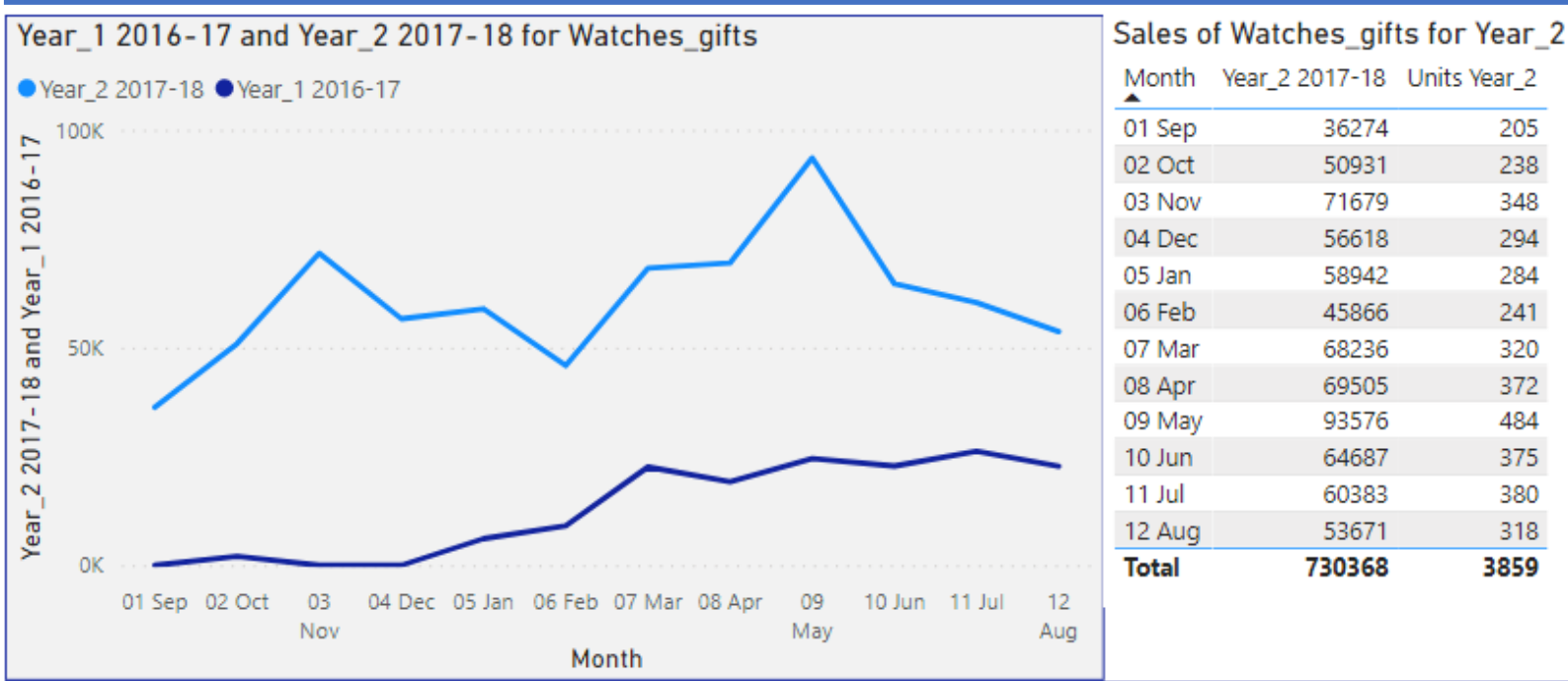
In graph "Product\_categories with biggest decline in Year\_2 & Product\_categories sales average" we see only these 4 product categories by month compared with the average sales of top 15 categories\*.

The sales of "toys" and "cool\_stuff" were above average until December 2017 but later they declined a lot

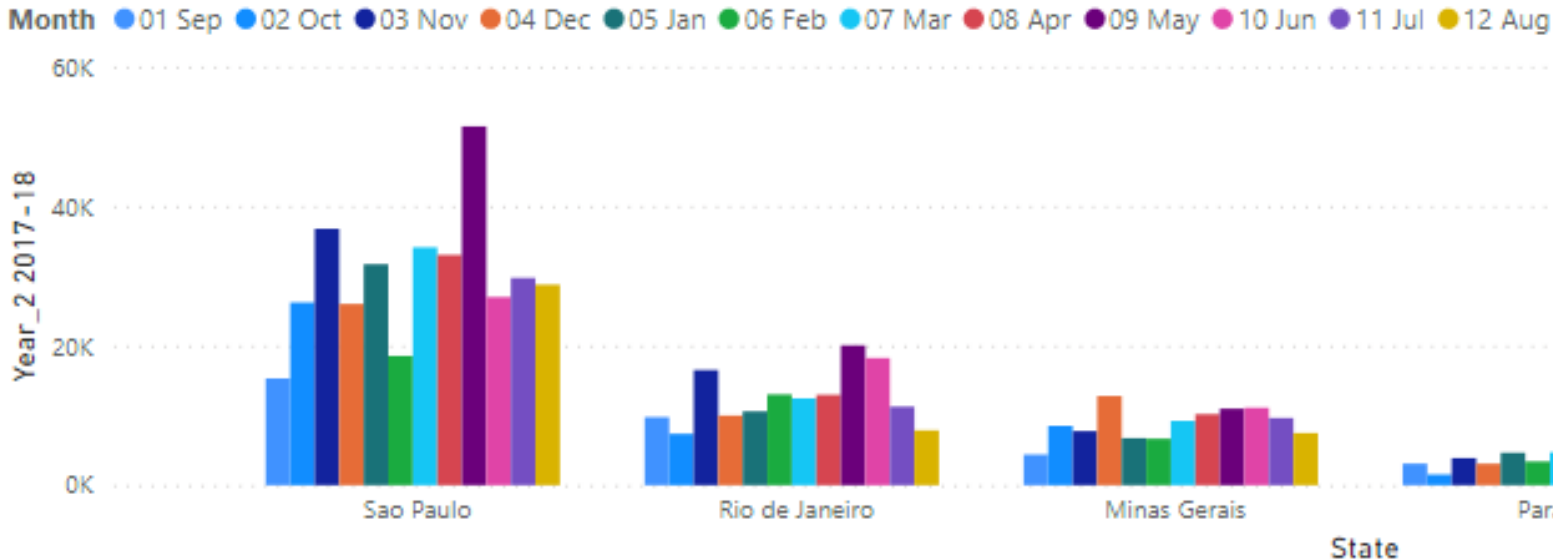
\*total sales / 15



# Watches\_gifts - Increase and drop in 2018



Sales for Watches\_gifts in Year\_2 by Month and State



For the products "watches\_gifts" we see 2 spikes, in November 2017 and in May 2018 but not for all States.

In November and in May, the spikes happened only in Sao Paulo and Rio de Janeiro. Parana had a spike only for November.

On the other hand the decline of February happened in Sao Paulo, Parana and Rio Grande do Sul.

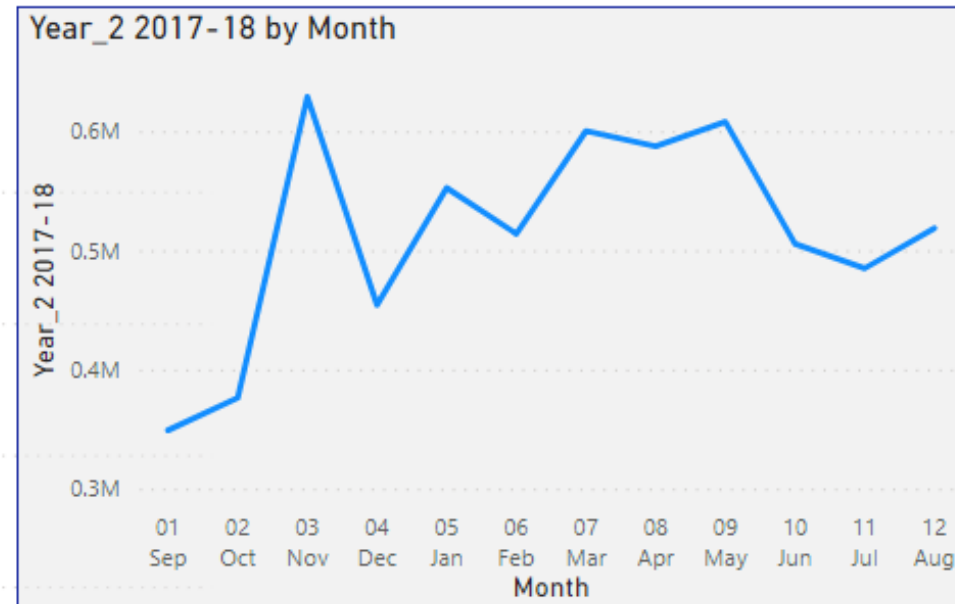
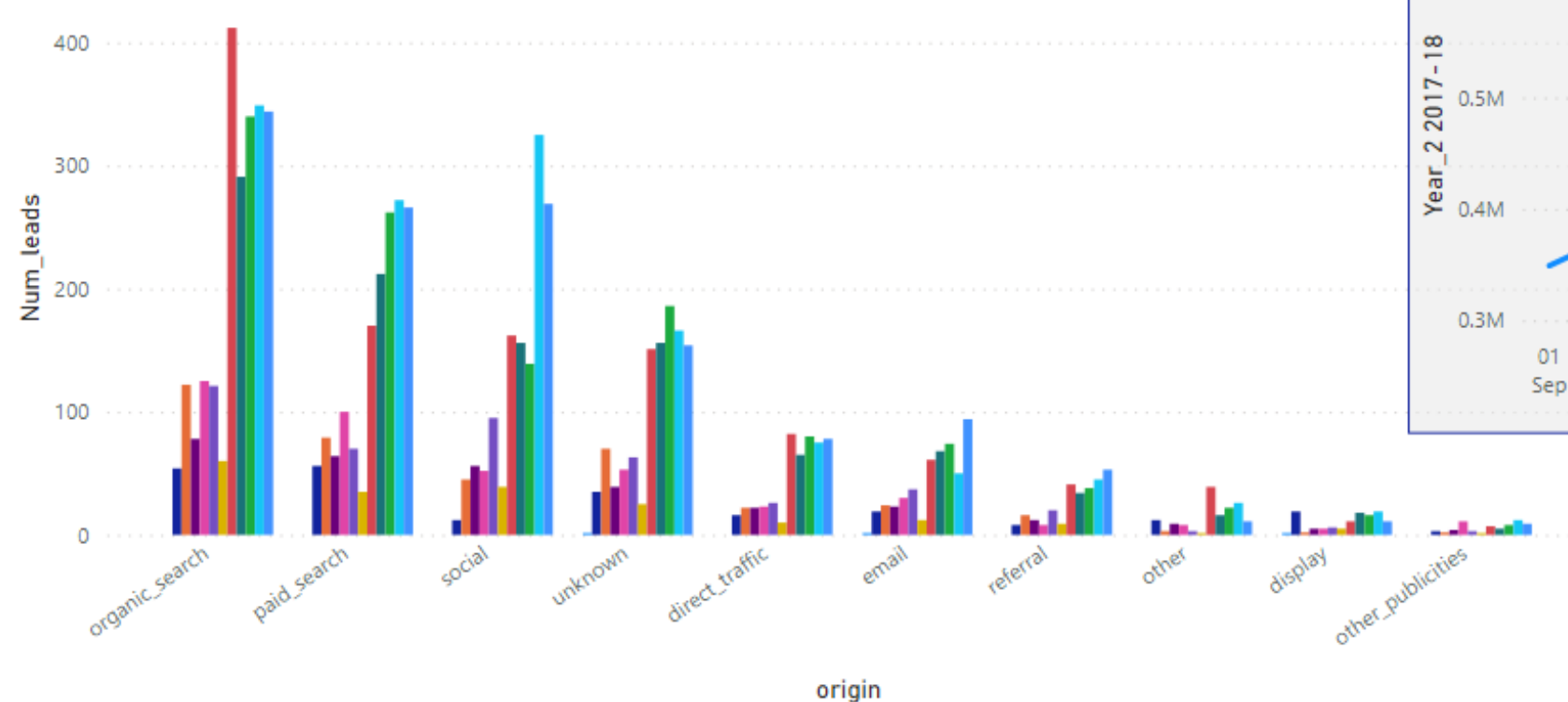
So far, we see that whatever happened to Sao Paulo affects the total growth, since in Sao Paulo there is the biggest portion of total sales, as we saw in graph "Sales in BR by State" in the page 7 "Sales in Top 6 States and Top 15 Products Categories".

We need to identify the reasons of increase and decline and react accordingly with a strategic plan.

# MARKETING - Origins for MRKT Qualified Leads per Month (data for period June 2017 - May 2018)

Num\_leads by origin and month Jun 2017 - May 2018

Months ● 01 Jun ● 02 Jul ● 03 Aug ● 04 Sep ● 05 Oct ● 06 Nov ● 07 Dec ● 08 Jan ● 09 Feb ● 10 Mar ● 11 Apr ● 12 May



We see in the graph "Num\_leads by origin and month Jun 2017 - May 2018" that there is a drop in all marketing channels in December 2017 which is reflected also in sales.

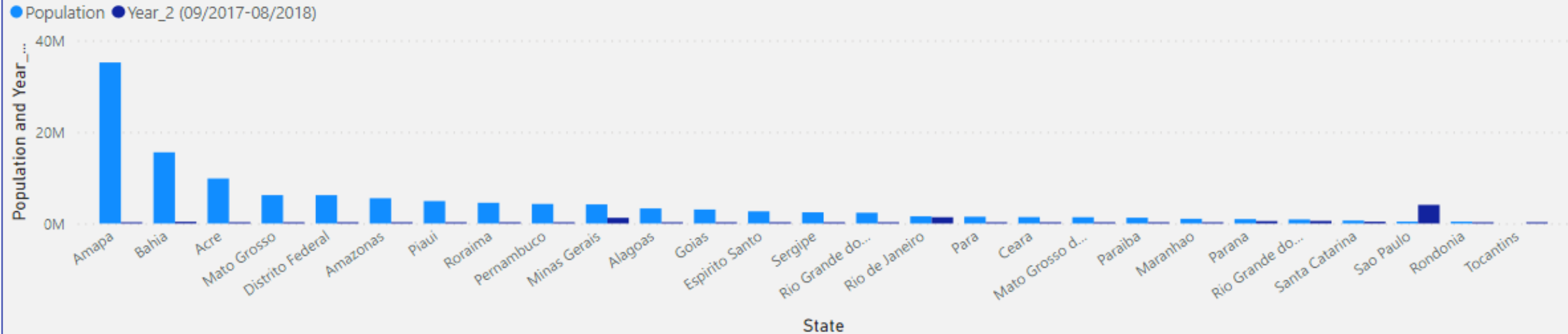
On the other hand there is significant increase in the marketing leads in January 2018 and in sales as well, but the increase in paid\_search, email and display of February didn't lead to increase in sales, but the opposite.

From March and on there was an increase in qualified leads from all marketing channels and an increase in sales as well.

Unfortunately, we don't have data regarding the marketing qualified leads after May 2018 to see if the decline in sales was related to marketing activities.

# POPULATION IN STATES vs SALES

Population and Year\_2 (09/2017-08/2018) by State



Ratio Sales/Population

State	Population	Year_2 (09/2017-08/2018)	sales/pop Year_2
Sao Paulo	419652	4,062,886.09	9,681.56
Rio de Janeiro	1561806	1,367,744.68	875.75
Rio Grande do Sul	881097	562,742.59	638.68
Santa Catarina	635998	395,809.28	622.34
Parana	955835	518,828.38	542.80
Minas Gerais	4169392	1,239,546.85	297.30
Ceara	1393408	171,762.43	123.27
Maranhao	1021557	94,388.82	92.40
Rondonia	413418	35,220.87	85.19
Para	1474859	121,949.91	82.69
Espírito Santo	2651458	212,864.69	80.28
Goiás	3055149	223,511.02	73.16
Paraíba	1258407	89,251.35	70.92
Mato Grosso do Sul	1366926	86,027.92	62.94
Pernambuco	4265247	206,122.13	48.33
<b>Total</b>	<b>121489598</b>	<b>10,428,712.46</b>	

In Sao Paulo, Olist's ratio sales vs population = 9681\*, which means that the revenue is 9,6 times more than the population of the State. Sao Paulo has the greatest sales/population ratio, which means the best performance compared to population

In the other top 5 States (Rio de Janeiro, Minas Gerais, Rio Grande do Sul, Parana and Santa Catarina) the ratio is from 297 to 875, which is not so bad.

But, there are many States, where the population is much higher than the top 6 States and the sales are very low. For example, at Amapa, Bahia, Acre, Mato Grosso the ratio sales/pop is 0,33 to 24. This means that either the habitats there are very poor or (if not) Olist should advertise more in these States because there is a big potential there.

\*The ratio sales/pop is : sales/population \* 1000



# Conclusion

---

- **Olist has 230% growth in the 2nd year of activity, in which the 76.73% of total sales were occurred**

- **In November 2017 there was a pick.**

Increase 52% in sales from October 2017, in all top 6 States and all top 15 product categories, without the qualified marketing leads to be increased accordingly (only 7% increase). The increase of sales units was about the same (62%). This means that the same number of people bought more items or more qualified leads were converted to clients in November 2017. Maybe it was a discount campaign, more products available in the platform or a deficit from competitors

- **After May 2018 and till July 2018 there was a drop of 20% in top 6 States and in top 15 Product\_categories.**

The drop was in almost all States. The highest drop was for States Sao Paulo, Rio Grande do Sul and Rio de Janeiro with 13.8%, 12.7% and 11.3% respectively. The drop was in all products except from health\_beauty. The biggest drop was in baby, toys and watches\_gifts, 25.7%, 22.3% and 18.8% respectively. The main reason for this drop could be the 9% decline of GDP per capita in 2018.

- **Less sales with more Sellers after May 2018**

There was a 58% increase in Sellers from May to July 2018 with a 10% decrease in sales for the same period. Also, the 69% of existing sellers from May had a decline in sales after May. We have to find out which was the reason of the drop (apart from GDP per capita) and create a strategic plan.

- **Big decline in watches\_gifts**

The product category "watches\_gifts" had the highest growth in whole 2nd Year (370%) but the 2nd biggest drop in June 2018 (32.5%) and the drop from May (93576 BR) to August (53671 BR) was 42.6%. The biggest drop from May to August was in Rio de Janeiro (60%) and in Sao Paulo (34%). We have to investigate the reasons and create a strategy.

- **The marketing affects the sales**

In December 2017, there was a drop in marketing qualified leads and this affected the sales for the same month. After February 2018 there was an increase in marketing qualified leads with an increase in sales. **On the contrary**, the increase in qualified leads from paid search and email campaigns in February 2018 didn't help the sales which were declined in the same month.

- **In many States the sales/population ratio is very small.** The fact that in July 2018 there was an overall 3.5% increase in sales but a 4% decrease in top 6 States (graph "Sales in Year\_2 for May to July 2018", page Sellers) confirms that there is potential in the rest 23 States.

# Suggestions

---

**We need to collect more data in order to get better insights. These data should be:**

- GDP per State in Brazil for the years 2016, 2017, 2018 to examine the growth of sales per State vs the market growth. That way we will identify in which states there is potential so as to advertise more.

Source: <https://tradingeconomics.com/> cost: \$199 monthly

- Retail market share for each product category, in order to compare Olist's sales vs competitors

Source: <https://www.reportlinker.com/> cost: € 1500 yearly

- Marketing cost for each marketing channel, number of clients who came from each marketing channel in order to calculate the CPC and the cost per customer. Then we can allocate the marketing budget properly.

Source: Olist

# Appendix

---

- **Datasets for Olist**

<https://www.kaggle.com/olistbr/brazilian-ecommerce>

- **Population for Brazil (2020)**

[https://en.wikipedia.org/wiki/List\\_of\\_Brazilian\\_states\\_by\\_population](https://en.wikipedia.org/wiki/List_of_Brazilian_states_by_population)

- **Brazil GDP Per Capita**

<https://www.macrotrends.net/countries/BRA/brazil/gdp-per-capita>

- **Brazilian States - Abbreviations**

[https://brazil-help.com/brazilian\\_states.htm](https://brazil-help.com/brazilian_states.htm)

EΥΧΑΡΙΣΤΩ TÄNAN HVALA GRACIAS DZIĘKUJĘ  
GRAZIE ありがとう MERCI TACK

**THANK YOU** DIAKUIU  
PALDIES

ACIU TACK DANKE DANK U WEL ДЗЯКУЮ  
СПАСИБО 谢谢 OBRIGADO diolch KIITOS  
TESEKKUR EDERIM