# Practicum Project: Ecommerce Dataset

11/13/2020

Sarah Hellmann, Eleanor Pulsipher, Ryan Boden, Isabel de Obaldia

# **Agenda**

- Background
- Interesting Finds
- Recommendation 1
- Recommendation 2
- Recommendation 3
- Recommendation 4
- Recommendation 5
- Recommendation 6
- Recommendation 7

# Background

## **About the Dataset**

- Data from the Brazilian e-commerce firm Olist Store
- Data was randomly selected info of 100,000 orders from 2016-2018 made at multiple marketplaces in Brazil







## **About Olist**

- Olist is a Brazillian online ecommerce site for sellers that connects merchants to a marketplace
- The largest department store in Brazilian marketplaces
- Merchants sell their products through Olist and ship them to customers using Olist logistics partners
- After a customer purchases the product from Olist the seller gets notified to fulfill that order
- Once the customer receives the product the customer gets a satisfaction survey where they can give a note for the purchase experience and write down some comments

# Interesting Finds

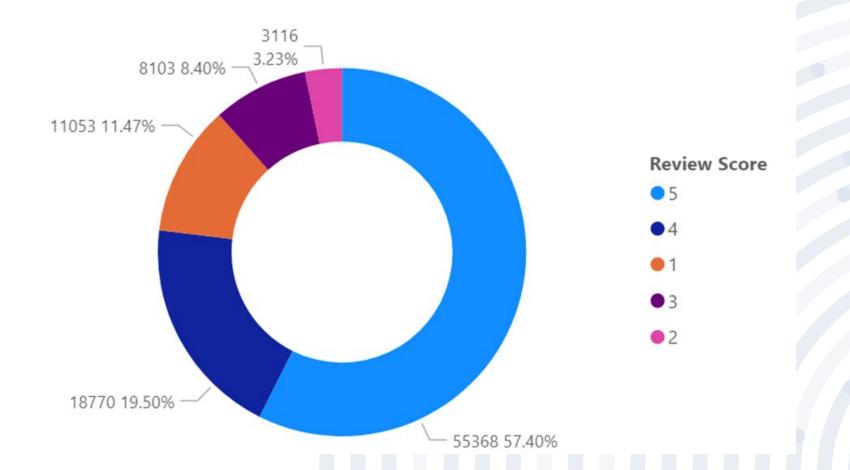
#### Product Categories %, by State Totals Cu.. SP 40.91% 44.49% 40.50% 39.56% 40.61% 41.29% 90K RJ 12.80% 12.98% 15.66% 12.64% 12.82% 14.83% 8,219 MG 12.43% 11.56% 10.12% 11.13% 11.63% 12.43% RS 5.72% 5.32% 4.97% 5.60% 5.42% 5.88% 1,04 80K PR 5.23% 4.90% 5.06% 5.75% 5.05% 5.17% SC 4.04% 3.49% 3.72% 3.58% 3.71% 3.39% 3.63% BA 3.26% 3.05% 3.34% 3.71% 3.67% 2.48% DF 2.13% 2.24% 2.20% 2.15% 1.89% 70K 23,183 ES | 2.01% 2.08% 1.91% 2.59% 2.05% 1.95% 1.96% 2.67% 1.65% 2.22% GO 1.65% 1.76% 1.47% 1.34% 2.01% 1.38% PE 1.46% 1.47% 60K CE 1.29% 1.04% 0.86% 1.61% 1.57% 1.34% PA 1.16% 0.75% 0.67% 0.87% 1.10% 1.28% 7,402 MA 0.89% 0.60% 0.48% 0.53% 0.88% 0.82% MS 0.83% 0.66% 0.76% 0.84% 0.84% 0.71% Value 50K MT 0.78% 0.86% 0.57% 0.94% 0.98% 1.09% PB 0.63% 0.43% 0.57% 0.44% 0.63% 0.47% PI 0.51% 0.42% 0.38% 0.77% 0.53% 0.53% 40K SE 0.50% 0.31% 0.29% 0.37% 0.39% 0.26% RN 0.45% 0.31% 2.10% 0.58% 0.57% 0.42% AL 0.44% 0.36% 0.38% 0.39% 0.49% 0.31% TO 0.27% 0.23% 0.29% 0.28% 0.33% 0.22% 30K 44,952 RO 0.24% 0.20% 0.19% 0.28% 0.26% 0.43% AM 0.13% 0.12% 0.10% 0.16% 0.17% 0.18% AC 0.10% 0.07% 0.29% 0.11% 0.08% 0.09% 20K AP 0.09% 0.06% 0.00% 0.09% 0.08% 0.01% RR 0.06% 0.05% 0.04% 0.00% 0.03% 0.08% 50% 0% 50% 50% 0% 0% 50% 0% 50% 50% 10K Auto/Industri.. Electronics % Books % Office % Home Goods % Fashion % OK

7

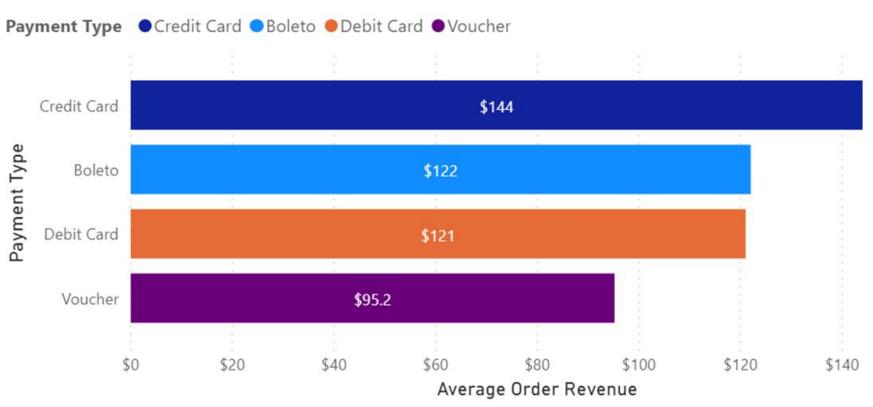
#### Number of Orders and Average Review Score by Product Category



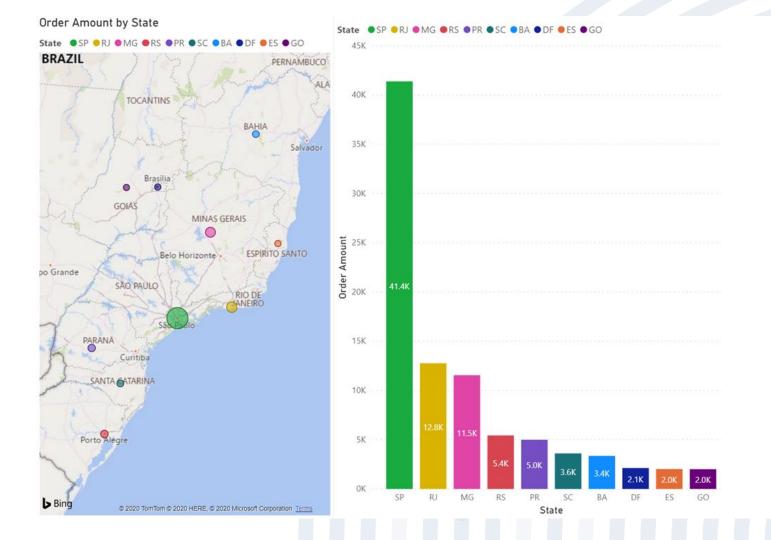
### Count and Percentage per Review Score



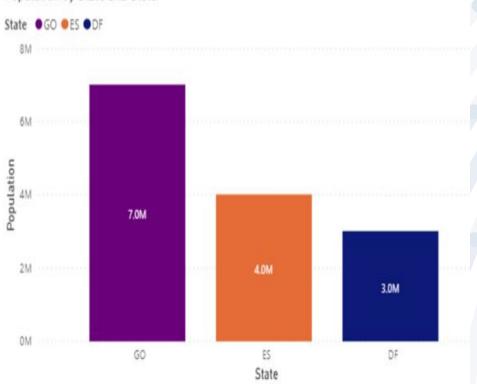
#### Average Order Revenue by Payment Type

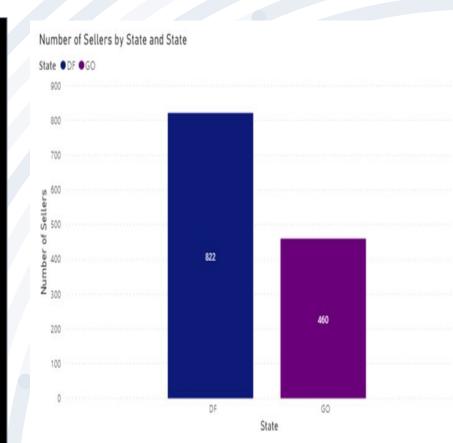


Reach out to more sellers in high population areas



# Population by State and State





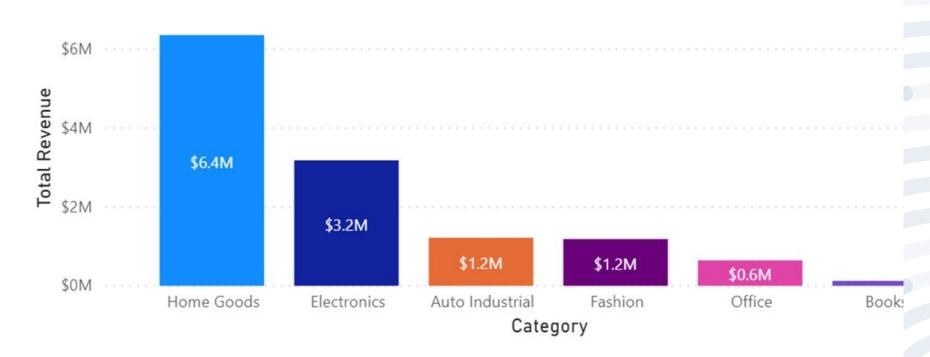
## **Our Recommendation:**

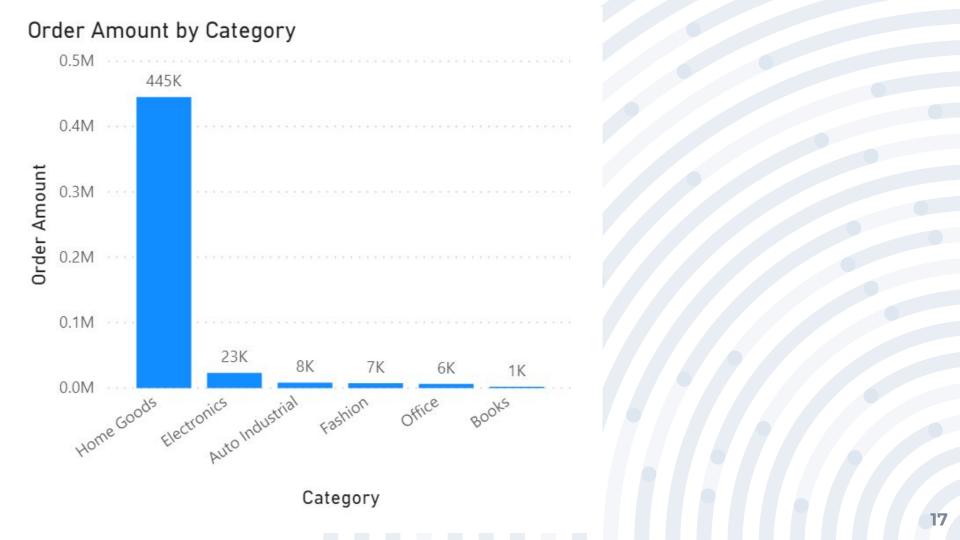
- We found out that even though Distrito Federal (DF)
  population is significantly smaller than that of Espirito Santo
  (ES) and Goias (GO), DF had a larger number of orders than
  both of these states.
- After doing extensive research, we noticed that this may be due to the fact that the amount of sellers Olist has in DF is greater than the amount of sellers it has in both ES and GO
- To increase revenue, Olist must take advantage of this opportunity and increase its sellers in high population areas.

# Recommendation 2: Diversify sources of income

#### Total Revenue by Category







## **Our Recommendation:**

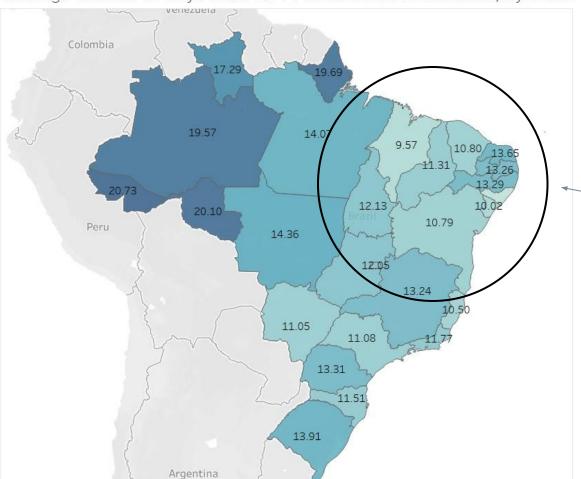
- After exploring the dataset, we noticed that a large portion of Olist's revenues relies heavily on the sale of Home Goods. We also noticed that the company is earning minimal revenue on its Books retailing services.
- We recommend that in the long term, Olist should consider to diversify its sources of income to not be so heavily dependent in just one product category.

Recommendation 3:
Improve Shipping to the Northeast Region

# **Flawed Shipping Estimation**

- In general, we found that the estimated arrival dates Olist provides their customers are highly inaccurate
- We created a new variable called "shipping disparity" that takes the difference between an order's estimated arrival and its actual arrival date (measured in days)
- We found that most orders are fulfilled 1-2 weeks before their estimated arrival, but high Standard Deviation and IQR suggest pervasive inconsistency
- However, customers in Northeastern states wait an average of 7-9 days longer for their deliveries than customers in other regions, presumably because the Northeast has lower population density and is geographically isolated by the Amazon Rainforest

Average Number of Days an Order Arrives Ahead of Schedule, by State

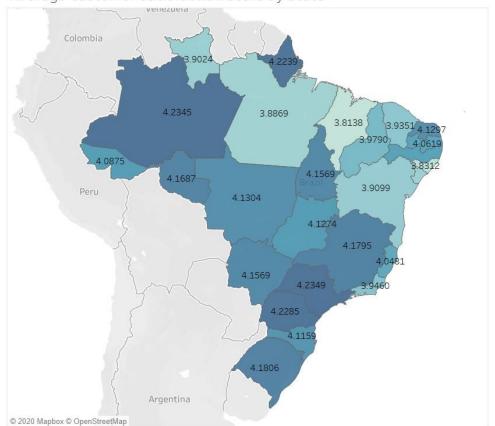


AVG(Shipping Disparity)	
8.71	20.73
Summary	
Count:	27
AVG(Shipping Disparity)	
Average:	13.33
Minimum:	8.71
Maximum:	20.73
Median:	12.19

Customers in the Northeast region (especially in the states Algoas, Maranhão, and Ceará) experience longer wait times and are less likely for their orders to arrive ahead of schedule.

# As a result, customers in the Northeast are less likely to be satisfied with their purchases

Average Customer Satisfaction Score by State



#### Our recommendation:

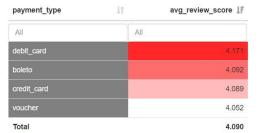
Olist should work towards improving the accuracy of their shipping estimates in general, and grow their logistical infrastructure in the Northeast region to better serve customers in these states. Though they currently make up a smaller proportion of Olist's customer base, improving accessibility to this region is a significant growth opportunity for the firm.

# Recommendation 4: Reduce the Distribution of Vouchers

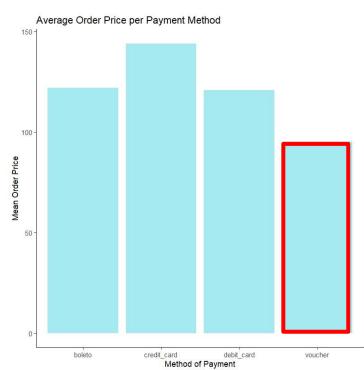
# **Vouchers and their Impact**

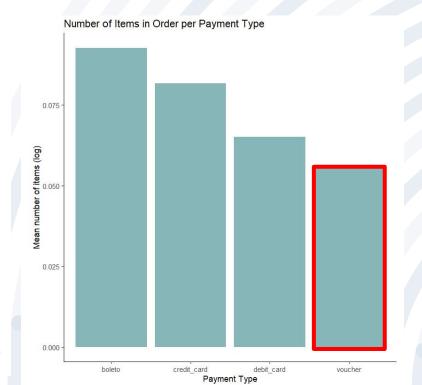
- Olist occasionally provides vouchers to some customers, which are similar to coupons
- One could assume that this would incentivize customers to spend more and increase their level of purchase satisfaction, but does it really?

#### Mean Satisfaction Score per Payment Type



### Voucher Impact on Average Order Price, Number of Items per Order, and Satisfaction Scores





## **Our Recommendation:**

- We found that customers who paid with vouchers spent less money per order, purchased less items, and had the lowest average satisfaction scores in comparison to other payment types.
- This validates the results of our linear regression, which showed voucher payment had a significant negative correlation with satisfaction scores.
- If Olist is incurring any expense by providing vouchers to their customers, the results of the program are likely not worth the cost. The firm should either distribute less vouchers in the future, or be more selective with the customers they are made available to.

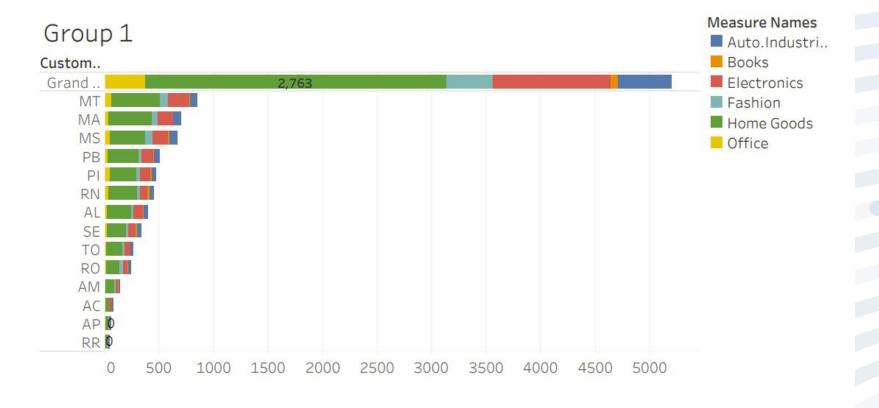
**Recommendation 5:** 

Focus business expansion efforts to Sao Paulo (SP), Rio de Janeiro (RJ), and Minas Gerais (MG) Regions

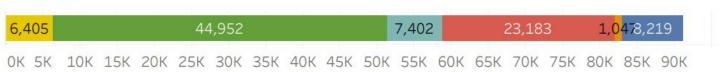
# Product Category Purchases by State with Cluster Analysis

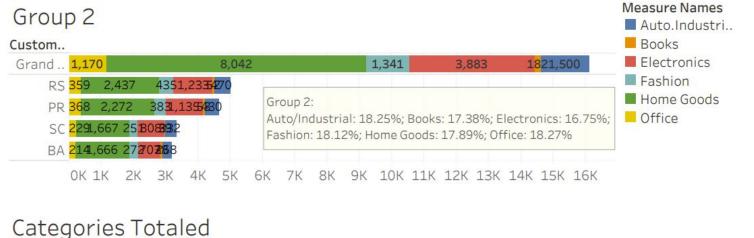
- We did a cluster analysis of Olist's purchases grouped by product category and state and found 5 groups
- OUTPUT: AC AL AM AP BA CE DF ES GO MA MG MS MT PA PB PE PI PR RJ RN
  1 1 1 1 2 3 3 3 3 1 4 1 1 1 1 3 1 2 4 1

  RO RR RS SC SE SP TO
  1 1 2 2 1 5 1
- After analyzing these groups' specific data per product category it becomes clear that most of their orders come from:
  - Home Goods (66%)
  - Sao Paulo, Rio de Janeiro, and Minas Gerais regions (65%)



#### Totals





#### categories rotales

1K

OK

2K

3K



## Group 3

#### Custom.. Grand .. 610 4,494 615 97 702 2,038 DF 141 967 140519267 Group 3: ES 166 922 1444822065 Auto/Industrial: 8.54%; Books: 9.26%; Electronics: 8.79%; GO 106 999 130455286 Fashion: 8.31%; Home Goods: 10.00%; Office: 9.52% PE 94 902 1034120 CE 103704 9241906

4K

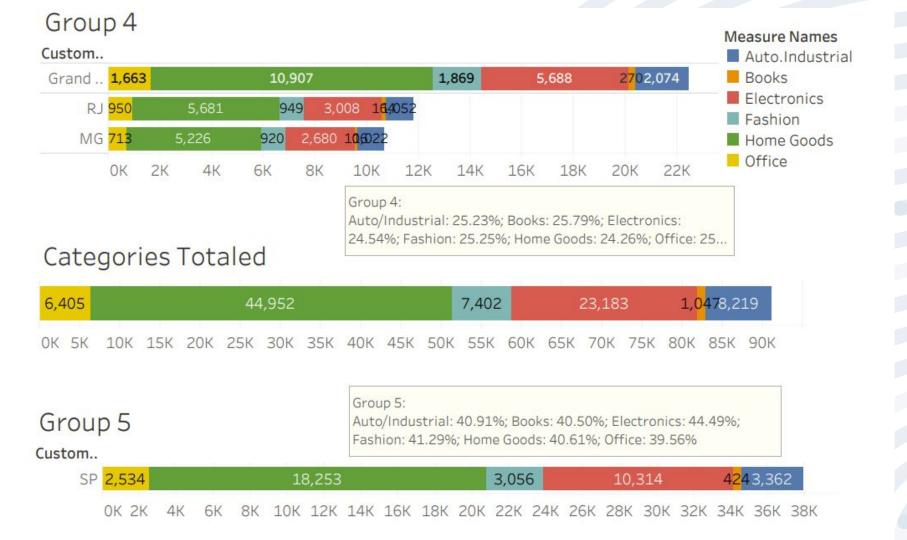
5K

6K

7K

8K

30



# **Product Categories %, by Group**

	Auto/ Industrial	Books	Electronics	Fashion	Home Goods	Office	TOTAL
Group 1:	5.91%	6.40%	4.68%	5.76%	6.15%	5.81%	5.70%
Group 2:	18.25%	17.38%	16.75%	18.12%	17.89%	18.27%	17.67%
Group 3:	8.54%	9.26%	8.79%	8.31%	10.00%	9.52%	9.38%
Group 4:	25.23%	25.79%	24.54%	25.25%	24.26%	25.96%	24.64%
Group 5:	40.91%	40.50%	44.49%	41.29%	40.61%	39.56%	41.60%
TOTAL:	9.01%	1.15%	25.42%	8.12%	49.29%	7.02%	100.00%

 Note the uniformity of product category proportions in Group 4 and 5, although nearly 50% of orders are from the Home Goods category

## **Our Recommendation**

- Short term: in order to keep up with their most profitable sectors, Olist should focus on improving their shipping, selling, and marketing efforts in two ways:
  - In a geographical sense, the regions in Groups 4 and 5 (SP, RJ, and MG)
  - In a product category sense, the Home Goods sector should be prioritized
- Long term: if Olist plans to expand their business into other product categories, these regions perform well across all categories, so there is a lot of potential for a new product category to succeed in these regions

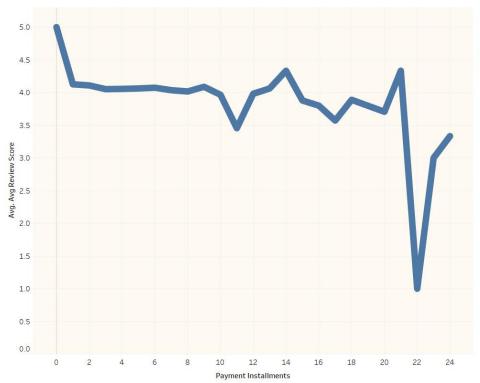
Recommendation 6:
Limit Payment Plan
Installments

# The "Pick Your Own" Payment Schedule

- Olist currently allows customers who make purchases on layaway to choose their own number of monthly payment installments
- The average customer pays in 2-3 installments with a standard deviation of ~2.7, suggesting that the vast majority of all customers pay off their purchases in less than 10 months
- Is there any benefit to the firm to continue offering customers the option to select more than 10 installments (up to 24 months in the maximum case)?



# The Relationship Between Number of Payments and Customer Satisfaction



```
Response variable : avg_review_score

Explanatory variables: payment_installments

Null hyp.: the effect of payment_installments on avg_review_score is zero

Alt. hyp.: the effect of payment_installments on avg_review_score is not zero

**Standardized coefficients shown (2 X SD)**

coefficient std.error t.value p.value

(Intercept) 0.000 0.002 0.000 1.000

payment_installments -0.033 0.003 -10.368 < .001 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

 The results of our linear regression show a significant negative correlation between number of installments and satisfaction scores - offering more installments does not provide material or sentimental benefit to the firm

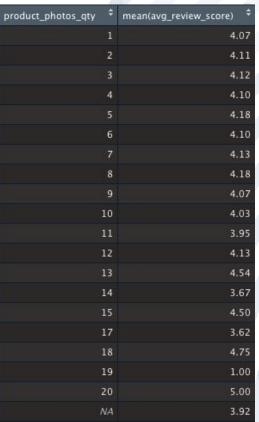
# **Our Recommendation**

 Because the vast majority of customers select payment plans under 10 months, and customers who choose more installments are more likely to be dissatisfied with their orders, the firm should limit the number of installments available to mitigate risk to accounts receivable

# Recommendation 7: Photo Quantity

## **Olist Product Photo Quantity & Review Score**

- Olist currently has products on their site with a photo range of no photos a maximum of 20 photos.
- Their review score breakdown can be seen in the table to the right:
  - Note: NAs (or no photos) have the lowest average rank



# The Relationship Between Photo Quantity and Average Review Score

- The results of our logistic regression show a significant correlation between the number of photos and average review score holding all else constant, including one additional photo means the odds of a higher review score increase by 1%.
- While 1% may seem insignificant, if each product sold with only one photo had two photos instead, the odds of the review score being increased would be up by with the odds of 480%.

1	2 3	4	5	6	7	8	9	10	11	12	13	14	15	17	18	19	20	
48028 1917	7 11183	7599	4981	3408	1411	685	289	318	62	46	26	6	12	8	4	2	1	

# Thanks!

Any questions?