**Abdullah Sharaf (2139533)**

E-Commerce Sentiment Analysis

Brazilian E-Commerce Public Dataset by Olist

Goal: The heatmap is designed to compare each day's sales and shows outliers when sales were unusual for the two most selling product categories, which are the product categories "cama_mesa_banho" (**bed, table, and bath**) and "moveis_decoracao" (**furniture and decoration**), on the Olist shopping site in Brazil during November 2017.
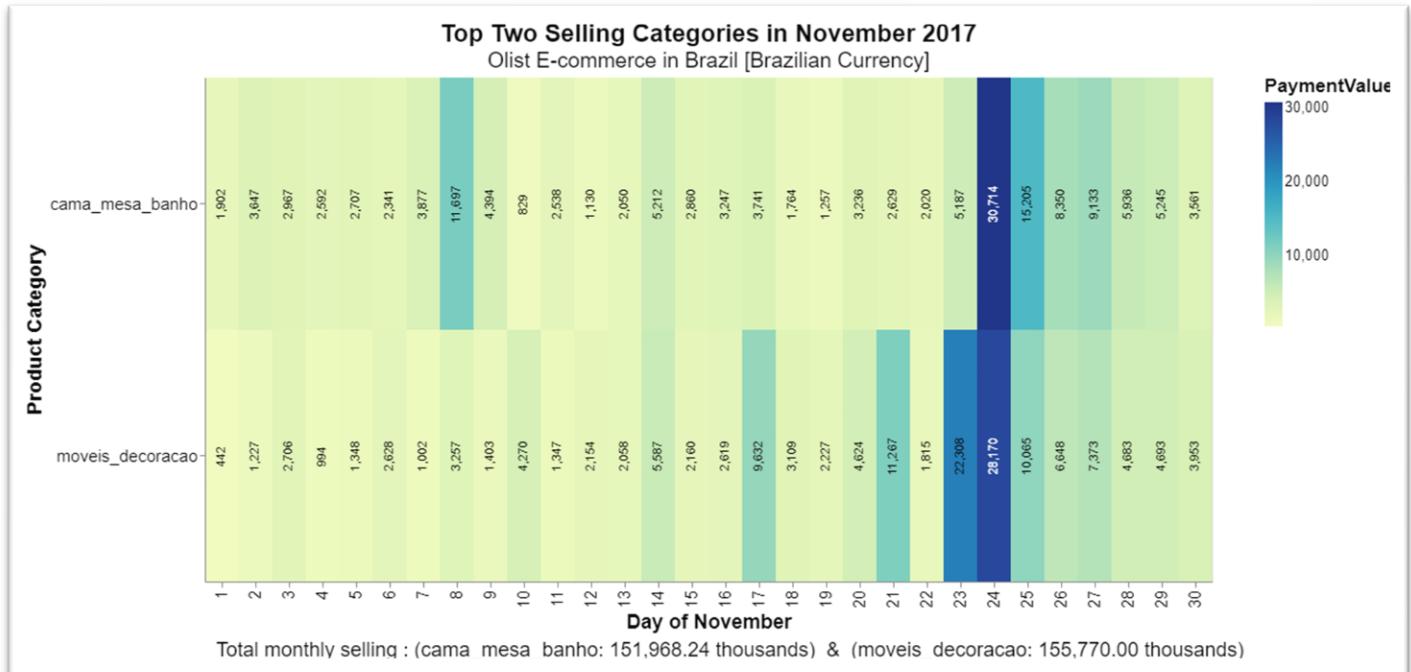
insight : The "cama_mesa_banho" category experienced a notable surge and outliers popout in sales at 30714 on November 24th during 2017, presumably coinciding with Black Friday, showcasing a more pronounced peak in comparison to the "moveis_decoracao" category, which also saw an increase in sales at 28170 on the same day but to a lesser extent. Throughout the month, sales volumes for both categories showed fluctuation. In Note under the chart, we can see total monthly sales figures, "moveis_decoracao" slightly outpaced "cama_mesa_banho," with totals of 155,770.00 thousand and 151,968.24 thousand in Brazilian currency, respectively.

Data Abstraction:

Dataset Type: The dataset is a <u>table</u> resulting from the merging of four tables (olist_order_payments_dataset.csv, olist_orders_dataset.csv, olist_products_dataset.csv, and olist_order_items_dataset.csv) From the merging table, I chose November because Black Friday in November was an outlier during the fluctuation that happened in November month.

1-olist_products_dataset

product_category_name: the name of the selling product categories .

2-olist_order_payments_dataset.

payment_value: the payment amount during the period

3-olist_orders_dataset.

order_purchase_timestamp: purchase time year, month, day, hour, minute, we used 2017 year and November month.

✓ order_id, product_id : to join four tables together (olist_order_items_dataset.csv, olist_orders_dataset.csv, olist_order_payments_dataset.csv,olist_products_dataset.csv)

Items: Each rectangular in the graph represents an item. In this case, items are the days of November 2017, and the corresponding value for each rectangular represents the total payment value for one of the top two selling product categories on that day.

Data Types of Attributes:

- Product Category   (Categorical).

- Day of Month   (Ordered, Quantitative, Cyclic).

- Payment Value   (Quantitative, Continuous).

Task Abstraction: The heatmap illustrates the daily sales data for the top two selling product categories and outliers day-selling on the Olist E-commerce platform in Brazil for November 2017. The product categories are "cama_mesa_banho" ("bed, table, and bath") and "moveis_decoracao" (" furniture decoration").

Marks: Rectangles

Channel:

- Payment Values   (Color Saturation: Ordered sequential).
- Day of Month   (Position: x-axis: Ordered by day number).
- product category   (Position: y-axis).

   Additional information :

- Payment Values   (Text labels).

Action and Target:

Level	Action	Target	
High-Level	Analyze(**Discover**)	**Outliers and trends**	**Discover** the sales **outliers** and day-on-day changes in payment value for the top two selling product categories during November 2017 on the Olist E-commerce platform in Brazil.
Mid-Level	Search(**Lookup**)	**Features**	**Data analysts lookup** consumer preferences by examining the payment value of the top two product categories over the days of November (**features**).
Low-Level	Query(**Identify&Compare**)	**Features**	**Stakeholders** such as the marketing manager, data analysts, and people **compare** payment values and **identify** outliers' dependencies between two product categories over the days of November (**features**).

Data source: Visualization tool (Python Altair).

URL: https://www.kaggle.com/code/thiagopanini/e-commerce-sentiment-analysis-eda-viz-nlp/input