



---

## RELATÓRIO

Análise de dados de vendas realizadas via canal Olist entre os anos de 2016 a 2018

Analista: Matilde da C. Pessoa

---

## Lista de Tabelas e Figuras

<b>Tabelas</b>	<b>PAGINA</b>
Tabela 1. Frequências absolutas, relativas e percentuais de crescimento no número de pedidos e entregas realizadas nos anos 2016, 2017 e 2018.....	4
Tabela 2. Tempo, em dias, correspondente as fases de processamento dos vendas realizadas via plataforma Olist e relação entre data real e data estimada de entrega.	5
Tabela 3. Número de itens comercializados via plataforma Olist e faturamento anual de acordo com as regiões brasileiras.....	6
Tabela 4. Top 5 categorias de produtos comercializados via plataforma Olist de maior faturamento nos anos de 2016, 2017 e 2018.....	7
 <b>Figuras</b>	
Figura 1. Visão geral da do status atual de pedidos realizados entre 04/09/2016 a 17/10/2018 via plataforma Olist.....	4
Figura 2. Número de produtos vendidos via plataforma Olist e faturamento alcançado entre setembro de 2016 a setembro de 2018.....	5
Figura 3. Categorias de maior destaque de vendas no período de 04/09/2016 a 17/10/2018 via plataforma Olist.....	7
Figura 4. Top 5 categorias de produtos comercializados via plataforma Olist de maior faturamento por regiões brasileiras.....	8
Figura 5. Frequência de escores indicadores da satisfação do cliente sobre a compra realizada via plataforma Olist entre 09/2016 a 09/2018.....	9

**Objetivo Geral:**

Apresentar o cenário de vendas do varejo brasileiro realizadas via plataforma Olist entre setembro de 2016 a outubro de 2018.

**Objetivos específicos:**

- Indicar e descrever os principais pontos relacionados a indicadores-chave de performance do e-commerce.
- Descrever a oscilação de vendas mensais nos anos avaliados.
- Apresentar o número de vendas/pedidos realizados via plataforma Olist segundo as diferentes regiões brasileiras.
- Descrever o faturamento alcançado pela e-commerce nos anos avaliados.
- Indicar as categorias de produtos que mais se destacaram no período avaliado e por regiões brasileiras
- Apresentar o número de lojistas participantes da e-commerce Olist por ano e região.
- Avaliar os tempos de processo entre compra e entrega ao consumidor final
- Descrever alguns dos KPIs importantes para o negócio.
- Descrever possíveis inconsistências encontradas no banco de dados e propor sugestões de melhoria.

## **1. ARQUIVOS DE DADOS, ESTRUTURAÇÃO E METODOLOGIA DE ANÁLISES DE DADOS**

Foram analisados oito arquivos de dados do varejo brasileiro que se conectam entre si por identificação e demais informações sobre pedidos, produtos, lojistas, clientes e locais de entrega referente às vendas realizadas entre setembro de 2016 a outubro de 2018 via plataforma de e-commerce Olist.

Estes arquivos foram inicialmente submetidos a uma análise de consistência no Software R com a finalidade de detectar possíveis erros ou informações discrepantes e verificar a compatibilidade entre arquivos. Em sequência foram preparadas algumas bases de dados para edição de gráficos e tabelas nos softwares Microsoft Excel e Microsoft Power BI ([https://github.com/matildecpessoa/olist\\_test/](https://github.com/matildecpessoa/olist_test/)) para se obter uma visão geral sobre o cenário de vendas e indicadores do negócio da empresa.

As análises de dados foram inicializadas pelos arquivos que continham informações sobre pedidos (`olist_orders_dataset`, `olist_order_items_dataset`). Por se conectarem em relação as identificações de pedidos e por conectarem com um maior número de arquivos que continham informações dos pontos chave do negócio (produtos, clientes, lojistas, pagamentos e localizações de entrega). Em seguida, foi gerado um arquivo de datas para melhor direcionar as análises e a descrição dos resultados. Feito isso, os arquivos foram analisados com foco em número de itens vendidos, categorias dos produtos, número de lojistas, tempo de entrega e faturamentos. Por fim, o arquivo que apresentava informações dos clientes foi analisado, por conter informações e feedbacks do nosso “ponto final” das vendas.

No item 2 são apresentadas descrições de cada um dos pontos chave do negócio para melhor visualização do panorama sobre as vendas ocorridas entre setembro de 2016 a outubro de 2018.

## **2. DESCRIÇÃO DE VARIÁVEIS ASSOCIADAS AOS INDICADORES-CHAVE DO NEGÓCIO**

### **2.1. DESCRIÇÃO DE PEDIDOS**

Foram analisadas informações de 98.666 pedidos realizados entre setembro de 2016 a setembro de 2018. Destes pedidos, 90% tinham apenas um item/pedido e 10% continham mais de um item/pedido totalizando 112.650 itens vendidos.

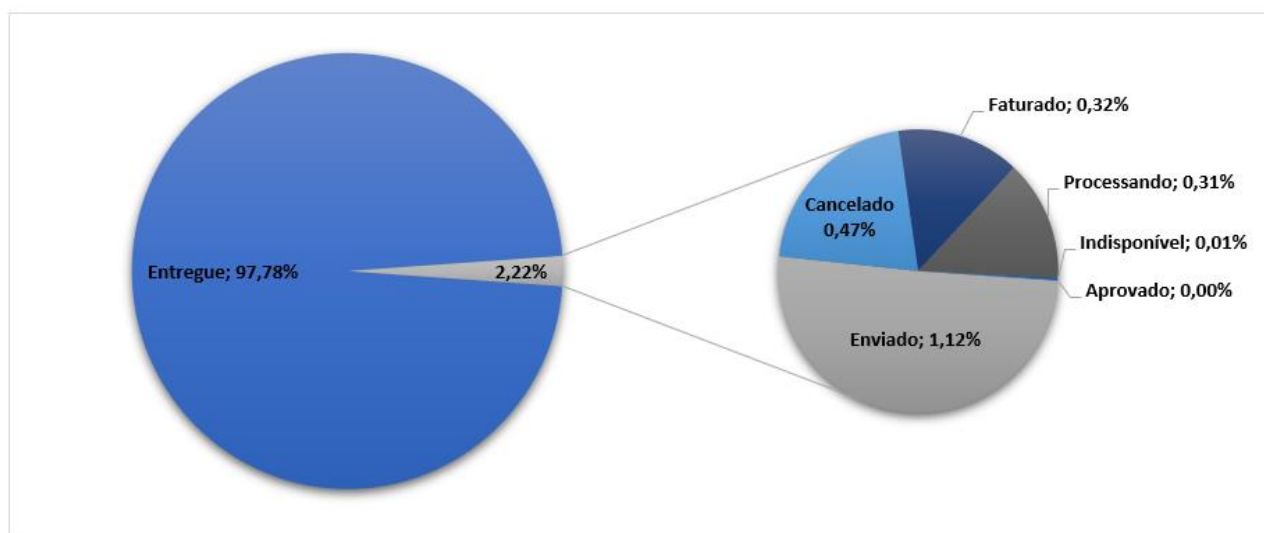
Observou-se um crescimento acentuado tanto no número de pedidos quanto no número de entregas de 2016 para 2017, representados por 14.188% e 16.165% respectivamente. Talvez este crescimento exacerbado se justifique pelo número reduzido de vendas no ano de 2016. Já de 2017 para 2018 houve um crescimento de 20,62% no número de pedidos e 21,54% no número de entregas (Tabela 1).

Do total de pedidos realizados neste período 97,78% foram entregues e os demais encontram-se em diferentes status de processamento entre data de compra e data de entrega ao cliente (Figura 1).

**Tabela 1. Frequências absolutas, relativas e percentuais de crescimento no número de pedidos e entregas realizadas nos anos 2016, 2017 e 2018**

Pedidos			Anos	Entregas		
Frequência Absoluta	Valor relativo (%)	Crescimento anual (%)		Frequência Absoluta	Valor relativo (%)	Crescimento anual (%)
312	0,31%		<b>2016</b>	267	0,27%	
44.579	45,18%	14.188%	<b>2017</b>	43.428	45,01%	16.165%
53.775	54,50%	20,62%	<b>2018</b>	52.783	54,70%	21,54%
<b>98.666</b>			<b>Total</b>	<b>96.478</b>		

**Figura 1. Visão geral da do status atual de pedidos realizados entre 04/09/2016 a 17/10/2018 via plataforma Olist**



O tempo mínimo, médio e máximo, em dias, referente a cada fase do processo de entrega dos produtos até o consumidor final e a relação entre data de entrega real e data de entrega estimada são apresentados na tabela 2. Pode-se destacar com estes resultados que o maior tempo para entrega do produto ao cliente (97,6%) está mais associado ao tempo entre distribuidora e cliente.

Um indicador de desempenho que merece destaque é o percentual de entregas realizadas abaixo ou acima do prazo de entregas estimado (Tabela 2.). Neste banco de dados observou-se que 91,9% dos pedidos são entregues antes do prazo estimado, 1,34% são entregues na data estimada e 6,77% fora deste prazo estimado. Este indicador pode estar relacionado com a satisfação ou insatisfação do cliente em realizar compras utilizando a plataforma Olist. Analisando a correlação entre tempo de entrega e escores de satisfação do cliente verificou-se que existe uma associação negativa e mediana (-0,33) entre as duas variáveis, mostrando que muitos dos clientes que recebem os produtos dentro do prazo indicaram escores maiores na pesquisa satisfação.

**Tabela 2. Tempo, em dias, correspondente as fases de processamento dos vendas realizadas via plataforma Olist e relação entre data real e data estimada de entrega.**

Processos	Tempo (dias)		
	Mínimo	Médio	Máximo
Da compra à aprovação	0	0,51	60
Da compra à entrega na distribuidora	*	3,22	126
Da distribuidora ao cliente	*	9,28	205
Da compra à entrega ao cliente	0	12,5	210

Data de entrega ao consumidor x Data de entrega estimada		
Data de entrega < data estimada	Data de entrega = data estimada	Data de entrega > data estimada
88649 pedidos (91,9%)	1292 pedidos (1,34%)	6535 pedidos (6,77%)

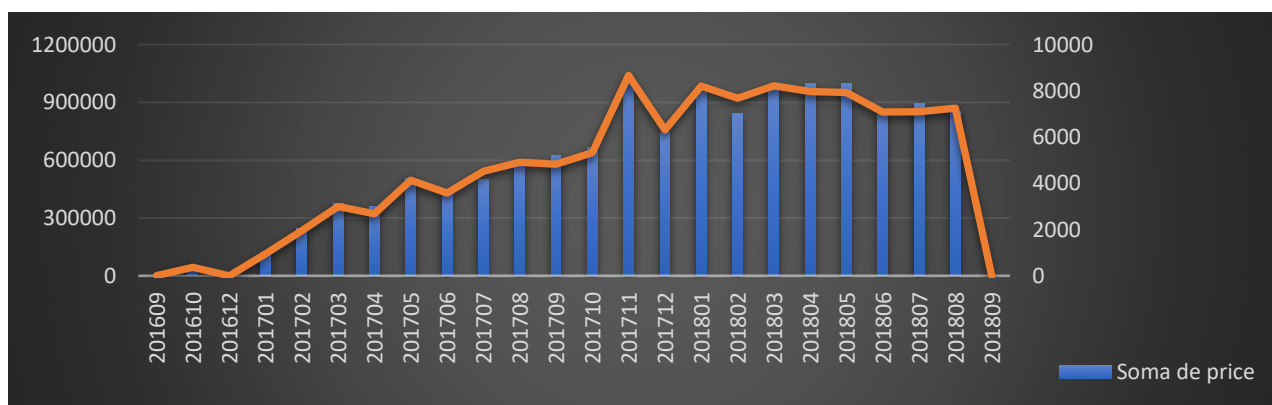
\*Valores negativos – Inconsistência de datas

## 2.2. DESCRIÇÃO DE VENDAS (ITENS VENDIDOS)

Na figura 2 está apresentado crescimento das variáveis número de itens vendidos e faturamento de acordo com os meses e anos avaliados. Pode-se perceber um crescimento acentuado tanto no número de vendas quanto no faturamento em outubro de 2017 com uma queda em dezembro do mesmo ano caracterizando a estação de compras para presentes de natal. Em 2018 não foi possível observar o mesmo comportamento neste período pois o banco de dados tinha informações apenas até outubro. Entretanto, observa-se uma certa estabilidade nos níveis de vendas e no faturamento entre janeiro a agosto de 2018.

A região Sudeste destacou-se como líder em número de itens vendidos e faturamento nos três anos avaliados, seguida da região Sul, Nordeste, Centro-oeste e Norte. De 2017 para 2018 houve um aumento de R\$1.230.243,72 no faturamento, correspondendo a 19,98% a mais que em 2017(Tabela 3).

**Figura 2. Número de produtos vendidos via plataforma Olist e faturamento alcançado entre setembro de 2016 a setembro de 2018**



**Tabela 3. Número de itens comercializados via plataforma Olist e faturamento anual de acordo com as regiões brasileiras**

Anos		Número de itens vendidos	Faturamento
2016	Sudeste	233	R\$ 31.686,27
	Sul	70	R\$ 8.530,19
	Nordeste	39	R\$ 6.013,32
	Centro – oeste	19	R\$ 2.355,95
	Norte	9	R\$ 1.200,19
Subtotal		370	R\$ 49.785,92
2017	Sudeste	34.251	R\$3.961.696,2
	Sul	7.502	R\$ 888.022,3
	Nordeste	5.041	R\$ 733.454,80
	Centro – oeste	2.992	R\$ 404.506,80
	Norte	1.058	R\$ 168.126,8
Subtotal		50.864	R\$ 6.155.806,98
2018	Sudeste	42.909	R\$ 4.894.010,6
	Sul	8.579	R\$ 1.057.388,6
	Nordeste	5.329	R\$ 806.025,6
	Centro – oeste	3.602	R\$ 463.599,3
	Norte	997	R\$ 165.026,7
Subtotal		61.416	R\$ 7.386.050,8
<b>Total</b>		<b>112.650</b>	<b>R\$ 13.591.643,7</b>

### 2.3. DESCRIÇÃO DE CATEGORIAS DE PRODUTOS E FATURAMENTOS

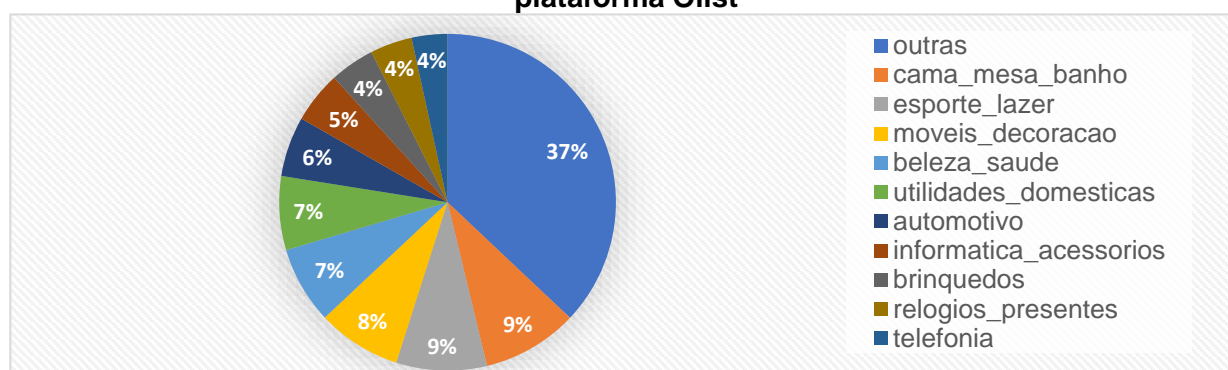
Foram analisadas informações de vendas de 32.951 diferentes produtos distribuídos em 73 categorias. Deste, 610 não apresentavam informações sobre categoria, descrições e fotos dos produtos e 2 produtos não apresentavam informações sobre medidas.

Os produtos mais comercializados nos anos de 2016, 2017 e 2018 pertencem às categorias perfumaria, cama-mesa-banho e móveis-decoração, nesta ordem. Entretanto, as categorias que mais se destacaram durante o período avaliado foram cama-mesa-banho e esporte-lazer (Figura 4).

As Top 5 categorias de produtos de maior faturamento nos anos de 2017 e 2018 foram cama-mesa-banho, relógios-presentes, beleza-saúde, esporte-lazer e informática e acessórios (Tabela 4).

Em relação às Top5 categorias mais vendidas nas diferentes regiões brasileiras, destacam-se beleza-saúde, esporte-lazer e informática, que estão entre as cinco mais vendidas em todas as regiões do país (Figura 5).

**Figura 3. Categorias de maior destaque de vendas no período de 04/09/2016 a 17/10/2018 via plataforma Olist**

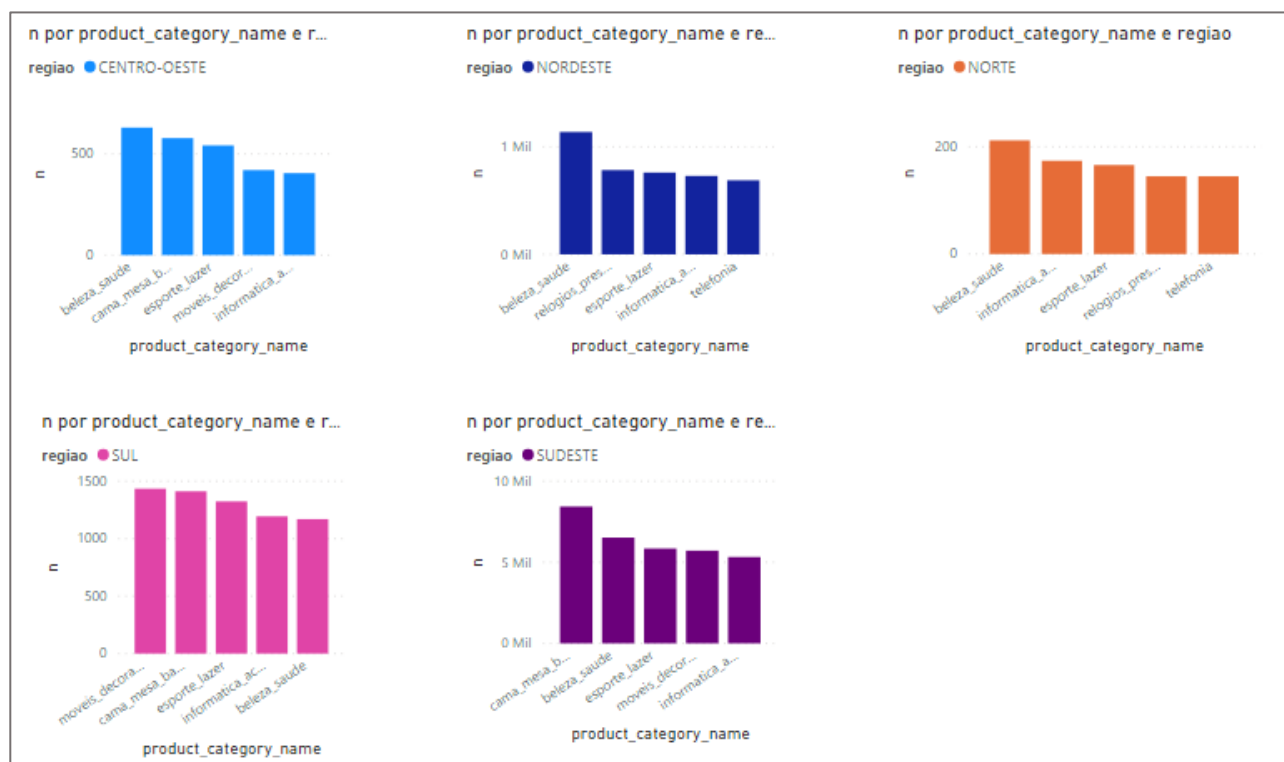


**Tabela 4. Top 5 categorias de produtos comercializados via plataforma Olist de maior faturamento nos anos de 2016, 2017 e 2018**

TOP5_CATEGORIAS			
Ano	Categoria	Faturamento absoluto	Faturamento relativo ao anual
2016	Móveis-decoração	5880.78	11.81%
2016	Perfumaria	5688.7	11.43%
2016	Beleza-saúde	4687.48	9.42%
2016	Brinquedos	4465.09	8.97%
2016	Consoles-games	3882.26	7.80%
	<b>TOTAL_2016</b>	<b>49785.92</b>	
2017	Cama-mesa-banho	498440.43	8.10%
2017	Relógios-presentes	492794.5	8.01%
2017	Beleza-saúde	481755.71	7.83%
2017	Esporte-lazer	452148.84	7.35%
2017	Informática-acessórios	405078.69	6.58%
	<b>TOTAL_2017</b>	<b>6155806.98</b>	
2018	Beleza-saúde	772238.15	10.46%
2018	Relógios-presentes	708850.94	9.60%
2018	Cama-mesa-banho	538069.26	7.28%
2018	Esporte-lazer	532566.49	7.21%
2018	Informática-acessórios	505476.31	6.84%
	<b>TOTAL_2018</b>	<b>7386050.8</b>	



**Figura 4. Top 5 categorias de produtos comercializados via plataforma Olist de maior faturamento por regiões brasileiras**



#### 2.4. DESCRIÇÃO DOS LOJISTAS

O banco de dados analisado continha um total de 3.095 lojistas distribuídos por todas as regiões brasileiras. O número de lojistas nos anos 2016, 2017 e 2018 foram 145, 1784 e 2383, nesta ordem. Do ano de 2017 para 2018 houve um crescimento de 33,57% no número de lojistas.

A região brasileira de maior concentração de lojistas é a região sudeste (73,90%), seguida da região sul (21,58%). As regiões centro-oeste, nordeste e norte concentram 2,55%, 1,80% e 0,16%, respectivamente.

O lojista que alcançou maior número de vendas no período avaliado vendeu 2.033 itens e faturou R\$ 123.304,83. Entretanto o lojista que atingiu maior faturamento (R\$ 229.472,63), neste mesmo período, vendeu 1.156 itens.

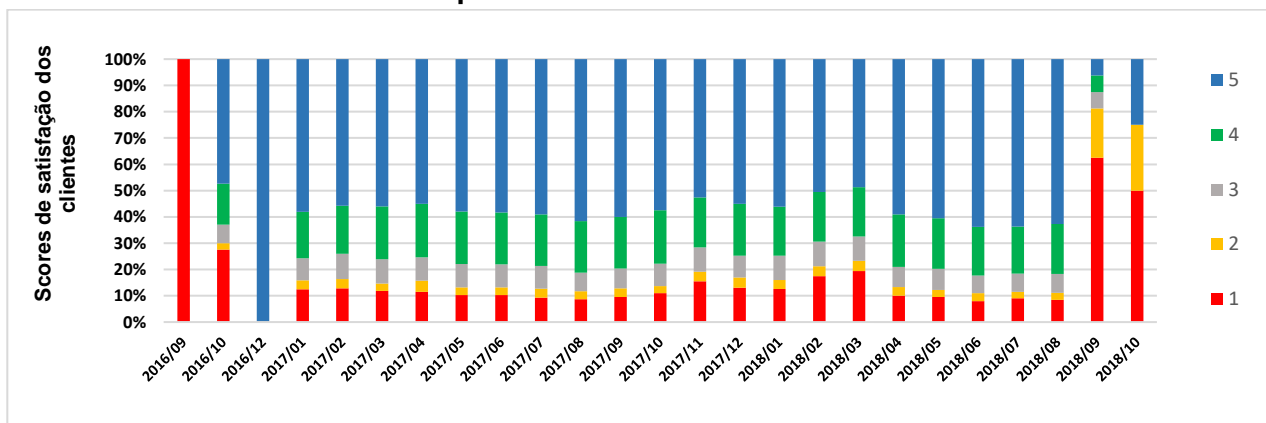
Entre os lojistas que obtiveram maior faturamento durante o período avaliado destacam-se aqueles que venderam produtos da categoria relógios-presentes, móveis-escritório, pcs e cama-mesa-banho.

#### 2.5. SATISFAÇÃO DOS CLIENTES

Entre os clientes que responderam à pesquisa de satisfação 57,42% e 19,2% indicaram escores iguais a 5 e 4, respectivamente. Já 11,85% e 3,24% indicaram escores iguais a 1 e 2, nesta ordem (Figura 6). Um dos fatores relacionados a insatisfação dos clientes é o tempo de entrega do produto. Esta variável apresenta correlação negativa e mediana (-0,33) com os escores fornecidos

pelos clientes na pesquisa de satisfação. Indicando que quanto menor o tempo de entrega maior é o escore indicado pelo cliente.

**Figura 5. Frequência de escores indicadores da satisfação do cliente sobre a compra realizada via plataforma Olist entre 09/2016 a 09/2018**



### 3. Indicadores-chave de desempenho - KPI's

Após análise e descrição dos dados observa-se que algumas variáveis são indicativas do desempenho do negócio e, portanto, devem ser bem monitoradas para gerar respostas satisfatórias e diagnosticar possíveis problemas. Entre estas variáveis podemos destacar o número de itens vendidos e o faturamento mensal, o número de pedidos cancelados, o tempo entre a data de compra e a data de entrega ao cliente e o escore de satisfação do cliente.

### 4. Inconsistências e incompatibilidades entre arquivos de dados

Das 99.441 identificações de pedidos presentes no arquivo `olist_orders_dataset`, 775 não apareciam no arquivo `olist_order_items_dataset`, indicando uma incompatibilidade entre arquivos e possível perda de informações importantes. Destes 775 pedidos, 77,80% estavam indisponíveis e 21,16% foram cancelados. Desta forma, o arquivo `olist_order_items_dataset` utilizado para análises continha 98666 identificações distintas de pedidos com informações de 112650 itens.

Para alguns pedidos listados no arquivo `olist_orders_dataset` foram detectadas datas de entrega à distribuidora menores que as datas de compras e, também, algumas datas de entrega ao cliente menores que a data de entrega na distribuidora. Este erro gera valores negativos ao se trabalhar com diferenças entre datas, como apresentado na tabela 5. Além disso e talvez o maior problema é alterar o valor de tempo médio de entrega calculado.

Existem muitas células vazias ou com informações do tipo "NA's" em muitas variáveis dos bancos de dados, como por exemplo a 610 produtos não tinham informações sobre categorias e muitas datas relacionadas a logística de entrega também estavam faltantes.

Outro ponto verificado que pode dificultar as análises e interpretações dos resultados é o grande número de classes das categorias dos produtos que indicam a mesma categoria.