

## Caso de Data Science – Empresa AliPaga

A empresa AliPaga é um super varejista chinês, que descobriu o eCommerce como forma de vender para todo o mundo, ao invés de vender apenas para a diminuta população de seu país de origem. A empresa veio crescendo mensalmente sem nenhuma preocupação. No entanto recentemente a Amazonense veio disputar o mesmo mercado de AliPaga, e esta viu que o que a trouxe até este momento não é o que vai leva-la para o futuro. E ela decidiu usar as mesmas armas da Amazonense para atingir seu público: a ciência de dados. E para tal ela contratou o melhor recurso de ciência de dados disponível em seu imenso banco de dados de cientistas de todo o mundo: você!

A AliPaga te disponibilizou seu banco de dados (disponível em <https://www.kaggle.com/jr2ngb/superstore-data>) e você terá 7 dias para apresentar suas conclusões para a diretoria da AliPaga. Há uma expectativa de que sejam apresentadas pelo menos as seguintes conclusões:

- 1 Uma segmentação dos países dos clientes da AliPaga, por valor de venda, e representatividade dos clusters dentro do faturamento da mesma
- 2 Uma previsão de vendas para o próximo período mensal
- 3 Qual o impacto do shipping cost e do discount sobre as vendas? Para aumentar as vendas em qual você recomendaria investir?

Além disso qualquer análise relevante de tendências, segmentação etc que venha a ser feita para que a diretoria possa manter as altíssimas taxas de crescimento da AliPaga será considerada para sua análise de desempenho, e ajudará a queimar a Amazonense no mercado.

A apresentação poderá ser em um power point ou qualquer ferramenta que você se sinta confortável, porém vamos querer ver o passo a passo efetuado (notebooks ou qualquer outra programação) para chegar às conclusões.

Boa Sorte!

## About this table

Let us understand the attributes of this data set:

**Row ID** - This is nothing but Serial No. (Isso nada mais é do que o número de série)

**Order ID** - ID created when a product order is placed. (ID criado quando um pedido de produto é feito)

**Order Date** - Date on which a customer places his/her order. (Data em que um cliente faz seu pedido)

**Ship Date** - Date on which the order is shipped. (Data em que o pedido é enviado)

**Ship Mode** - Mode of shipment of each order. (Modo de envio de cada pedido)

**Customer ID** - ID assigned to each customer who places an order. (ID atribuído a cada

cliente que faz um pedido)

**Customer Name** - Name of Customer. (Nome do cliente)

**Segment** - Section from where the order is placed. (Seção de onde o pedido é feito)

**Country**(**ACHO QUE TÁ ERRADA – Acho que são todos os países**) - Country details of this data set. We are looking only for US store data. (Detalhes do país deste conjunto de dados. Estamos procurando apenas os dados da loja dos EUA.)

**City**(**ACHO QUE TÁ ERRADA – Acho que são todas as cidades**) - Cities of US are listed here. (Os estados dos EUA estão listados aqui)

**State**(**ACHO QUE TÁ ERRADA – Acho que são todos os Estados**) - States of US are listed here. (Os estados dos EUA estão listados aqui)

**Postal Code** - pin code. (CEP)

**Region** - grouped into region wise. (agrupados em regiões)

**Product ID** - Product ID of each product. (ID do produto de cada produto)

**Category** - Category to which each product belongs to. (Categoria à qual cada produto pertence)

**Sub-Category** - Sub-Category of each Category. (Sub-categoria de cada categoria)

**Product Name** - Name of products. (Nome dos produtos)

**Sales** - Selling Price of each product. (Preço de venda de cada produto)

**Quantity** - number of quantity available for a particular product. (número de quantidade disponível para um determinado produto)

**Discount** - Discount available on each product. (Desconto disponível em cada produto)

**Profit** - Profit gained on each product. (Lucro ganho em cada produto)

Colunas que não vieram – {'Market' - Mercado, 'Order Priority' - Prioridade de pedido, 'Shipping Cost' - Frete}

<https://www.kaggle.com/juhi1994/superstore>

## Planejamento

08/04 – Análise exploratória dos dados

09/04 – Análise exploratória dos dados

10/04 – Construção do modelo

11/04 – Montar a Apresentação