

# Desafio

Seu desafio é extrair dados de duas bases diferentes - que serão destacadas a seguir - e montar um modelo de dados onde seja possível plugarmos alguma ferramenta de visualização de dados e tirarmos análises simples desses dados.

---

## Bases

Você receberá por email os endereços e informações complementares para acessar as bases abaixo.

- MySQL
    - Será uma base de pedidos contendo informações pertinentes à pedidos
  - [Kinesis Stream](#)
    - É um [stream](#) de dados que recebe eventos relacionados a email, sendo que os eventos esperados são: envio de um email, abertura de um email e click em um email
    - É possível ler desse stream usando diferentes linguagens
      - [Python](#)
      - [Java](#)
      - [Outros](#)
- 

## Exemplos

Abaixo seguem exemplos de linhas (MySQL) e eventos (Kinesis) que você pode encontrar em seu processo de extração desses dados.

### MySQL

- Tabela de produtos

```
desc products;
```

Field	Type	Null	Key	Default	Extra
product_id	varchar(40)	NO	PRI	NULL	
category	varchar(255)	NO		NULL	
sub_category	varchar(255)	NO		NULL	
brand	varchar(255)	NO		NULL	
seller	varchar(255)	NO		NULL	

- Tabela de pedidos

```
desc orders;
```

Field	Type	Null	Key	Default	Extra
order_id	varchar(40)	NO	PRI	NULL	
customer_id	varchar(40)	NO		NULL	
total	decimal(10,5)	NO		NULL	
shipping_price	decimal(10,5)	NO		NULL	
city	varchar(255)	NO		NULL	
state	varchar(255)	NO		NULL	
status	varchar(255)	NO		NULL	
order_date	datetime	NO		NULL	
last_update	datetime	NO		NULL	

- Tabela de itens por pedido

```
desc orderitem;
```

Field	Type	Null	Key	Default	Extra
id	int(11)	NO	PRI	NULL	auto_increment
order_id	varchar(40)	NO	MUL	NULL	
product_id	varchar(40)	NO	MUL	NULL	
selling_price	decimal(10,5)	NO		NULL	

## Kinesis

- Envio de um email

```
{
  "event_id": "2145d4b3-d2be-4659-a0dd-f063da79034a",
  "event_type": "sent",
  "datetime": 14567898765,
  "algorithms": [
    "trends",
    "abandoned-cart"
  ],
  "layout": "material-design-layout.html",
  "partner": "4999",
  "user": 1,
  "subject": "abandonado-com-ofertas",
  "bounce": false,
  "complaint": false,
  "email_provider": "hotmail.com"
}
```

- Abertura de um email

```
{
  "event_id": "2145d4b3-d2be-4659-a0dd-f063da79034a",
  "event_type": "opened",
  "datetime": 14567909765,
  "device": "desktop",
  "os": "mac os"
}
```

- Click em um email

```
{
  "event_id": "2145d4b3-d2be-4659-a0dd-f063da79034a",
  "event_type": "clicked",
  "datetime": 14567909765,
  "device": "mobile",
  "os": "android"
}
```

---

## Análises que podemos fazer

- Quantidade de pedidos fechados por dia / categorias / sub-categorias / marcas
- Venda captada no dia / categorias / sub-categorias / marcas
- Lucro bruto / líquido por dia / categorias / sub-categorias / marcas

- Quantidade de emails enviados / abertos / clicados em um dia

## Requisitos

- Todos os seus processos de ETL devem estar documentados e disponíveis em um repositório como [Github](#), [Bitbucket](#), [Gitlab](#) e etc.
- Documento simples indicando quais as dependências para executarmos os seus procedimentos
- Descrição do ambiente de saída desses dados