

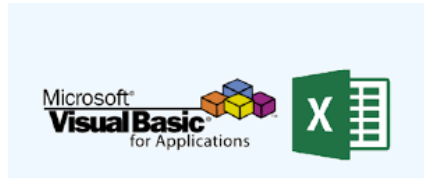
Case OLX - Analista de Dados



Candidato: Raphael da Silva Barradas

Pessoal, esse documento foi elaborado com objetivo de informar as tecnologias/ferramentas utilizados na resolução do desafio proposto.

- **Tecnologias e ferramentas**



Excel + VBA para criar as análises/visualizações.

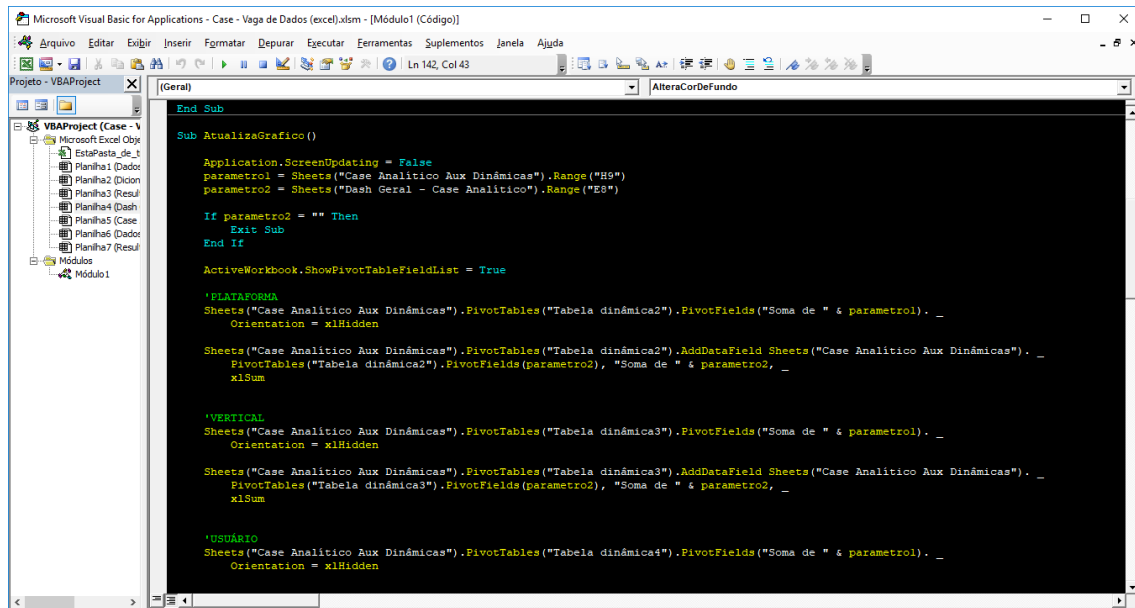


Git para versionar o projeto

GitHub para ser o repositório

- VBA para auxiliar validações de campos, filtros e gráficos dinâmicos.

Os scripts realizados no vba podem ser acessados diretamente na planilha “Case - Vaga de Dados (excel).xslm”, pressionando Alt + F11.



- Excel para visualização dos dados e análises

Foi criada uma aba “Dash Geral - Case Analítico”, para ter uma visão geral das informações e adquirir alguma proximidade com os dados.

Foram criados filtros que podem ser manipulados de forma dinâmica, para obter outros insights sobre os dados.



Como temos 4 métricas, coloquei uma lista de opções para que seja selecionada qual das métricas será exibida nas visualizações:

Anúncios aprovados

- Visitantes
- Anunciantes
- Anúncios aprovados**
- Receita

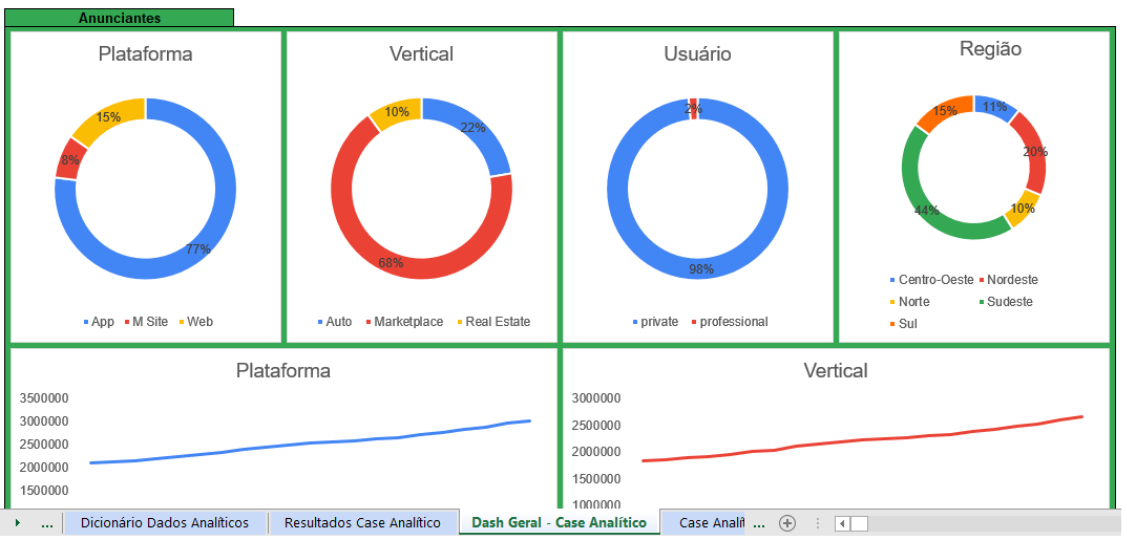
Ao selecionar uma métrica, na lista de opções, a cor de fundo do dashboard também é alterado para a cor referente à métrica.

Ex:

Totalizadores com as cores de cada métrica:

Visitantes	Anunciantes	Anúncios Aprovados	Receita
7.046.911.748	68.566.752	188.897.357	R\$ 179.469.545,12

Métrica “Anunciantes” selecionada:



- **Questões do Case – Dados analíticos**

1. Essa empresa precisa acompanhar a saúde da plataforma periodicamente. Quais três métricas você julgaria mais importantes para acompanhamento?

Acho interessante deixar qualquer tipo de visualização a mais dinâmica possível, com possibilidade de colocar parâmetros ou filtros, como região, plataforma, vertical, etc..

Como a empresa precisa acompanhar periodicamente a saúde da plataforma, julgaria importante utilizar visões em função do tempo, como por exemplo:

- *Volume de receita por período (Dia/Semana/Mês/Etc..)*
- *Volume de anunciantes por período (Dia/Semana/Mês/Etc..)*
- *Volume de visitantes por período (Dia/Semana/Mês/Etc..)*

Obs: Todos essas visões existem na aba “Dash Geral - Case Analítico”.

2. A partir de Setembro/2019, houve uma queda na receita de produtos digitais da empresa. Um engenheiro extraiu para você os dados apresentados na aba "Dados analíticos", que envolvem algumas dimensões e métricas da empresa no período de Janeiro/2018 a Setembro/2019. A aba "Dicionário dados analíticos" contém a descrição da planilha. Com base nesses dados, responda: o que causou a queda?

A partir do dash geral criado, foi possível identificar uma queda no número de anunciantes/anúncios aprovados no mês de setembro/2019, o que refletiu na queda da receita.

Essa queda foi identificada apenas para o cluster abaixo:

Usuário = Profissional

Vertical = Auto

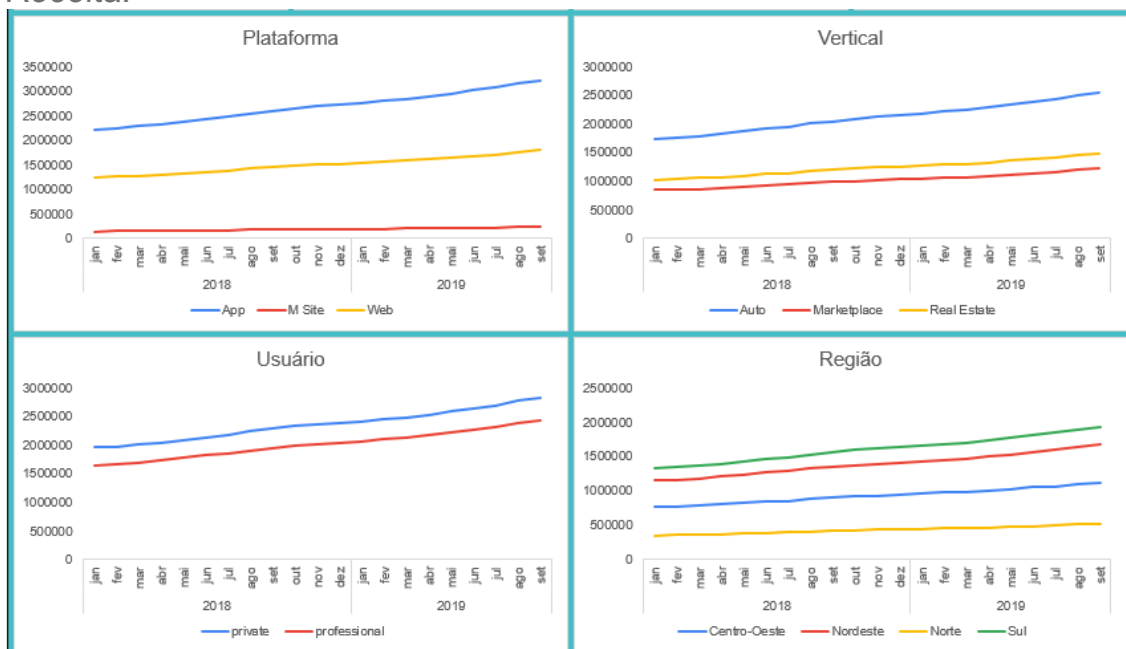
Região = Sudeste

3. Baseado no que você descobriu, prepare uma apresentação para a empresa (a ser realizada em um evento envolvendo desde o estagiário até o diretor) para recomendar os próximos passos.

Como informado na questão 2, a queda na receita ocorreu em um cluster específico (Usuário = Profissional, Vertical = Auto e Região = Sudeste), devido à uma redução no número de anunciantes, que gerou uma quantidade de anúncios aprovados menor e, por consequência, uma queda da receita, para usuário.

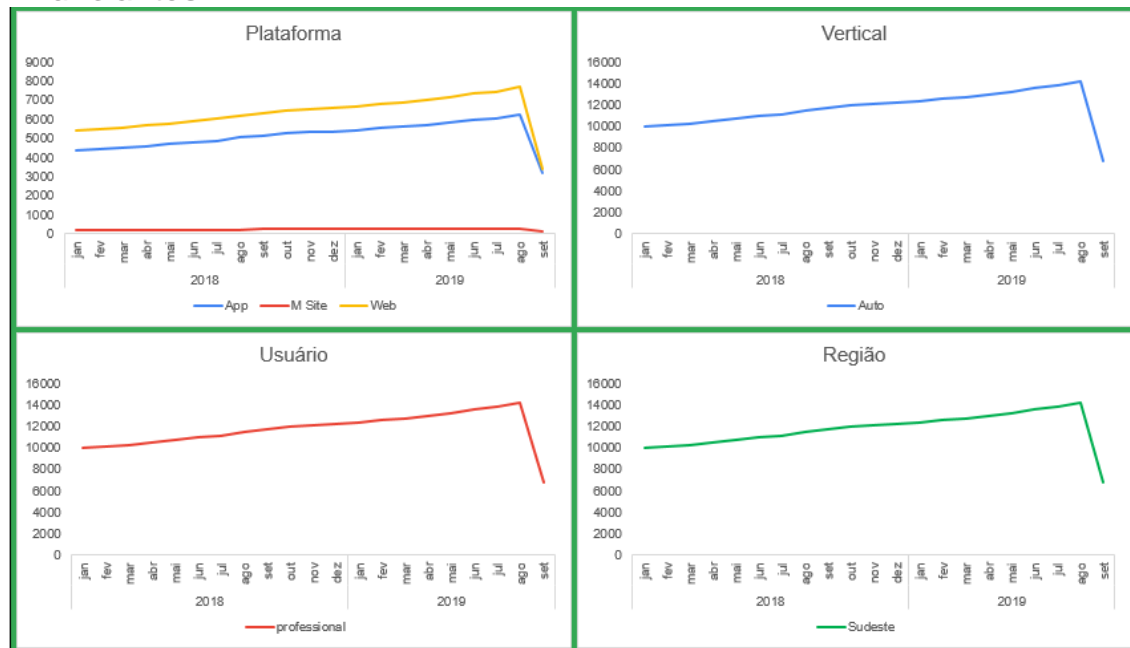
Quando retiramos este cluster e analisamos os demais, não percebemos essa queda de receita, no mês de setembro/2019:

Receita:

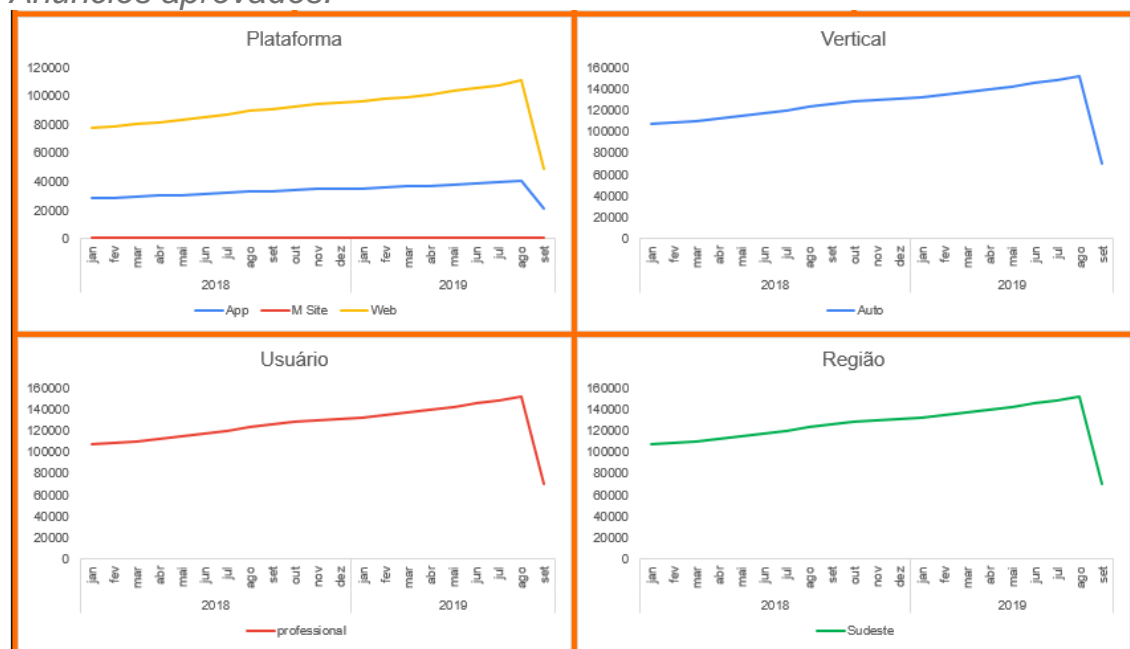


Agora vamos observar filtrando apenas o cluster que causou a queda:

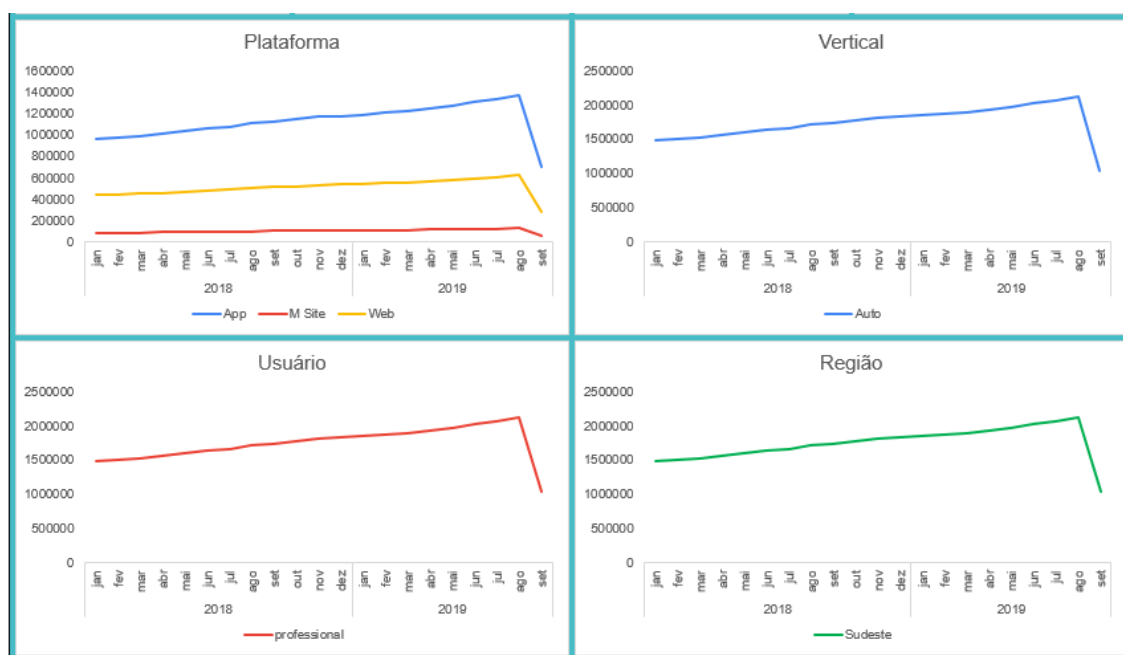
Anunciantes:



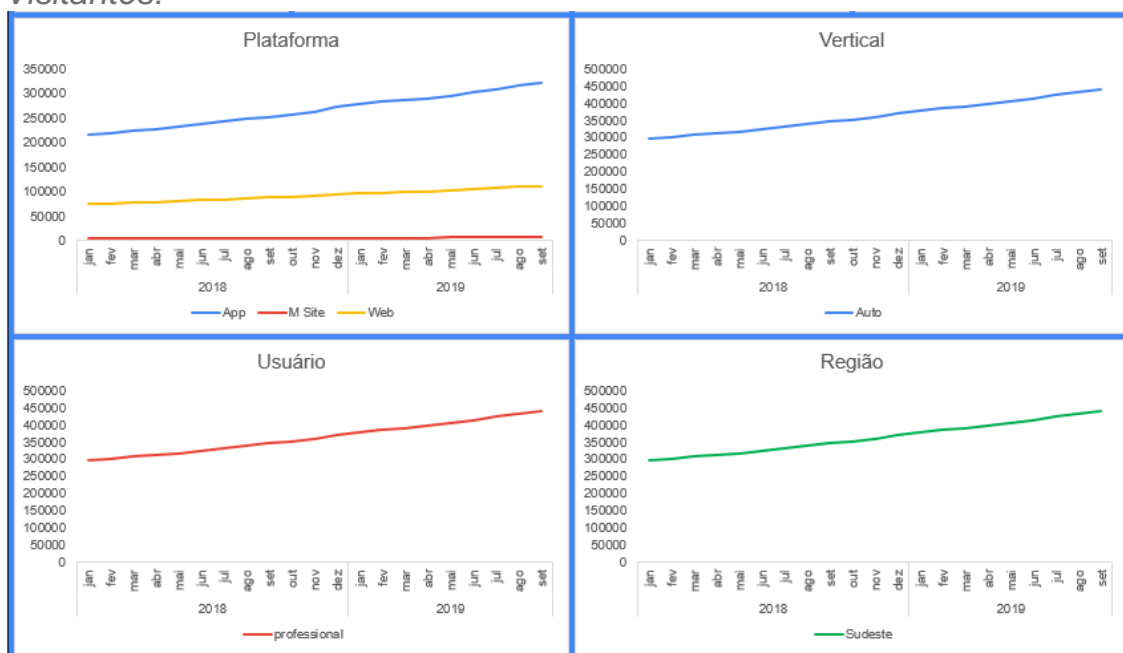
Anúncios aprovados:



Receita:



Visitantes:



Podemos ver que em todas as plataformas (App, M Site e Web), desse cluster, tivemos redução de receita. Porém foi apenas na categoria de “Automóveis e peças de automóveis”, da região Sudeste.

Não houve redução no volume de visitantes.

Como “Próximos passos”, devemos investigar mais profundamente o que ocorreu na região sudeste, para a categoria de automóveis, neste período de setembro/2019, para justificar essa queda nos anúncios e na receita.

Uma sugestão seria enviar uma notificação no App ou e-mail dos clientes do sudeste, que já anunciaram alguma vez nessa categoria de automóveis, com algum tipo de pesquisa, para tentar entender o real motivo.

- **Questões do Case – Dados SQL**

1. A partir desses dados, gostaríamos de ter uma visão do volume de anúncios aprovados nos últimos 60 dias, com preços entre R\$10.000 e R\$100.000, que ainda estão ativos na plataforma, segmentados por estado, DDD, categoria principal e categoria. Apresente a query SQL que retornaria essa visão.

```
SELECT
    T5.state_name AS ESTADO,
    T4.area_code AS DDD,
    T3.category_main_name AS CATEGORIA_PRINCIPAL,
    T2.category_name AS CATEGORIA,
    COUNT(T1.ad_id_pk) AS VOLUME_ANUNCIOS
FROM
    ad AS T1
    LEFT JOIN category AS T2
        ON T2.category_id_pk = T1.category_id_fk
    LEFT JOIN category_main AS T3
        ON T3.category_main_id_pk = T2.category_main_id_fk
    LEFT JOIN location AS T4
        ON T4.location_id_pk = T1.location_id_fk
    LEFT JOIN state AS T5
        ON T5.state_id_pk = T4.state_id_fk
WHERE
    T1.price BETWEEN 10000 AND 100000
    AND T1.approval_date IS NOT NULL
    AND T1.deletion_date IS NULL
GROUP BY
    T5.state_name,
    T4.area_code,
    T3.category_main_name,
    T2.category_name
```

2. E se quiséssemos descobrir o preço médio do primeiro anúncio aprovado de cada anunciante? Apresente a query SQL.

```
WITH primeiro_anuncio AS (  
    SELECT  
        ROW_NUMBER() OVER(PARTITION BY lister_id_nk ORDER BY approval_date)  
    AS NUMBER,  
        lister_id_nk,  
        price,  
        approval_date  
    FROM  
        ad  
    WHERE  
        approval_date IS NOT NULL  
)  
  
SELECT  
    AVG(price) AS MEDIA_PRIMEIRO_ANUNCIO_APROVADO  
FROM  
    primeiro_anuncio  
WHERE  
    NUMBER = 1
```

- **GitHub**

Link do repositório => <https://github.com/raphaeldsb/CaseOLX>

Rio de Janeiro, 08 de fevereiro de 2020

Raphael da Silva Barradas