

Descrição do Projeto

Você trabalha como analista para a empresa de telecomunicações Megaline. TA empresa oferece aos seus cliente planos pré-pagos, Surf e Ultimate. O departamento comercial quer saber quais dos planos dão mais receita para ajustar o orçamento de publicidade.

Você vai realizar uma primeira análise dos planos baseados em uma pequena seleção de clientes. Você terá dados de 500 clientes da Megaline: que clientes são, de onde eles são, qual plano usam, o número de chamadas que eles fizeram e mensagens que eles enviaram em 2018. O seu trabalho é analisar o comportamento dos clientes e determinar quais planos pré-pagos dão mais receita.

Descrição dos planos

Perceba: A Megaline arredonda segundos para minutos, e megabytes para gigabytes. Para **chamadas**, cada chamada individual é arredondada para cima: mesmo se uma chamada tenha durado apenas um segundo, será contado como um minuto. Para **tráfego de web**, sessões individuais de web não são arredondadas para cima. Ao invés disso, o total do mês é arredondado para cima. Se alguém usar 1025 megabytes esse mês, eles serão cobrados por 2 gigabytes.

Surf

1. Preço mensal: \$20
2. 500 minutos mensais, 50 mensagens de texto, e 15 GB de dados
3. Depois de exceder os limites do pacote:
 - 1 minuto: 3 centavos
 - 1 mensagem de texto: 3 centavos
 - 1 GB de dados: \$10

Ultimate

1. Preço mensal: \$70
2. 3000 minutos mensais, 1000 mensagens de texto, e 30 GB de dados
3. Depois de exceder os limites do pacote:
 - 1 minuto: 1 centavo
 - 1 mensagem de texto: 1 centavo
 - 1 GB de dados: \$7

Instruções para completar o projeto

Passo 1. Abra o arquivo de dados e estude as informações gerais

Caminho do arquivo:

`/datasets/megaline_calls.csv` [Baixe o conjunto de dados](#)

[/datasets/megaline_internet.csv](#) Baixe o conjunto de dados
[/datasets/megaline_messages.csv](#) Baixe o conjunto de dados
[/datasets/megaline_plans.csv](#) Baixe o conjunto de dados
[/datasets/megaline_users.csv](#) Baixe o conjunto de dados

Passo 2. Prepare os dados

- Converta os dados para os tipos necessários
- Encontre e elimine erros nos dados

Explique quais erros você encontrou e como você os eliminou. Nota: muitas chamadas têm uma duração de 0,0 minutos. Estas podem ser chamadas perdidas. Fica ao seu critério se você deve pré-processar ou não esses valores; avalie o quanto a ausência deles afetaria os resultados de sua análise.

Para cada usuário, encontre:

- O número de chamadas feitas e minutos usados por mês.
- O número de mensagens de texto enviadas por mês.
- O volume de dados por mês.
- A receita mensal para cada usuário (subtraia o limite do pacote grátis do número total de chamadas, mensagens de texto, e dados; multiplique o resultado pelos valores do plano; adicione o preço mensal dependendo do plano de chamada).

Passo 3. Analise os dados

Descreva o comportamento dos clientes. Encontre os minutos, mensagens de texto e volume de dados que usuários de cada plano necessitam por mês. Calcule a média, variância e o desvio padrão. Construa histogramas. Descreva as distribuições.

Passo 4. Teste as hipóteses

- A receita média dos usuários dos planos Ultimate e Surf são diferentes.
- A receita média dos usuários da área de NY-NJ é diferente dos usuários de outras regiões.

Você decide quais valores alfa usar.

Explique:

- Como você formulou as hipóteses alternativas e nulas.
- Qual critério você usou para testar as hipóteses e porquê.

Passo 5. Escreva uma conclusão geral

Formato. Complete a tarefa em um notebook Jupyter. Coloque o código de programação nas células `code` e explicações de texto nas células `markdown` e aplique a formatação e cabeçalhos.

Descrição dos dados

Lembre-se! A Megaline arredonda segundos para minutos e megabytes para gigabytes. Para **chamadas**, cada chamada individual é arredondada para cima: mesmo se uma chamada tenha durado apenas um segundo, será contado como um minuto. Para **tráfego de web**, sessões individuais de web não são arredondadas para cima.. Ao invés disso, o total do mês é arredondado para cima. Se alguém usar 1025 megabytes esse mês, eles serão cobrados por 2 gigabytes.

A tabela `users` (dados sobre usuários):

- *user_id* — identificação do usuário
- *first_name* — nome do usuário
- *last_name* — último sobrenome do usuário
- *age* — idade do usuário (em anos)
- *reg_date* — data da inscrição (dd, mm, aa)
- *churn_date* — a data que o usuário parou de usar o serviço (se o valor for ausente, o plano estava sendo usado quando esse dado foi gerado)
- *city* — cidade de residência do usuário
- *plan* — nome do plano

A tabela `calls` (dados sobre as chamadas)

- *id* — identificador de chamada unívoco
- *call_date* — data da chamada
- *duration* — duração da chamada (em minutos)
- *user_id* — o identificador do usuário que faz a chamada

A tabela `messages` (dados nas mensagens de texto):

- *id* — identificador unívoco de mensagem de textos
- *message_date* — data da mensagem de texto
- *user_id* — o identificador do usuário que envia a mensagem de texto

A tabela `internet` (dados sobre sessões web):

- *id* — identificador de sessão unívoco

- *mb_used* — o volume de dados gasto durante a sessão (em megabytes)
- *session_date* — data da sessão web
- *user_id* — identificador do usuário

A tabela `plans`(dados sobre os planos):

- *plan_name* — o nome do plano de chamadas
- *usd_monthly_fee* — preço mensal em dólares dos EUA
- *minutes_included* — pacote de minutos mensal
- *messages_included* — pacote de mensagens de texto mensal
- *mb_per_month_included* — pacote de volume de dados (em megabytes)
- *usd_per_minute* — preço por minuto depois de exceder o limite do pacote (por exemplo, se o pacote inclui 100 minutos, o primeiro minuto excedente será cobrado)
- *usd_per_message* — preço por mensagem de texto depois de exceder o limite do pacote
- *usd_per_gb* — preço por gigabyte extra de dados após exceder o limite do pacote (1 GB = 1024 megabytes)

Como o meu projeto será avaliado?

Nós elaboramos alguns critérios de avaliação dos projetos. Leia-os com cuidado antes de você começar a trabalhar.

Aqui está o que os revisores de projeto irão ver quando forem avaliar o seu projeto:

- Como você explica os problemas identificados nos dados.
- Como você prepara os dados para análise.
- Quais gráficos você fez para as distribuições.
- Como você interpreta os gráficos resultantes.
- Como você calcula o desvio padrão e variância.
- Se você criou hipóteses alternativas e nulas.
- Quais métodos você usa para testar hipóteses.
- Se você interpreta os resultados dos seus testes de hipóteses.
- Se você segue a estrutura do projeto e mantém o código organizado.
- As conclusões que você tirou.
- Se você deixa comentários a cada passo.

As folhas de conclusões e os resumos das lições anteriores têm tudo o que você precisa para concluir o projeto.

Boa sorte!