



Descrição do projeto

Você trabalha como analista para a empresa de telecomunicações **Megaline**. A empresa oferece aos clientes dois planos pré-pagos: **Surf** e **Ultimate**. O departamento comercial quer saber qual dos planos gera mais receita para ajustar o orçamento de publicidade.

Você vai realizar uma análise preliminar dos planos com base em uma pequena seleção de clientes. Você terá dados de 500 clientes da Megaline: quais clientes são, de onde eles são, qual plano usam e o número de chamadas e mensagens realizadas em 2018. Seu trabalho é:

- Analisar o comportamento dos clientes e determinar qual plano pré-pago gera mais receita. Mais tarde, nas instruções do projeto, você verá exatamente quais aspectos do comportamento dos clientes precisa analisar;
- Determinar qual plano, em média, gera mais receita é uma tarefa que pode ser resolvida usando testes estatísticos. Você vai receber mais informações sobre isso na seção de instruções do projeto;

Descrição dos planos:

Surf:

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

Ultimate:

- Preço mensal: \$70;
- 3.000 minutos mensais, 1.000 mensagens de texto e 30 GB de dados;
- Após exceder os limites do pacote:
 - 1 minuto: 1 centavo
 - 1 mensagem de texto: 1 centavo
 - 1 GB de dados: \$7

Observação: a Megaline arredonda segundos para minutos e megabytes para gigabytes. Para chamadas, cada chamada individual é arredondada para cima: mesmo que uma chamada tenha durado apenas um segundo, um minuto será contado. Para tráfego da web, sessões individuais da web não são arredondadas para cima. Ao invés disso, o total do mês é arredondado para cima. Se alguém usar 1.025 megabytes no mês, a cobrança será de 2 gigabytes.

Neste projeto, você vai trabalhar com cinco tabelas separadas.

Dicionário de dados:

Tabela users (dados sobre usuários):

- *user_id* — identificador exclusivo do usuário;
- *first_name* — nome do usuário;
- *last_name* — 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 estiver ausente, isso significa que o plano estava em uso quando o banco de dados foi extraído);
- *city* — cidade de residência do usuário;
- *plan* — nome do plano;

Tabela calls (dados sobre as chamadas):

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

Tabela messages (dados sobre mensagens de texto):

- *id* — identificador exclusivo da mensagem de texto;
- *message_date* — data da mensagem de texto;
- *user_id* — identificador do usuário que envia a mensagem de texto;

Tabela internet (dados sobre sessões web):

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

Tabela plans (dados sobre os planos):

- *plan_name* — nome do plano;
- *usd_monthly_fee* — preço mensal em dólares americanos;
- *minutes_included* — pacote mensal de minutos;
- *messages_included* — pacote mensal de mensagens de texto;
- *mb_per_month_included* — volume do pacote 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 = 1.024 megabytes);