

# Projeto:

# Análise da Saúde Financeira da Empresa Netflix

Walter Trevisan

The Netflix logo, consisting of the word "NETFLIX" in a bold, white, sans-serif font with a slight 3D effect, set against a solid red rectangular background.

# Perfil da Empresa

**Netflix** é o maior serviço de entretenimento de streaming do mundo, com mais de **195 milhões (novembro/2020)** de *clientes pagantes* em **mais de 190 países**, desfrutando de séries de TV, documentários e filmes de longa-metragem numa grande variedade de géneros e línguas. Os clientes podem assistir o quanto quiserem, a qualquer hora, em qualquer lugar, em qualquer tela ligada à Internet. Os clientes podem reproduzir, pausar e recomeçar a assistir, tudo sem anúncios ou compromissos.

The Netflix logo, consisting of the word "NETFLIX" in white, bold, sans-serif capital letters, set against a red rectangular background.

# Problema de Negócio

O **objetivo deste projeto** é preparar um resumo estatístico dos dados financeiros da empresa “**Netflix**”, demonstrando sua saúde financeira e outras informações relevantes, que poderão ajudar potenciais investidores a decidir sobre investir ou não na companhia.

Analizamos o volume de “**Clientes**”, as “**Receitas**”, os “**Custos**”, os “**Rendimentos**” e os “**Lucros/Margens**” no período compreendido de **2012** até **2018** para o segmento “*Domestic Streaming*” (Transmissão Doméstica).



# Solução do Projeto

Neste projeto utilizamos a **linguagem R** para o desenvolvimento da solução que foi executada em **4 fases**, a saber:

1. **Fase 1:** Problema de Negócio
2. **Fase 2:** Obter os Dados
3. **Fase 3:** Análise Exploratória dos Dados
4. **Fase 4:** Inferência Estatística sobre os Dados



# Solução do Projeto

## Fase 1: Problema de Negócio

- Nesta fase explicamos do “**problema de negócio**” a ser resolvido, assim como os **objetivos** a serem alcançados para o projeto.

## Fase 2: Obter os Dados

- Nesta fase obtemos e preparamos os dados que foram utilizados durante o desenvolvimento e execução do projeto, ou seja, nesta fase fizemos todas as atividades relacionadas ao **pré-processamento** dos dados, como **limpeza** e **transformação** (“*Feature Engineering*”).

## Fase 3: Análise Exploratória dos Dados

- Nesta fase realizamos uma “**análise exploratória**” dos dados, ou seja, estudamos cada atributo (variável) e suas características, e criamos algumas visualizações (gráficos) dos dados.

## Fase 4: Inferência Estatística

- Nesta fase realizamos alguns **testes estatísticos** para fazermos algumas **inferências** sobre os dados. Estas inferências nos auxiliarão a responder algumas questões sobre a saúde financeira da empresa Netflix.

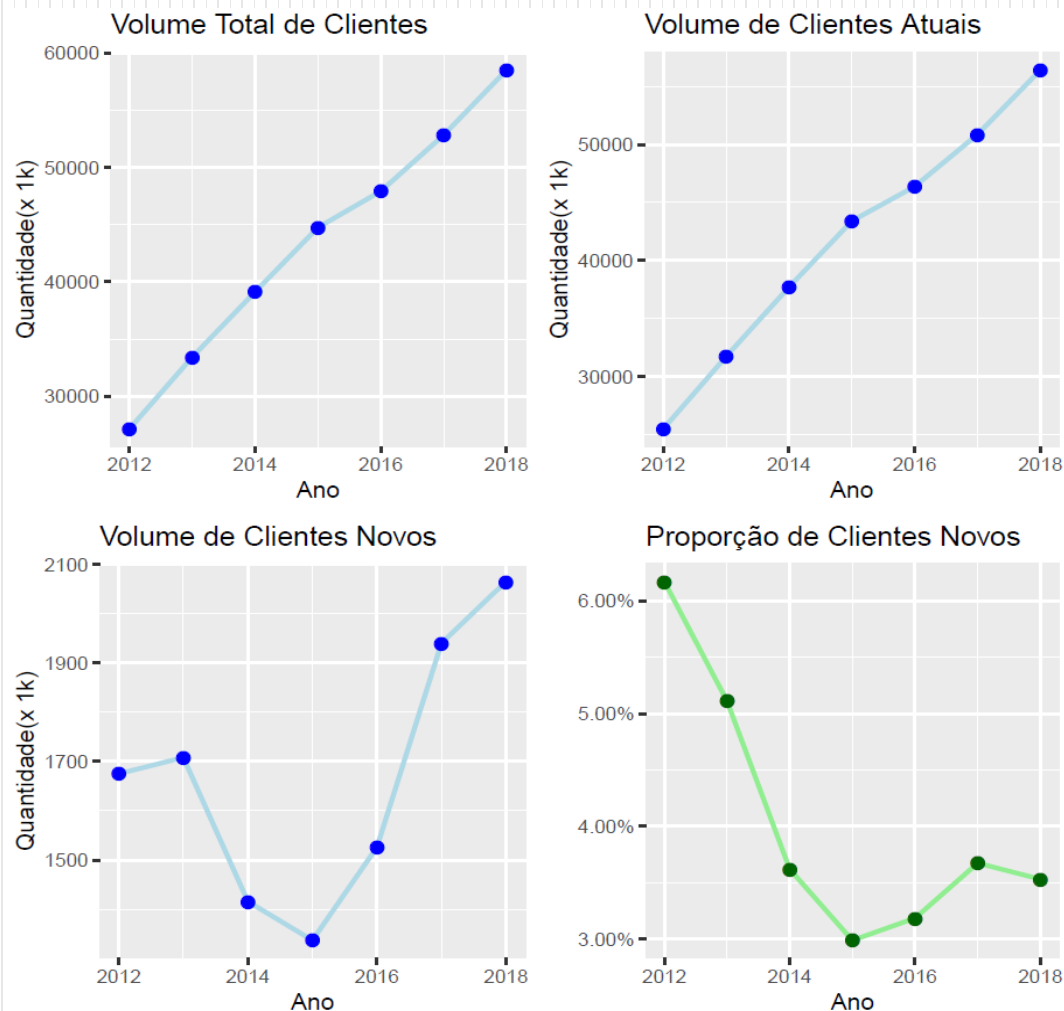
# Análise Exploratória dos Dados (A.E.D.)

A base de dados (***dataset***) utilizada possui um total de **17** atributos (**variáveis**), com informações relacionadas aos “**clientes**”, as “**receitas**”, os “**custos**”, os “**rendimentos**” e aos “**lucros/margens**” da empresa, referente ao **período de 7 anos**, ou seja, de “**2012**” até o ano de “**2018**”. Nesta fase analisamos as informações por “**ano**” e por “**trimestre**”.

As informações sobre os “**clientes**” foram separadas em **3 variáveis**:

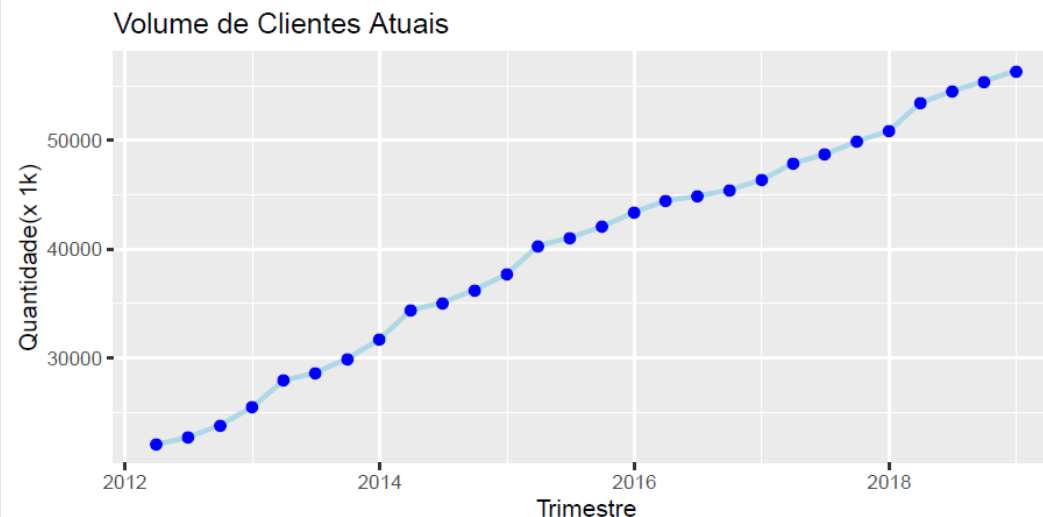
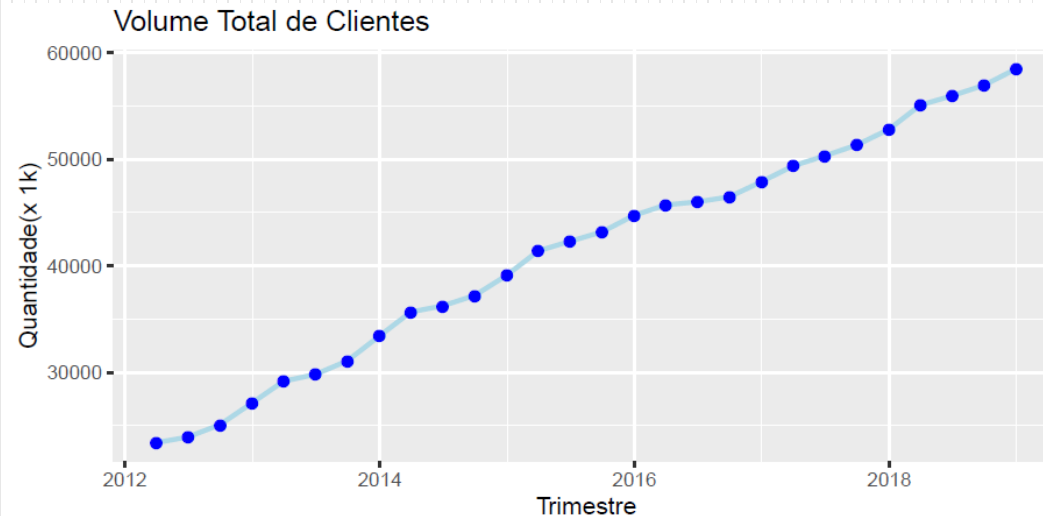
- **Volume Total de Clientes:** é a soma do volume de **clientes atuais** (que pagam suas assinaturas) com o volume de **clientes novos** (que não pagam suas assinaturas);
- **Volume de Clientes Atuais:** são os clientes que pagam suas assinaturas;
- **Volume de Clientes Novos:** são os clientes que não pagam as suas assinaturas, porque ainda estão dentro do período de avaliação gratuita do plano escolhido.

# A. E. D.: Clientes por Ano



- O **volume total de clientes** *aumentou* em todo o período, o mesmo ocorrendo com o **volume de clientes atuais**, sendo que neste caso o aumento foi de “~127%” em **7 anos** (2012 até 2018), ou seja, de ~25 milhões de clientes em **2012** para ~56 milhões de clientes em **2018**;
- O **volume de clientes novos** teve um aumento acentuado nos anos de 2017 e 2018, provavelmente devido alguma campanha de marketing para captar mais clientes. Entretanto, a **proporção dos clientes novos** em relação ao volume total de clientes, *caiu* de ~6% em **2012** para ~3.5% em **2018**. Isto ocorreu, provavelmente, devido ao crescimento progressivo do volume de clientes atuais (*pagantes*), ou seja, entendemos que a maioria dos clientes novos continuam com o plano após o término do período de “avaliação gratuita”.

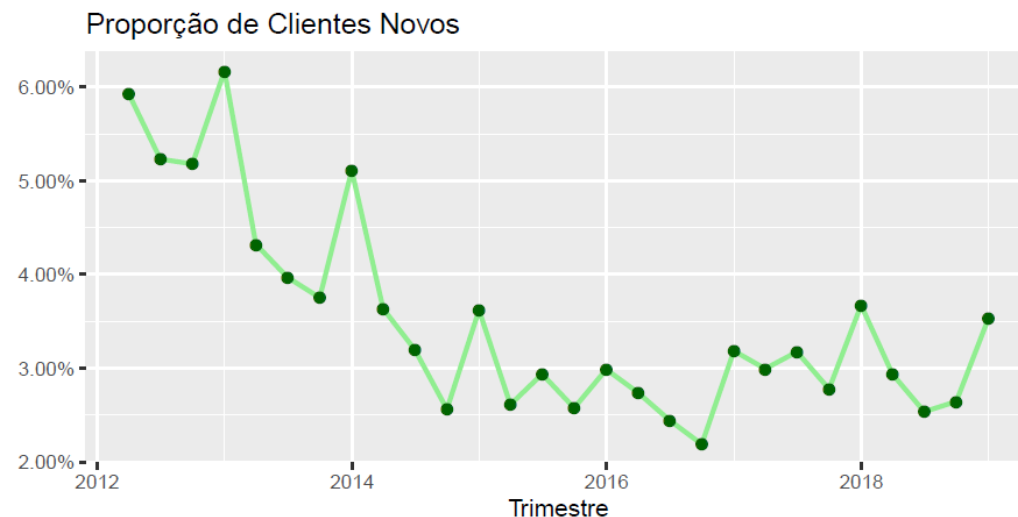
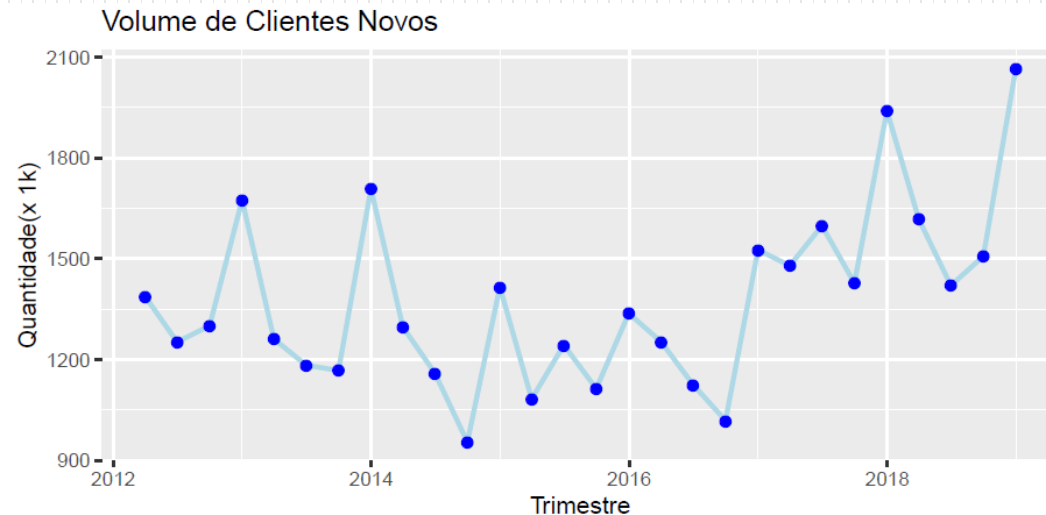
# A. E. D.: Clientes por Trimestre



- Os gráficos mostram que tanto o **volume total de clientes**, quanto o **volume de clientes atuais** (que pagam suas assinaturas) cresceram também em cada trimestre, ou seja, mantiveram um crescimento progressivo em todo o período **(2012-2018)**.

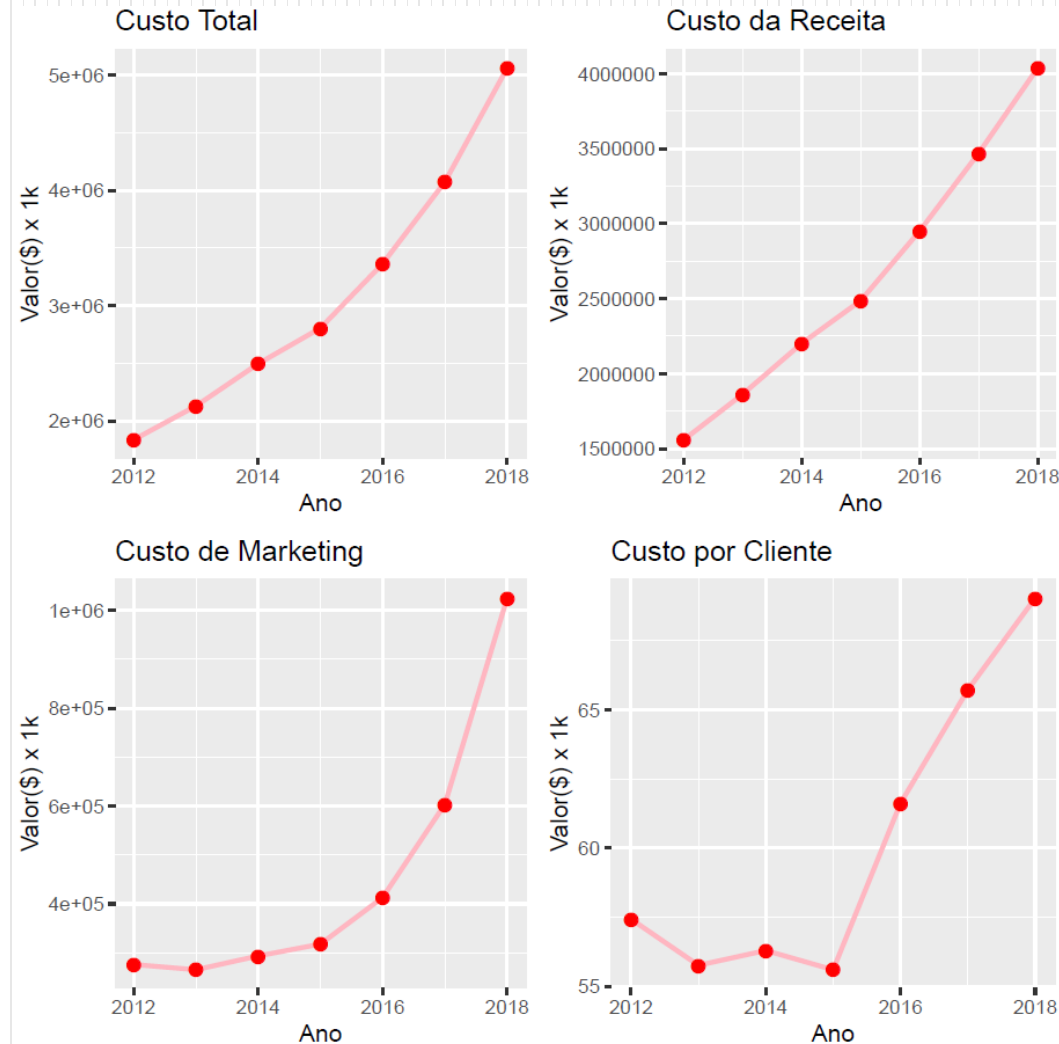


# A. E. D.: Clientes por Trimestre



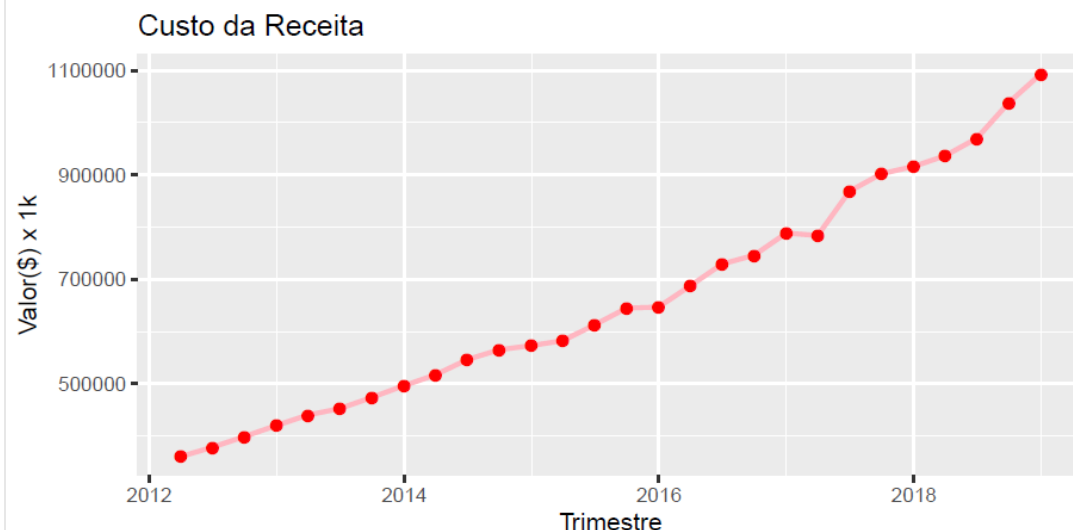
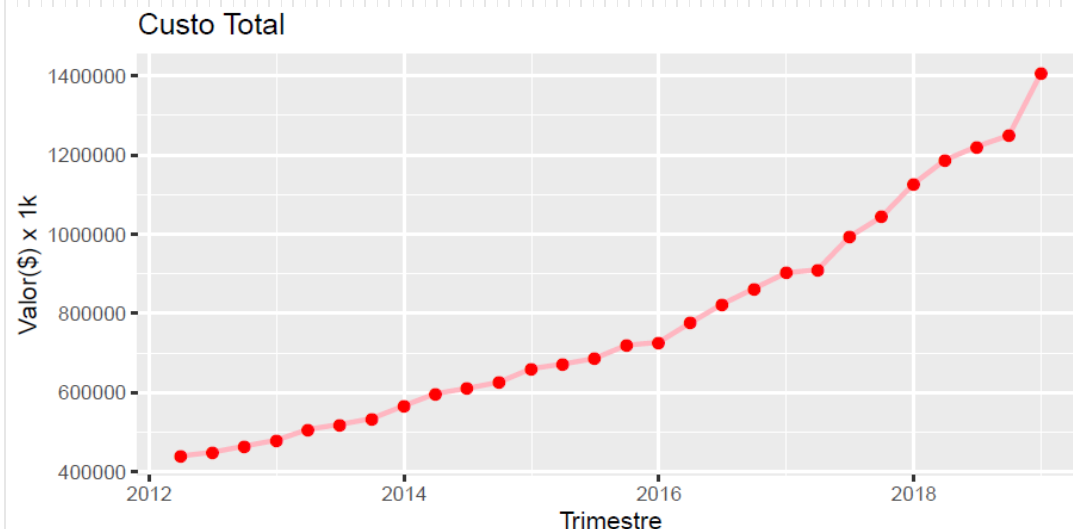
- Podemos observar que o **volume de clientes novos sempre *aumentou*** no último trimestre (Q4) em todos os anos;
- O mesmo ocorreu com a **proporção de clientes novos**, exceto durante o ano de **2015** em que a *proporção se manteve estável* entre os trimestres, ou seja, aproximadamente entre **2.5% e 3.0%**.

# A. E. D.: Custos por Ano



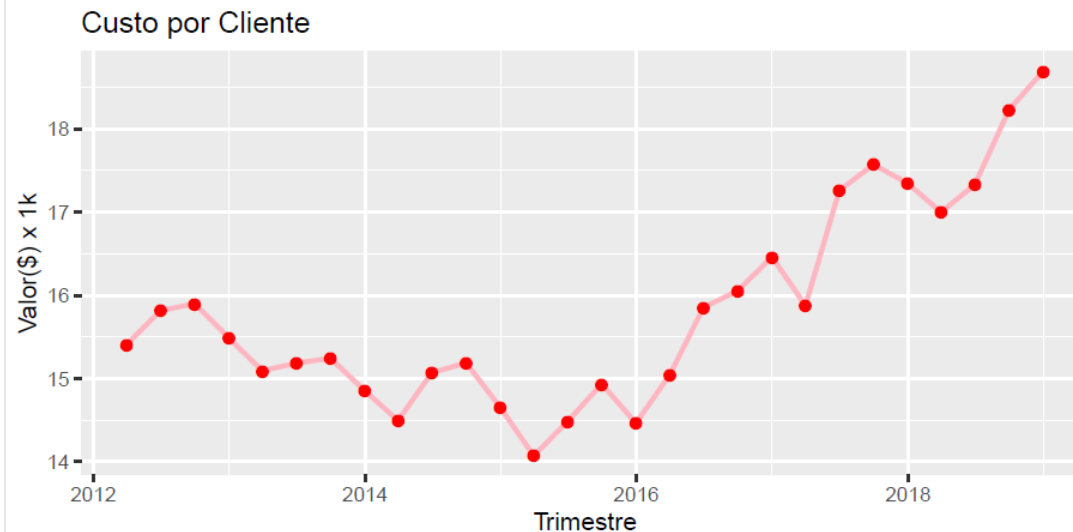
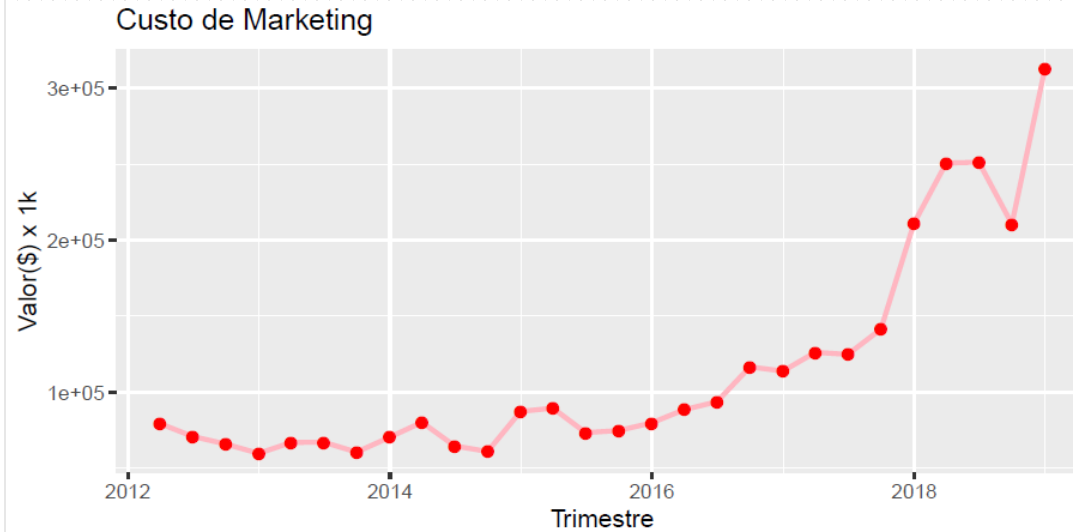
- O **custo total** e o **custo da receita** aumentaram com o passar anos, muito provavelmente devido ao aumento do volume de clientes que cresceu no mesmo período. Entretanto, notamos um aumento mais acentuado a partir de **2016**;
- Observamos que o **custo de marketing** se manteve constante até o ano de **2015**. Entretanto, notamos um aumento mais acentuado a partir de **2016**. Isto também aconteceu com o **custo por cliente**.

# A. E. D.: Custos por Trimestre



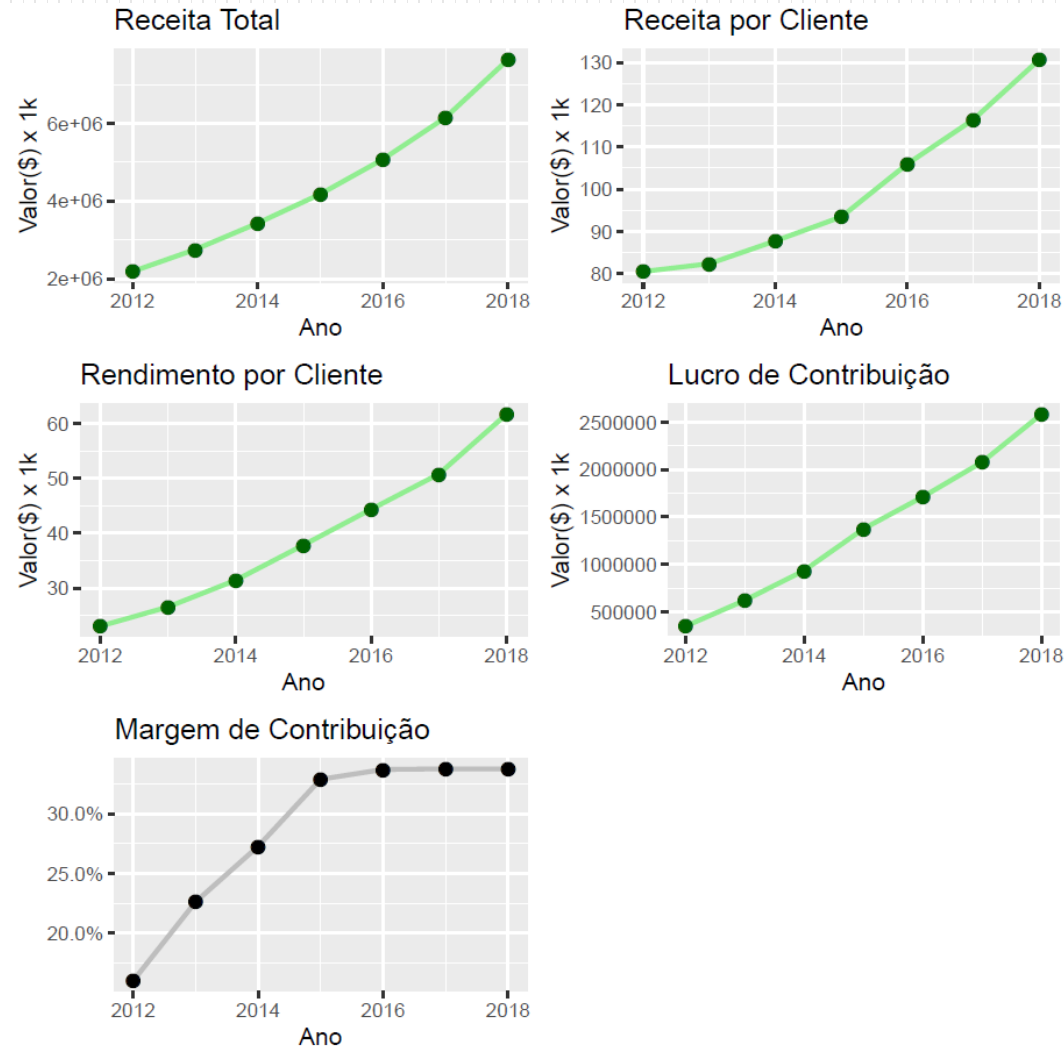
- O **custo total** e o **custo da receita** aumentaram a cada trimestre, muito provavelmente devido ao aumento do volume de clientes que cresceu no mesmo período. Entretanto, notamos um aumento mais acentuado a partir de **Q1/2016** e **Q2/2017**.

# A. E. D.: Custos por Trimestre



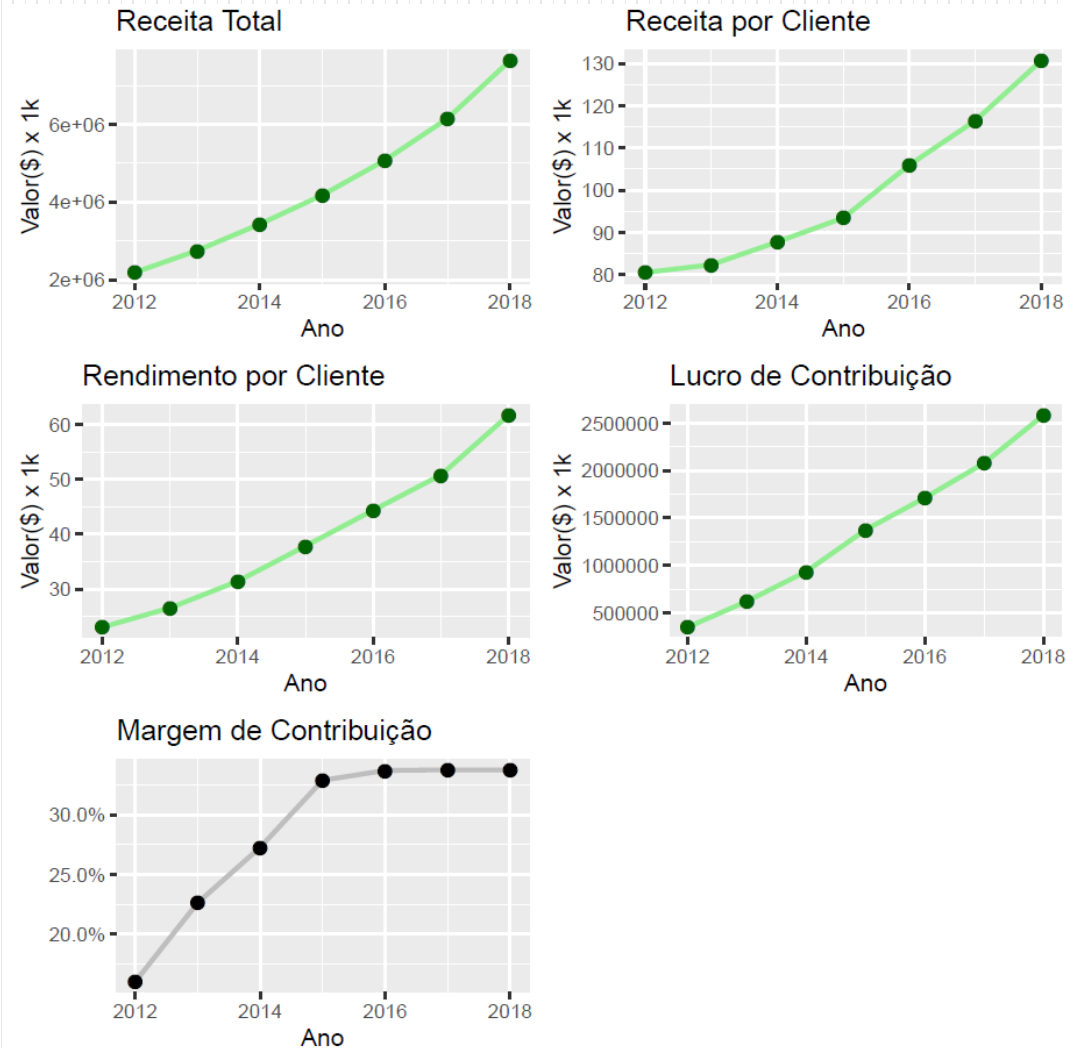
- A partir de **Q1/2017** os **custos de marketing** tiveram um *aumento acentuado* até o final do período (**Q4/2018**). Isto coincide com o aumento do volume de clientes novos no mesmo período, ou seja, provavelmente a empresa aumentou suas campanhas de marketing para captar novos clientes no mercado;
- Apesar de observarmos um aumento progressivo no volume de clientes (total e atuais), o **custo por cliente** se manteve praticamente constante de **Q1/2012** a **Q1/2017**, com um aumento acentuado dos custos a partir de **Q2/2017**. No cálculo do custo por cliente não é considerado os custos com marketing.

# A. E. D.: Receitas/Rendimentos/Lucros por Ano



- Podemos observar nos primeiros 4 gráficos que as **receitas**, os **rendimentos** e os **lucros** *aumentaram* com o passar dos anos, muito provavelmente devido ao aumento do volume de clientes que cresceu no mesmo período. Entretanto, a **margem de contribuição** *aumentou de forma acentuada* nos primeiros 3 anos (entre **2012** e **2015**) e depois se manteve estável e constante até o final do período (entre **2016** e **2018**).

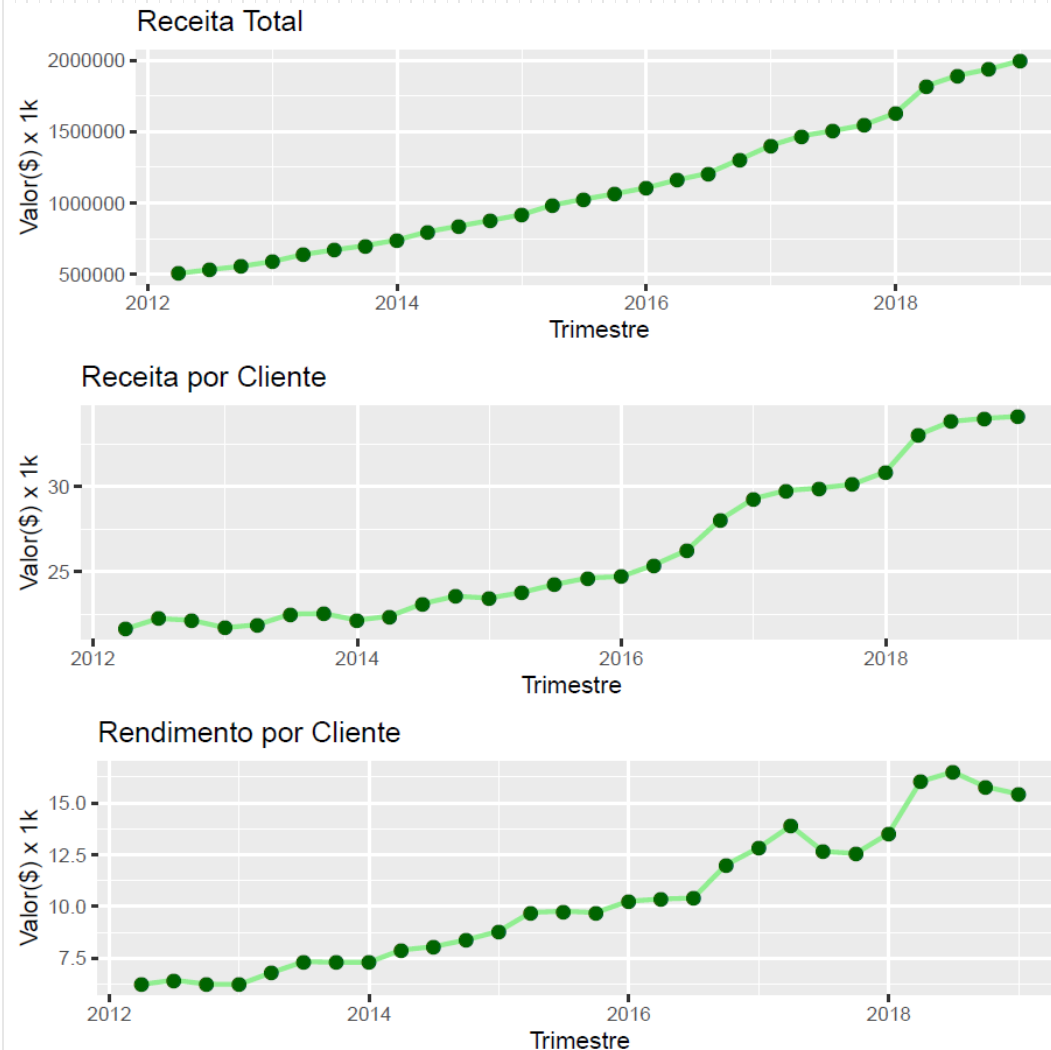
# A. E. D.: Receitas/Rendimentos/Lucros por Ano



Considerando o período de **7 anos**, temos:

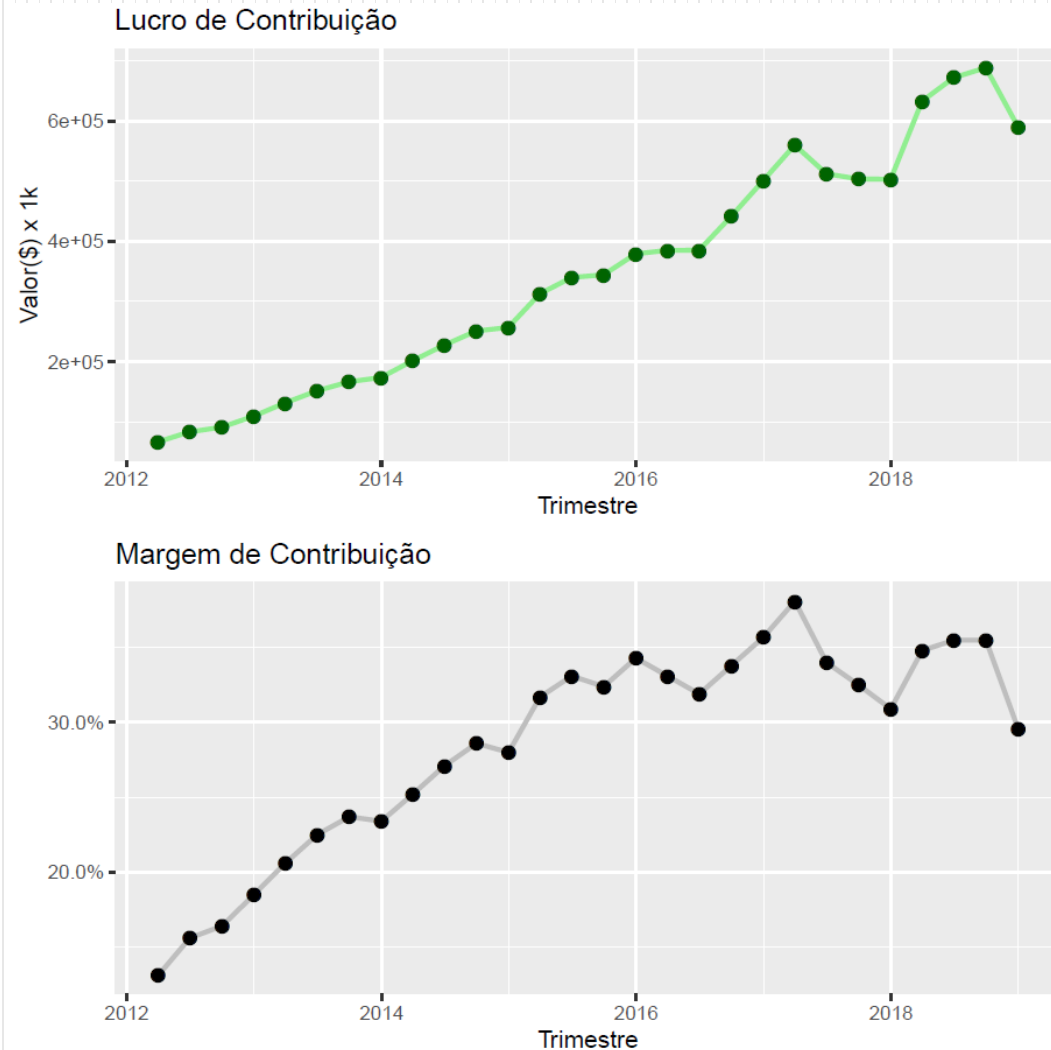
- **Receita Total:** *aumento* de ~**250%**, ou seja, de ~\$**2.2 bilhões** em **2012** para ~\$**7.6 bilhões** em **2018**;
- **Receita por Cliente:** *aumento* de ~**62%**, ou seja, de ~\$**80 mil** em **2012** para ~\$**131 mil** em **2018**;
- **Rendimento por Cliente:** *aumento* de ~**167%**, ou seja, de ~\$**23 mil** em **2012** para ~\$**62 mil** em **2018**;
- **Lucro de Contribuição:** *aumento* de ~**638%**, ou seja, de ~\$**350 milhões** em **2012** para ~\$**2.6 bilhões** em **2018**;
- **Margem de Contribuição:** *aumento* de ~**111%**, ou seja, de ~**16%** em **2012** para ~**34%** em **2018**;

# A. E. D.: Receitas/Rendimentos por Trimestre



- A **receita total** *aumentou* a cada trimestre (em todos os anos), muito provavelmente devido ao aumento do volume de clientes que cresceu no mesmo período;
- A **receita por cliente** *aumentou lentamente* a cada trimestre até **Q2/2016**. A partir de **Q3/2016** até o final do período (**Q4/2018**) o *aumento foi mais acentuado*;
- O **rendimento por cliente** também *aumentou progressivamente* até **Q1/2017**. Em seguida, os rendimentos *caíram um pouco* em **Q2/2017** e **Q3/2017**, crescendo novamente a partir de **Q4/2017** até o final do período.

# A. E. D.: Lucros/Margens por Trimestre



- Os **lucros e margens de contribuição** também *cresceram progressivamente* em todo o período (2012 a 2018);
- Em **Q1/2017** a **margem de contribuição** atingiu o seu *melhor resultado* em todo o período analisado, ou seja, **~37%**.



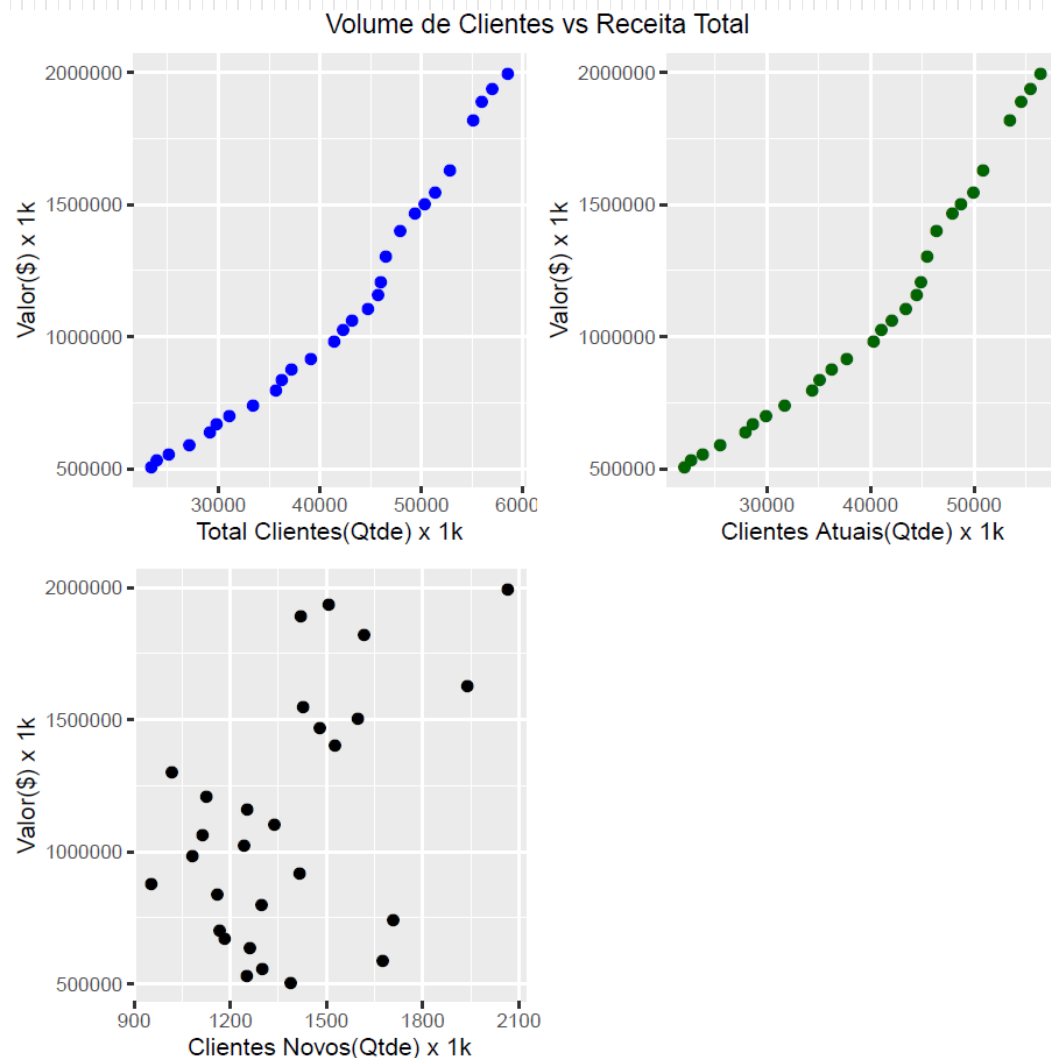
# Inferência Estatística (I.E.)

Nesta fase, realizamos alguns **testes estatísticos** para respondermos algumas questões sobre os dados, por exemplo:

- Se existe alguma **correlação**, ou seja, alguma **relação de dependência** entre as **variáveis**;
- Se os **volumes de clientes** *aumentaram e/ou diminuíram*;
- Se os **custos** *aumentaram e/ou diminuíram*;
- Se as **receitas e rendimentos** *aumentaram e/ou diminuíram*;
- Se os **lucros e margens** *aumentaram e/ou diminuíram*.

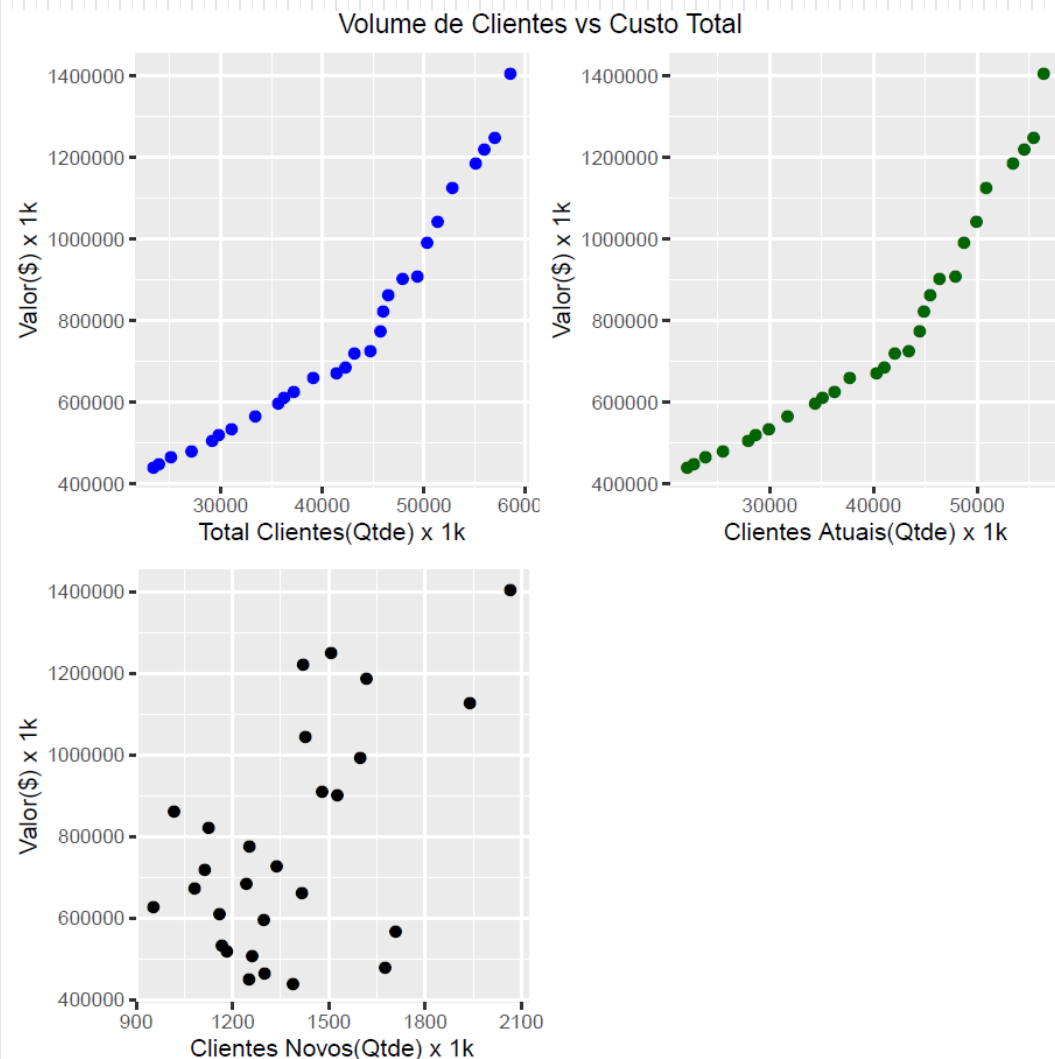
Em cada etapa apresentaremos os **resultados** e as **conclusões** (*insights*) sobre os dados.

**I. E.:** Existe alguma relação de dependência (considerando todo o período) entre a **receita total** da empresa e o **volume total de clientes**, o **volume de clientes atuais** e o **volume de clientes novos**?



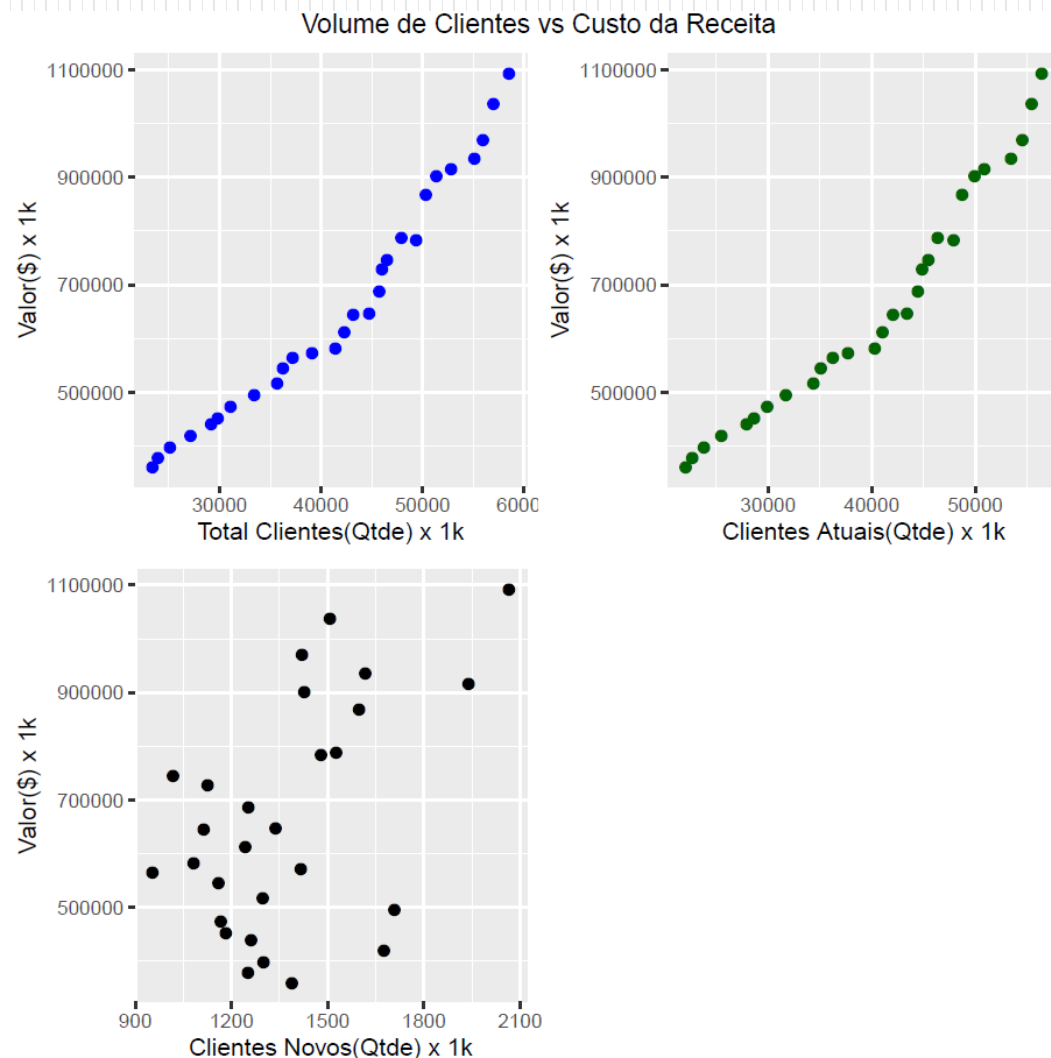
- **Total Clientes x Receita Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume total de clientes aumenta a receita total da empresa. Portanto, a receita total é dependente do volume total de clientes;
- **Clientes Atuais x Receita Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes atuais aumenta a receita total da empresa. Portanto, a receita total é dependente do volume de clientes atuais;
- **Clientes Novos x Receita Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes novos aumenta a receita total da empresa. Portanto, a receita total é dependente do volume de clientes novos.

**I. E.:** Existe alguma relação de dependência (considerando todo o período) entre o **custo total** da empresa e o **volume total de clientes**, o **volume de clientes atuais** e o **volume de clientes novos**?



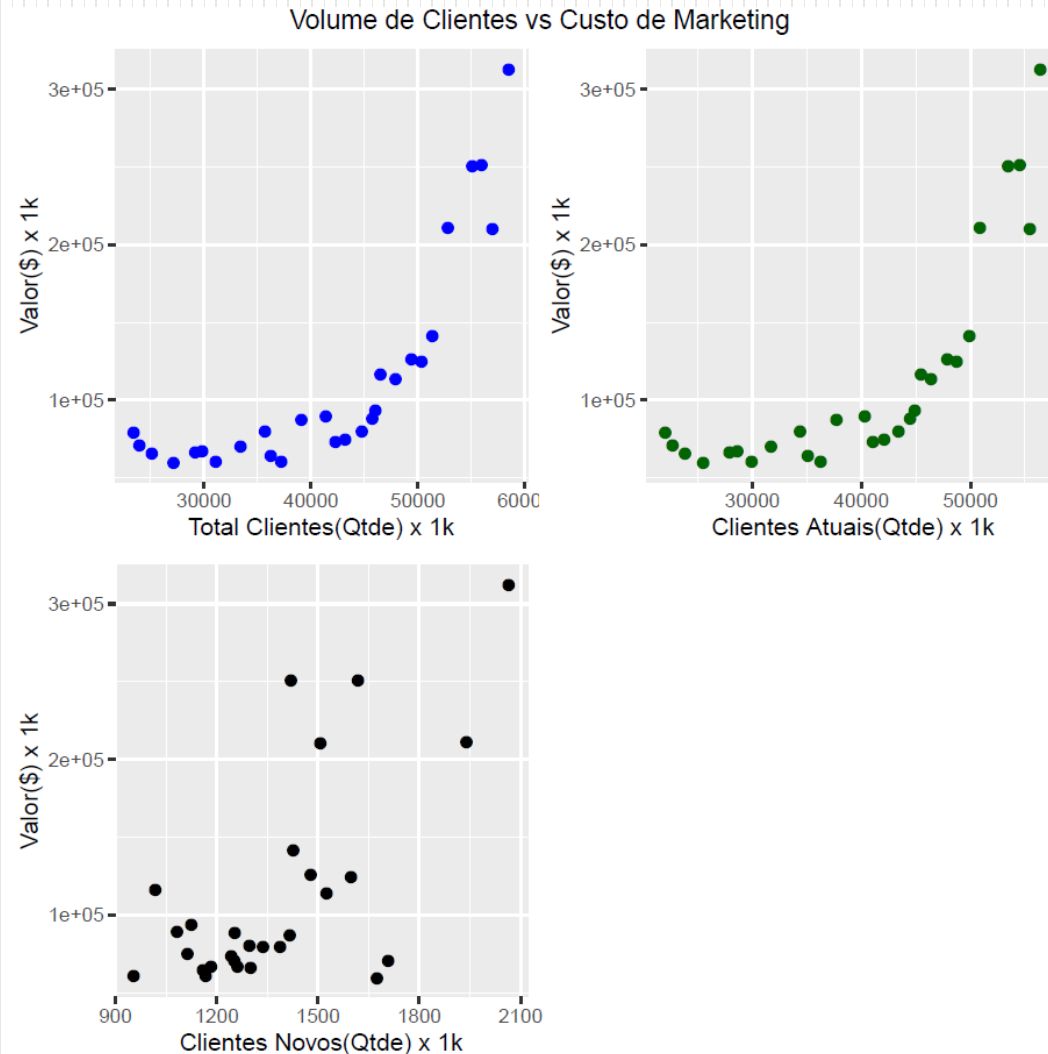
- **Total Clientes x Custo Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume total de clientes aumenta o custo total da empresa. Portanto, o custo total é dependente do volume total de clientes;
- **Clientes Atuais x Custo Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes atuais aumenta o custo total da empresa. Portanto, o custo total é dependente do volume de clientes atuais;
- **Clientes Novos x Custo Total:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes novos aumenta o custo total da empresa. Portanto, o custo total é dependente do volume de clientes novos.

**I. E.:** Existe alguma relação de dependência (considerando todo o período) entre o **custo da receita** da empresa e o **volume total de clientes**, o **volume de clientes atuais** e o **volume de clientes novos**?



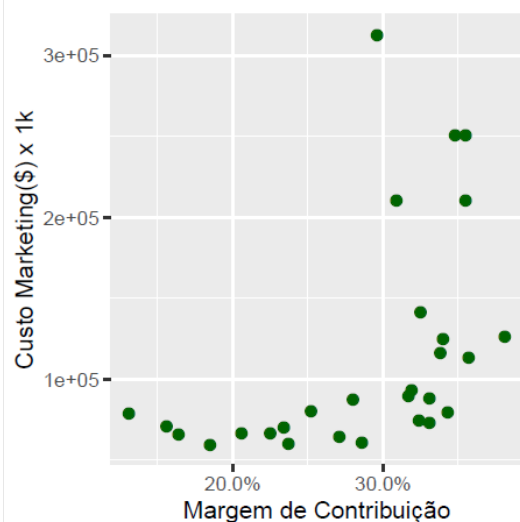
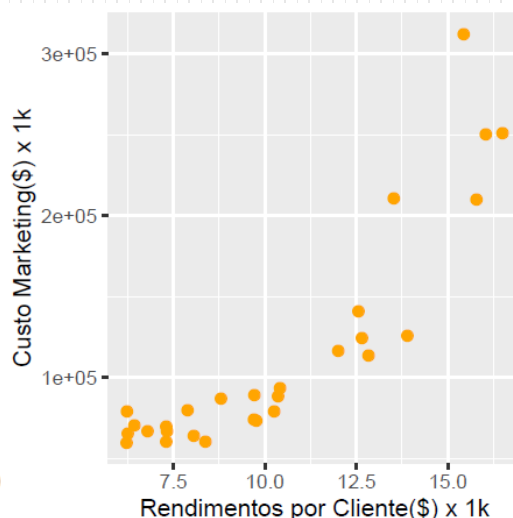
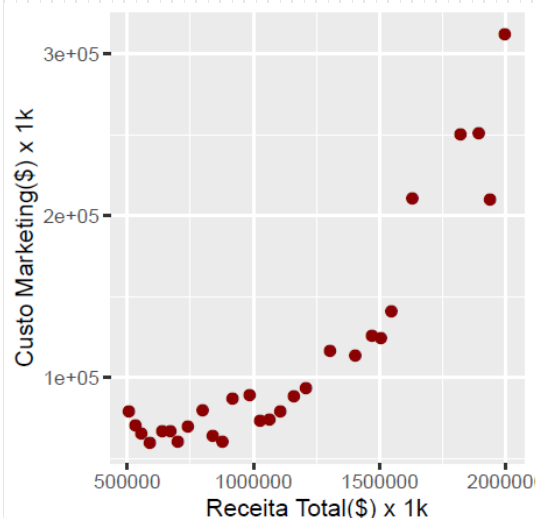
- **Total Clientes x Custo da Receita:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume total de clientes aumenta o custo da receita. Portanto, o custo da receita é dependente do volume total de clientes;
- **Clientes Atuais x Custo da Receita:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes atuais aumenta o custo da receita. Portanto, o custo da receita é dependente do volume de clientes atuais;
- **Clientes Novos x Custo da Receita:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o volume de clientes novos aumenta o custo da receita. Portanto, o custo da receita é dependente do volume de clientes novos.

**I. E.:** Existe alguma relação de dependência (considerando todo o período) entre o **custo de marketing** da empresa e o **volume total de clientes**, o **volume de clientes atuais** e o **volume de clientes novos**?



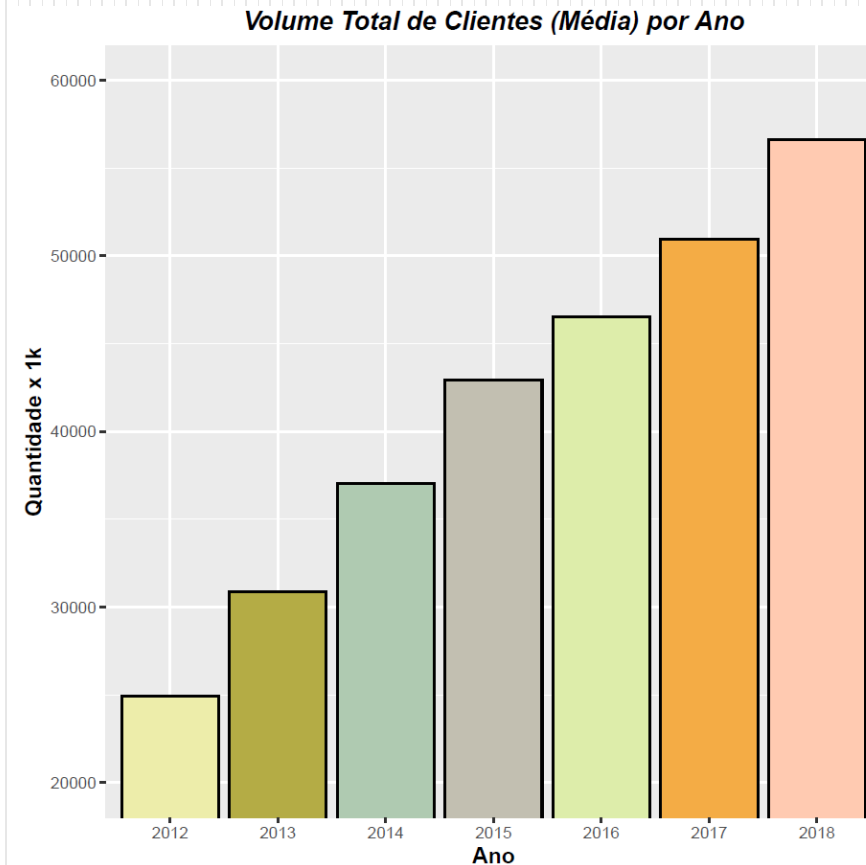
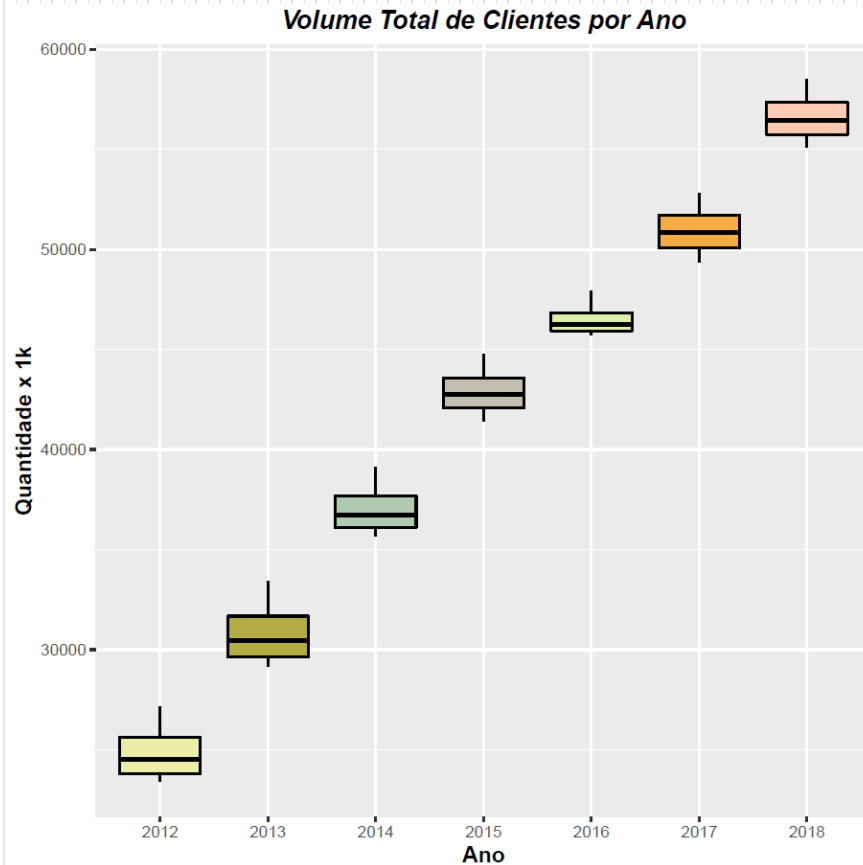
- **Total Clientes x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o **volume total de clientes** aumenta o custo de marketing. Portanto, o custo de marketing é dependente do **volume total de clientes**;
- **Clientes Atuais x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o **volume de clientes atuais** aumenta o custo de marketing. Portanto, o custo de marketing é dependente do **volume de clientes atuais**;
- **Clientes Novos x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta o **volume de clientes novos** aumenta o custo de marketing. Portanto, o custo de marketing é dependente do **volume de clientes novos**.

**I. E.:** Existe alguma relação de dependência (considerando todo o período) entre o **custo de marketing** da empresa e a **receita total**, a **margem de contribuição** e o **rendimento por cliente**?



- **Receita Total x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta a **receita total** aumenta o **custo de marketing**. Portanto, o **custo de marketing** é dependente da **receita total**;
- **Rendimentos por Cliente x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta os **rendimentos por cliente** aumenta o **custo de marketing**. Portanto, o **custo de marketing** é dependente dos **rendimentos por cliente**;
- **Margem de Contribuição x Custo de Marketing:** confirmamos que existe uma *correlação positiva* entre as variáveis, ou seja, a medida que aumenta a **margem de contribuição** aumenta o **custo de marketing**. Portanto, o **custo de marketing** é dependente da **margem de contribuição**.

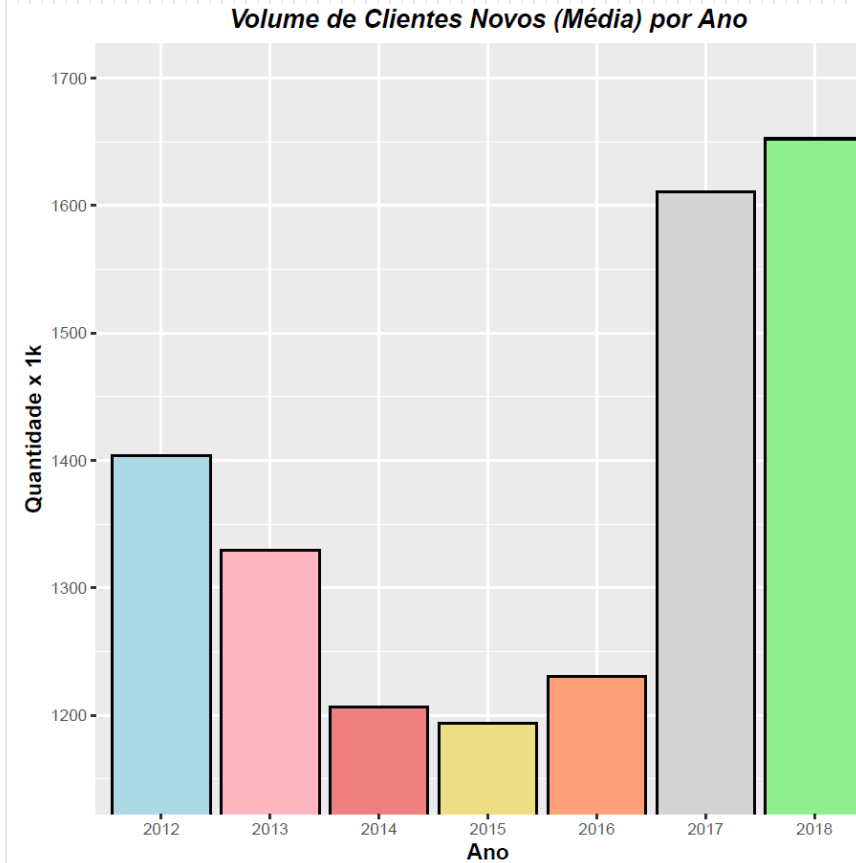
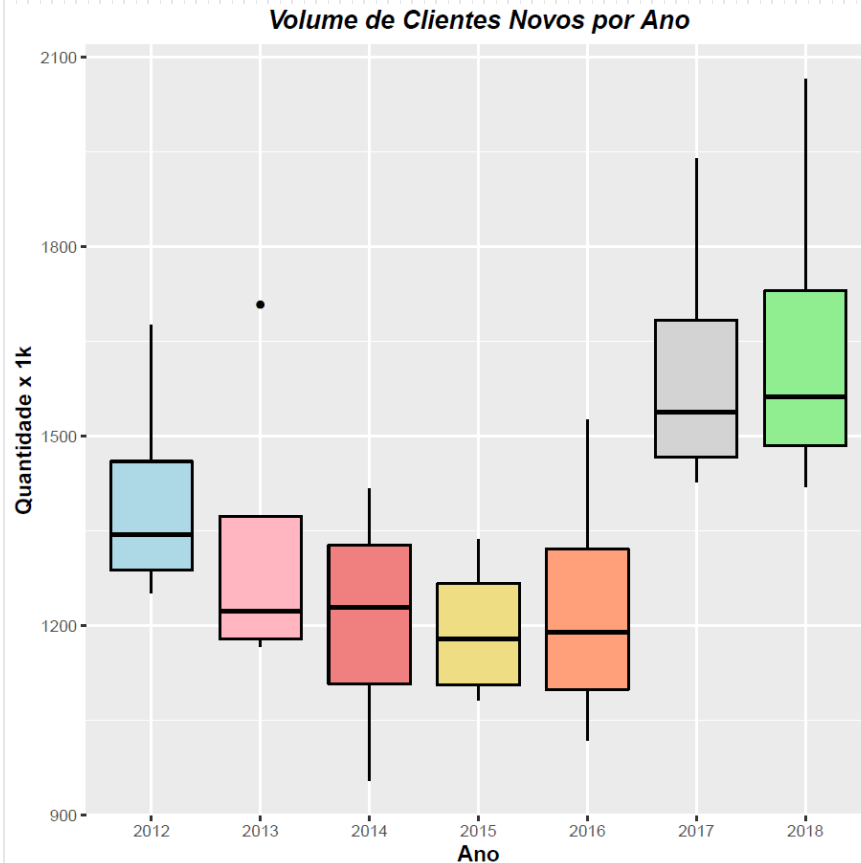
I. E.: Considerando todo o período (2012 a 2018), o **volume total de clientes** *aumentou* e/ou *diminuiu*?



- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que o **volume total de clientes *aumentou significativamente* em todos os anos**;
- O **volume total de clientes *aumentou* (em média, por ano) ~127.4% em 7 anos**, ou seja, de **~25 milhões em 2012 para ~56.6 milhões em 2018**.



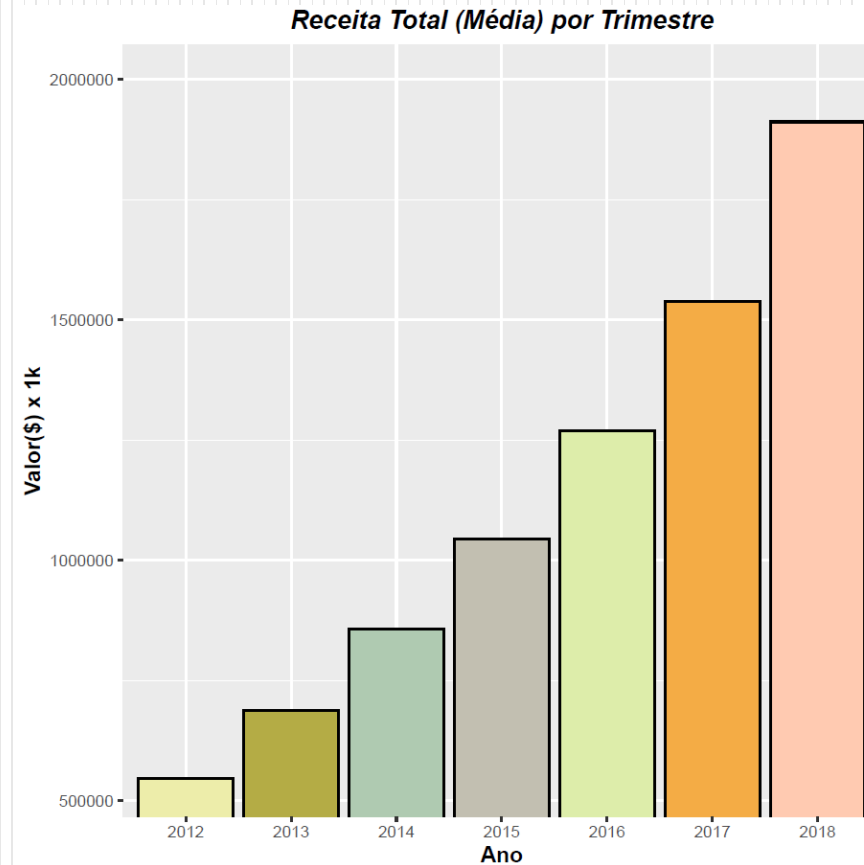
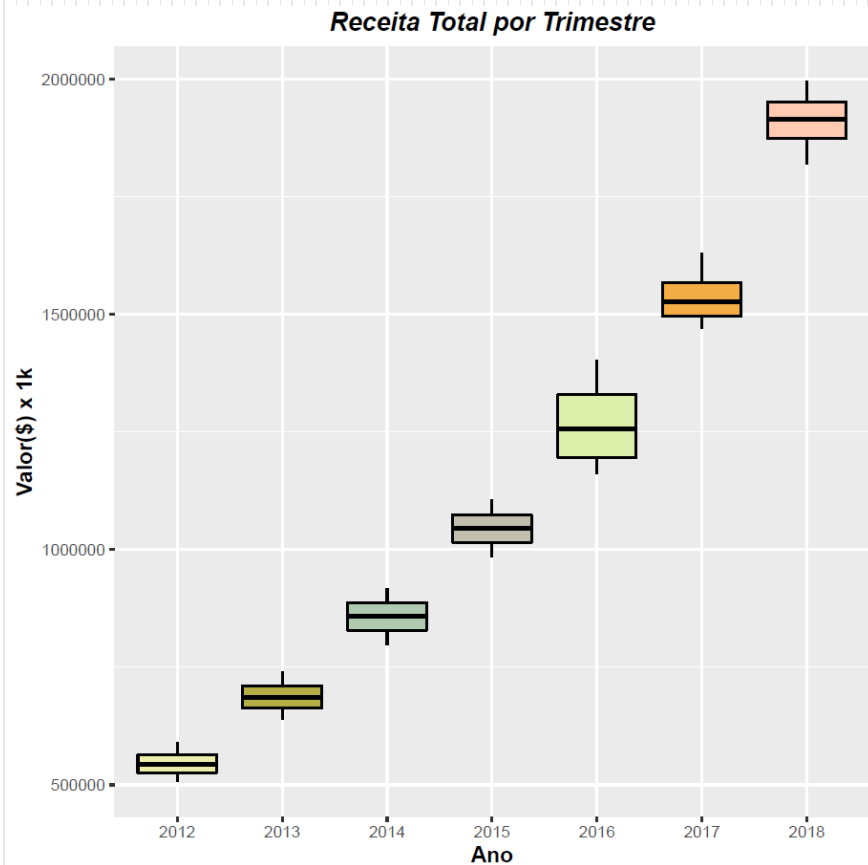
I. E.: Considerando todo o período (2012 a 2018), o **volume de clientes novos** *aumentou* e/ou *diminuiu*?



- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que o **volume de clientes novos aumentou significativamente a partir de 2016 e não diminuiu significativamente no período entre 2012 e 2015**;
- O **volume de clientes novos aumentou (em média, por ano) ~38.0% em 3 anos**, ou seja, de **~1.2 milhões em 2015 para ~1.65 milhões em 2018**.

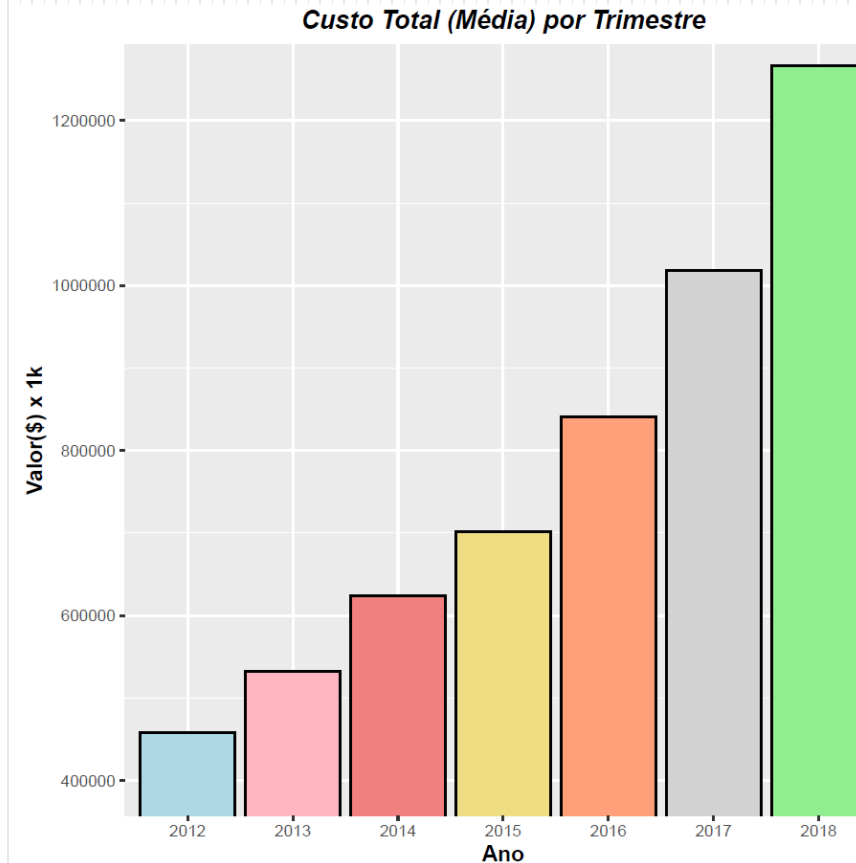
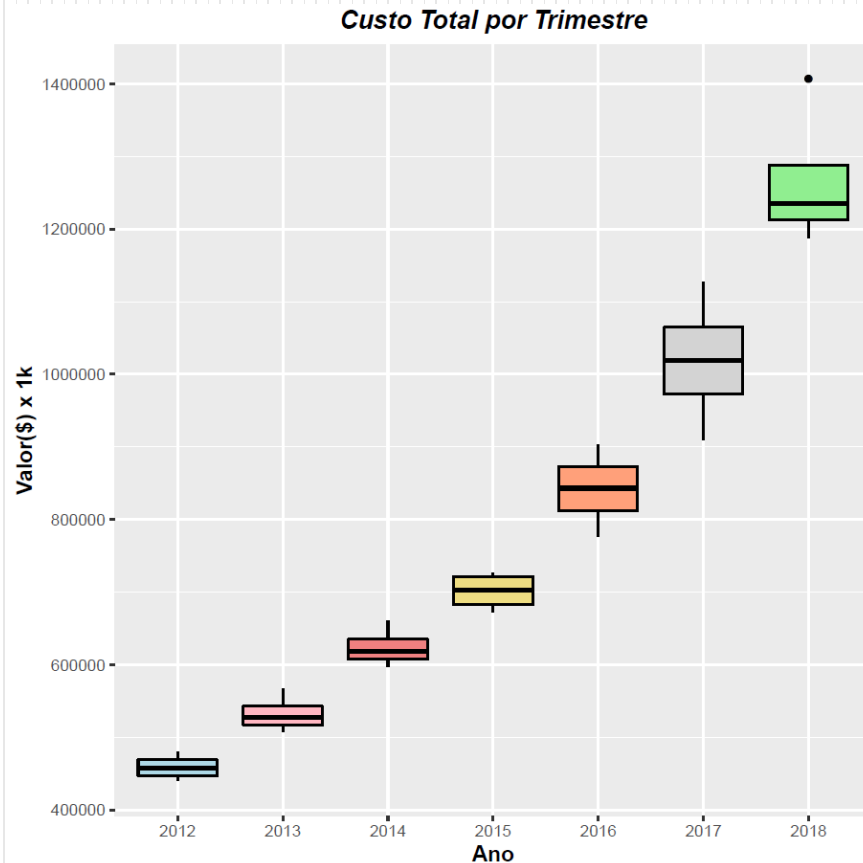


I. E.: Considerando todo o período (2012 a 2018), a **receita total** da empresa *aumentou e/ou diminuiu*?



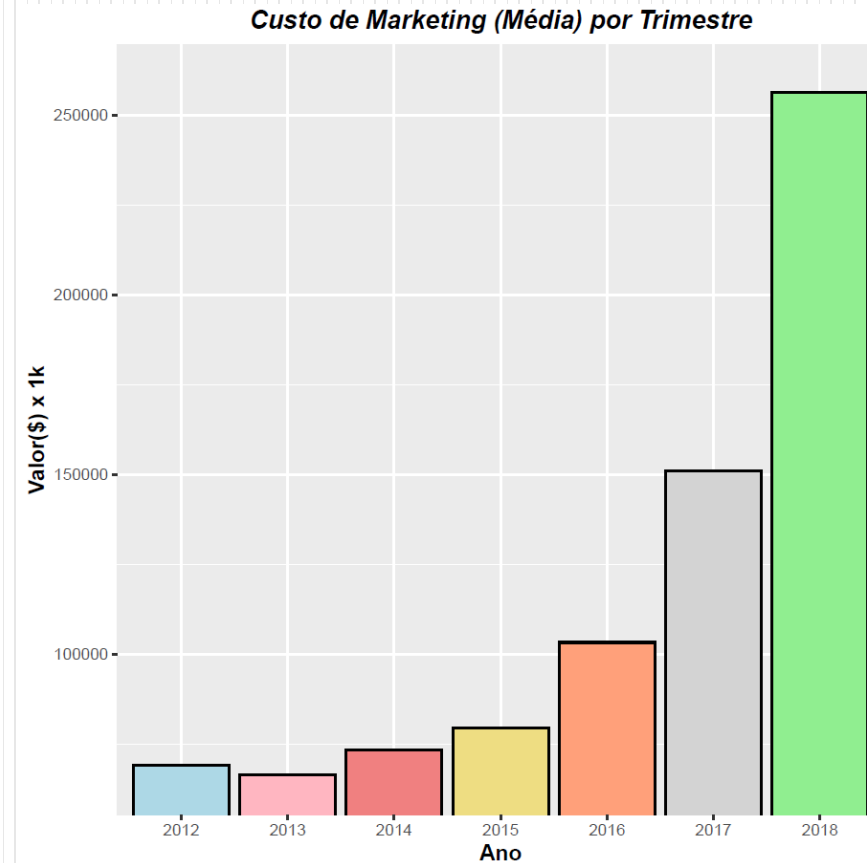
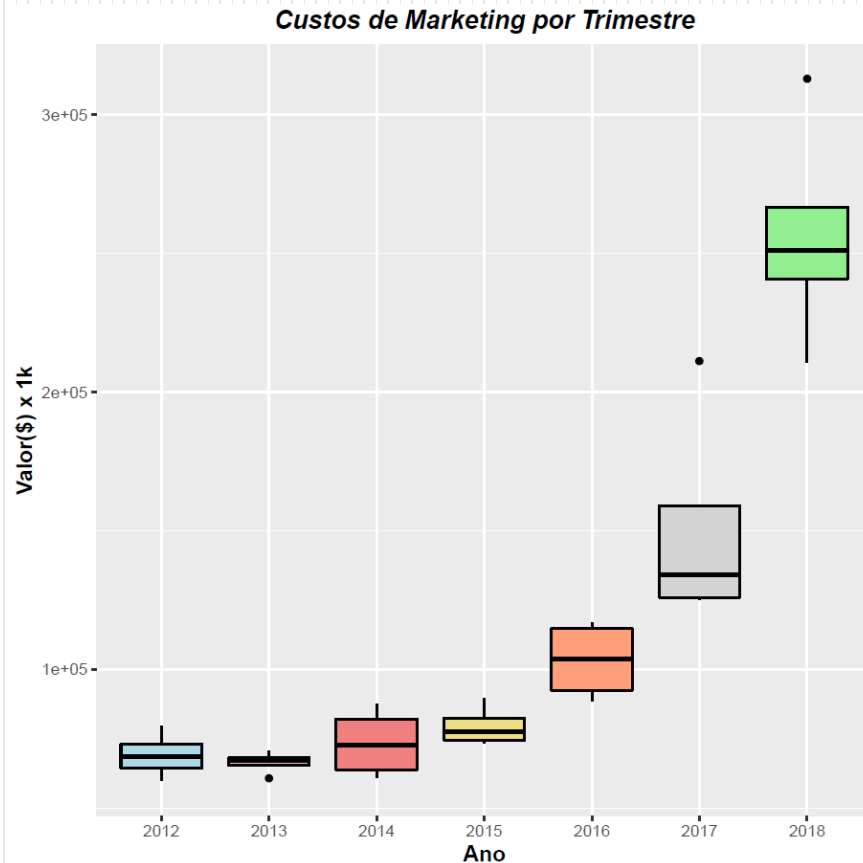
- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que a **receita total aumentou significativamente em todos os anos no período entre 2012 e 2018**;
- A **receita total aumentou (em média, por trimestre) ~250.0%** em **7 anos**, ou seja, de **~\$546 milhões em 2012 para ~\$1.9 bilhões em 2018**.

I. E.: Considerando todo o período (2012 a 2018), o **custo total** da empresa *aumentou e/ou diminuiu*?



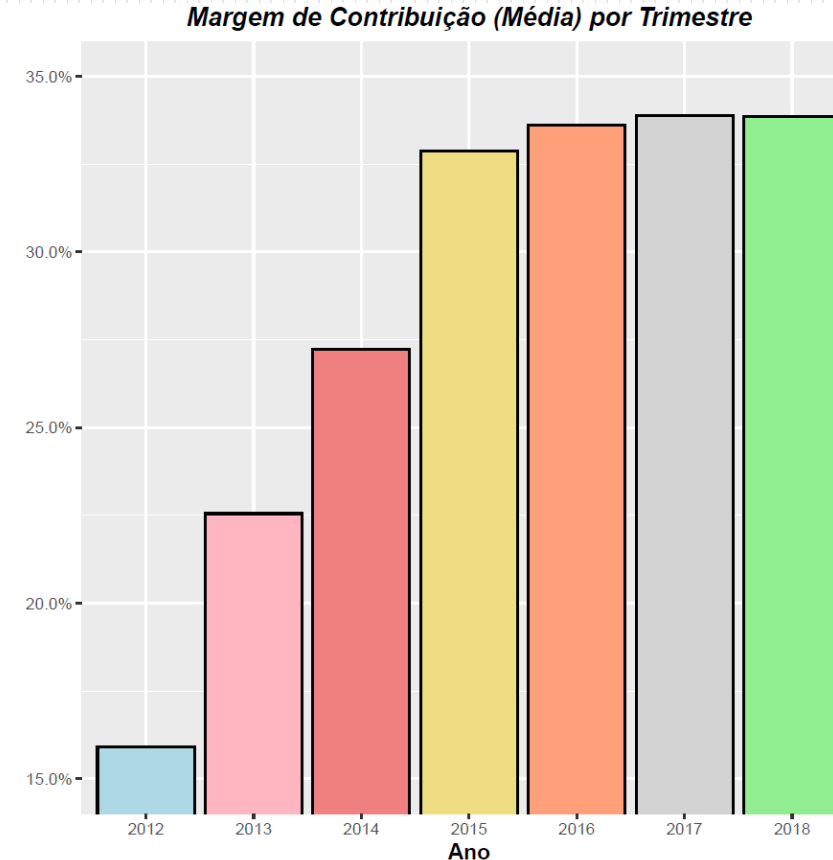
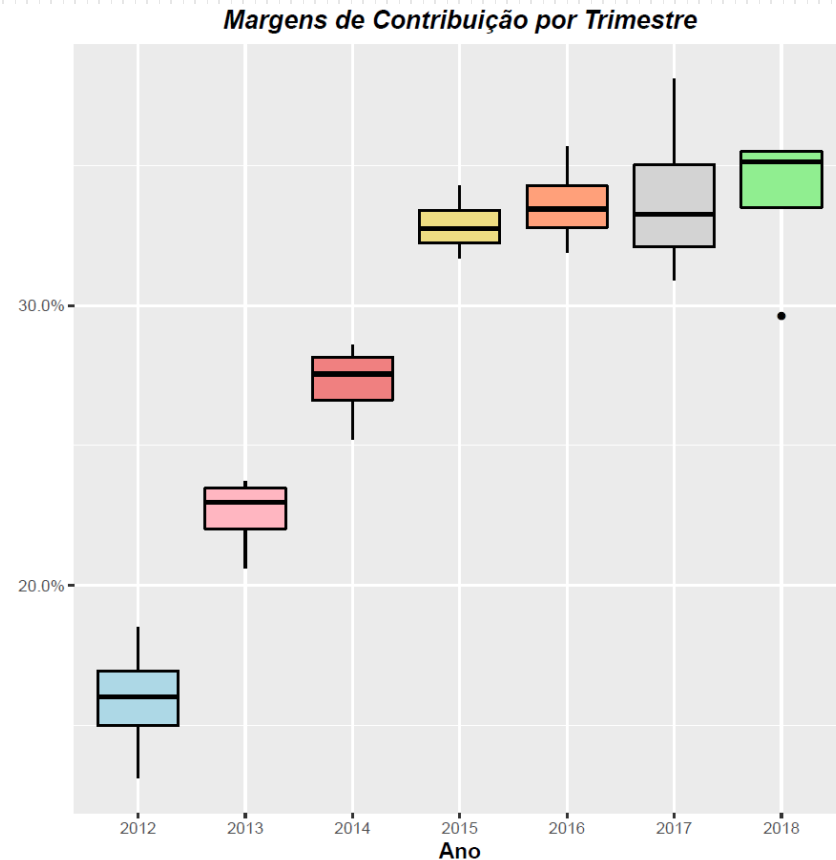
- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que o **custo total aumentou significativamente em todos os anos** no período entre 2012 e 2018;
- O **custo total aumentou (em média, por trimestre) ~176.0%** em **7 anos**, ou seja, de **~\$460 milhões** em **2012** para **~\$1.3 bilhões** em **2018**.

I. E.: Considerando todo o período (2012 a 2018), o **custo de marketing** da empresa *aumentou e/ou diminuiu?*



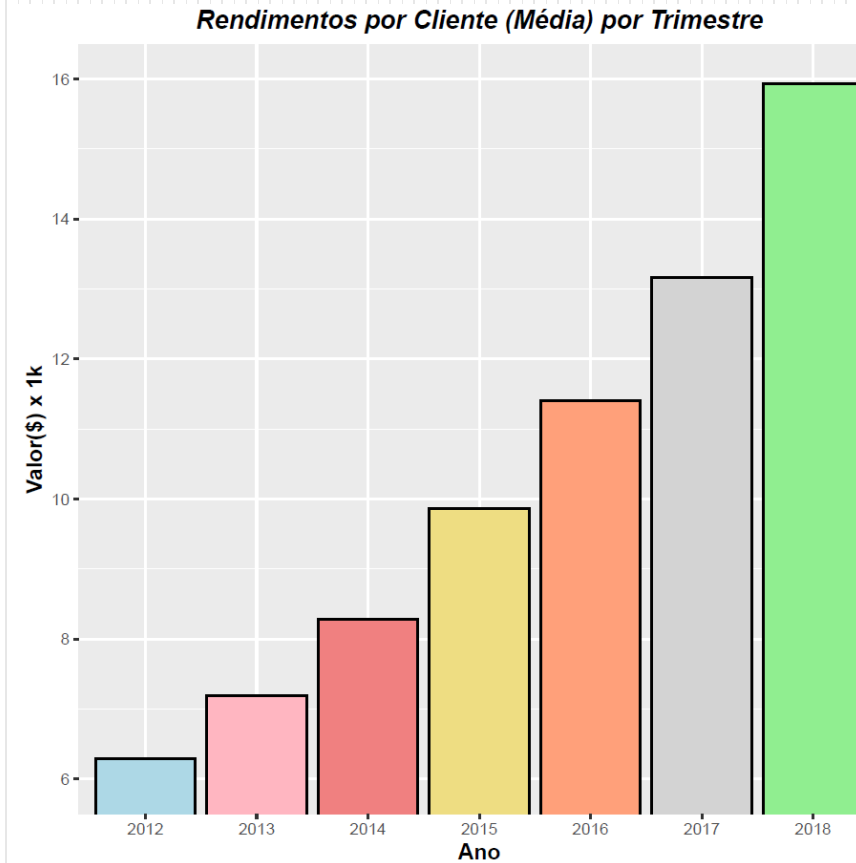
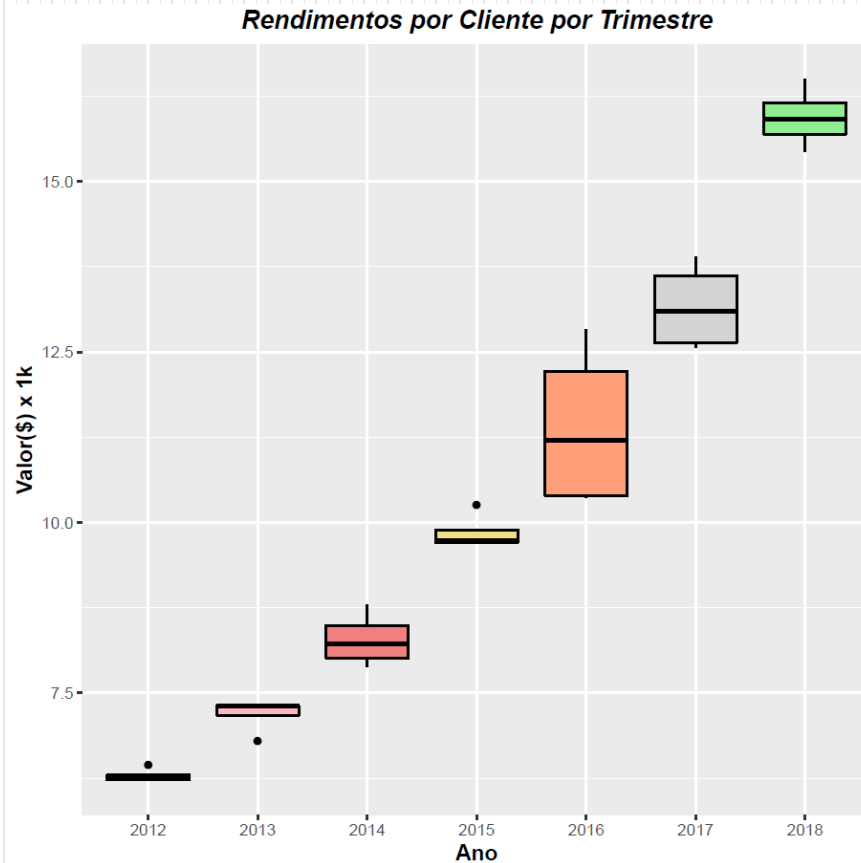
- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que o **custo de marketing aumentou significativamente no período entre 2015 e 2018 e não diminuiu significativamente no período entre 2012 e 2014**;
- O **custo de marketing aumentou (em média, por trimestre) ~271.4% em 7 anos**, ou seja, de **~\$69 milhões em 2012 para ~\$256 milhões em 2018**.

I. E.: Considerando todo o período (2012 a 2018), a **margem de contribuição** da empresa *aumentou e/ou diminuiu*?



- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que a **margem de contribuição aumentou significativamente** no período entre 2012 e 2015 e **não diminuiu significativamente** no período entre 2016 e 2018;
- A **margem de contribuição aumentou** (em média, por trimestre) **~113.0%** em 7 anos, ou seja, de **~\$16%** em 2012 para **~\$34%** em 2018.

I. E.: Considerando todo o período (2012 a 2018), o **rendimento por cliente** da empresa *aumentou e/ou diminuiu?*



- O gráfico acima (**Boxplot**) confirma todos os **testes estatísticos** realizados, ou seja, que o **rendimento por cliente *aumentou significativamente em todos os anos*** no período entre 2012 e 2018;
- O **rendimento por cliente *aumentou (em média, por trimestre)*** ~153.5% em 7 anos, ou seja, de ~\$6.3 mil em 2012 para ~\$16 mil em 2018.

# F I M

Walter Trevisan