

DRAFT EXECUTIVO

DELIVERY ON COMMITMENT

Análise Estatística
2021

AGOSTO / 2021

*Confidencial *

Criado por: Seu Nome

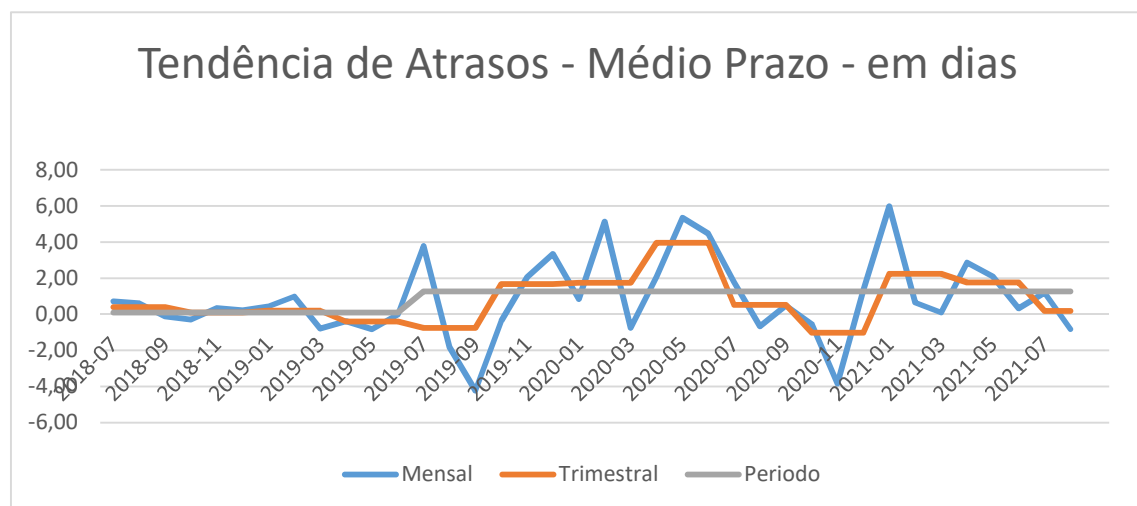
Análise Estatística – Delivery on Commitment

Contexto:

A empresa define DOC – Delivery On Commitment como o indicador que mostra a performance de entrega e a capacidade da empresa cumprir com o compromisso assumido em entregar os produtos encomendados pelo cliente. A área comercial identificou uma variação nos prazos de entrega previamente negociados com o cliente. Em uma análise de Pareto, a causa mais frequente de reclamação de clientes dos últimos 12 meses se tornou – Atraso na entrega dos Produtos.

Metodologia:

Identificar os pedidos de vendas registrados no ERP, obtendo a data do primeiro compromisso assumido com o cliente e contrastando com a data de faturamento efetiva. A análise exploratória buscou os pedidos emitidos entre julho de 2018 e agosto de 2021. Para cada linha de pedido foi calculada a diferença em dias entre a data [Primeira Data Aprovada PCP] e a data [Emissão da NF-e].



Em análise prévia, identifica-se que o comportamento dos atrasos sofreu alteração brusca na metade do ano 2018. Os registros da área de operações indicam que entre maio/2018 e julho/2018 ocorreram alguns fatos relevantes: a) mudança do layout da fábrica; b) mudança no modelo de gestão (kanban); c) implementação de novo módulo do ERP na fábrica. [

Foi extraída uma amostra aleatória de 9607 itens de pedido faturados até o dia 30/06/2019 e [ANTES] amostra de igual tamanho de pedidos faturados após 30/06/2019 [DEPOIS]. Os dados foram gerados a partir da tabela de itens de pedidos no ERP, exportados para Microsoft Excel e o trabalho de análise foi executado na ferramenta R.

Hipótese:

Hipótese Nula (H0) – Não há diferença significativa na performance de entrega após as mudanças na fábrica. a mudança no gráfico foi "por acaso".

Hipótese Alternativa (H1) – Há diferença significativa na performance de entrega após as mudanças na fábrica.

A condição preliminar para se aplicar o teste t é a existência de distribuição normal dos dados em ambos grupos de dados.

Exploração da Hipótese – ANTES x DEPOIS

Exploração dos Dados

```
> summary(dados)
      antes      depois
Min.   :-98.0000  Min.   :-91.00
1st Qu.:  0.0000  1st Qu.:  0.00
Median :  0.0000  Median :  3.00
Mean   :  0.6139  Mean   :  7.64
3rd Qu.:  0.0000  3rd Qu.: 11.00
Max.   :116.0000  Max.   : 99.00
> |
```

O intervalo interquartílico variou consideravelmente

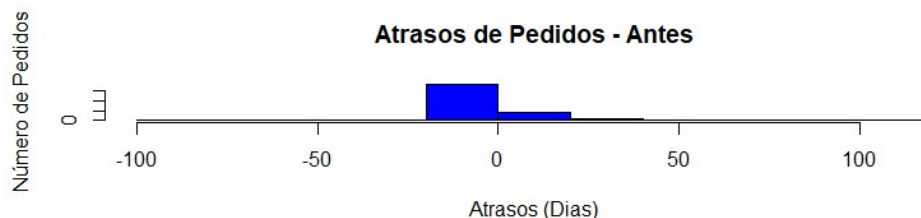
Antes -> Era irrelevante e os valores máximos eram provavelmente outliers

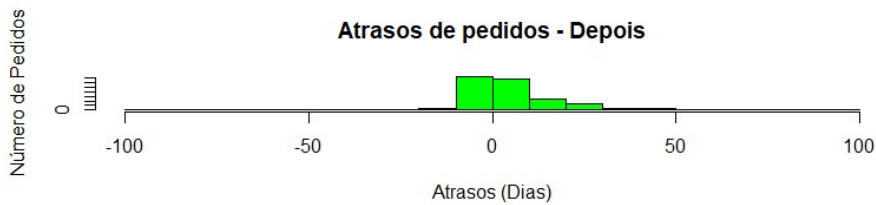
Depois -> Os atrasos se concentram nos quartis superiores

	V1	V2	V3	V4
nomes_stats	Média	Desvio Padrão	Variância	Atrasos do Período
dados_stats_antes	0.61	12.7	161.37	Pedidos ate 06/2019
dados_stats_depois	7.64	15.25	232.68	Pedidos apos 06/2019

Antes e Depois da data de corte, ocorriam antecipações de até três meses, e ocorreram atrasos. Entretanto, a média dos atrasos saltou de 0.6 dias para 7,6 dias.

Distribuição dos atrasos:





Skewness (distorção):

Uma distribuição equilibrada de pedidos entregues com atraso x pedidos entregues em dia, apresentariam um skewness próximo de zero. Um skewness positivo, indica que os pedidos estão sendo entregues com atraso e skewness negativo, indica que os pedidos estão sendo antecipados.

Antes:

```
> skewness(dados$antes)
[1] 0.397903
> summary(dados$antes)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-98.0000  0.0000  0.0000  0.6139  0.0000 116.0000
```

Depois:

```
> skewness(dados$depois)
[1] 1.557964
> summary(dados$depois)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -91.00   0.00   3.00   7.64  11.00   99.00
```

Anderson-Darling normality test

Se você estiver executando um teste estatístico paramétrico em seus dados (por exemplo, uma ANOVA), o uso de dados altamente inclinados para a direita ou esquerda pode levar a resultados enganosos. Portanto, se você deseja executar um teste nesse tipo de dados, execute uma transformação de log e execute o teste nos números transformados.

Teste de Normalidade - Anderson-Darling

Hipótese Nula (H0): Os dados são normalmente distribuídos.

Hipótese Alternativa (H1): Os dados não são normalmente distribuídos.
> `ad.test(dados$antes)`

Anderson-Darling normality test

data: dados\$antes
A = 1524.4, p-value < 2.2e-16

> `ad.test(dados$depois)`

Anderson-Darling normality test

data: dados\$depois
A = 756.64, p-value < 2.2e-16

As duas amostras respeitam uma distribuição normal.

Verificando a correlação entre as duas amostras

Para realização do teste de hipótese, as duas amostras devem ser independentes. O teste de correlação mostra o grau de independência entre as duas amostras - Antes e depois. Este coeficiente varia entre -1 e 1. O coeficiente negativo indica que as amostras são totalmente independentes, o coeficiente igual a 1, indica que são totalmente dependentes.

```
> cor(dados$antes, dados$depois)
[1] 0.01247518
> |
```

O coeficiente encontrado $c = 0.01$ indica que as amostras tem baixíssimo índice de correlação.

Teste de Hipóteses

O Teste de Hipóteses é uma forma de inferência estatística que usa dados de uma amostra para tirar conclusões sobre um parâmetro populacional ou uma distribuição de probabilidade populacional. Primeiro, é feita uma suposição provisória sobre o parâmetro ou distribuição. Essa suposição é chamada de hipótese nula e é indicada por H_0 . Uma hipótese alternativa (denotada H_1), é o oposto do que é declarado na hipótese nula. O procedimento de teste de hipóteses envolve o uso de dados de amostra para determinar se H_0 pode ou não ser rejeitada. Se H_0 for rejeitada, a conclusão estatística é que a hipótese alternativa H_1 é verdadeira.

Formulação da Hipótese

Hipótese Nula (H_0) – Não há diferença significativa na performance de entrega após as mudanças na fábrica a mudança no gráfico foi "por acaso".

H_0 : $\mu_{antes} - \mu_{depois}$ igual a 0 (zero)

Hipótese Alternativa (H_1) – Há diferença significativa na performance de entrega após as mudanças na fábrica.

H1: muantes - mudepois diferente de 0 (zero)

Todavia, existe um impasse! Há três tipos de test t:

teste t de uma amostra

teste t de amostras independentes

teste t de amostras relacionadas (pareadas)

Usamos o teste t de uma amostra para verificar os valores da variável em relação a média conhecida de uma população. Para realizarmos os testes de igualdade de variâncias e os testes de médias, precisamos que as duas populações sejam independentes. Esse é um teste de amostras independentes.

O teste t pareado é útil para analisar o mesmo conjunto de itens que foram medidos sob duas condições diferentes, as diferenças nas medições feitas sobre o mesmo assunto antes e depois de um tratamento, ou diferenças entre dois tratamentos dados ao mesmo assunto. Como estamos medindo conjuntos diferentes de dados (pedidos distintos) coletados em situações distintas, Por isso paired = F.

```
> teste_hipo <- t.test(dados$antes, dados$depois,  
f.level = 0.95, alternative = "t")  
> teste_hipo
```

welch Two sample t-test

```
data: dados$antes and dados$depois  
t = -34.69, df = 18603, p-value < 2.2e-16
```

Interpretação resultado - p-value muito baixo

Rejeitamos a hipótese nula e aceitamos a hipótese alternativa

- ✓ **Valor-p Baixo: Forte evidência empírica contra H0.**
- X Valor-p Alto: Pouca ou nenhuma evidência empírica contra H0

Determinando a “Força do Teste” – Power of the Test

A força do teste de uma hipótese indica a probabilidade de que o teste rejeitou a hipótese nula corretamente.

Calculando o tamanho ideal da amostra – Uma amostra a partir de 105 pedidos já pode oferecer uma alta probabilidade de acertar em rejeitar a hipótese nula.

Two-sample t test power calculation

```
n = 104.9279
```

```
d = 0.5
```

```
sig.level = 0.05
```

```
power = 0.95
```

```
alternative = two.sided
```

NOTE: n is number in *each* group

```
> power_teste <- pwr.t.test(n = 9607, d = size_effect, sig.level  
= 0.05, alternative = "t")  
> power_teste
```

Two-sample t test power calculation

```
n = 9607
```

```
d = 0.356113
```

```
sig.level = 0.05
```

```
power = 1
```

```
alternative = two.sided
```

NOTE: n is number in *each* group

Power of the test = 1 -> Alta probabilidade de que o teste rejeitou corretamente a hipótese nula.

Conclusão:

Os resultados sugerem que a amostra ANTES e a amostra DEPOIS apresentam comportamento totalmente distinto. Os testes de hipótese bem como a verificação da Força do Teste indicam que existe alta probabilidade das mudanças implementadas em 2018 tenham interferido na performance de entrega dos pedidos.