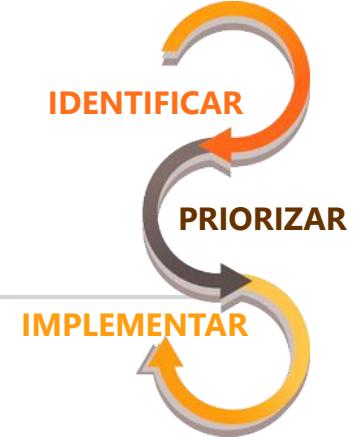


# Treinamento BB – Modelo Híbrido



**ANALYZE**  
Bem-vindos!!!

# CICLO ITERATIVO – DEFINE HÍBRIDO



## 1. Identificar

Estruturar o Projeto de Melhoria  
Empatia e Voz do Cliente  
Entender/enxergar o Fluxo de Valor  
Análise de Valor Lean



## 2. Priorizar

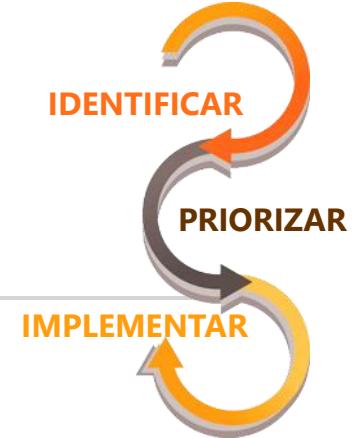
Estruturação do Backlog de Melhorias  
Priorização das Melhorias  
Composição da Sprint



## 3. Implementar

Planejamento da Sprint  
Sprint de Melhorias  
Daily Scrum  
Revisão da Sprint  
Retrospectiva da Sprint

# CICLO ITERATIVO – MEASURE HÍBRIDO



## 1. Identificar

Mapear Variáveis do processo  
Validar Sistema de medição  
Análise Exploratória de Dados  
Variáveis chaves do processo



## 2. Priorizar

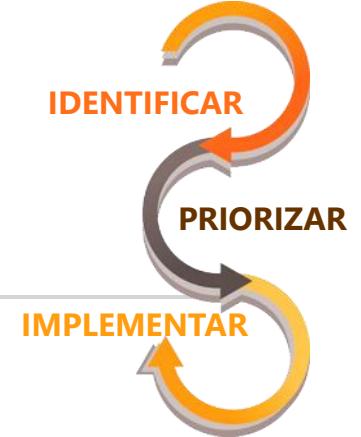
Estruturação do Backlog de Melhorias  
Priorização das Melhorias  
Composição da Sprint



## 3. Implementar

Planejamento da Sprint  
Sprint de Melhorias  
Daily Scrum  
Revisão da Sprint  
Retrospectiva da Sprint

# CICLO ITERATIVO – ANALYZE HÍBRIDO



## 1. Identificar

Fatores Críticos & Impacto

Riscos potenciais & Causas Raiz

Melhorias relacionadas com a consolidação das variáveis críticas



## 2. Priorizar

Estruturação do Backlog de Melhorias

Priorização das Melhorias

Composição da Sprint



## 3. Implementar

Planejamento da Sprint

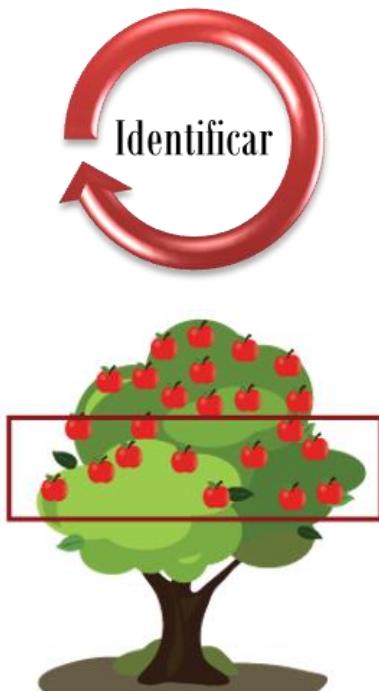
Sprint de Melhorias

Daily Scrum

Revisão da Sprint

Retrospectiva da Sprint

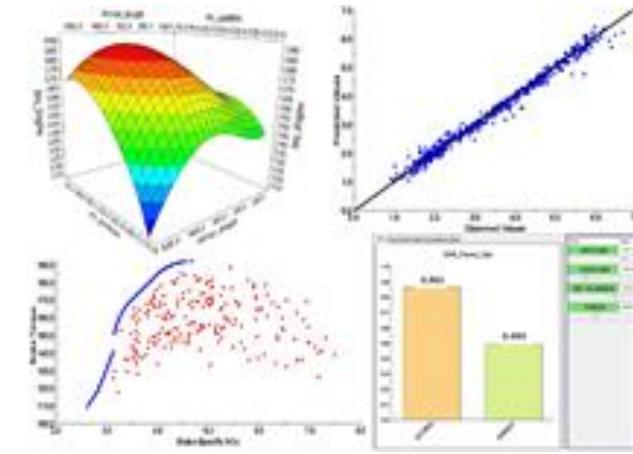
# Atividade A.1 – Ciclo Iterativo Analyze



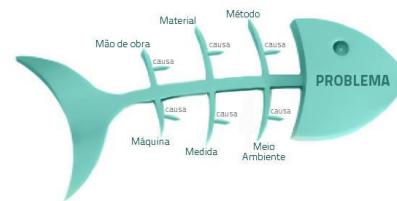
Analyze Híbrido

## A) Fatores Críticos & Impacto Ferramentas Gráficas e Analíticas

```
Results for: Mood's Median Test
Mood Median Test: A&E Processing time versus Time of week
Mood median test for A&E Processing time
Chi-Square = 3.91 DF = 1 P = 0.048
Individual 95.0% CIs
Time of week N< N> Median Q3-Q1
Weekdays 20 10 2.75 2.31 (-*-+---)
Weekends 10 15 4.25 1.75 (----*---)
2.80 3.50 4.20 4.90
Overall median = 3.75
A 95.0% CI for median(Weekdays) - median(Weekends): (-2.27, -0.23)
```



## B) Riscos Potencias e Causas Raízes de Falhas Ferramentas Qualitativas



5 Por quês?

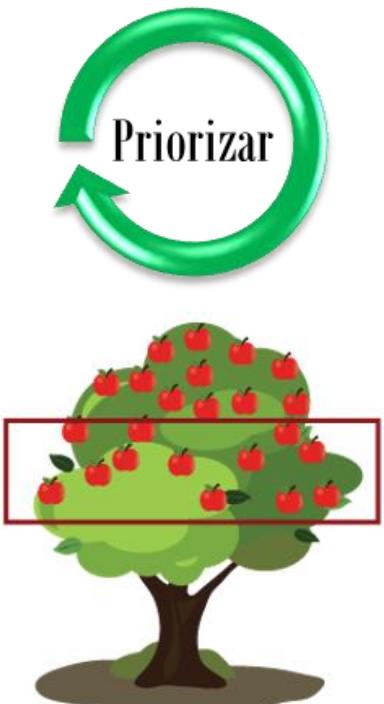
FMEA

Etapas / Entrada do Processo	Modo de Falha Potencial	Efeito da Falha	S V	Causas Potenciais	O C C	Controles Atuais	D E T	R P N
Qual a etapa / entrada do processo sobre investigação?	Em quais caminhos a entrada prejudica?	Qual é o impacto sobre as variáveis de saída (Exigências do Cliente) ou exigências internas?	Qual grave é o efeito para o cliente?	Quais as causas para a entrada atuar prejudicialmente?	Com que frequência as causas ocorrem?	Quais são os controles e procedimentos existentes e inspecionar se estes que podem detectar a Causa ou o Modo de Falha? Deveria incluir um número SOP.	Caixa bem vinda pode detectar a causa ou o efe	Ser/Occ/Def
								0

## C) Melhorias Relacionadas com a consolidação das Variáveis Críticas Matriz Causa e Efeito

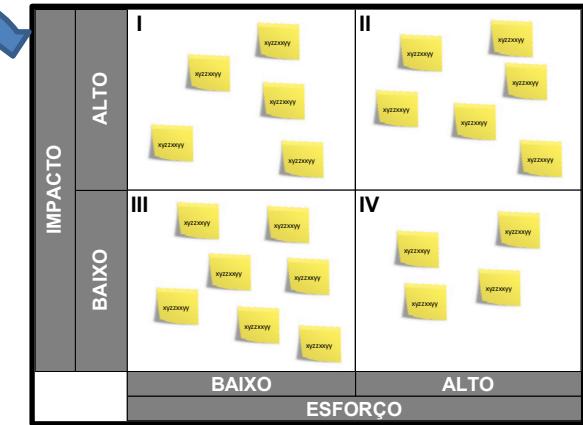
Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida
	CTQ 1	CTQ 2	CTQ N		
	Peso	Peso	Peso		
X1					
X2					
X3					
X4					
X5					
X6					
XN					

# Atividade A.2 – Ciclo Iterativo Analyze



## A) Matriz Esforço/Impacto: Estruturar Backlog de Melhorias

Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida	Esforço
	CTQ 1	CTQ 2	CTQ N			
	Peso	Peso	Peso			
X1						
X2						
X3						
X4						
X5						
X6						
XN						



## B) Matriz Esforço/Impacto: Priorização das Melhorias

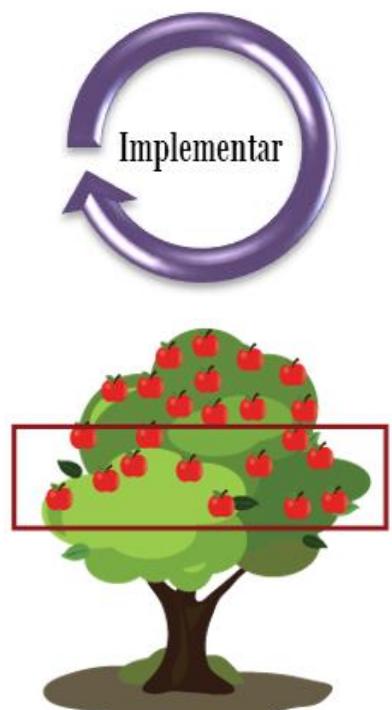


## C) Sprint Board: Composição dos Sprints

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D					
M					
A					
I					
C					

Analyze Híbrido

# Atividade A.3 – Ciclo Iterativo Analyze



Analyze Híbrido

## A) Planejamento do Sprint: Detalhamento das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D	5 yellow sticky notes				12 blue squares (3x4 grid)
M	4 yellow sticky notes				10 blue squares (2x5 grid)
A	3 yellow sticky notes	6 blue squares (2x3 grid)			
I	2 yellow sticky notes				
C	2 yellow sticky notes				

## B) Sprint Analyze: Implantação das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
A	2 yellow sticky notes	3 blue squares (3x1 grid)	1 blue square (1x1)	2 blue squares (2x1 grid)	3 blue squares (3x1 grid)

## C) Daily Scrum: Reuniões Diárias

## D) Revisão da Sprint

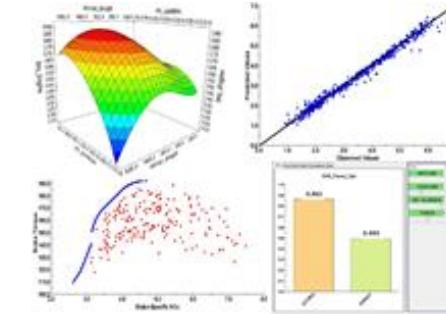


## E) Retrospectiva da Sprint

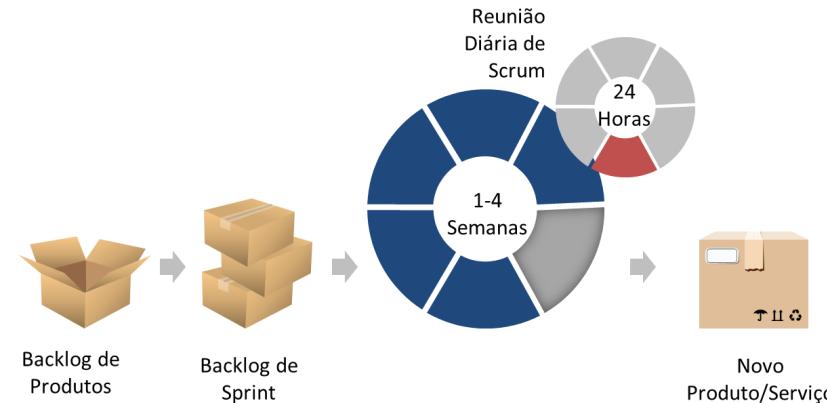
Com análise de Objetivos,  
Metas e Gestão à Vista

# Metodologias do Analyze

## LEAN SIX SIGMA



+ FMEA



# CICLO ITERATIVO – ANALYZE HÍBRIDO



## 1. Identificar

Fatores Críticos & Impacto

Riscos potenciais & Causas Raiz

Melhorias relacionadas com a  
consolidação das variáveis críticas



# Matriz de Seleção de Ferramentas



# Recordando... Tipo de Dados



## Variáveis ou Contínuos:

São dados obtidos através da medição de uma característica, os quais podem assumir qualquer valor dos números reais.

Exemplos: tempo, volume, preço, distância, peso, juros, etc.

## Atributos ou Discretos:

São dados obtidos através da contagem de uma característica (números naturais), ou atribuição de um nome.

Exemplos: número de clientes, número de reclamações, cor, Turno, Ok/Não OK, Certo/Errado, Região A/B..., etc.

É muito importante distinguir dados Contínuos de Discretos, pois as ferramentas e gráficos utilizados para Análises Estatísticas são distintas para cada tipo de dado.

# Identificação do Tipo de Dado

## Indique no seu Plano de Coleta de Dados!

**Discreto** **Contínuo**

- Diagrama de Dispersão
- Regressão Linear Simples

- Regressão Logística

**X**

**Contínuo**

**Discreto**

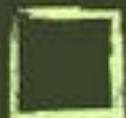
- Box Plot
- Teste de Hipóteses:  
Testes t, ANOVA,  
Mood's, Variância

- Pareto
- Teste de Proporção
- Chi Quadrado

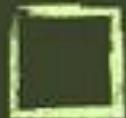


Matriz de  
Seleção de  
Ferramentas

# TO DO LIST



Identifique o Y



Indique se o Y é Discreto (D) ou Contínuo (C)



Identifique o X



Indique se o X é Discreto (D) ou Contínuo (C)



Selecione a Ferramenta Estatística a utilizar

# Exercício 1

A área de Marketing de uma Empresa de Bebidas deseja saber quanto o preço de venda do produto A afeta o Market Share daquele produto.

$Y =$	Market Share
Y	Contínuo
$X =$	Preço de Venda
X	Contínuo
Ferramenta(s): Regressão Diagrama de Dispersão	



## Exercício 2

A área de Customer Service deseja saber se o número de reclamações está mais concentrado em uma determinada Região do país.

$Y = \text{ Número de Reclamações}$

$Y$  Discreto

$X = \text{ Região}$

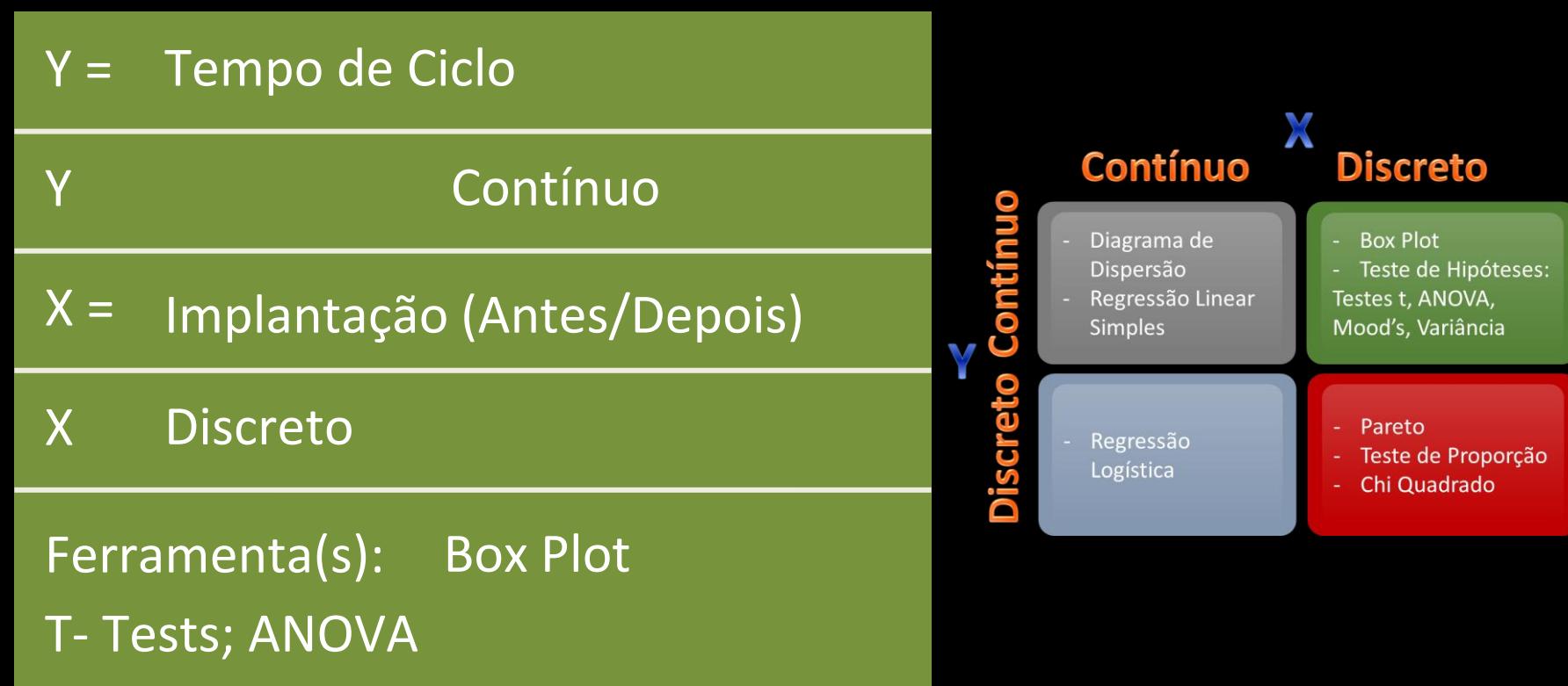
$X$  Discreto

Ferramenta(s): Pareto  
Chi Quadrado



# Exercício 3

Um engenheiro deseja comparar o tempo de ciclo de operação de um processo antes e depois de efetuar a implantação de algumas alterações.



# Exercício 4

Um desenvolvedor de um site de compras deseja verificar como o tempo gasto por um cliente no site afeta a quantidade de itens comprados.

$Y = \text{Quantidade de itens comprados}$

$Y$  Discreto

$X = \text{Tempo gasto no site}$

$X$  Contínuo

Ferramenta(s): Regressão Logística



# Exercício 5

A área de Qualidade deseja saber se a quantidade de defeitos está mais concentrada em um determinado turno de produção.

$Y = \text{Quantidade de Defeitos}$

$Y$  Discreto

$X = \text{Turno}$

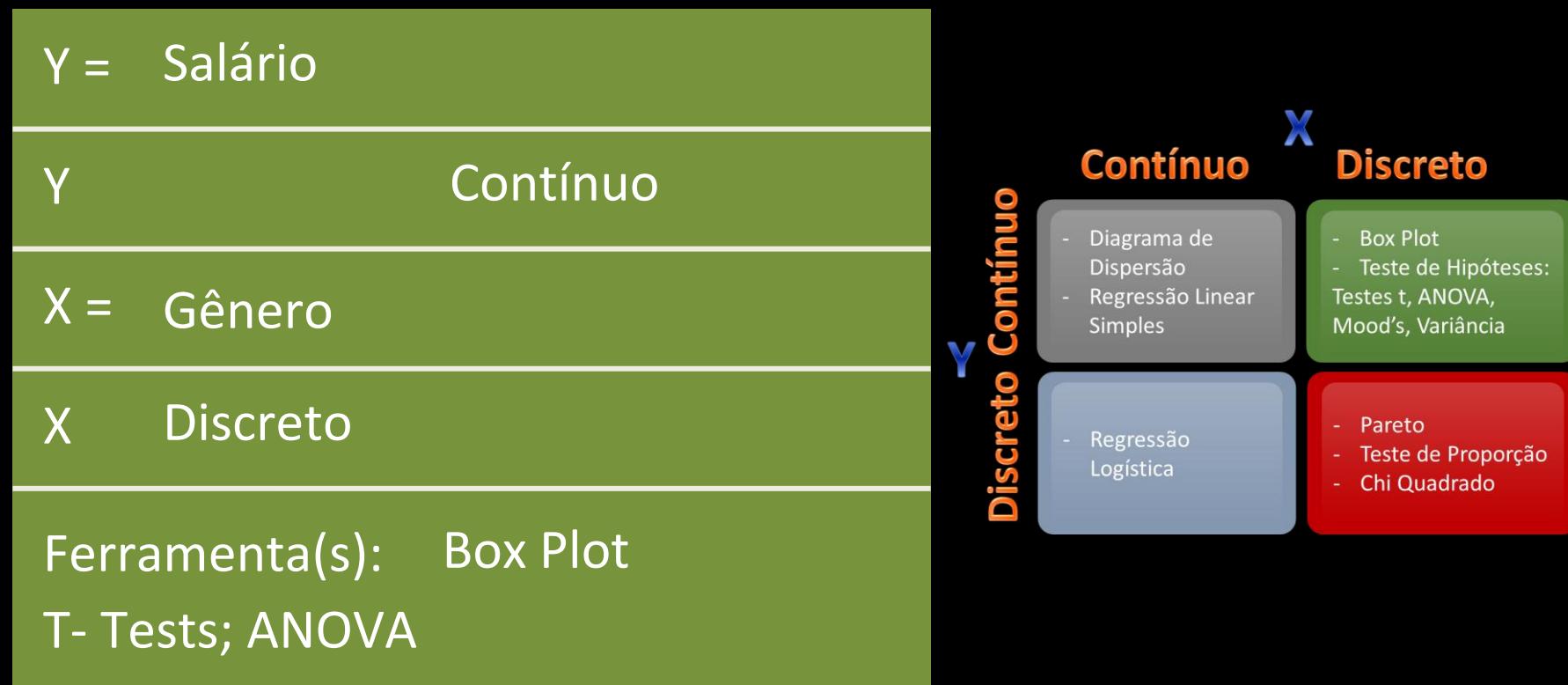
$X$  Discreto

Ferramenta(s): Pareto  
Chi Quadrado



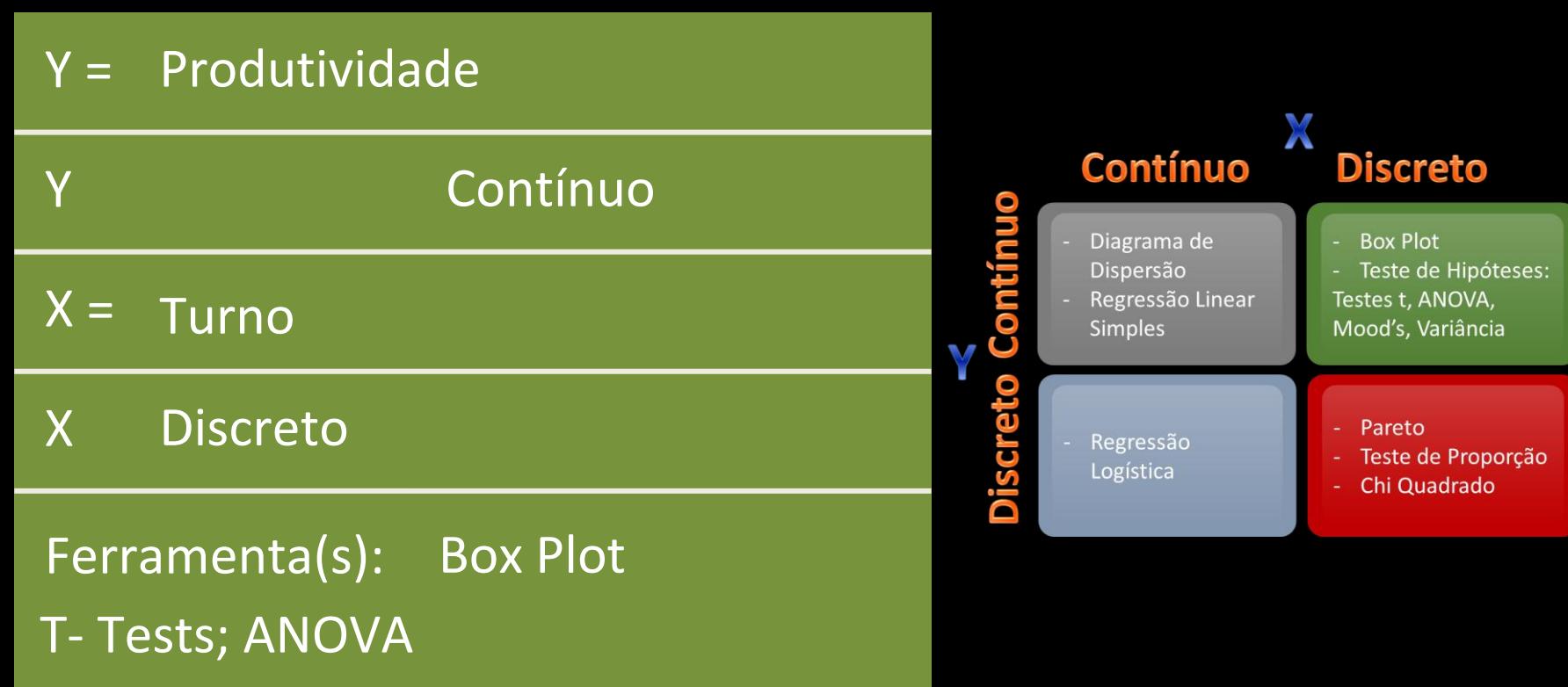
# Exercício 6

Uma Empresa quer verificar se o salário pago para funções de mesmo nível sofre alterações de acordo com o gênero.



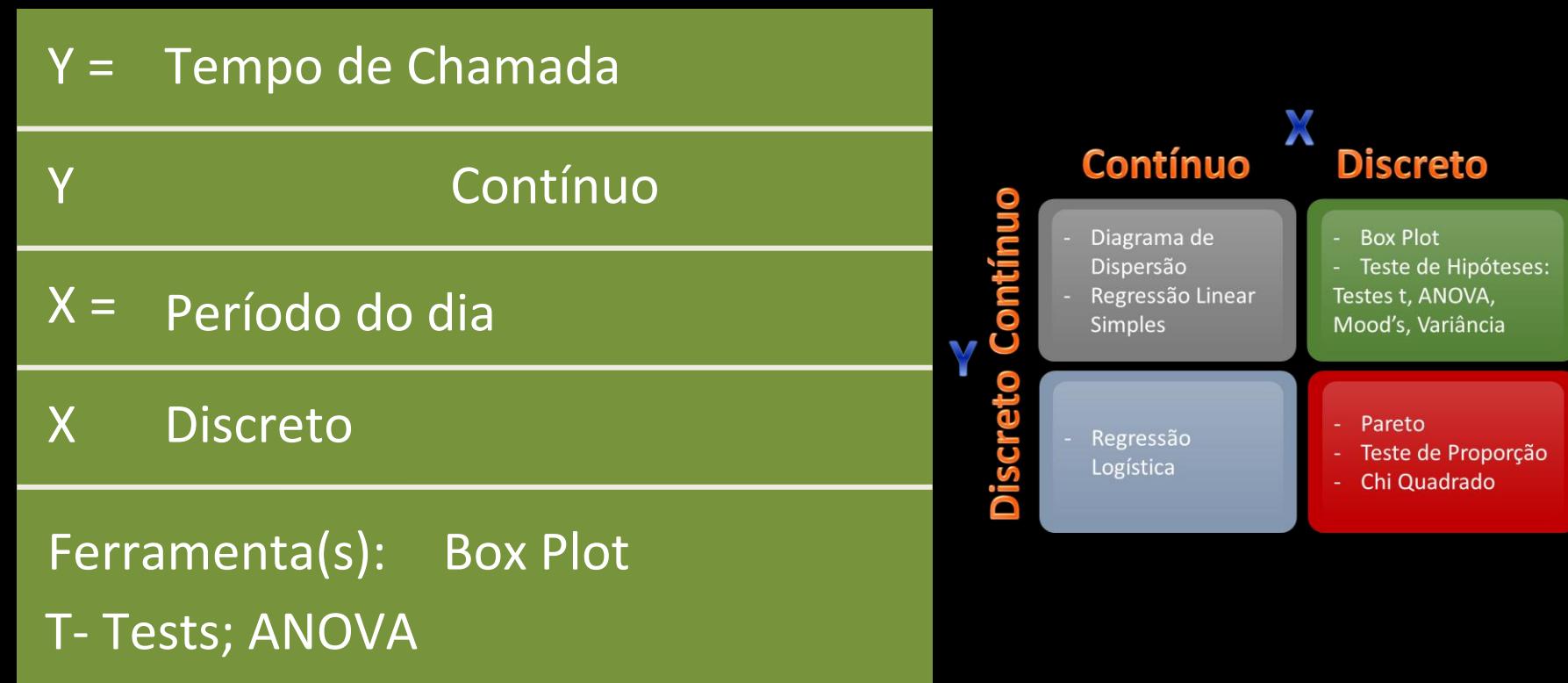
# Exercício 7

Uma Gerente de um Centro de Distribuição quer verificar se um turno tem melhor produtividade em termos de toneladas carregadas por hora.

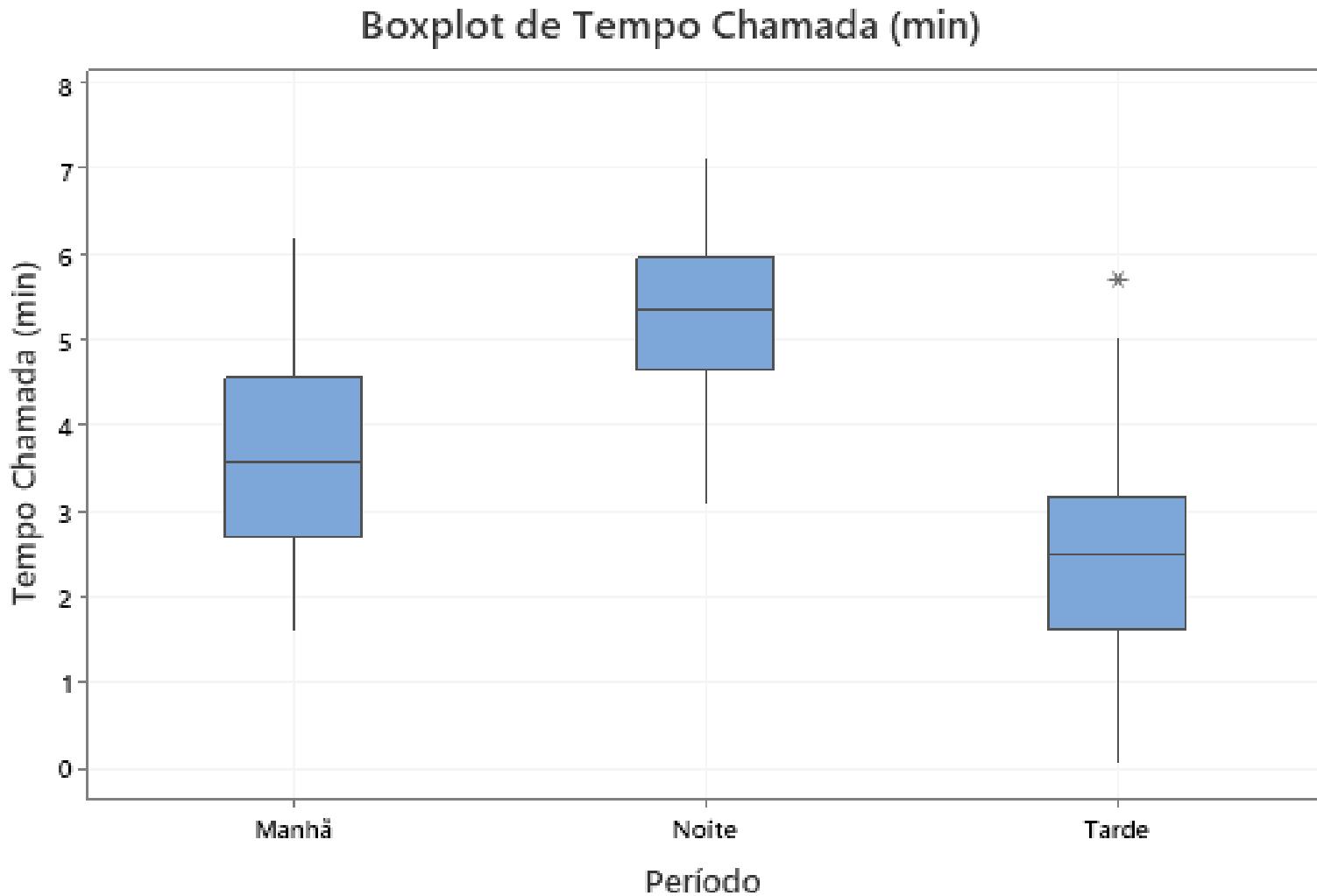


# Exercício 8

Um gerente de Customer Service quer saber em qual período do dia (manhã, tarde ou noite) o tempo gasto com chamadas é maior. Após selecionar ferramenta utilize o arquivo “Chamadas Customer Service.xlsx”.



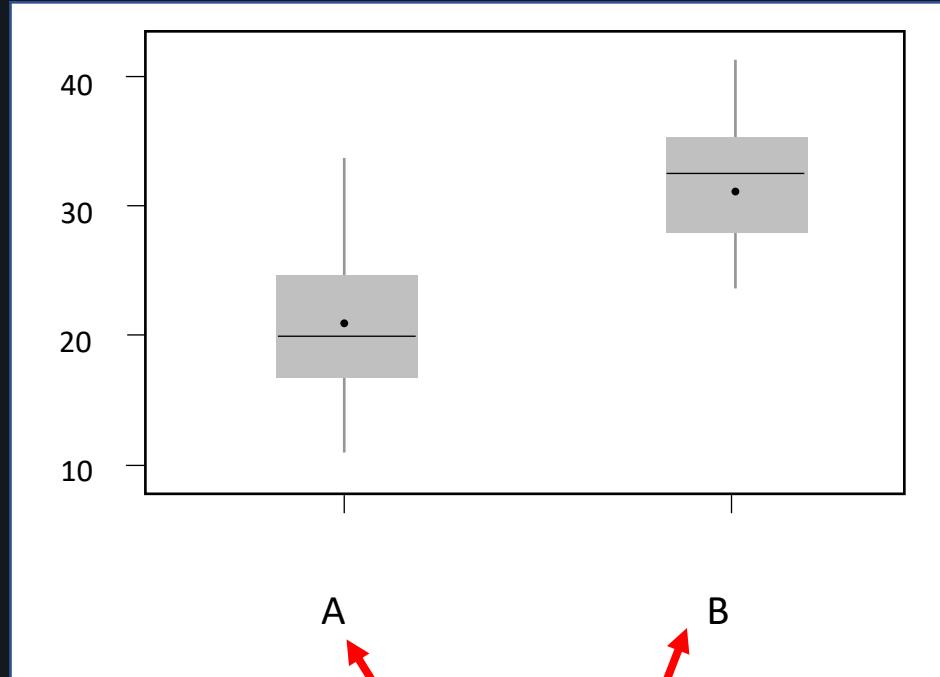
Box Plot  
Chamadas  
por Período



# Teste de Hipótese – Visão Geral

# O que é um Teste de Hipótese?

- Um teste de hipótese é usado para identificar diferenças entre dois ou mais grupos que podem ser importantes para o negócio
- **Devido à variação nunca duas coisas serão exatamente iguais**
  - A questão é se a diferença identificada entre amostras, grupos, processos etc., são randômicas, devido à causas comuns de variação, ou se realmente existe uma diferença significativa
  - **Teste de Hipótese responde esta questão**



**Estatisticamente falando: A e B são realmente diferentes?**

# Por que utilizar um teste de hipótese?

- Para identificar se uma mudança em um input de processo (X) altera de forma significativa o output (Y) do processo.
- Determinar estatisticamente se existe diferenças entre dois ou mais grupos de saídas (outputs) do processo
- Resumindo: determinar quais fatores do processo são as causas da variação nas saídas

$$Y = f(X_1, X_2, X_3 \dots, X_n)$$

mmm....será que é o X1 ou o X2 que está causando problemas?



# Formulação das Hipóteses

- O padrão em Testes de Hipóteses é estabelecer duas possíveis conclusões. A análise indicará qual delas é a mais provável de ser verdadeira com base nos dados obtidos.
- Estabelecer as hipóteses pode parecer uma etapa simples que pode ser pulada, mas ela é CRÍTICA!

## Hipótese Nula ( $H_0$ )

- Não há diferença/não há alteração
- Fator estatisticamente não significativo

## Hipótese Alternativa ( $H_a$ )

- Há diferença/alteração
- Fator estatisticamente significativo
- O que se deseja verificar

Assuma  $H_0$  ser verdadeira até prova em contrário. Ônus da prova fica com  $H_a$

# Formulação das Hipóteses

- **Hipótese Nula** =  $H_0$ : Não há diferença ou mudanças
  - Para testes de Média:  $H_0: \mu_1 = \mu_2$
  - Para testes de Variância:  $H_0: \sigma^2_1 = \sigma^2_2$
  - Para testes de Proporções:  $P_1 = P_2$
- **Hipótese Alternativa** =  $H_a$ : Há diferença ou mudanças
  - Para testes de Média:  $H_a: \mu_1 \neq \mu_2$
  - Para testes de Variância:  $H_a: \sigma^2_1 \neq \sigma^2_2$
  - Para testes de Proporções:  $H_a: P_1 \neq P_2$
- Hipóteses também podem conter os sinais  $>$  ou  $<$

$H_0$  contém sempre o sinal de igual ( $=$ )

# Estabeleça o Nível de Tolerância do Risco

- O Teste Hipótese identifica o **risco** de efetuar uma conclusão errada com base nos dados.
- A conclusão do teste será a escolha de uma das duas hipóteses formuladas ao princípio
- Antes de continuar devemos determinar **qual o nível de risco** de fazer uma conclusão errada **que estamos dispostos a assumir**
  - Considere a criticidade de fazer uma conclusão errada
    - Riscos relativos à segurança ou regulatórios?
    - Riscos relativos à resultados financeiros ou de negócio?
  - **Alpha ( $\alpha$ )** é o risco primário: relacionado com estar errado por rejeitar a  $H_0$
  - **Beta ( $\beta$ )** é o segundo tipo de risco: relacionado com estar errado por aceitar a  $H_0$

Nível de Tolerância do Risco deve ser definido antes da execução do teste

# Estabeleça o Nível de Tolerância do Risco

- Risco  $\alpha$ 
  - Risco de incorretamente concluir que existe uma diferença
  - Também chamado de nível de significância
  - Tipicamente definido como 5%
  - Similar a Intervalo de Confiança
- $\beta$ -risk
  - Risco de incorretamente concluir que não há diferença
  - Frequentemente relacionado com a não existência de dados suficientes
  - Tipicamente definido como 20%

Nível de Tolerância do Risco deve ser definido antes da execução do teste

# Calcule o Valor-P

- Valor-P é utilizado para identificar qual hipótese escolher
  - Risco real de estar errado se rejeitar  $H_0$
  - Comparado com risco Alpha para tomada de decisão
  - Calculado com base nos dados
  - Range de 0,0 to 1,0 (de 0% de chance até 100% de chance)
  - Minitab® ou outros aplicativos estatísticos devem ser utilizados para calcular o Valor-P
- CUIDADO:
  - O cálculo somente sabe o que você diz para ele em forma de dados
  - Se os dados não representam o processo, então as conclusões do teste de hipótese podem ser errôneas
  - Entra lixo; sai lixo!

P-value quantifica o risco de estarmos errados por rejeitar  $H_0$

# Calcule o Valor-P

- $H_0$  não pode ser rejeitado quando
  - Teste estatístico ( $T$ ) é pequeno
  - Valor-P é alto
- $H_0$  pode ser rejeitado (e inferir que  $H_a$  é verdade) quando
  - Teste estatístico ( $T$ ) é elevado
  - Valor-P é pequeno

Minitab® calculará o Valor-P

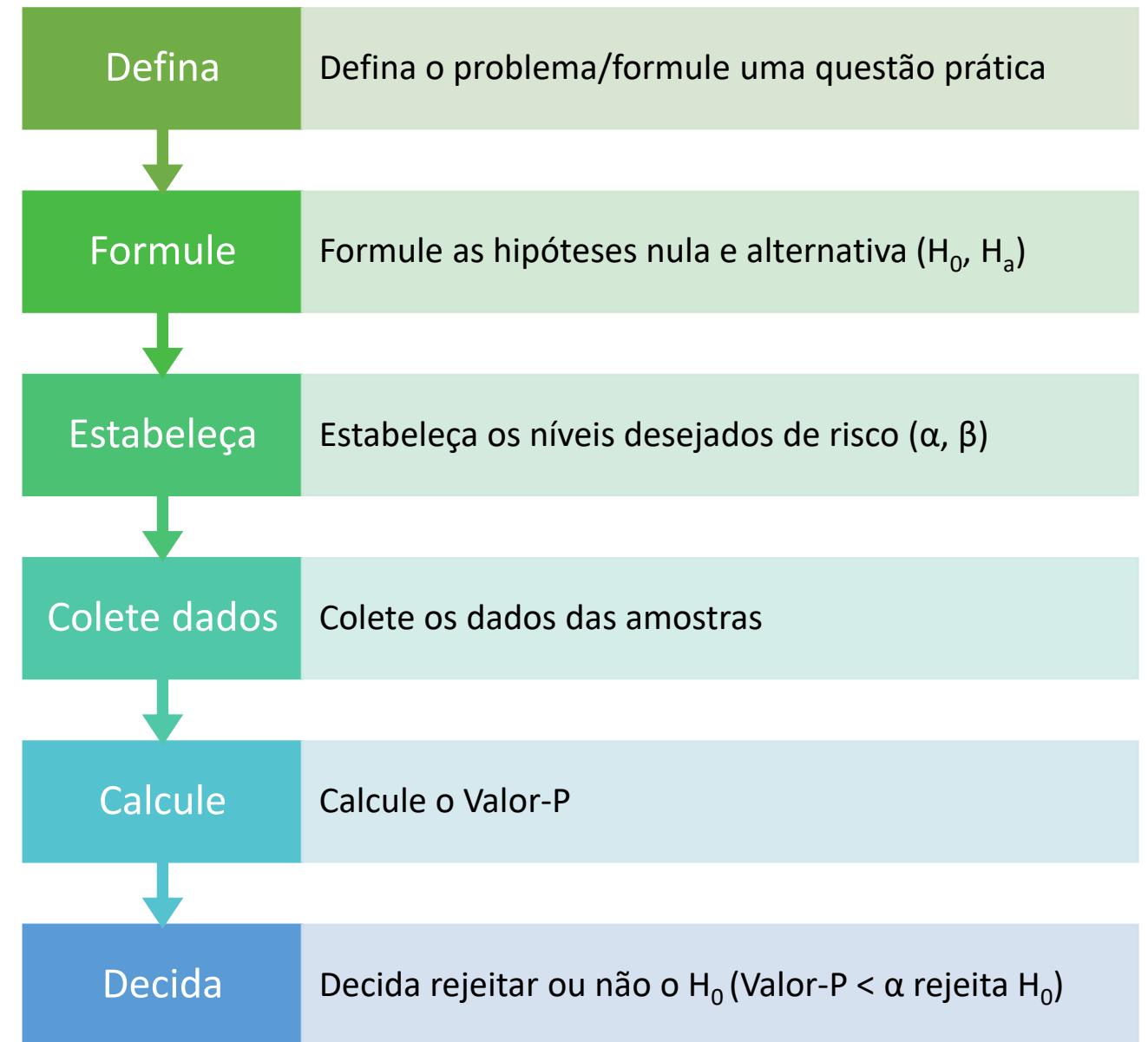
Rejeitar  
ou  
Não Rejeitar  $H_0$   
("≈ aceitar"  $H_0$ )

- Compare Valor-P com risco- $\alpha$ :
  - Se Valor-P > risco- $\alpha$  (tipicamente = 5%)
    - Não rejeitar  $H_0$
    - Não pode assumir diferença
  - Se Valor-P < risco- $\alpha$ 
    - Rejeite  $H_0$
    - Grupos são diferentes

Se Valor-P é baixo,  $H_0$  deve ser rejeitada!



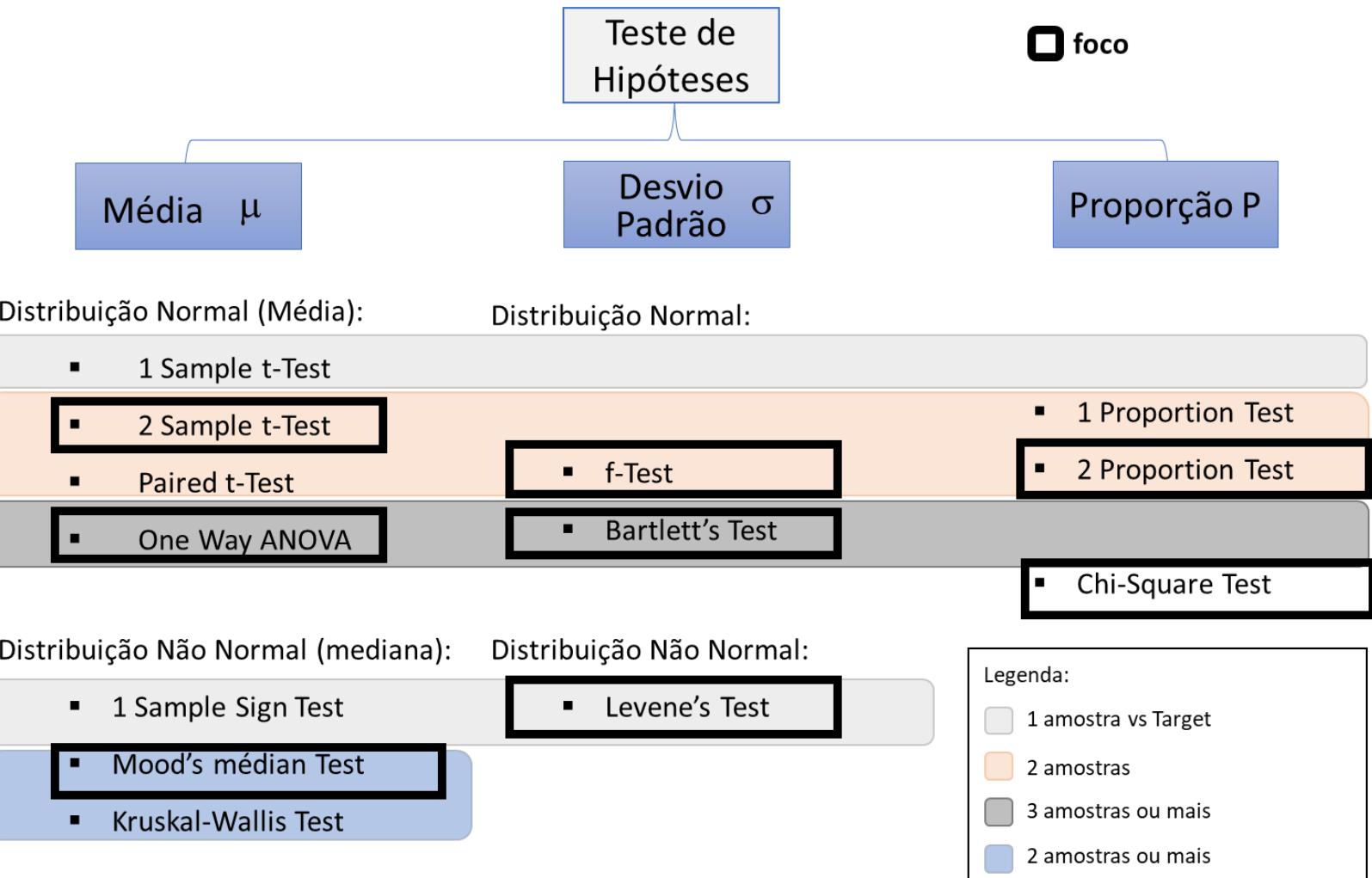
# Resumo do Processo de Teste de Hipóteses



O que quer verificar	A Hipótese Nula	A Hipótese Alternativa
Tempo médio de chamada para call centers distintos	<b>Não há</b> diferença no tempo médio de chamada entre os call centers distintos	<b>Há</b> diferença no tempo médio de chamada entre os call centers distintos
Níveis de satisfação do cliente para produtos distintos	<b>Não há</b> diferença no nível de satisfação dos clientes entre os produtos A e B	<b>Há</b> diferença no nível de satisfação dos clientes entre os produtos A e B
A variação na dimensão de um produto em relação à máquinas diferentes	<b>Não há</b> diferença na dimensão do produto vindo de máquinas diferentes	<b>Há</b> diferença na dimensão do produto vindo de máquinas diferentes
Tempo de espera de processo entre diferentes regiões	<b>Não há</b> diferença no tempo de espera entre regiões diferentes	<b>Há</b> diferença no tempo de espera entre regiões diferentes

## Exemplos de Formulação de Hipóteses

# Testes de Hipóteses



# Teste de Hipótese – Testes para Médias & Mediana

# Matriz de Seleção de Ferramentas

Discreto Contínuo

Contínuo

X

Discreto

- Diagrama de Dispersão
- Regressão Linear Simples

- Box Plot
- Teste de Hipóteses.
- Testes t, ANOVA,  
Mood's, Variância

- Regressão Logística

- Pareto
- Teste de Proporção
- Chi Quadrado

# Testes de Hipóteses

focus



Distribuição Normal (Média):

- 1 Sample t-Test
- 2 Sample t-Test
- Paired t-Test
- One Way ANOVA

Distribuição Normal:

- f-Test
- Bartlett's Test
- 1 Proportion Test
- 2 Proportion Test
- Chi-Square Test

Distribuição Não Normal (Mediana):

- 1 Sample Sign Test
- Mood's Median Test
- Kruskal-Wallis Test

Distribuição Não Normal:

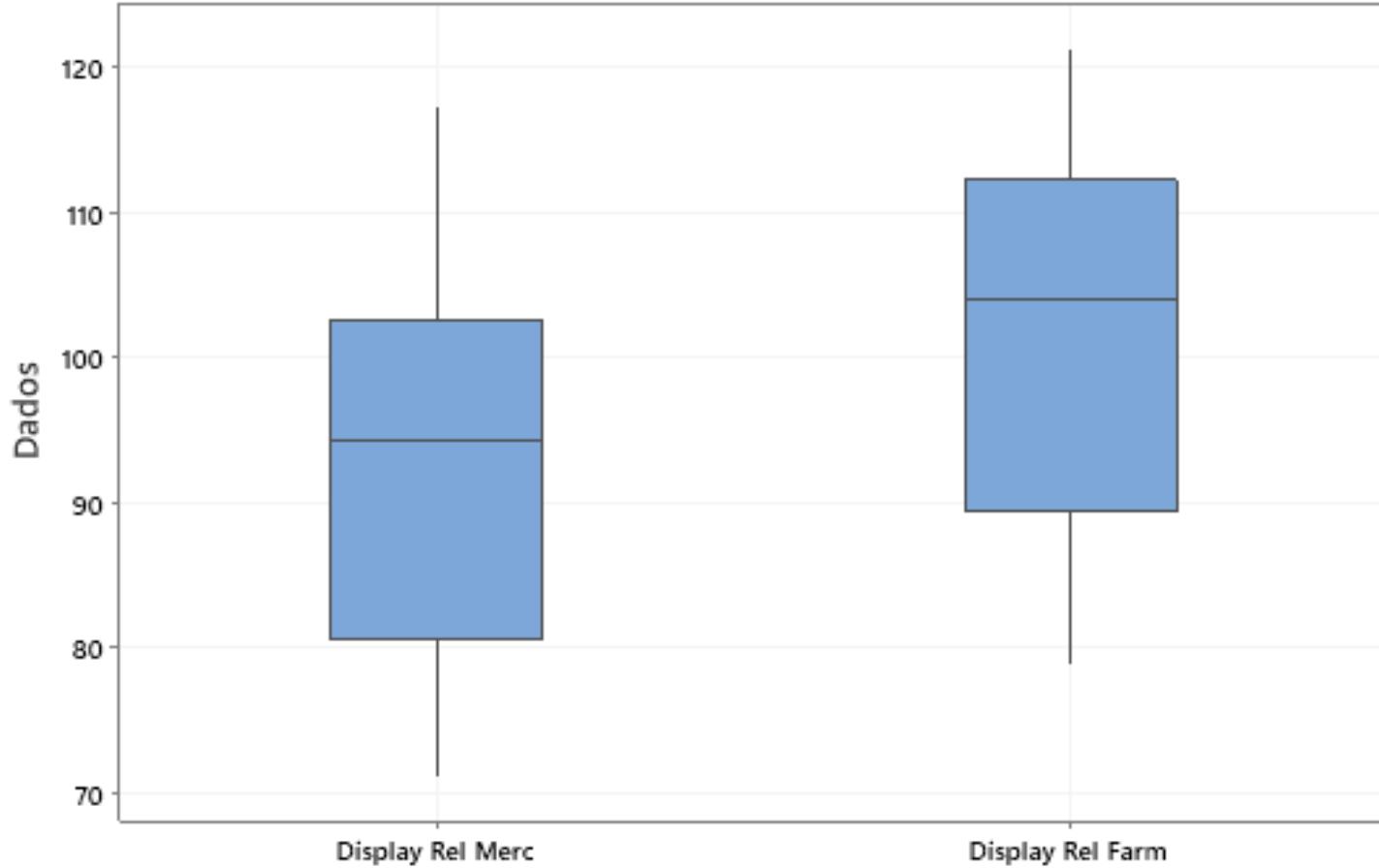
- Levene's Test

Legenda:

- 1 amostra vs Target
- 2 amostras
- 3 amostras ou mais
- 2 amostras ou mais

# Há diferença entre os grupos?

Boxplot de Display Rel Merc; Display Rel Farm



Boxplot  
Share of  
Display tipo  
de PDV

# t-Test – para comparação de duas médias

## O t-test

- É um teste de hipóteses para comparação entre duas médias
- A hipótese nula é que a média dos dois grupos é a mesma
  - Diferença entre os grupos = 0
- Se o Valor-P é < 0,05, rejeite a hipótese nula
- Isto é formulado da seguinte forma:

Hipótese Nula

$$H_0: \text{média}_A = \text{média}_B$$

Hipótese Alternativa

$$H_a: \text{média}_A \neq \text{média}_B$$

**$H_0: \text{Não há diferenças (Valor-P} > 0,05)$**

**$H_a: \text{Há diferenças (Valor-P} \leq 0,05)$**

# t-Test – para comparação de médias

## 1 Sample t-Test

Para comparar a média de uma amostra com um valor alvo ou um histórico.

## 2 Sample t-Test

Para comparar as médias de duas amostras uma contra a outra.

## Paired t-Test

Para comparar as médias das duas amostras que contêm dados vinculados aos pares (i.e. relativos ao mesmo período de tempo).

*Neste curso nos focaremos em um dos mais utilizados t-Tests : o 2 Sample t-Test.*

# Como realizar um t-Test

- Use o software estatístico de sua escolha para calcular a estatística-*t* dos dados e encontrar o Valor-P
- Se **Valor-P < 0,05** *rejeite a hipótese nula* (“as médias dos grupos é mesma”) e *aceite a hipótese alternativa* (“as médias dos grupos são diferentes”) com 95% de confiança
- Se  $\text{Value-P} \geq 0,05$  conclua que não há suficiente evidência para rejeitar a hipótese nula. Interpretações possíveis:
  - Os grupos são o mesmo ou talvez iguais
  - Não temos uma amostra do tamanho suficiente para ser capaz de detectar uma diferença



# Exemplo de aplicação do 2 Sample t-Test

## Teoria:

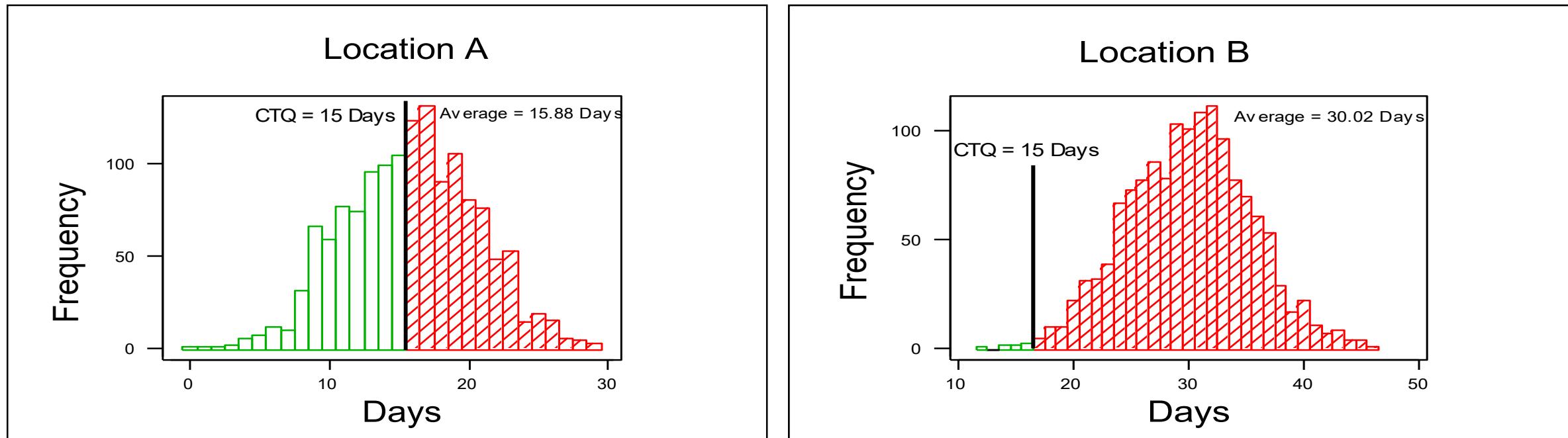
Um líder de projeto deseja validar a sua teoria de que o tempo médio de tratamento de reclamação no site A é estatisticamente diferente do da site B.

**Tipo de Dado:** X discreto X (site), Y Contínuo (tempo de ciclo do processo)

## Análise Gráfica:

Os histogramas indicam que as duas médias foram diferentes:

- ❖ Média do site A: 15,88 dias
- ❖ Média do site B: 30,02 dias



# Exemplo de aplicação do 2 Sample t-Test

A equipe verificou que um teste de hipóteses poderia lhes dar maior confiança em relação à essa teoria...

## Questão a responder com o t-Test:

A diferença entre as médias do tempo de ciclo dos sites A e B é estatisticamente significativa ou é devido à uma de variação por causas comuns?

## Hipótese nula:

- $H_0$ : O tempo médio de processamento de reclamação no site A é o mesmo que o do site B.

## Test:

- A equipe rodou um 2 Sample t-Test com um software estatístico.

## Resultado:

- O Valor-P foi menor que 0,05

## Conclusão:

- A equipe rejeitou a hipótese nula e concluiu que há diferença estatisticamente significativa entre as médias de processamento de reclamações dos sites A e B.

# Exercício de aplicação do 2 Sample t- Test

## Questão a responder com o t-Test:

A diferença entre a média de tempo de ciclo de entrega da Rota Antiga e da Rota Nova é estatisticamente significativa?

### Hipótese Nula:

- $H_0$ : A média do tempo de ciclo do processo de entrega na Rota Antiga é a mesma que na Nova Rota.

### Teste:

- Vamos rodar um 2 Sample t-Test com um software estatístico.

### Resultado:

- O Valor-P foi menor que 0,05?

### Conclusão:

- A equipe rejeitou a hipótese nula e concluiu que há uma diferença significativa entre as médias de tempo de ciclo de entrega das duas rotas?

# Exercício de aplicação do 2 Sample t-Test

Two-Sample-T-Shipping.xls

Minitab - Hypothesis Tests - Mean\_Median.MPJ

#1

File Edit Data Calc Stat Basic Statistics ►

- Regression
- ANOVA
- DOE
- Control Charts
- Quality Tools
- Reliability/Survival
- Multivariate
- Time Series
- Tables
- Nonparametrics
- Equivalence Tests
- Power and Sample Size
- 2-Sample t...**
- $\mu-\mu$  Paired t...
- 1 Proportion...
- 2 Proportions...
- $\lambda$  1-Sample Poisson Rate...
- $\lambda$  2-Sample Poisson Rate...
- $\sigma^2$  1 Variance...

2-Sample t  
Determine whether the mean differs significantly between two groups.

C6	C7-T
Summarised data	
4.5	Sample Size
1.3	
4.0	
5.7	

Two-Sample t: Options

Difference = (sample 1 mean) - (sample 2 mean)

Confidence level:

Hypothesized difference:

Alternative hypothesis: Difference ≠ hypothesized difference

Assume equal variances

**#3**

Help OK Cancel

Two-Sample t for the Mean

**#2**

Both samples are in one column

Samples: Time (days)

Sample IDs: Route

C1 Time (days)  
C2 Route  
C4 New  
C5 Old  
C7 Summarised data  
C8 New\_1  
C9 Old\_1

Select Options... Graphs...

Help OK Cancel

Two-Sample t: Graphs

Individual value plot

Boxplot

Help OK Cancel



# Exercício de aplicação do 2 Sample t-Test

## Results for: 2 Sample t Test

### Two-Sample T-Test and CI: Time (days), Route

Two-sample T for Time (days)

Route	N	Mean	StDev	SE Mean
New	20	32.12	3.46	0.77
Old	25	34.35	3.12	0.62

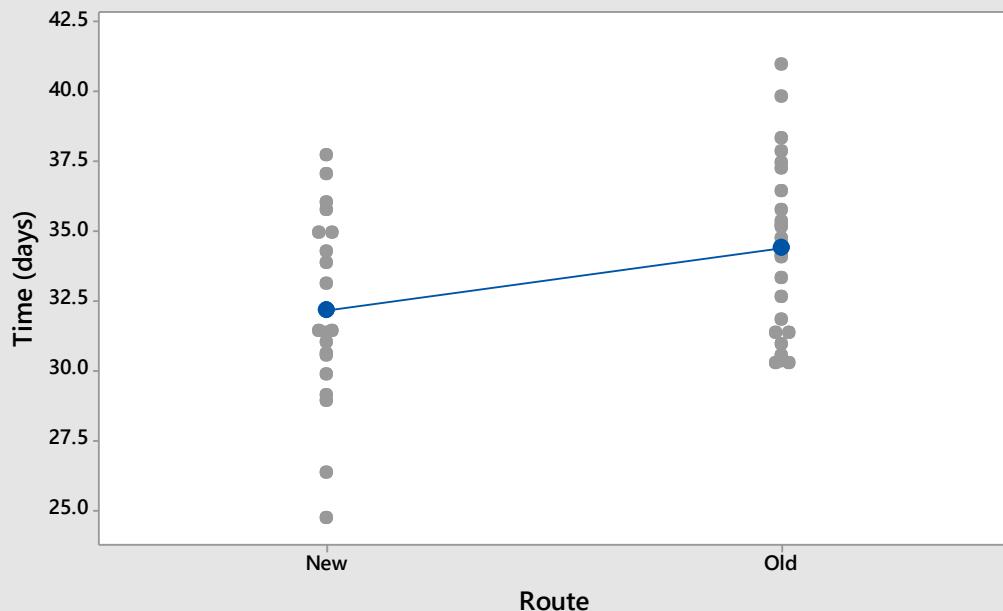
Difference =  $\mu$  (New) -  $\mu$  (Old)

Estimate for difference: -2.237

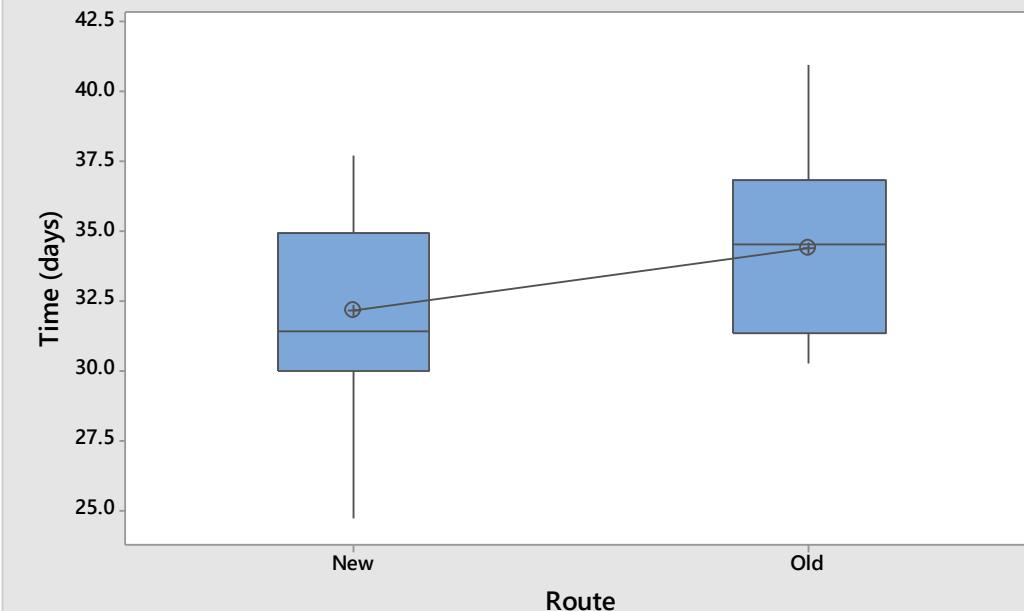
95% CI for difference: (-4.250, -0.224)

T-Test of difference = 0 (vs ≠): T-Value = -2.25 P-Value = 0.030 DF = 38

Individual Value Plot of Time (days) vs Route



Boxplot of Time (days)





# Uso do Assistente

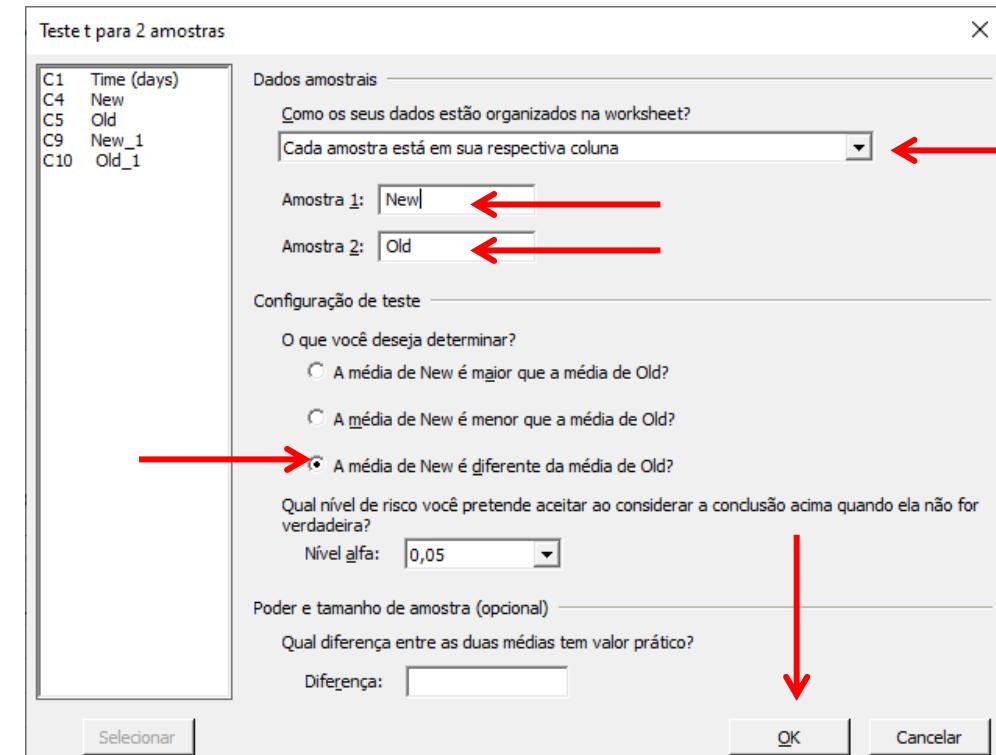
#1

Assistente Módulo de análise preditiva Ferramentas Adicionais

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
**Testes de Hipótese...**  
Regressão...  
Testes de Hipóteses  
Usar para localizar diferenças entre uma amostra e um alvo, entre duas amostras ou entre múltiplas amostras.  
Cartas de Controle...



#2



#3

# 2-Sample-t Shipping

## Teste t para 2 amostras para a Média de New e Old

### Relatório Resumo

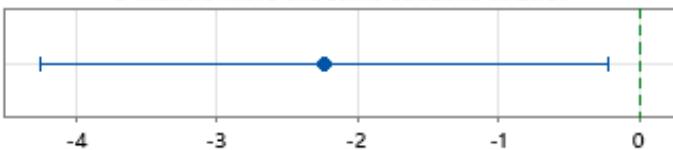
As médias diferem?



A média de New é significativamente diferente da média de Old ( $p < 0,05$ ).

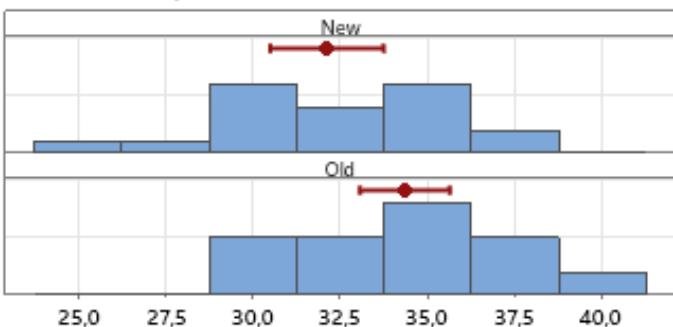
IC de 95% para a Diferença

O intervalo inteiro está acima ou abaixo de zero?



Distribuição de Dados

Compare os dados e as médias das amostras.



### Amostras Individuais

Estatísticas	New	Old
Tamanho amostral	20	25
Média	32,115	34,352
IC de 95%	(30,49; 33,74)	(33,066; 35,638)
Desvio padrão	3,4648	3,1153

### Diferença Entre Amostras

Estatísticas	*Diferença
Diferença	-2,237
IC de 95%	(-4,2496; -0,22435)

\*Diferença = New - Old

### Comentários

- Teste: você pode concluir que as médias diferem no nível de significância de 0,05.
- IC: quantifica a incerteza associada à estimativa da diferença nas médias dos dados amostrais. Você pode ter 95% de confiança de que a verdadeira diferença está entre -4,2496 e -0,22435.
- Distribuição dos Dados: compare a localização e as médias das amostras. Procure dados atípicos antes de interpretar os resultados do teste.

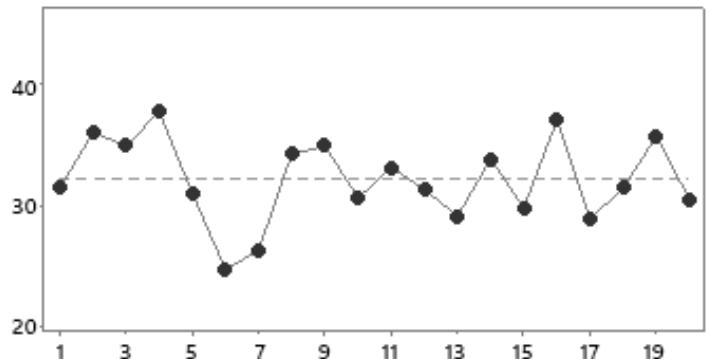
# 2-Sample-t Shipping

## Teste t para 2 amostras para a Média de New e Old Relatório de Diagnóstico

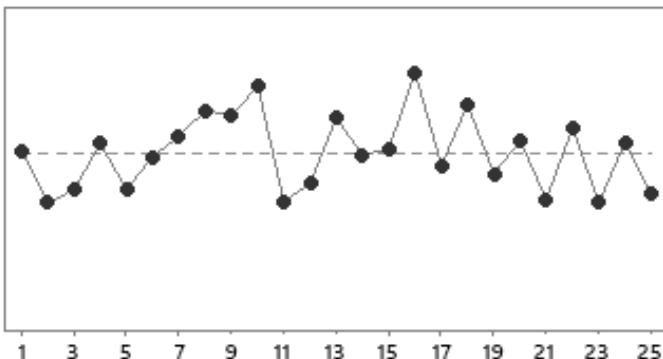
Dados na Ordem da Worksheet

Investigar todos os outliers (marcados em vermelho).

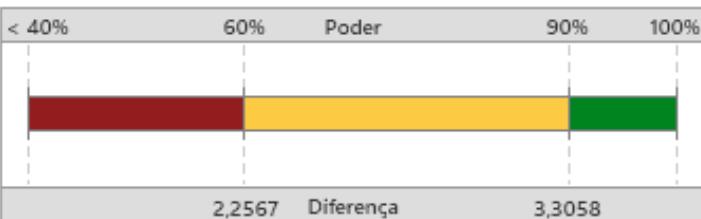
New



Old



Qual é a chance de detectar uma diferença?



Para  $\alpha = 0,05$  e tamanhos amostrais = 20; 25:

Se as médias verdadeiras diferissem em 2,2567, você teria uma chance de 60% de detectar a diferença. Se elas diferissem em 3,3058, você teria uma chance de 90%.

Que diferença você pode detectar com os tamanhos amostrais de 20 e 25?

Diferença	Poder
2,2567	60%
2,5332	70%
2,8569	80%
3,3058	90%

Diferença observada = -2,237

O poder é uma função dos tamanhos amostrais e dos desvios padrão. Para detectar diferenças menores, considere aumentar os tamanhos amostrais.

Teste t para 2 amostras para a Média de New e Old  
Cartão de Relatório

Verificar	Status	Descrição
Dados Atípicos		Não há pontos de dados atípicos. Os dados atípicos podem ter uma forte influência nos resultados.
Normalidade		Como os tamanhos de ambas as amostras são, no mínimo, de 15, a normalidade não é um problema. O teste é exato com dados não-normais quando os tamanhos amostrais são grandes o suficiente.
Tamanho da Amostra		A amostra é suficiente para detectar a diferença entre as médias.
Variância Igual		O Assistente do Minitab usa o método de Welch, que não considera nem exige que as duas amostras tenham variâncias iguais. A pesquisa mostra que o teste tem bom desempenho com variâncias desiguais, mesmo quando os tamanhos amostrais não são iguais.

2-Sample-t  
Shipping

# Business Case Logística Álcool Gel - Exemplo

Foi efetuado um acompanhamento durante alguns meses para avaliar como se comporta o OTIF.

- Há diferença significativa entre a média de On Time e In Full?
- Utilize o BoxPlot e 2-sample-t para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_Logística.xlsx Aba OTIF semanal

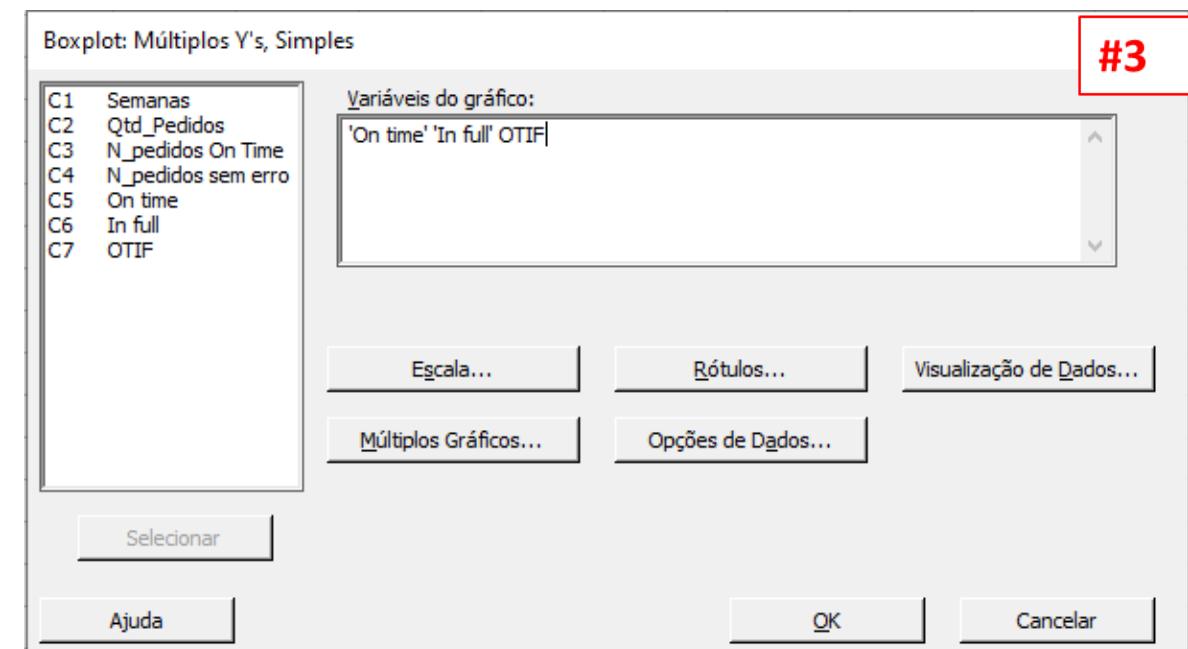
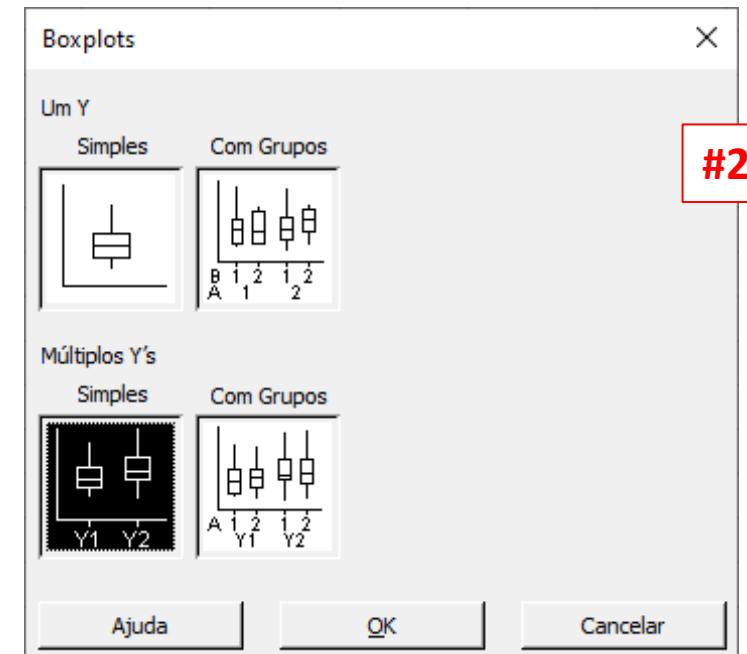


Gráfico Visualizar Ajuda Assistente Módulo de análise preditiva Ferramentas Adicionais

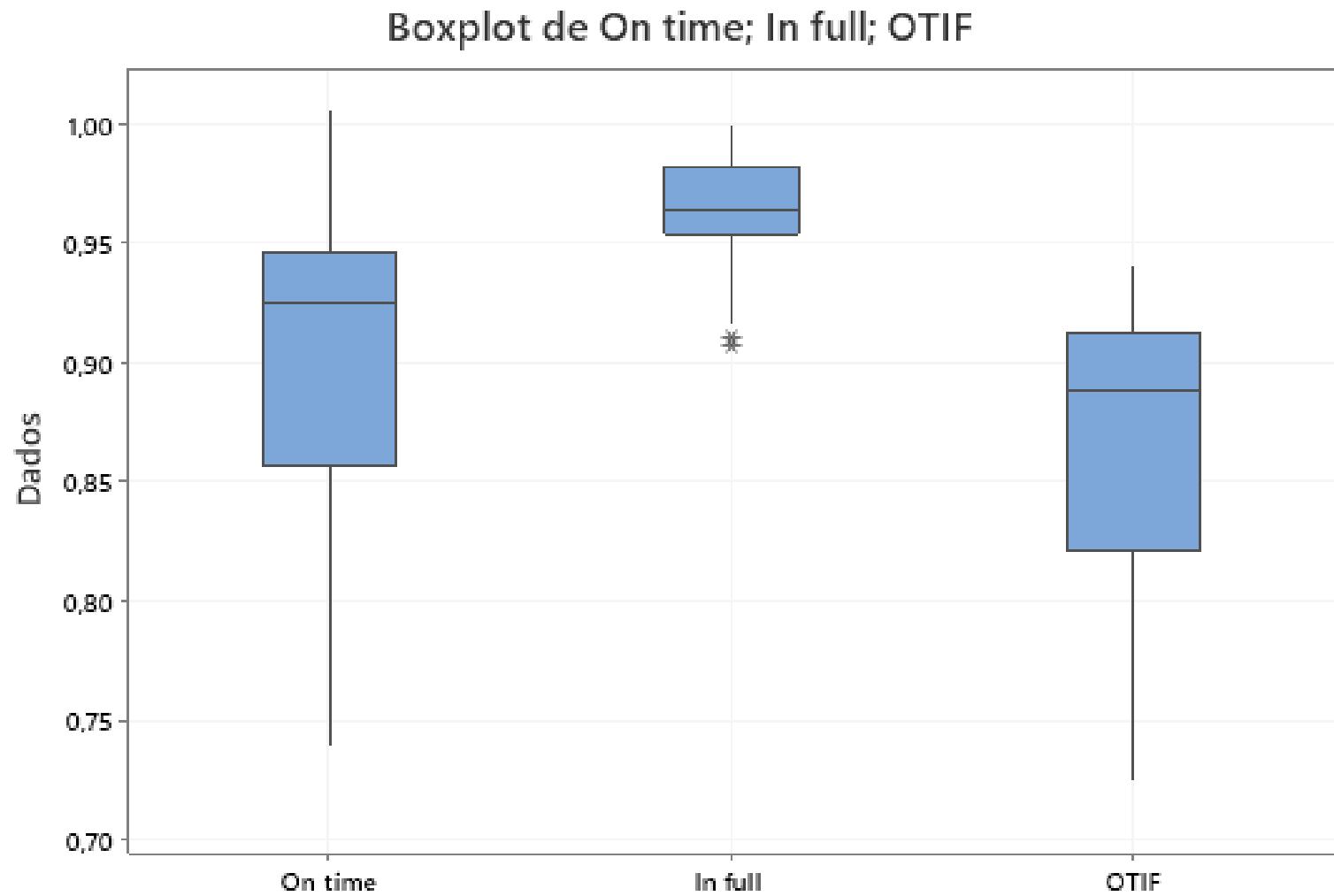
**#1**

- Construtor de Gráficos...
- Gráfico de Dispersão...
- Gráfico de Dispersão em Caixas...
- Matriz de Dispersão...
- Correlograma...
- Gráfico de Bolhas...
- Gráfico de Dispersão Marginal...
- Histograma...
- Diagrama de Pontos...
- Ramo-e-Folhas...
- Gráfico de Probabilidade...
- FDA Empírica...
- Gráfico de Distribuição de Probabilidade...
- Boxplot...**
- Gráfico de Intervalos...
- Gráfico de Boxplot
- Gráfico de Boxplot...  
Examinar o centro e a dispersão dos dados distribuindo a mediana, os quartis e quaisquer outliers. Funciona melhor com amostras de tamanho, no mínimo, moderado ( $n > 20$ ).
- Gráfico de Barras...

	C5	C6	C7	C8
On time	In full	OTIF		
0,89048	0,96667	0,860794		
0,91103	0,91103	0,829979		
0,93989	0,96995	0,911643		
0,93761	0,97302	0,912307		
0,90062	0,95652	0,861464		
0,93289	0,93960	0,876537		
0,96235	0,96000	0,923859		
0,95833	0,95673	0,916867		
0,90909	0,98485	0,895317		
0,93727	0,94834	0,888850		
0,92265	0,96133	0,886969		
0,93559	0,98305	0,919736		
0,92506	0,97403	0,910778		
	5	0,885369		
	4	0,918401		
	5	0,914999		
	0,91414	0,98485	0,900291	



Box Plot  
On Time  
In Full  
OTIF



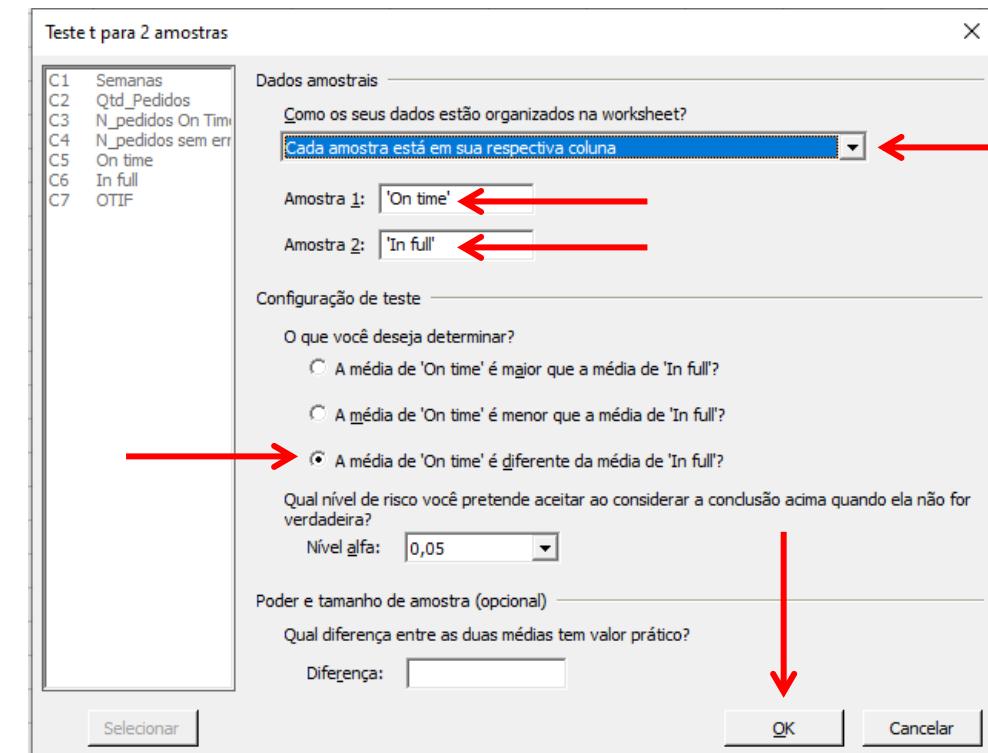


# Uso do Assistente

#1

Assistente Módulo de análise preditiva Ferramentas Adicionais

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
**Testes de Hipótese...**  
Testes de Hipóteses  
Usar para localizar diferenças entre uma amostra e um alvo, entre duas amostras ou entre múltiplas amostras.  
Cartas de Controle Antes/Depois...  
Cartas de Controle...



#2

#3

## Teste t para 2 amostras para a Média de On time e In full

### Relatório Resumo

As médias diferem?



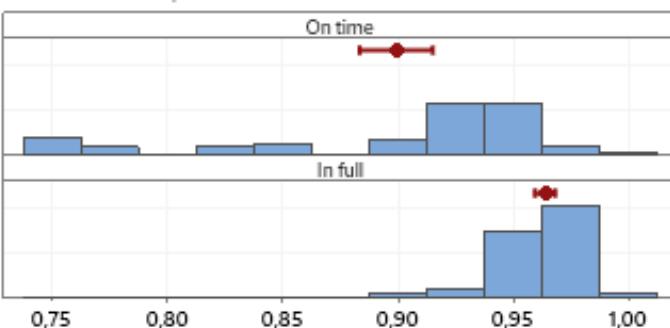
IC de 95% para a Diferença

O intervalo inteiro está acima ou abaixo de zero?



Distribuição de Dados

Compare os dados e as médias das amostras.



Amostras Individuais

Estatísticas	On time	In full
Tamanho amostral	79	79
Média	0,89906	0,96361
IC de 95%	(0,8832; 0,9149)	(0,95925; 0,96798)
Desvio padrão	0,070683	0,019486

Diferença Entre Amostras

Estatísticas	*Diferença
Diferença	-0,064553
IC de 95%	(-0,080944; -0,048162)

\*Diferença = On time - In full

Comentários

- Teste: você pode concluir que as médias diferem no nível de significância de 0,05.
- IC: quantifica a incerteza associada à estimativa da diferença nas médias das amostras. Você pode ter 95% de confiança de que a verdadeira diferença está entre -0,080944 e -0,048162.
- Distribuição dos Dados: compare a localização e as médias das amostras. Procure dados atípicos antes de interpretar os resultados do teste.

# 2-Sample-t On Time In Full

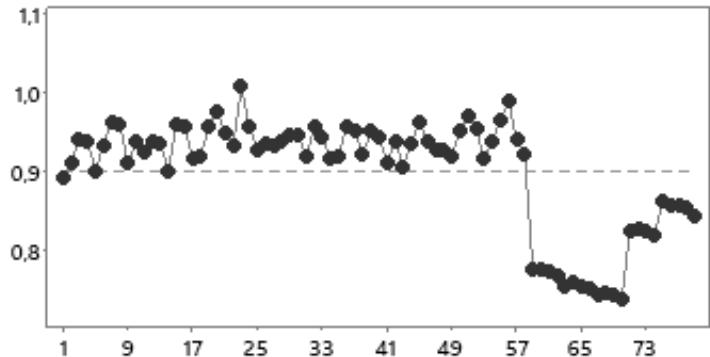
## Teste t para 2 amostras para a Média de On time e In full

### Relatório de Diagnóstico

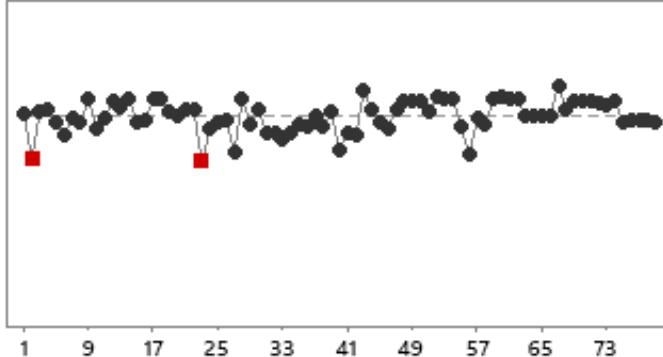
Dados na Ordem da Worksheet

Investigar todos os outliers (marcados em vermelho).

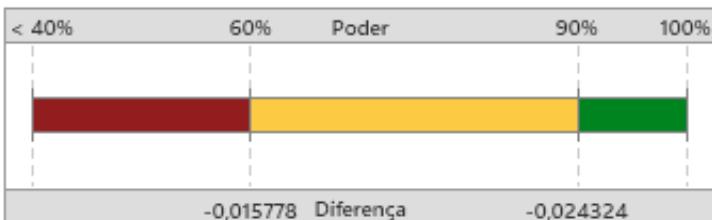
On time



In full



Qual é a chance de detectar uma diferença?



Para  $\alpha = 0,05$  e tamanhos amostrais = 79:

Se a média verdadeira de On time fosse 0,015778 menor do que In full, você teria uma chance de 60% de detectar a diferença. Se On time fosse 0,024324 menor do que In full, você teria uma chance de 90%.

Que diferença você pode detectar com tamanhos amostrais de 79?

Diferença	Poder
-0,015778	60%
-0,018031	70%
-0,020668	80%
-0,024324	90%

Diferença observada = -0,064553

# 2-Sample-t On Time In Full

O poder é uma função dos tamanhos amostrais e dos desvios padrão. Para detectar diferenças menores, considere aumentar os tamanhos amostrais.

# 2-Sample-t On Time In Full

Teste t para 2 amostras para a Média de On time e In full Cartão de Relatório		
Verificar	Status	Descrição
Dados Atípicos		Alguns dos pontos de dados são atípicos quando comparados aos outros na mesma amostra. Como os dados atípicos podem ter uma forte influência nos resultados, você deve tentar identificar a causa de sua natureza atípica. Esses pontos estão marcados em vermelho no Relatório de Diagnóstico. Você pode passar o cursor sobre um ponto ou usar o recurso da Função Brush do Minitab para identificar a linha da worksheet. Corrija quaisquer erros de entrada de dados ou medições. Considere remover os dados que estão associados a causas especiais e repetir a análise.
Normalidade		Como os tamanhos de ambas as amostras são, no mínimo, de 15, a normalidade não é um problema. O teste é exato com dados não-normais quando os tamanhos amostrais são grandes o suficiente.
Tamanho da Amostra		A amostra é suficiente para detectar a diferença entre as médias.
Variância Igual		O Assistente do Minitab usa o método de Welch, que não considera nem exige que as duas amostras tenham variâncias iguais. A pesquisa mostra que o teste tem bom desempenho com variâncias desiguais, mesmo quando os tamanhos amostrais não são iguais.

# Business Case Logística Álcool Gel - Exercício

A área financeira disse que houve um aparente aumento de inadimplência de clientes e isto pode ter sido devido ao aumento de demanda e/ou atraso nas entregas. Foram coletados dados de tempo em dias para pagamento.

- Será que há diferença no prazo de pagamentos entre os dois meses analisados?
- Utilize o BoxPlot e 2-sample-t para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_Logística.xlsx Aba Pagamentos

# Business Case Mercado – PDV Álcool Gel – Exercício

As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado no período de abr/2018 – mar/2020):

- Espaço relativo (Relative space)
- Distribuição Relativa (Relative distribution)
- Exibições Relativas (Relative displays)

Será que o tipo de PDV afeta esses resultados? Utilize Box Plot e 2-sample-t para suas conclusões

# Business Case Mercado – PDV Álcool Gel – Exercício

Foi efetuado um acompanhamento para avaliar como se comporta o Índice de Positivação implantado durante o Sprint Define.

- Há diferença significativa no valor deste Índice quando se compara o mês 1 e o 2?
- Utilize o BoxPlot e 2-sample-t para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_PDV.xlsx na aba Índice de Positivação

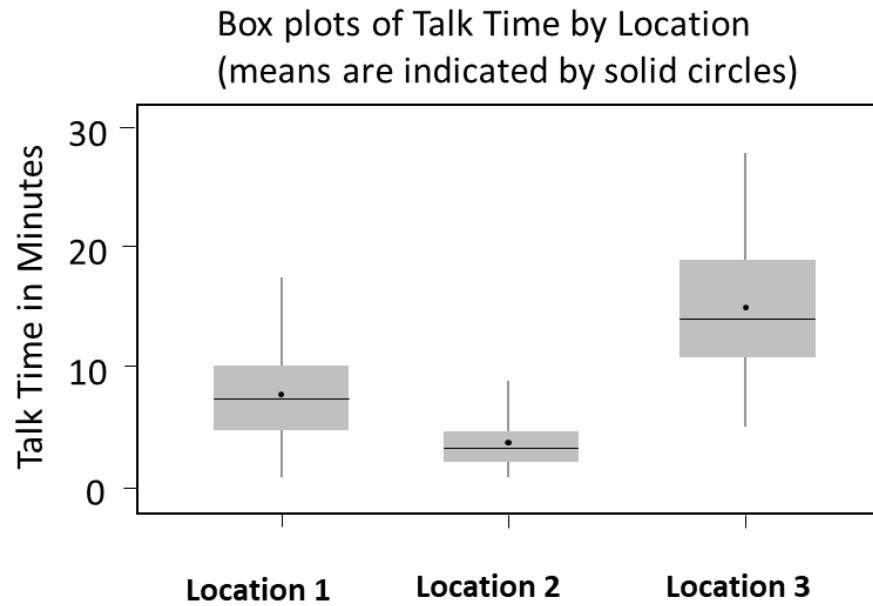
# ANOVA (Analysis of Variance – Análise de Variância)

- É um **teste de hipóteses** utilizado para comparar a média de três ou mais grupos de dados simultaneamente.
- Diferente do **t-Test**, ela compara a variância **entre** grupos em relação a variância **dentro** dos grupos
  - A variância entre grupos é obtida pela variância ( $\sigma^2$ ) da **média dos grupos**
  - A variância dentro dos grupos é obtida a partir das variâncias ( $\sigma^2$ ) entre os valores dentro de cada grupo e, em seguida, **combinada entre os grupos**
- Se a variância entre os grupos é a mesma que a variância entre os grupos, então, pode ser concluído que não há diferença entre as médias dos grupos

# ANOVA

- A hipótese nula é que as médias dos três (ou mais) grupos são a mesma
- A hipótese alternativa é que pelo menos uma média é diferente ou
  - $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \dots = \mu_k;$
  - $H_a:$  pelo menos uma  $\mu_k$  é diferente
- Se Valor-P < 0,05 rejeite a hipótese nula, conclua que pelo menos um grupo é diferente
- Se Valor- P  $\geq 0,05$  não rejeite a hipótese nula

# Exemplo de aplicação da ANOVA



## Teoria:

- O tempo médio das chamadas de atendimento ao cliente é estatisticamente diferente entre os call centers dos locais 1, 2 e 3.

**Tipo de dado:** X discreto X (local), Y Contínuo (tempo de chamada)

## Análise gráfica:

- O box plot indica que as três medias são diferentes
  - Média Local 1: 8 minutos
  - Média Local 2: 5 minutos
  - Média Local 3: 16 minutos

## Abordagem:

- Foram obtidos os dados de 100 atendimentos durante uma semana em cada local.

# Exemplo de aplicação da ANOVA

A equipe sabia que um teste de hipóteses poderia dar mais confiança para a sua teoria....

## Questão a responder com a ANOVA:

A diferença no tempo médio de chamada dos atendimentos efetuados pelos call centers do locais 1, 2 e 3 é significativa ou é apenas uma função da variação devido à causas comuns?

### Hipótese Nula:

- $H_0$ : As médias de tempo de atendimento nos call centers dos locais 1, 2 e 3 são a mesma.

### Premissas:

- Confirme que o processo é estável e que os subgrupos são normais

### Teste:

- A equipe rodou uma ANOVA com um software estatístico

### Resultado:

- O Valor-P foi maior que 0,05

### Conclusão:

- A equipe aceitou a hipótese nula e concluiu que não há diferença significativa entre as medias de atendimento dos call centers nas três diferentes localidades

# Exercício de aplicação da ANOVA

## Questão a responder com a ANOVA:

A diferença na média de vendas entre os diferentes Centros de Vendas é significativa?

### Hipótese Nula:

- $H_0$ : A média de vendas entre os diferentes Centros é a mesma

### Teste:

- Vamos rodar uma ANOVA em um programa estatístico

### Resultado:

- O P-Valor é menor que 0,05?

### Conclusão:

- A equipe rejeitou a hipótese nula e concluiu que há pelo menos um Centro com média de vendas diferente da dos demais?

# Exercício de aplicação da ANOVA

ANOVA-Sales.xls

Minitab - Hypothesis Tests - Mean\_Median.MPJ

#1

Stat > ANOVA > One-Way...

**One-Way**  
Determine whether the means of two or more groups differ.

#2

One-Way Analysis of Variance

Response data are in one column for all factor levels

Response: 'Sales (GBP)'  
Factor: Subscripts

#3

One-Way Analysis of Variance: Options

Assume equal variances  
Confidence level: 95  
Type of confidence interval: Two-sided

One-Way Analysis of Variance: Graphs

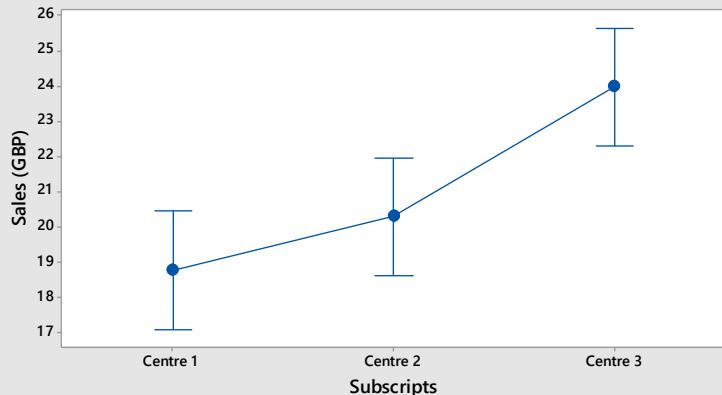
Data plots:  
 Interval plot  
 Individual value plot  
 Boxplot of data

Residual plots:  
 Individual plots  
 Histogram of residuals  
 Normal probability plot of residuals  
 Residuals versus fits  
 Residuals versus order  
 Four in one

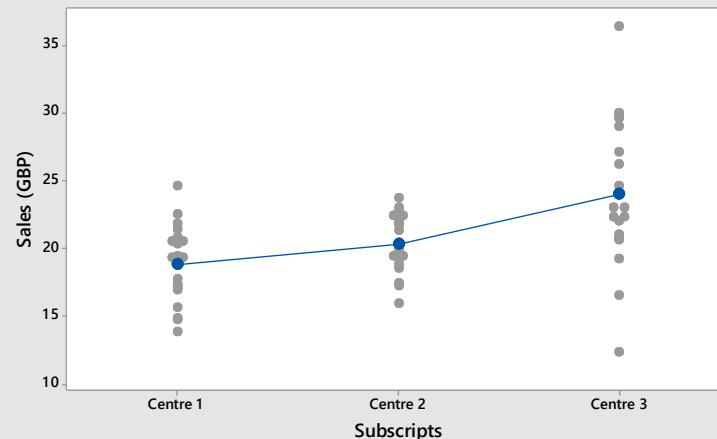
Residuals versus the variables:

OK Cancel Select Results... Storage... Help OK Cancel

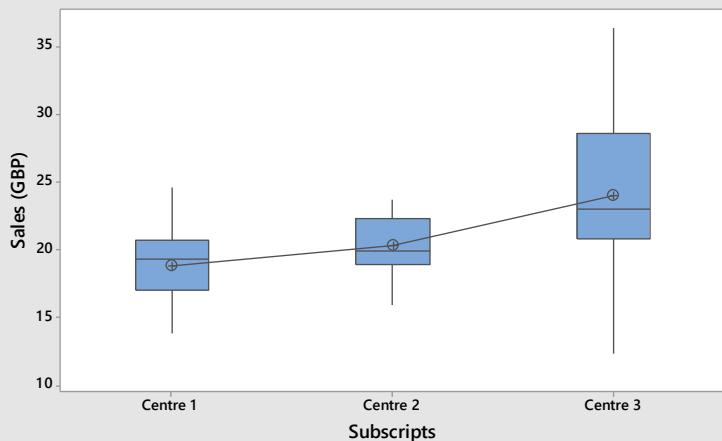
Interval Plot of Sales (GBP) vs Subscripts  
95% CI for the Mean



Individual Value Plot of Sales (GBP) vs Subscripts



Boxplot of Sales (GBP)



#### Factor Information

Factor	Levels	Values
Subscripts	3	Centre 1, Centre 2, Centre 3

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Subscripts	2	285.8	142.90	10.25	0.000
Error	57	794.8	13.94		
Total	59	1080.6			

# Exercício de aplicação da ANOVA



# Uso do Assistente

#1

Assistente Módulo de análise preditiva Ferramentas Adicionais

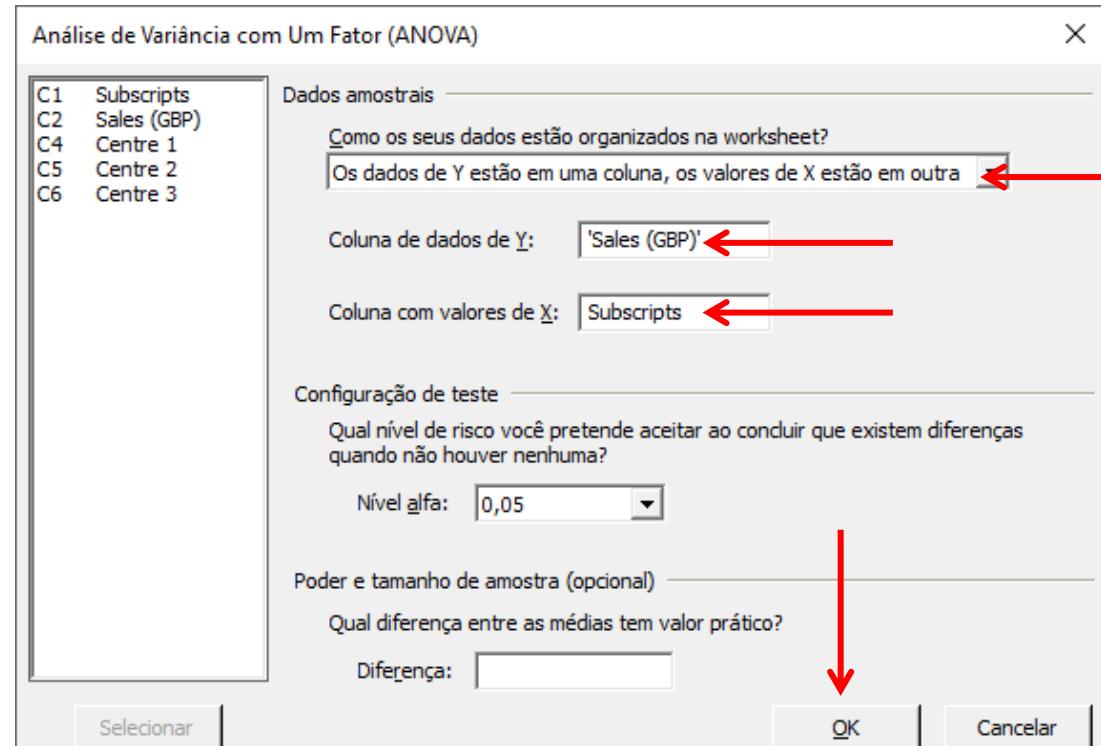
Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
**Testes de Hipótese...**  
Regressão...  
**Testes de Hipóteses**

Usar para localizar diferenças entre uma amostra e um alvo, entre duas amostras ou entre múltiplas amostras.

Cartas de Controle...



#2



#3

## ANOVA com um Fator para Sales (GBP) por Subscripts

### Relatório Resumo

As médias diferem?



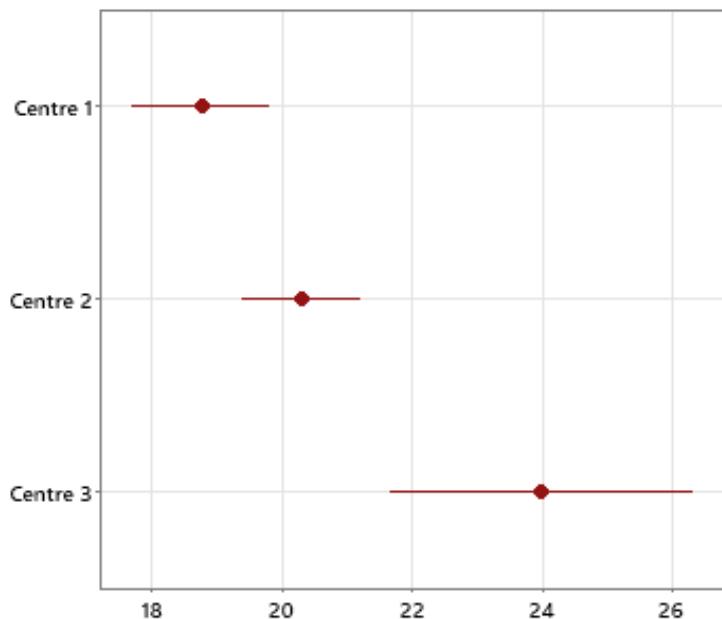
As diferenças entre as médias são significativas ( $p < 0,05$ ).

Quais médias diferem?

#	Amostra	Difere de
1	Centre 1	3
2	Centre 2	3
3	Centre 3	1 2

#### Carta de Comparação de Médias

Intervalos vermelhos que não se sobrepõem, diferem.



#### Comentários

- Teste: Você pode concluir que existem diferenças entre as médias no nível 0,05 de significância.
- Gráfico de Comparação: Procure os intervalos de comparação em vermelho que não se sobrepõem para identificar as médias que diferem entre si. Considere o tamanho das diferenças para determinar se elas têm implicações práticas.

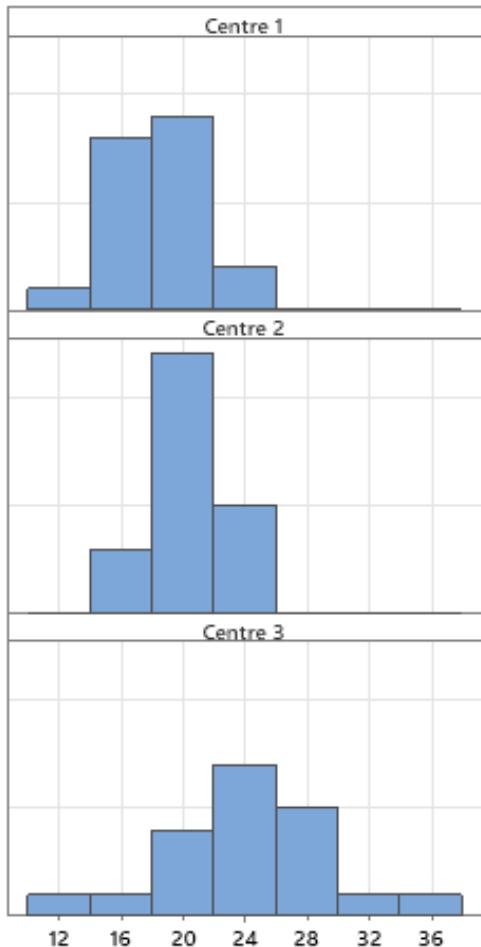
# ANOVA Sales

## ANOVA com um Fator para Sales (GBP) por Subscripts

### Relatório de Diagnóstico

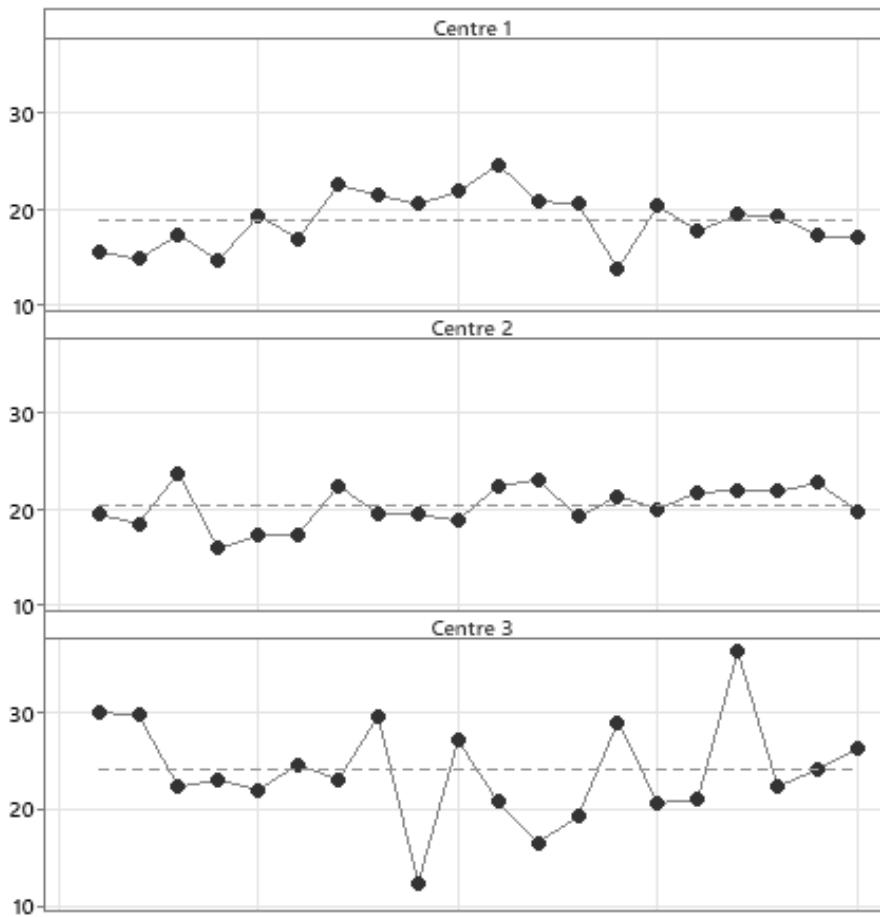
#### Distribuição de Dados

Compare a localização e a dispersão.



#### Dados na Ordem da Worksheet

Investigar todos os outliers (marcados em vermelho).

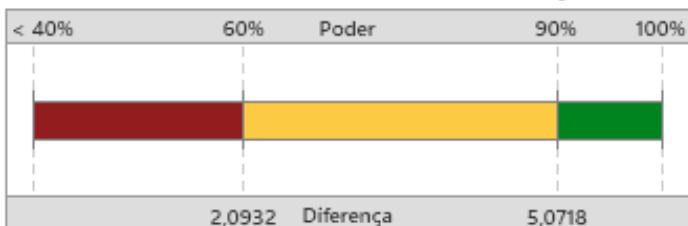


# ANOVA Sales

## ANOVA com um Fator para Sales (GBP) por Subscripts

### Relatório de Poder

Qual é a chance de detectar uma diferença?



Com base nas amostras e no  $\alpha$  (0,05), você tem, no mínimo, uma chance de 90% de detectar uma diferença de 5,0718 e, no máximo, uma chance de 60% de detectar uma diferença de 2,0932.

Que diferença você pode detectar com os tamanhos amostrais?

Diferença	Poder
2,0932	24,5 - 60,0%
3,5419	60,0 - 97,6%
3,9519	70,0 - 99,5%
4,4275	80,0 - 100,0%
5,0718	90,0 - 100,0%

O poder é uma função dos tamanhos amostrais e dos desvios padrão. Para detectar diferenças menores do que 4,4275, considere aumentar os tamanhos amostrais.

#### Estatísticas

Subscripts	Tamanho da Amostra	Média	Desvio Padrão	IC de 95% Individual para a Média
Centre 1	20	18,775	2,8703	(17,432; 20,118)
Centre 2	20	20,3	2,1506	(19,293; 21,307)
Centre 3	20	23,975	5,3820	(21,456; 26,494)

# ANOVA Sales

## ANOVA com um Fator para Sales (GBP) por Subscripts

### Cartão de Relatório

Verificar	Status	Descrição
Dados Atípicos		Não há pontos de dados atípicos. Os dados atípicos podem ter uma forte influência nos resultados.
Tamanho da Amostra		A amostra é suficiente para detectar diferenças entre as médias.
Normalidade		Como todos os tamanhos amostrais são, no mínimo, de 15, a normalidade não é um problema. O teste é exato com dados não-normais quando os tamanhos amostrais são grandes o suficiente.
Variância Igual		O Assistente do Minitab usa o método de Welch, que não considera nem exige que as amostras tenham variâncias iguais. A pesquisa mostra que o teste tem bom desempenho com variâncias desiguais, mesmo quando os tamanhos amostrais não são iguais.

# ANOVA Sales

# Business Case Produtividade Álcool Gel - Exemplo

Foram coletados dados semanais do desempenho da eficiência operacional geral e por turnos para entender o problema de variabilidade entre turnos.

- Há diferença significativa na eficiência operacional geral entre os turnos?
- Utilize o BoxPlot e ANOVA para suas conclusões

Os dados estão no arquivo Case  
CoolGel\_Dados\_Produtividade na aba Eficiência



Gráfico Visualizar Ajuda Assistente Módulo de análise pr...

#1

- Construtor de Gráficos...
- Gráfico de Dispersão...
- Gráfico de Dispersão em Caixas...
- Matriz de Dispersão...
- Correlograma...
- Gráfico de Bolhas...
- Gráfico de Dispersão Marginal...
- Histograma...
- Diagrama de Pontos...
- Ramo-e-Folhas...
- Gráfico de Probabilidade...
- FDA Empírica...
- Gráfico de Distribuição de Probabilidade...
- Boxplot...  
Boxplot  
Examinar o centro e a dispersão dos dados distribuindo a mediana, os quartis e quaisquer outliers. Funciona melhor com amostras de tamanho, no mínimo, moderado ( $n > 20$ ).  
Selecionar

Boxplots

Um Y

Simples Com Grupos

C5

Geral

60,0

55,7

59,3

57,0

55,7

58,3

54,3

58,7

57,7

58,0

54,7

59,0

50,3

56,3

58,7

56,3

Boxplots

Múltiplos Y's

Simples Com Grupos

Y1 Y2

A1 A2 Y1 Y2

Ajuda OK Cancelar

Boxplot: Múltiplos Y's, Simples

Variáveis do gráfico:

C1 Semana  
C2 Turno1  
C3 Turno2  
C4 Turno3  
C5 Geral

Turno1 Turno2 Turno3 Geral

Escala... Rótulos... Visualização de Dados...

Múltiplos Gráficos... Opções de Dados...

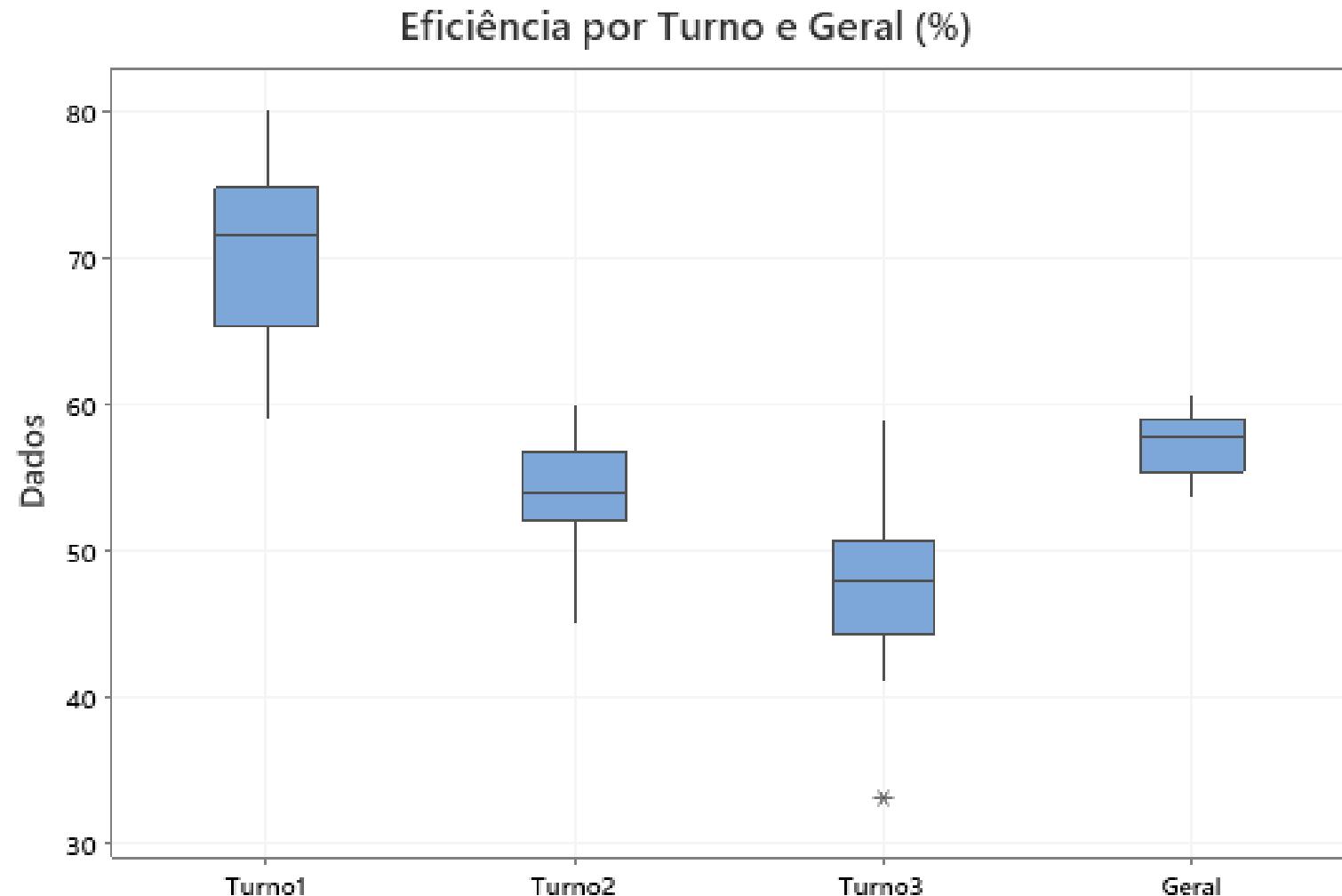
Selecionar

Ajuda OK Cancelar

#2

#3

Box Plot  
Eficiência  
por Turno e  
Geral





# Uso do Assistente

#1

Assistente Módulo de análise preditiva Ferramentas Adicionais

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
**Testes de Hipótese...**  
Testes de Hipóteses  
Usar para localizar diferenças entre uma amostra e um alvo, entre duas amostras ou entre múltiplas amostras.  
Cartas de Controle Antes/Depois...  
Cartas de Controle...



#2

Análise de Variância com Um Fator (ANOVA)

Dados amostrais  
Como os seus dados estão organizados na worksheet?  
Os dados de Y para cada valor de X estão em colunas separadas

Colunas de dados de Y:  
Turno1 Turno2 Turno3

Configuração de teste  
Qual nível de risco você pretende aceitar ao conduir que existem diferenças quando não houver nenhuma?  
Nível alfa: 0,05

Poder e tamanho de amostra (opcional)  
Qual diferença entre as médias tem valor prático?  
Diferença: \_\_\_\_\_

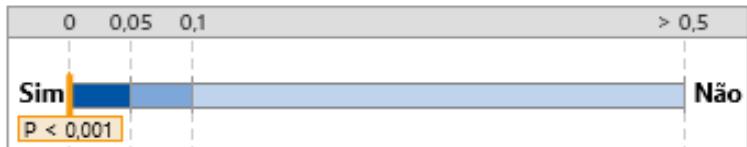
OK Cancelar

#3

## ANOVA com um Fator para Turno1; Turno2; Turno3

### Relatório Resumo

As médias diferem?



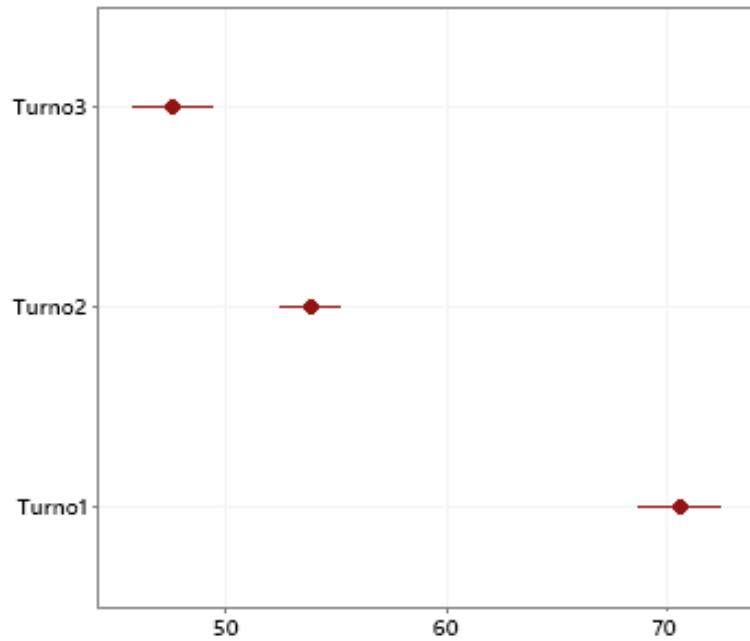
Quais médias diferem?

#	Amostra	Difere de
1	Turno3	2 3
2	Turno2	1 3
3	Turno1	1 2

As diferenças entre as médias são significativas ( $p < 0,05$ ).

#### Carta de Comparação de Médias

Intervalos vermelhos que não se sobrepõem, diferem.



#### Comentários

- Teste: Você pode concluir que existem diferenças entre as médias no nível 0,05 de significância.
- Gráfico de Comparação: Procure os intervalos de comparação em vermelho que não se sobrepõem para identificar as médias que diferem entre si. Considere o tamanho das diferenças para determinar se elas têm implicações práticas.

# ANOVA

# Eficiência

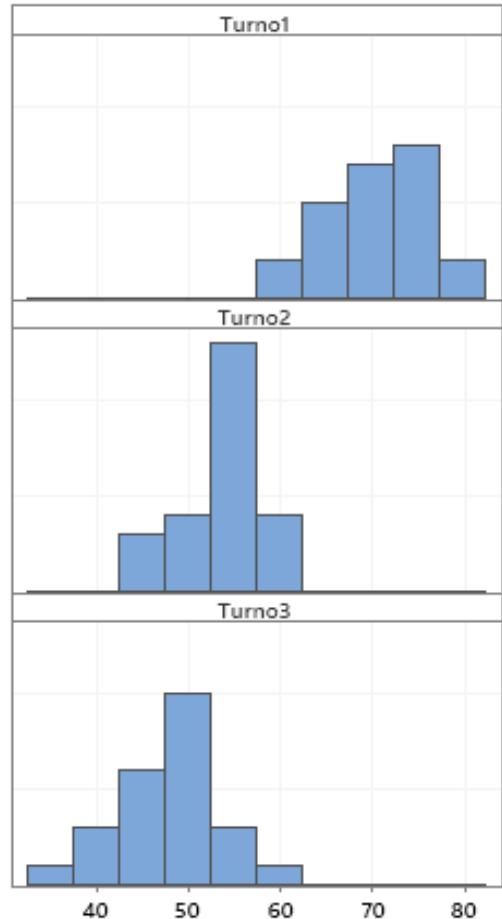
# por Turno

## ANOVA com um Fator para Turno1; Turno2; Turno3

### Relatório de Diagnóstico

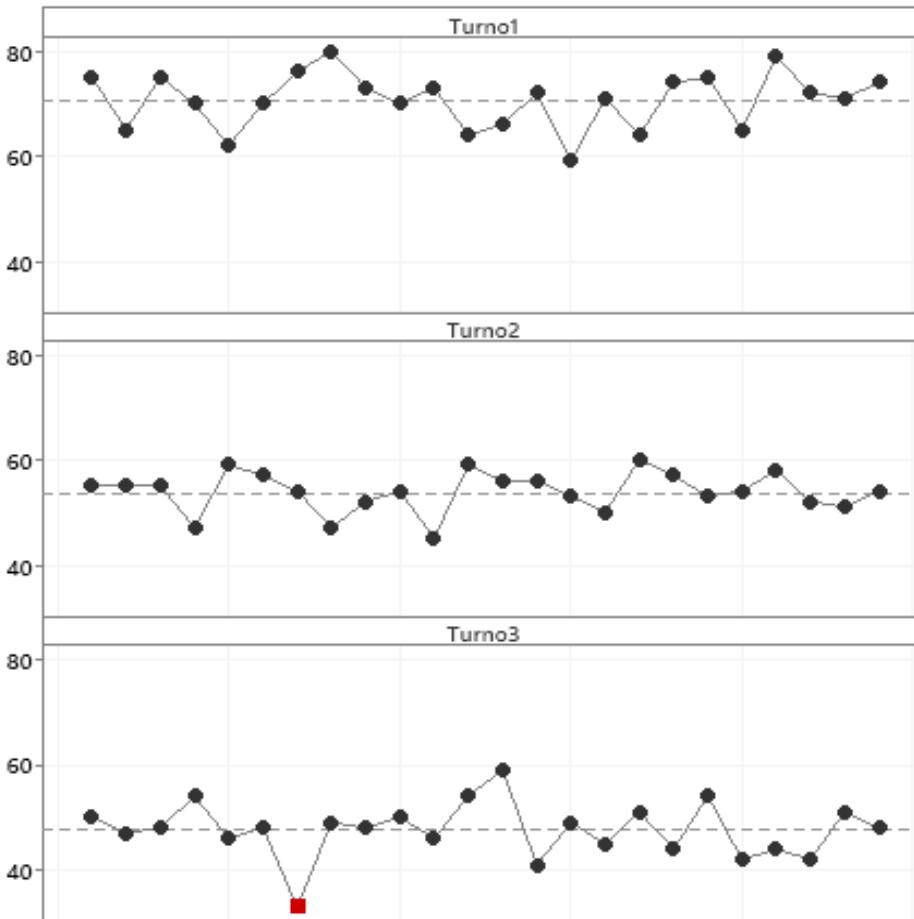
#### Distribuição de Dados

Compare a localização e a dispersão.



#### Dados na Ordem da Worksheet

Investigar todos os outliers (marcados em vermelho).

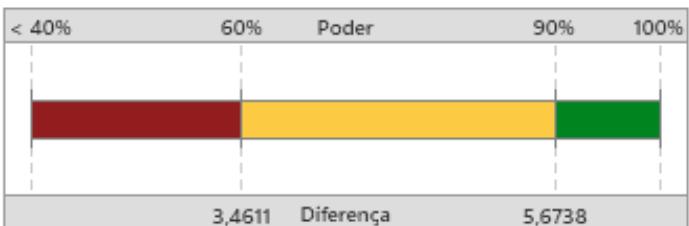


ANOVA  
Eficiência  
por Turno

## ANOVA com um Fator para Turno1; Turno2; Turno3

### Relatório de Poder

Qual é a chance de detectar uma diferença?



Com base nas amostras e no  $\alpha$  (0,05), você tem, no mínimo, uma chance de 90% de detectar uma diferença de 5,6738 e, no máximo, uma chance de 60% de detectar uma diferença de 3,4611.

O poder é uma função dos tamanhos amostrais e dos desvios padrão. Para detectar diferenças menores do que 4,9506, considere aumentar os tamanhos amostrais.

Que diferença você pode detectar com os tamanhos amostrais?

Diferença	Poder
3,4611	48,1 - 60,0%
3,9687	60,0 - 72,6%
4,4225	70,0 - 82,0%
4,9506	80,0 - 90,1%
5,6738	90,0 - 96,6%

# ANOVA

# Eficiência

# por Turno

## ANOVA com um Fator para Turno1; Turno2; Turno3

### Cartão de Relatório

Verificar	Status	Descrição
Dados Atípicos		Um ponto de dados (linha 7) é atípico quando comparado aos outros em Turno3. Como os dados atípicos podem ter uma forte influência nos resultados, tente identificar a causa de sua natureza atípica. Corija quaisquer erros de entrada de dados ou medições. Considere remover os dados que estão associados com causas especiais e repetir a análise.
Tamanho da Amostra		A amostra é suficiente para detectar diferenças entre as médias.
Normalidade		Como todos os tamanhos amostrais são, no mínimo, de 15, a normalidade não é um problema. O teste é exato com dados não-normais quando os tamanhos amostrais são grandes o suficiente.
Variância Igual		O Assistente do Minitab usa o método de Welch, que não considera nem exige que as amostras tenham variâncias iguais. A pesquisa mostra que o teste tem bom desempenho com variâncias desiguais, mesmo quando os tamanhos amostrais não são iguais.

# ANOVA Eficiência por Turno

# Business Case Logística Álcool Gel - Exercício

Foi efetuado um acompanhamento durante alguns meses para avaliar como se comporta o tempo de carregamento dos caminhões.

- Há diferença significativa na média do tempo de carregamento entre os meses?
- Utilize o BoxPlot, ANOVA e 2-sample-t (caso necessário) para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_Logística.xlsx Aba Tempo Carreg\_Pos

# Business Case Mercado – PDV Álcool Gel – Exercício

As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado no período de abr/2018 – mar/2020):

- Espaço relativo (Relative space)
- Distribuição Relativa (Relative distribution)
- Exibições Relativas (Relative displays)

Será que há diferença significativa entre esses indicadores?  
Utilize Box Plot e ANOVA para suas conclusões

# Business Case Mercado – PDV Álcool Gel – Exercício

Foi efetuado um acompanhamento para avaliar como se comporta o Índice de Positivação implantado durante o Sprint Define.

- Há diferença significativa no valor deste Índice em relação à semana do mês ou dia da semana?
- Utilize o BoxPlot e ANOVA para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_PDV.xlsx na aba Índice de Positivação

# Business Case Mercado – PDV Álcool Gel – Exercício

- Houve variação de Market Share ao longo dos anos?
- Volume de Vendas é afetado pelo número de SKU's por PDV?

Utilize Box Plot e ANOVA para as suas conclusões

Os dados estão no arquivo Coolgel\_Banco de Dados\_PDV.xlsx nas abas “Market Share Decline” e “Volume de Vendas x SKU”

# Business Case Mercado – PDV Álcool Gel – Exercício

Foi efetuado um acompanhamento para avaliar a variação de PTC dentro de um cinco grupos diferentes de PDV's em dias diferentes.

- O PTC médio dos grupos é diferente?
- A dispersão de PTC é diferente entre esses grupos?
- Utilize Box Plot, ANOVA e teste de Variância para as suas conclusões

Os dados estão no arquivo Coolgel\_Banco de Dados\_PDV.xlsx na aba “Variação PTC”

# Mood's Median Test – Para comparar duas medianas

## Mood's Median Test (dados Não Normais + outliers)

- É um teste de hipótese para comparação entre duas medianas
- A hipótese nula é que a mediana dos dois grupos é a mesma
  - Diferença entre as medianas = 0
- Se Valor-P < 0,05 rejeite a hipótese nula

Isto é expresso por:

### Hipótese Nula

$$H_0: \text{mediana}_A = \text{mediana}_B$$

### Hipótese Alternativa

$$H_a: \text{mediana}_A \neq \text{mediana}_B$$

$H_0: \text{Não há diferenças (Valor-P} > 0,05 \text{ )}$

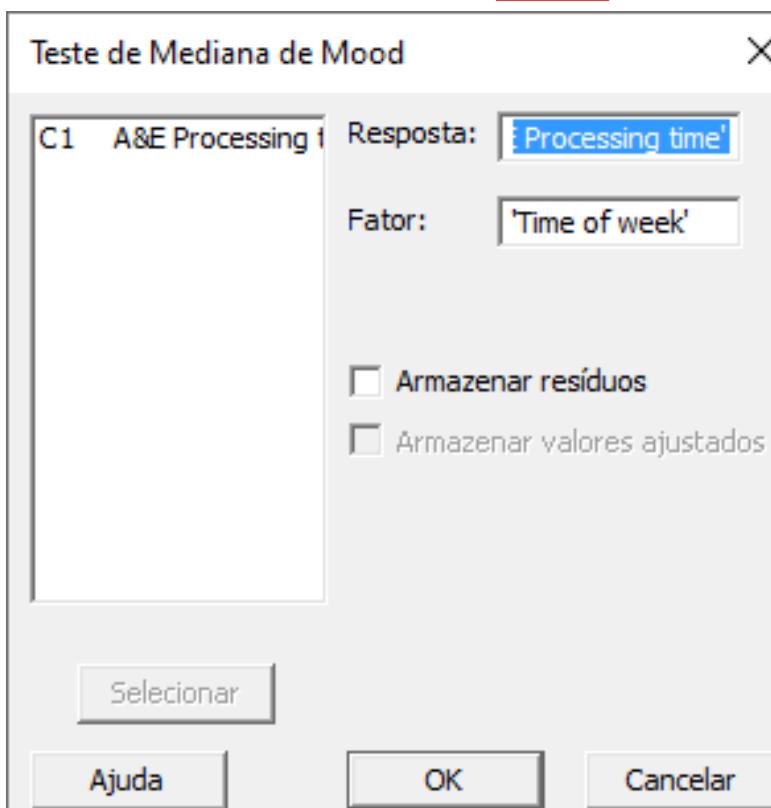
$H_a: \text{Há diferenças (Valor-P} \leq 0,05 \text{)}$

# Exercício de aplicação do Mood's Median Test

- **Pergunta para responder com Mood's Median-Test:**  
A diferença entre o tempo de processamento nos dias de semana e finais de semana é significativa?
- **Hipótese:**  
 $H_0$ : A mediana do tempo para o atendimento de pacientes no pronto socorro nos dias de semana é a mesma que para nos finais de semana
- **Teste:**  
Vamos rodar um Mood's médian Test com um software estatístico.  
Lembre-se: dados não normais
- **Resultado:**  
O Valor-P é menor que 0,05?
- **Conclusão:**  
A equipe rejeitou a hipótese nula e concluiu que há uma diferença significativa entre as medianas do tempo para o atendimento nos dois períodos?

# Exercício de aplicação do Mood's Median Test

Dados Moods  
median Test.xlsx



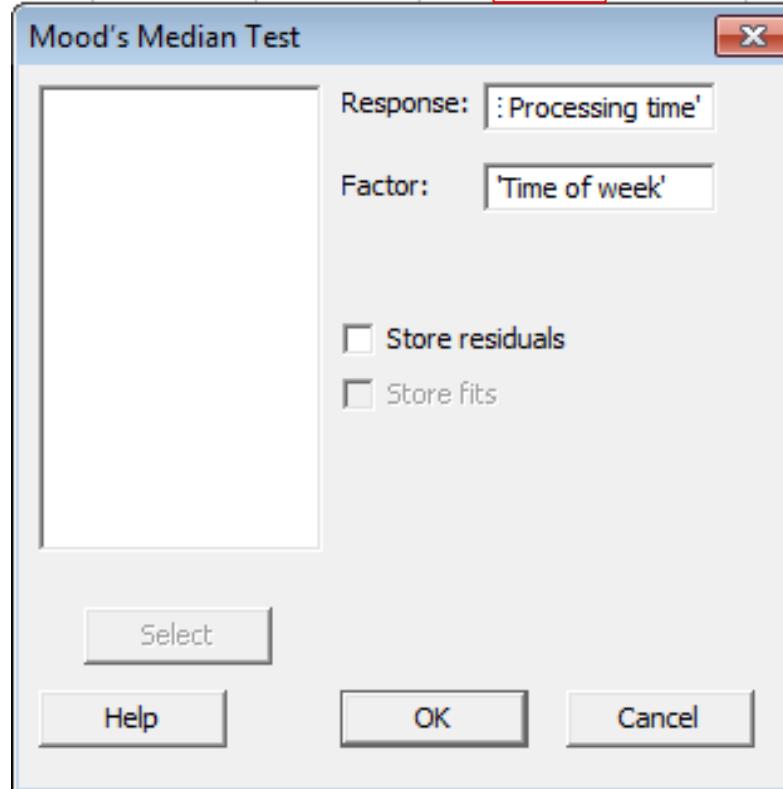
#1

#2

The menu path is: Estat > Não-Paramétricos > Teste de Mood para a Mediana. The 'Teste de Mood para a Mediana' option is highlighted. A tooltip for this option is displayed: "Determinar se as medianas de dois ou mais grupos diferem. Usar quando você não tiver certeza de que os grupos têm distribuições com formas semelhantes." Below the tooltip are two other options: "Diferenças Pareadas..." and "Inclinações Pareadas...".

# Exercício de aplicação do Mood's Median Test

Dados Moods  
médian Test.xlsx



#1

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Basic Statistics

Regression

ANOVA

DOE

Control Charts

Quality Tools

Reliability/Survival

Multivariate

Time Series

Tables

Nonparametrics

Equivalence Tests

Power and Sample Size

1-Sample Sign...

1-Sample Wilcoxon...

Mann-Whitney...

Kruskal-Wallis...

Mood's Median Test...

Friedman...

Runs Test...

Pairwise Averages...

Pairwise Differences...

Pairwise Slopes...

Mood's Median Test

Determine whether the medians of two or more groups differ. Use when you are not confident that the groups have similarly shaped distributions.

## Estatísticas Descritivas

Time of week	Mediana	Média geral	N <=	Média geral	N >	Q3 – Q1	IC de 95% da mediana
Weekdays	2,75		20		10	2,3125	(2,5; 4,13564)
Weekends	4,25			10	15	1,7500	(3,75; 5)
Global	3,75						

IC de 95,0% para mediana(Weekdays) - mediana(Weekends): (-2,27134; -0,228664)

## Teste

Hipótese nula  $H_0$ : as medianas da população são todas iguais

Hipótese alternativa  $H_1$ : as medianas da população não são todas iguais

GL	Qui-Quadrado	Valor-p
1	3,91	0,048

Exercício de aplicação do Mood's Median Test

## Results for: Mood's Median Test

### Mood Median Test: A&E Processing time versus Time of week

Mood median test for A&E Processing time

Chi-Square = 3.91 DF = 1 P = 0.048

Time of week	N≤	N>	Median	Q3-Q1	Individual 95.0% CIs			
					-----+-----+	-----+-----+	-----+-----+	-----+-----+
Weekdays	20	10	2.75	2.31	(---*-----)			
Weekends	10	15	4.25	1.75		(-----*-----)		

2.80      3.50      4.20      4.90

Overall median = 3.75

A 95.0% CI for median(Weekdays) - median(Weekends) : (-2.27, -0.23)

Exercício de  
aplicação do  
Mood's  
Median Test

# Teste de Hipótese – Teste para Dispersão

**Discreto** **Contínuo**

**Contínuo**

**X**

**Discreto**

- Diagrama de Dispersão
- Regressão Linear Simples

- Regressão Logística

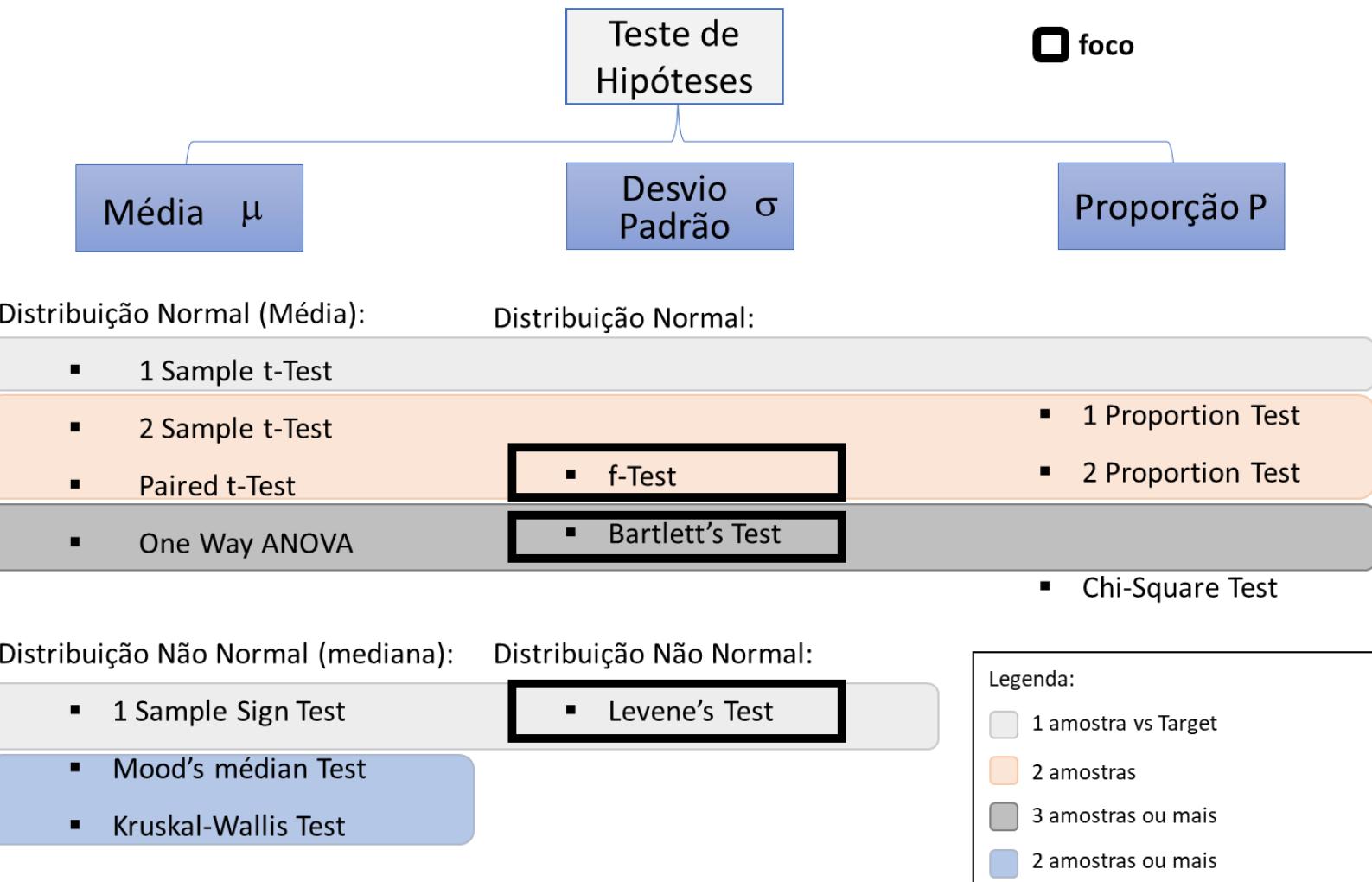
- Box Plot
- Teste de Hipóteses:  
Testes t, ANOVA,  
Mood's, Variância

- Pareto
- Teste de Proporção
- Chi Quadrado



Matriz de Seleção de Ferramentas

# Testes de Hipóteses



# Testes para comparação de Variâncias

Três testes que podem ser feitos para estudar variâncias:

## F-Test

- Para comparar o desvio padrão de **duas amostras** que são **normalmente distribuídas**

## Bartlett's Test

- Para comparer o desvio padrão de **três ou mais amostras** que são **normalmente distribuídas**

## Levene's Test

- Para comparar o desvio padrão de **duas ou mais amostras** que são **não normalmente distribuídas.**

**Todos esses testes podem ser obtidos por meio da mesma função: teste para variâncias iguais**

# Test for Equal Variances – para comparar variâncias

## Test for Equal Variance (Dados não normais & normais)

- É um **teste de hipóteses que compara duas ou mais variâncias**
- A hipótese nula é que as variâncias são a mesma
  - Diferença = 0
  - Se Valor-P < 0,05 rejeite a hipótese  $H_0$

Isto é expresso por:

### Hipótese Nula

$$H_0: \text{variância}_A = \text{variância}_B$$

Variâncias são iguais

### Hipótese Alternativa

$$H_a: \text{variância}_A \neq \text{variância}_B$$

Variâncias não são iguais

$H_0$ : Não há diferencias (Valor-P > 0,05)

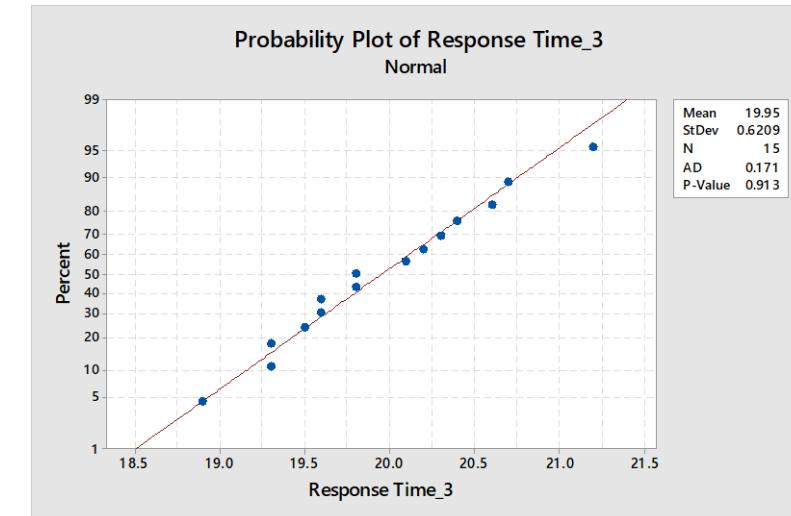
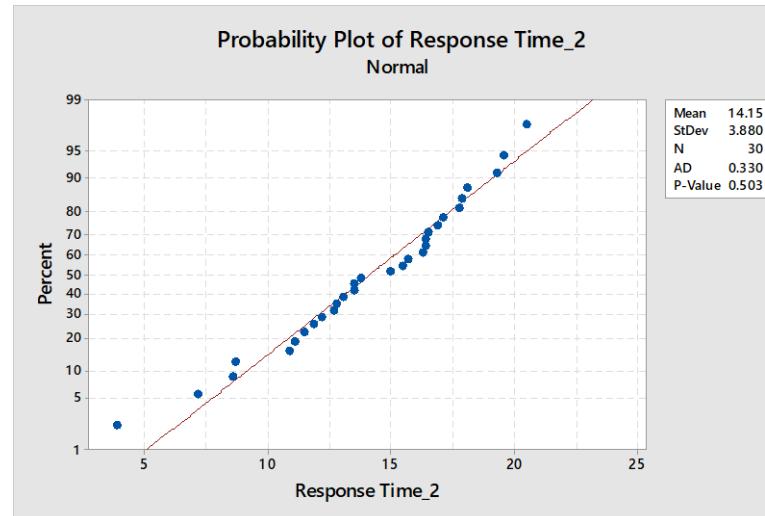
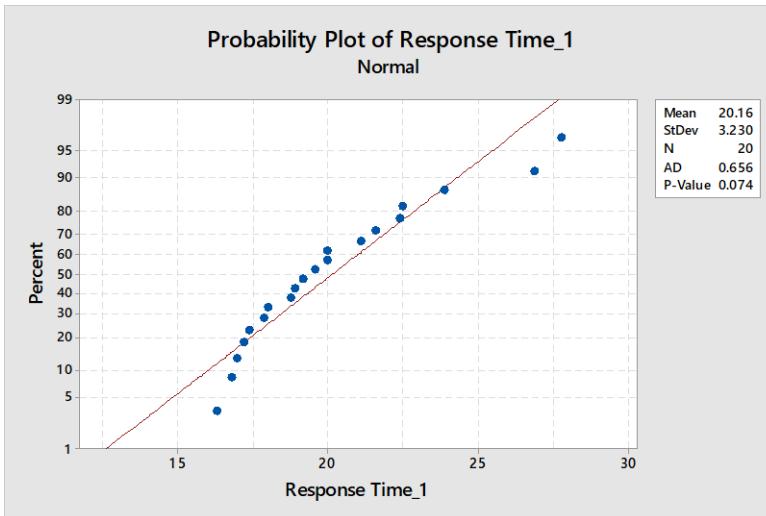
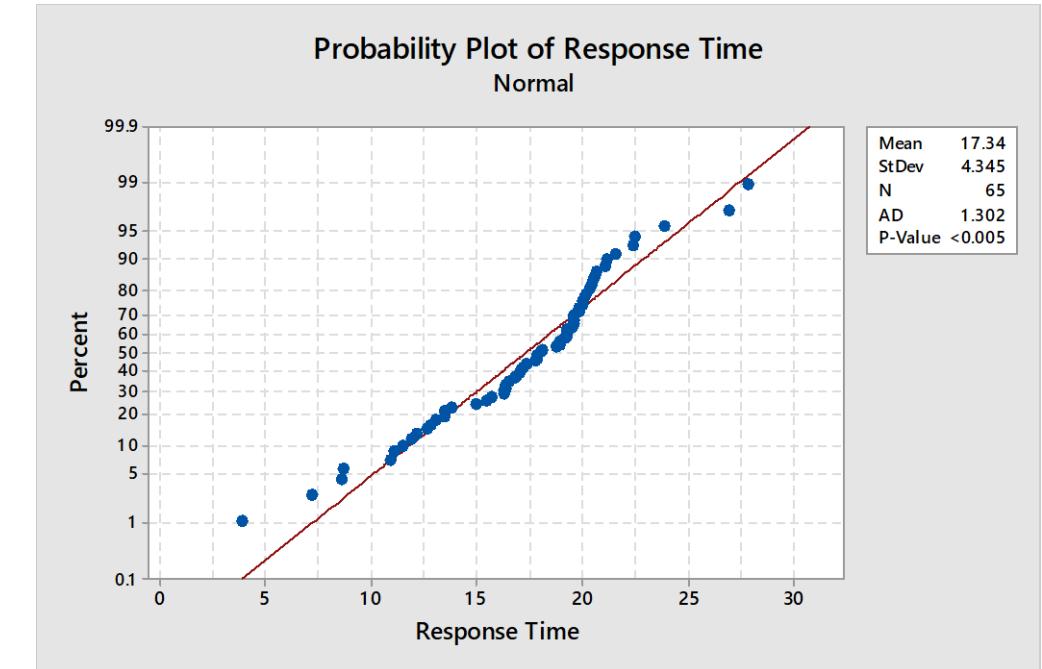
$H_a$ : Há diferencias (Valor-P ≤ 0,05)

# Exercício de aplicação de Test for Equal Variance

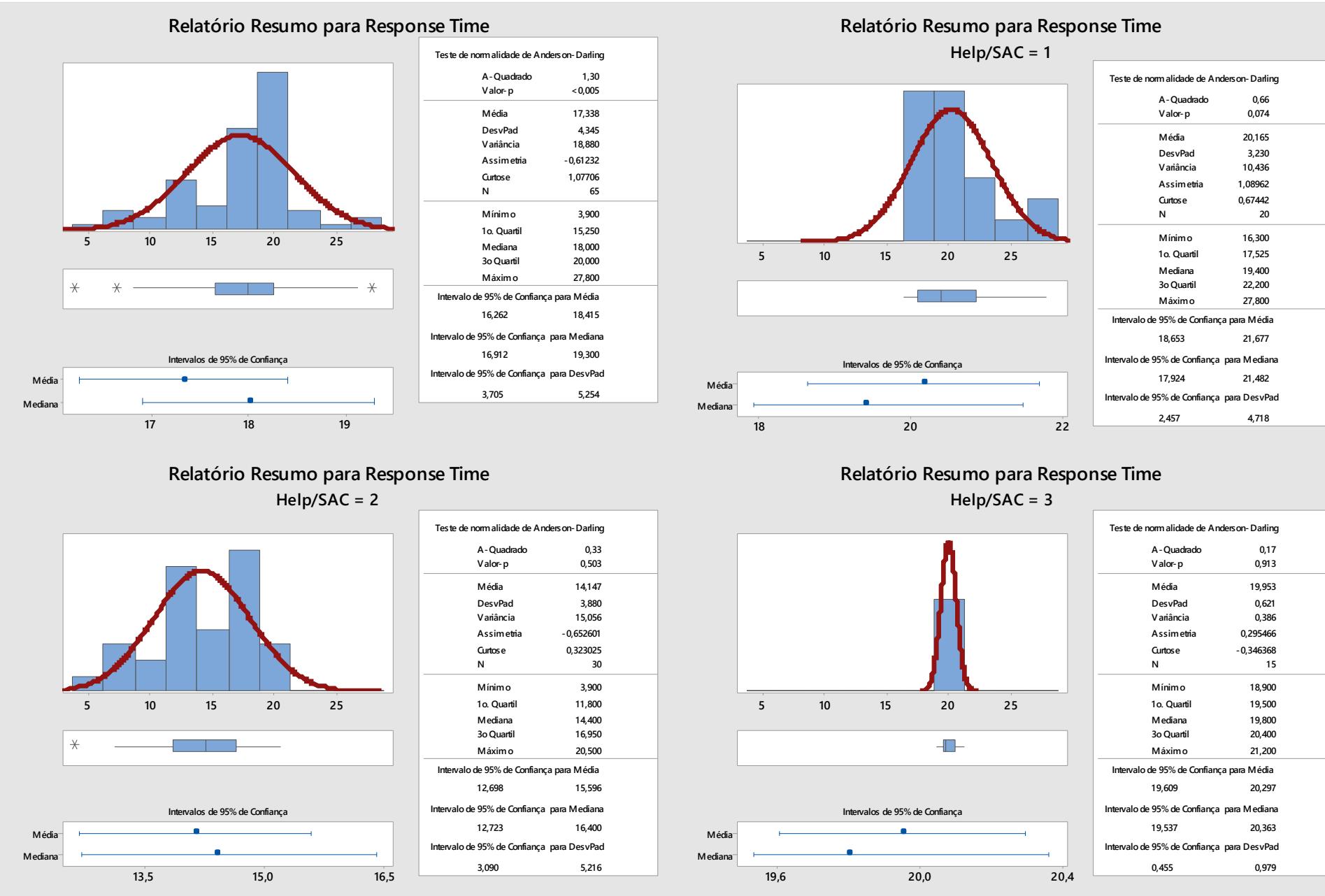
- **Questão a responder com o Test for Equal Variance:**  
Há uma diferença significativa entre a variação do tempo de resposta dos diferentes times de serviço de atendimento ao cliente (SAC)?
- **Hipótese Nula:**  
 $H_0$ : A variação do tempo de resposta é a mesma para os diferentes times do SAC.
- **Teste:**  
Vamos rodar um Test for Equal Variance com um software estatístico.  
Lembre-se: antes disso devemos verificar a normalidade dos dados
- **Resultado:**  
O Valor-P é menor que 0,05?
- **Conclusão:**  
A equipe rejeitou a hipótese nula e concluiu que há uma diferença significativa entre as variâncias apresentadas pelos diferentes times do SAC?



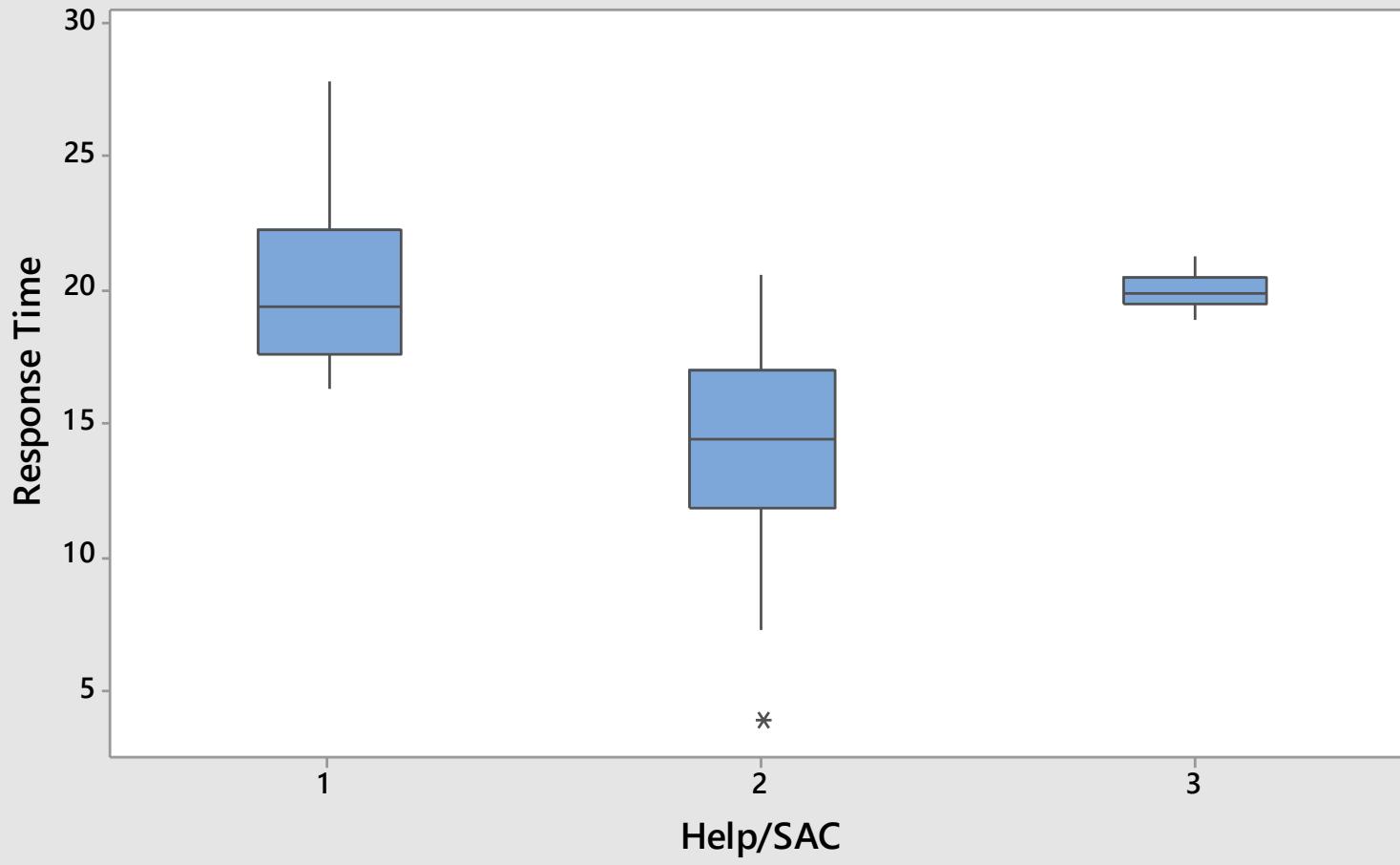
# Test for Equal Variance



# Test for Equal Variance



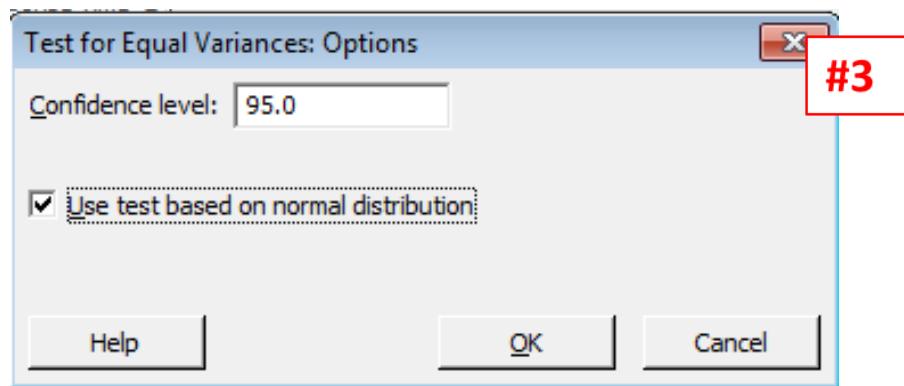
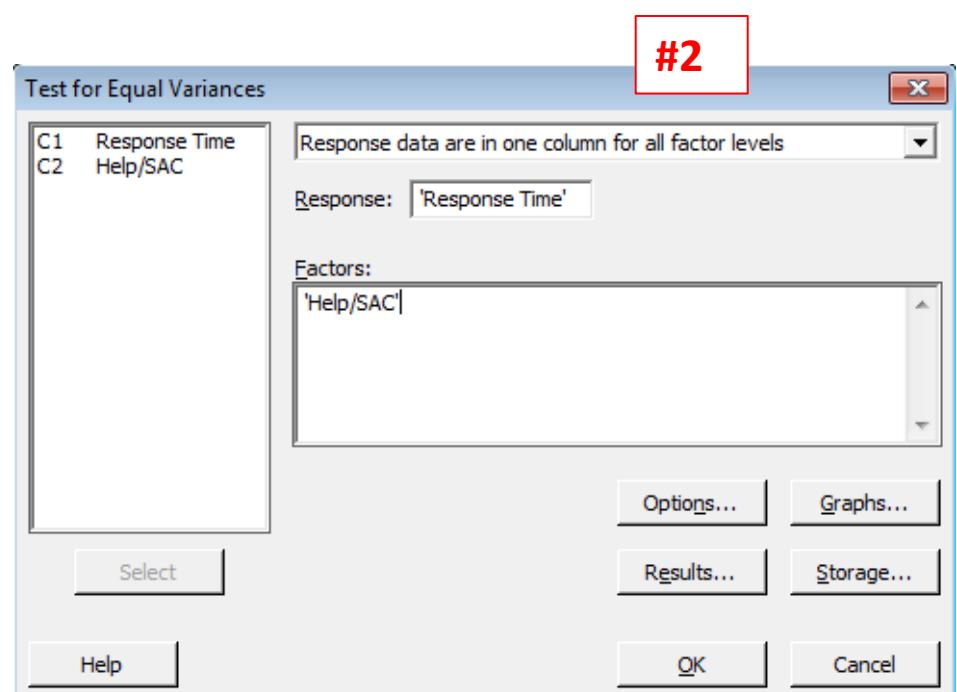
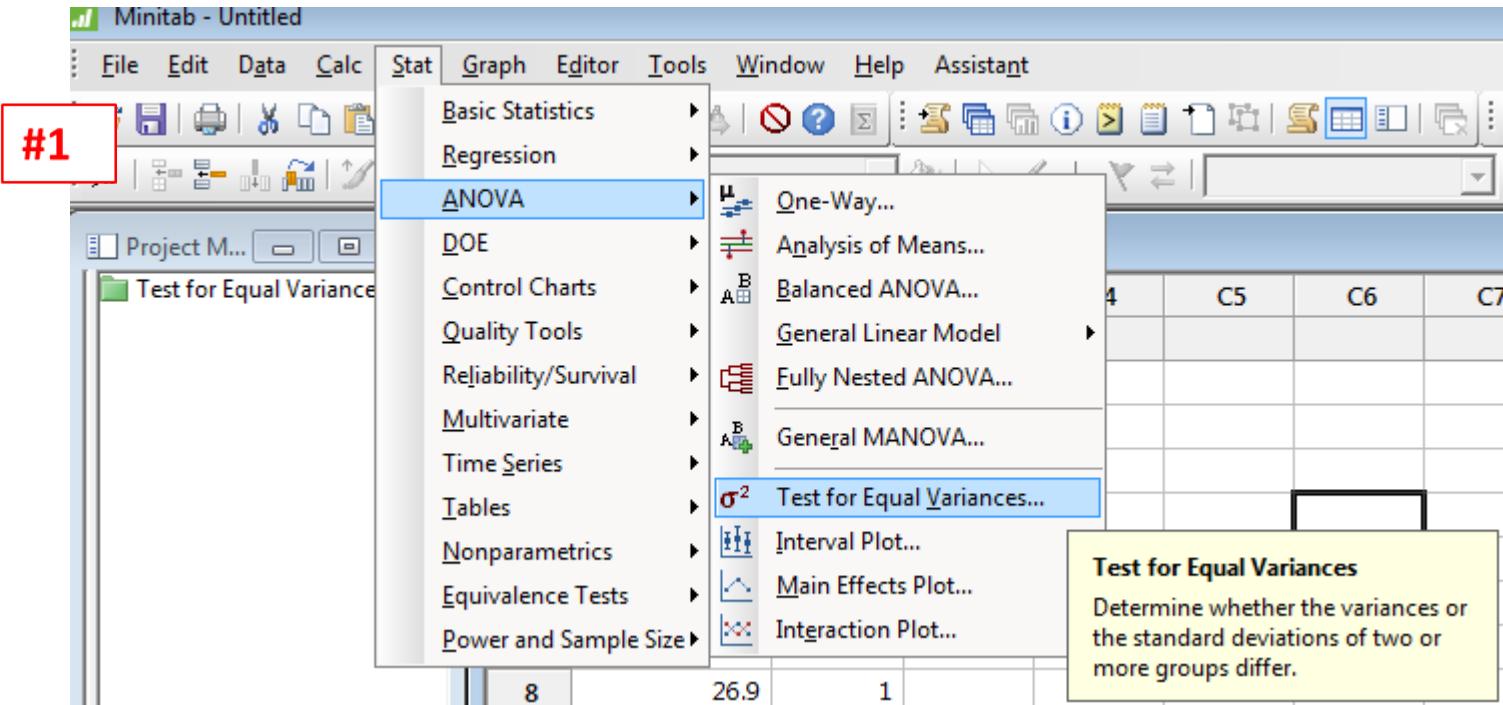
### Boxplot of Response Time



Test for  
Equal  
Variance

# Test for Equal Variance

Test for Equal Variances.xlsx





# Test for Equal Variance

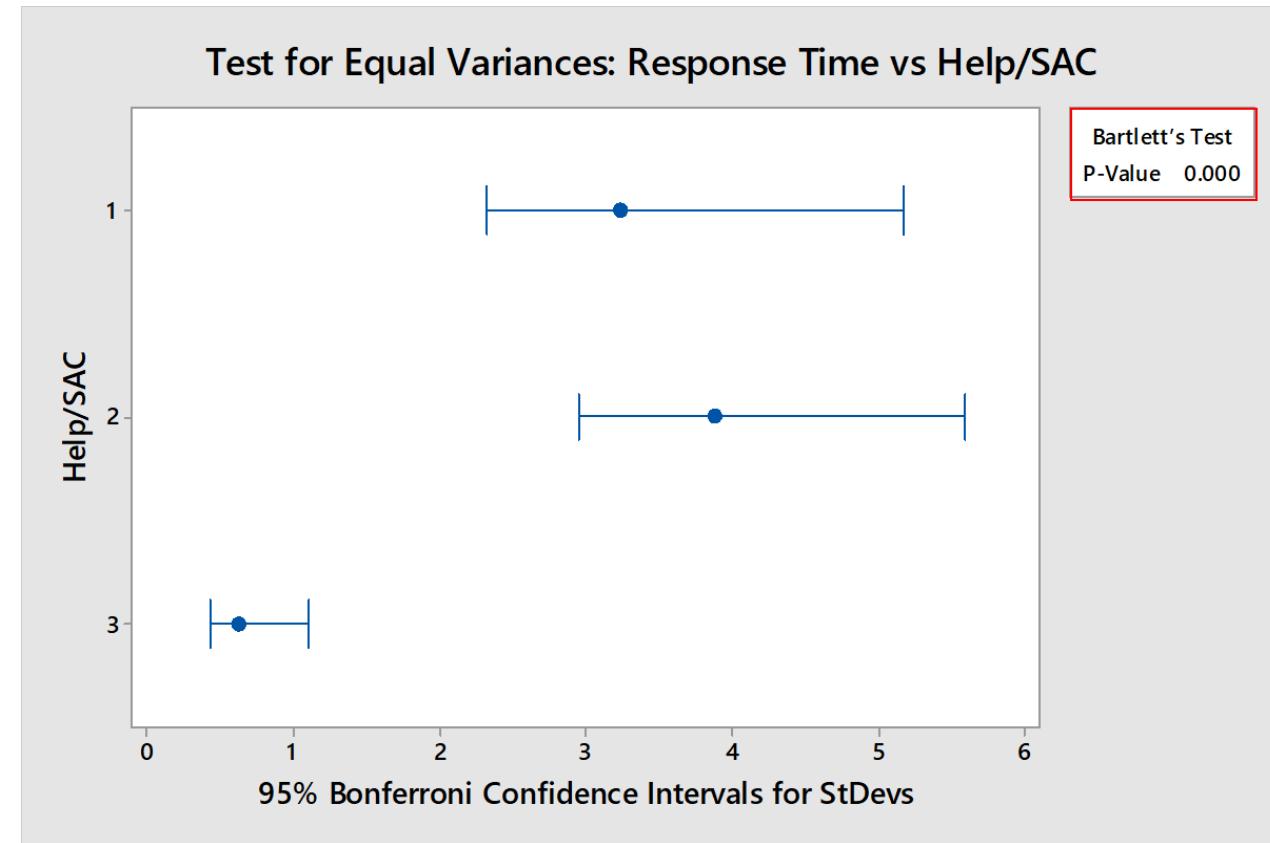
95% Bonferroni Confidence Intervals for Standard Deviations			
Help/SAC	N	StDev	CI
1	20	3.23049	(2.32031, 5.17197)
2	30	3.88025	(2.94560, 5.59623)
3	15	0.62091	(0.42614, 1.09567)

Individual confidence level = 98.3333%

Tests

Method	Statistic	Test	P-Value
Bartlett	34.09		0.000

Bartlett's method is used. This method is accurate for normal data only.





# Test for Equal Variance – usando o Assistente

Test for Equal  
Variances.xlsx

#2

Standard Deviations Test

Sample data

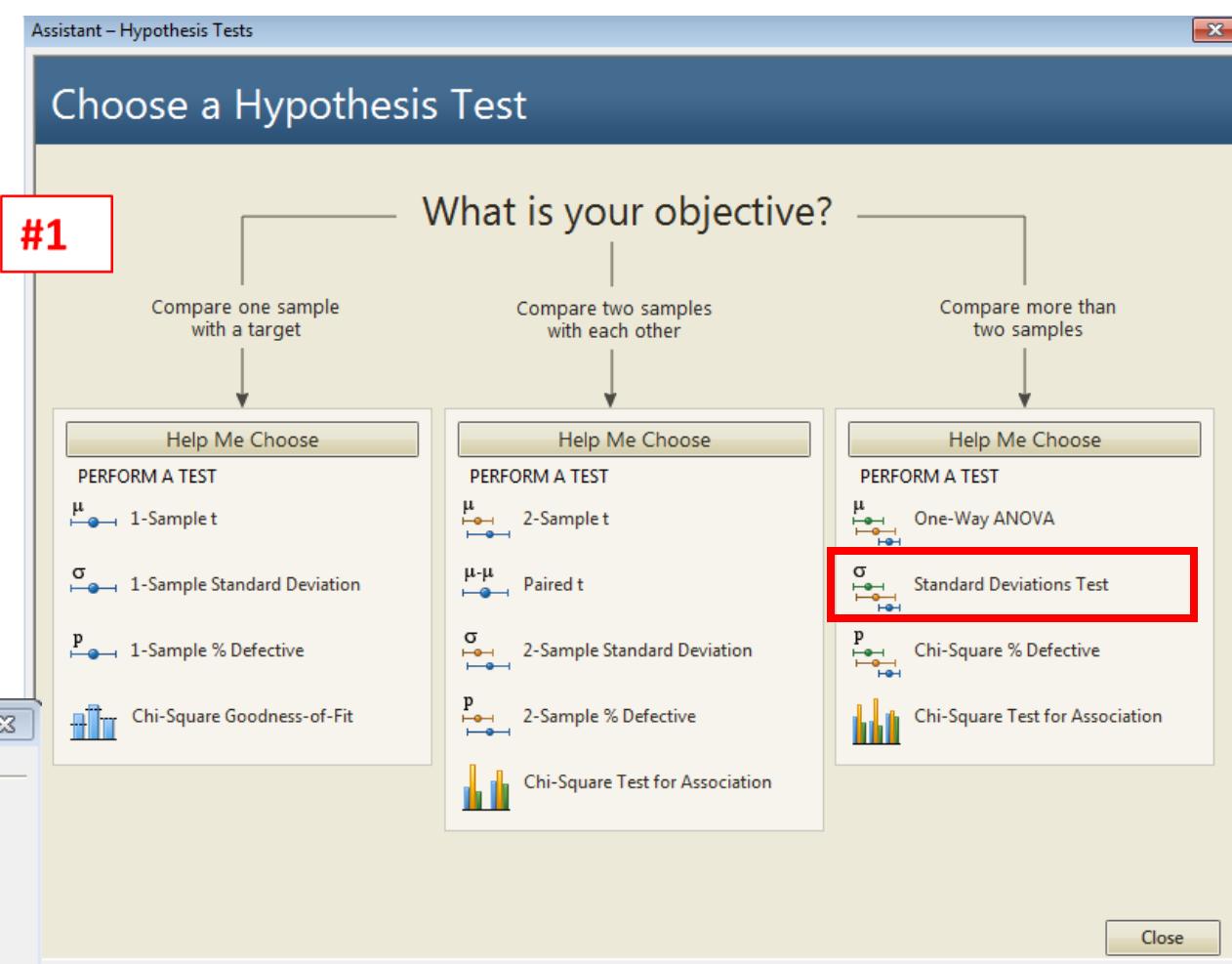
How are your data arranged in the worksheet?  
Y data are in one column, X values in another column

Y data column: 'Response Time'  
X values column: 'Help/SAC'

Test setup

How much risk are you willing to accept of concluding there are differences when there are none?  
Alpha level: 0.05

Select      OK      Cancel



# Test for Equal Variance – usando o Assistente

Standard Deviations Test for Response Tim by Help/SAC  
Summary Report

Do the standard deviations differ?

0	0.05	0.1	> 0.5
Yes	No		
P < 0.001			

Differences among the standard deviations are significant ( $p < 0.05$ ).

Which standard deviations differ?

Help/SAC	Differs from
3	2 1
2	3
1	3

Standard Deviations Comparison Chart  
Red intervals that do not overlap differ.

Comments

- Test: You can conclude that there are differences among the standard deviations at the 0.05 level of significance.
- Comparison Chart: Look for red comparison intervals that do not overlap to identify standard deviations that differ from each other. Consider the size of the differences to determine if they have practical implications.

# Business Case Logística Álcool Gel - Exercício

Foi efetuado um acompanhamento durante alguns meses para avaliar como se comporta o tempo de carregamento dos caminhões.

- Há diferença significativa na dispersão do tempo de carregamento entre os meses?
- Utilize o BoxPlot e 2-sample-standard deviation para suas conclusões

Os dados estão no arquivo CoolGel\_Banco de Dados\_Logística.xlsx na aba Tempo Carg\_Pos

# Teste de Hipótese – Teste para Proporções

**Discreto**

**Contínuo**

**X**

**Discreto**

**Discreto**

- Diagrama de Dispersão
- Regressão Linear Simples

- Regressão Logística

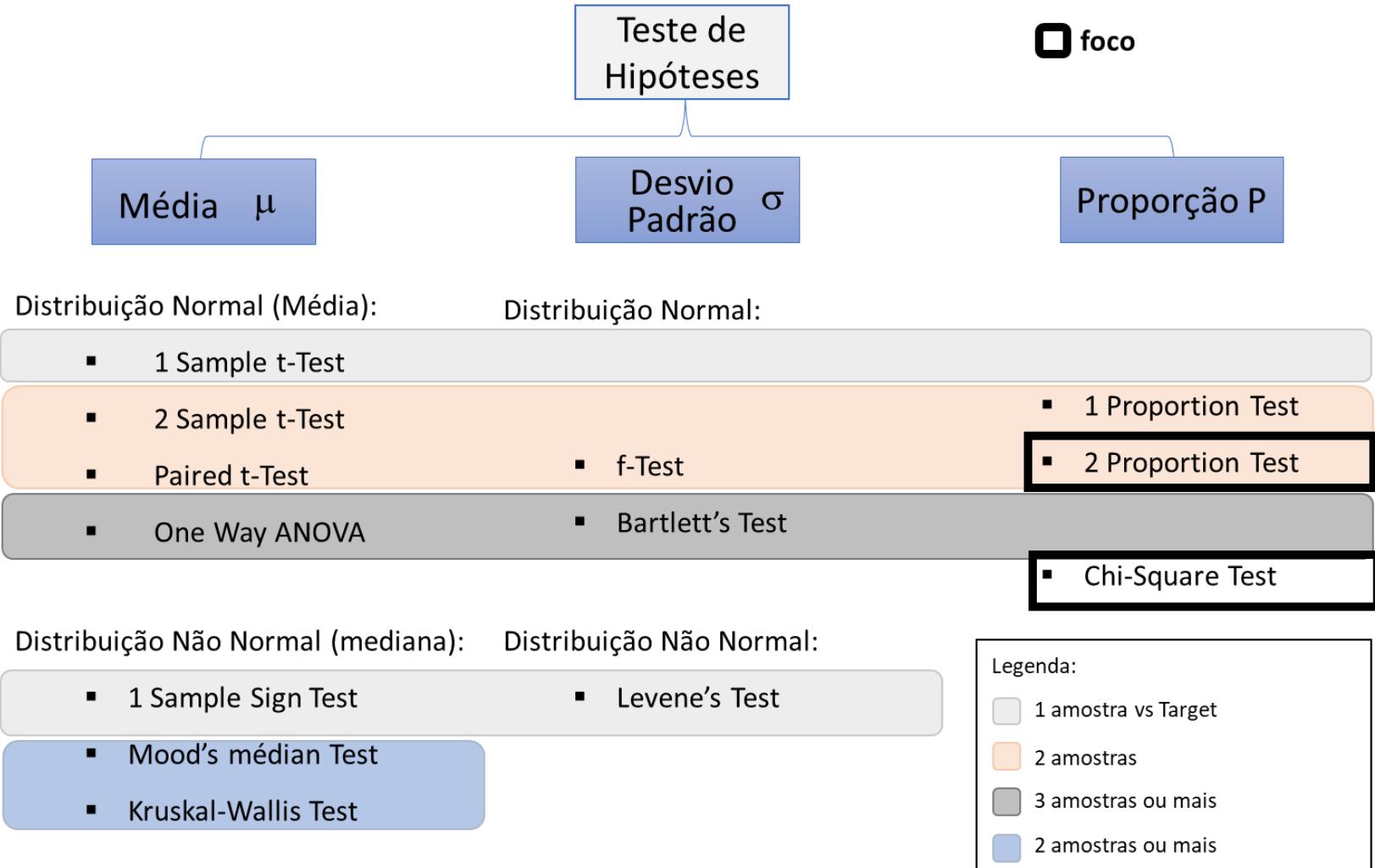
- Box Plot
- Teste de Hipóteses: Testes t, ANOVA, Mood's, Variância

- Pareto
- Teste de Proporção
- Chi Quadrado



Matriz de Seleção de Ferramentas

# Testes de Hipóteses



# Testes de comparação para Proporções

1 Proportion Test

Compara uma proporção contra uma outra proporção alvo ou histórica.

2 Proportion Test

Compara as proporções de duas amostras uma contra a outra.

Chi-Square Test

Compara as proporções de três ou mais amostras uma contra a outra.

# 2 Proportion Test

---

## —

### Compara Proporções & Porcentagens

O Test for 2 Proportions é um **teste de hipóteses que compara duas proporções**

- A hipótese nula é que as proporções são iguais
  - Diferença entre as proporções = 0
- Se Valor-P < 0,05 rejeite a hipótese  $H_0$

Isso é expresso por:

**Hipótese Nula**

$H_0: \text{proporção}_A = \text{proporção}_B$

Proporções são iguais

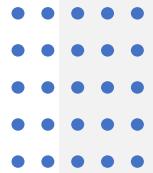
**Hipótese Alternativa**

$H_a: \text{proporção}_A \neq \text{proporção}_B$

Proporções não são iguais

$H_0: \text{Não há diferenças} \ (\text{Valor-P} > 0,05)$

$H_a: \text{There are Differences} \ (\text{Valor-P} \leq 0,05)$

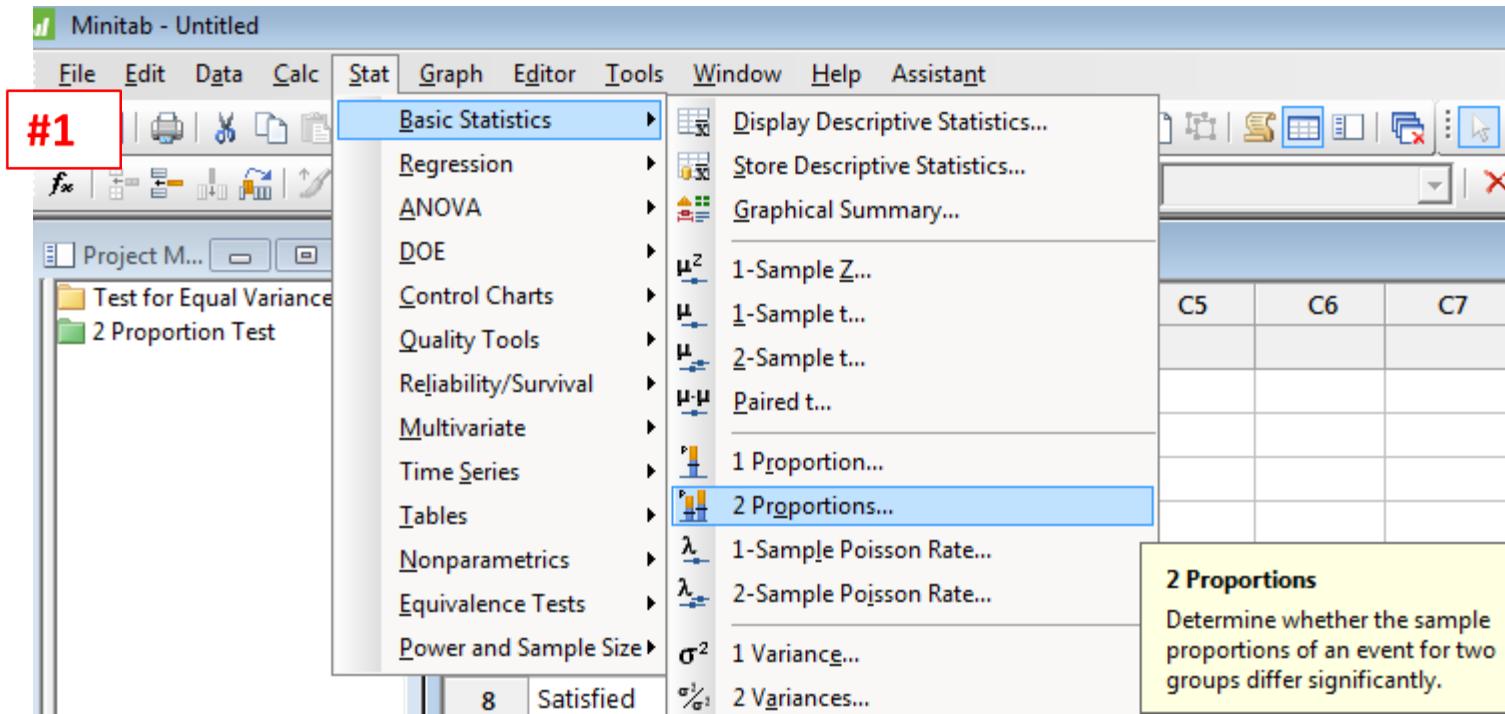


# Exercício de aplicação de 2 Proportion Test

- **Questão a responder com o p-Test:**  
Há uma diferença significativa entre o nível de satisfação das diferentes áreas?
- **Hipótese Nula:**  
 $H_0$ : A proporção de satisfação é a mesma para as Áreas
- **Teste:**  
Vamos rodar um 2 Proportion Test com um software estatístico
- **Resultado:**  
O Valor-P é menor que 0,05?
- **Conclusão:**  
A equipe rejeitou a hipótese nula e concluiu que há uma diferença significativa entre o nível de satisfação das áreas?

# 2 Proportion Test

2 Proportion Test.xlsx



Variable	X	N	Sample p
Area A	72	80	0.900000
Area B	79	100	0.790000

Difference = p (Area A) - p (Area B)  
Estimate for difference: 0.11  
95% CI for difference: (0.00658520, 0.213415)  
Test for difference = 0 (vs  $\neq$  0): Z = 2.08 P-Value = 0.037  
Fisher's exact test: P-Value = 0.065

**#1**

**#2**

**#3**

Two-Sample Proportion

Each sample is in its own column

Sample 1: 'Area A'  
Sample 2: 'Area B'

Select Options... OK Cancel Help

Two-Sample Proportion

Both samples are in one column

Both samples are in one column  
Each sample is in its own column  
Summarized data  
Sample IDs:

Select Options... OK Cancel Help

# Business Case Logística Álcool Gel - Exercício

Uma semana crítica sugerida pelo time, é a 2ª semana de Abril, onde houve um aumento na participação de 2 CD's : Recife e Piracicaba

Os dados foram coletados e estão no arquivo CoolGel\_Banco de Dados\_Logística, Aba OTIF WK2 Abr2020 Pirac\_Recife

Responda às seguintes perguntas:

- Podemos afirmar que a proporção de In Full é diferente entre esses CD's?
- Podemos afirmar que a proporção de On Time é diferente entre esses CD's ?

Business Case  
– PDV  
Minitab  
Exemplo uso  
do Assistente



Foi efetuado uma amostragem para verificar se o uso de Cartaz ocorre na mesma proporção tanto em Mercados como em Farmácias. Os resultados estão na tabela abaixo. Existe diferença no uso de cartaz em função do tipo de loja?

Loja	Cartaz Loja - Sim	Cartaz Loja - Não	Total de Lojas
Farmácia	174	70	244
Mercado	307	141	448

Qual é o seu objetivo?

The flowchart starts with 'Qual é o seu objetivo?' (What is your objective?). It branches into three categories: 'Comparar uma amostra a um alvo' (Compare one sample to a target), 'Comparar duas amostras entre si' (Compare two samples against each other), and 'Comparar mais de duas amostras' (Compare more than two samples). Each category leads to a 'Ajude-me a Escolher' (Help me choose) box containing 'REALIZAR UM TESTE' (Perform a test) options. The 'Comparar uma amostra a um alvo' box includes 'Teste t para 1 amostra' (t-test for 1 sample), 'Teste de desvio padrão para 1 amostra' (Test of standard deviation for 1 sample), 'Teste de % Defeituosos para 1 amostra' (Test of % Defective for 1 sample), and 'Qualidade de Ajuste Qui-Quadrado' (Goodness of Fit Chi-Square). The 'Comparar duas amostras entre si' box includes 'Teste t para 2 amostras' (t-test for 2 samples), 'Teste t pareado' (Paired t-test), 'Teste de desvio padrão para 2 amostras' (Test of standard deviation for 2 samples), and 'Teste de % Defeituosos para 2 amostras' (Test of % Defective for 2 samples). The 'Comparar mais de duas amostras' box includes 'ANOVA com um fator' (ANOVA with one factor), 'Teste de desvios padrão' (Test of standard deviations), 'Teste Qui-Quadrado para % de Defeituosos' (Chi-Square test for % Defective), and 'Teste Qui-Quadrado para Associação' (Chi-Square test for Association). The 'Teste de % Defeituosos para 2 amostras' option is highlighted with a red box.

Assistente Ferramentas Adicionais

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
Testes de Hipótese...

Testes de Hipóteses

Usar para localizar diferenças entre uma amostra e um alvo, entre duas amostras ou entre múltiplas amostras.

Cartas de Controle Antes/Depois...  
Cartas de Controle...

REALIZAR UM TESTE

$\mu$  Teste t para 1 amostra  
 $\sigma$  Teste de desvio padrão para 1 amostra  
 $p$  Teste de % Defeituosos para 1 amostra  
Qualidade de Ajuste Qui-Quadrado

$\mu$  Teste t para 2 amostras  
 $\mu - \mu$  Teste t pareado  
 $\sigma$  Teste de desvio padrão para 2 amostras  
 $p$  Teste de % Defeituosos para 2 amostras

$\mu$  ANOVA com um fator  
 $\sigma$  Teste de desvios padrão  
 $p$  Teste Qui-Quadrado para % de Defeituosos  
Teste Qui-Quadrado para Associação

Clique para realizar este teste. a  
Associação

Business Case  
– PDV  
Minitab  
Exemplo uso  
do Assistente



Foi efetuado uma amostragem para verificar se o uso de Cartaz ocorre na mesma proporção tanto em Mercados como em Farmácias. Os resultados estão na tabela abaixo. Existe diferença no uso de cartaz em função do tipo de loja?

Loja	Cartaz Loja - Sim	Cartaz Loja - Não	Total de Lojas
Farmácia	174	70	244
Mercado	307	141	448

Teste de % de Defeituosos para 2 Amostras

Dados amostrais

Insira seus nomes de amostra ou use os padrões. Digite os dados para as duas amostras.

Nome da Amostra	Total de Números Testados	Número de Defeituosos
Farmacia	244	70
Mercado	448	141

Configuração de teste

O que você deseja determinar?

O % de defeituosos de Farmacia é maior que o % de defeituosos de Mercado?

O % de defeituosos de Farmacia é menor que o % de defeituosos de Mercado?

O % de defeituosos de Farmacia é diferente do % de defeituosos de Mercado?

Qual nível de risco você pretende aceitar ao considerar a conclusão acima quando ela não for verdadeira?

Nível  $\alpha$ : 0,05

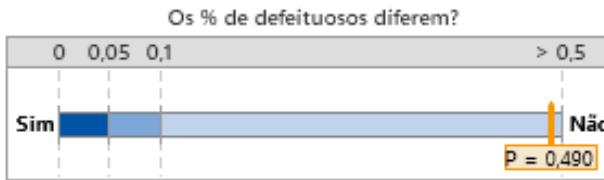
Poder e tamanho de amostra (opcional)

Qual diferença entre os dois % de defeituosos tem valor prático?

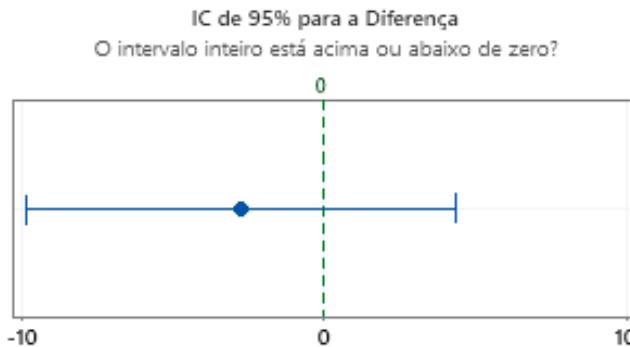
Diferença: \_\_\_\_\_

## Teste de % de Defeituosos para 2 Amostras para Farmacia vs Mercado

### Relatório Resumo



O % de defeituosos de Farmacia não é significativamente diferente do % de defeituosos de Mercado ( $p > 0,05$ ).



Estatísticas	Amostras Individuais	
	Farmacia	Mercado
Número total testado	244	448
Número de defeituosos	70	141
% de Defeituosos	28,69	31,47
IC de 95%	(23,10; 34,80)	(27,20; 36,00)

Estatísticas	*Diferença	Diferença Entre Amostras
		Diferença
IC de 95%	(-9,91; 4,34)	-2,78

\*Diferença = Farmacia - Mercado

Comentários
<ul style="list-style-type: none"><li>• Teste: não há evidências suficientes para concluir que o % de defeituosos diferem no nível de significância de 0,05.</li><li>• IC: quantifica a incerteza associada à estimativa da diferença de dados amostrais. Você pode ter 95% de confiança de que a verdadeira diferença está entre -9,91% e 4,34%.</li></ul>

## Business Case – PDV

### Minitab Exemplo uso do Assistente

## Teste de % de Defeituosos para 2 Amostras para Farmacia vs Mercado

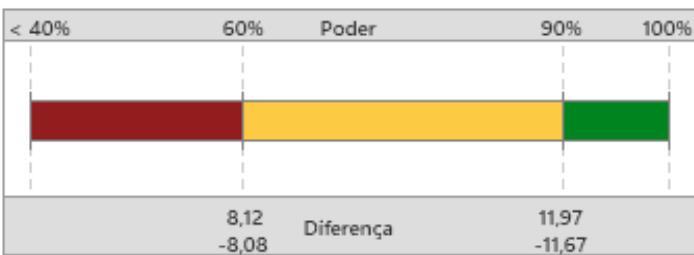
### Relatório de Diagnóstico

Número de Itens Defeituosos e Não-defeituosos

	Farmacia	Mercado
Defeituoso	70	141
Não-defeituoso	174	307

Deve haver, no mínimo, 5 itens defeituosos e 5 não-defeituosos em cada amostra para assegurar a validade do intervalo de confiança para a diferença. Todas as violações estão indicadas em vermelho.

Qual é a chance de detectar uma diferença?



Para  $\alpha = 0,05$  e tamanhos amostrais = 244; 448:

Se o verdadeiro % de defeituosos de Farmacia fosse 8,12 maior do que Mercado ou 8,08 menor, você teria uma chance de 60% de detectar a diferença. Se ele fosse 11,97 maior ou 11,67 menor, você teria uma chance de 90%.

Que diferença você pode detectar com os tamanhos amostrais de 244 e 448?

Diferença	Poder
8,12	60%
9,14	70%
10,32	80%
11,97	90%

Diferença observada = -2,78

## Business Case – PDV

## Minitab Exemplo uso do Assistente

O poder é uma função dos tamanhos amostrais e do % de defeituosos. Para detectar diferenças menores, considere aumentar os tamanhos amostrais.

Teste de % de Defeituosos para 2 Amostras para Farmacia vs Mercado  
Cartão de Relatório

Verificar	Status	Descrição
Validade do IC		Como ambas as amostras têm, no mínimo, 5 defeituosos e 5 não-defeituosos, o intervalo de confiança da diferença deve ser exato.
Tamanho da Amostra		Os dados não fornecem evidências suficientes para concluir que o % de defeituosos de Farmacia difere de Mercado. Isso pode resultar de tamanhos amostrais muito pequenos. Com base nos tamanhos amostrais e no $\alpha$ , você teria uma chance de 90% de detectar uma diferença de 11,97 ou -11,67. Para determinar o quanto grandes as amostras precisariam ser para detectar uma diferença que tem implicações práticas, repita a análise e insira um valor para a diferença.

Business Case –  
PDV  
Minitab Exemplo  
uso do Assistente

# Teste de Hipótese – Chi-Square

# Chi-Square— Compara Proporções & Porcentagens

Os dados estão em um tabela  
muito simples

Status	Area A	Area B	Area C	Area D
Satisfied	72	79	35	9
Not Satisfied	8	21	15	1

O Chi-Square Test é um **teste de hipóteses para comparar proporções (três ou mais)**

- A hipótese nula é que as proporções são iguais
  - Diferença entre as proporções = 0
- Se Valor-P < 0,05 rejeite a hipótese  $H_0$

Isso é expresso por:

## Hipótese Nula

$H_0$ : Proporções são iguais

## Hipótese Alternativa

$H_a$ : Pelo menos uma proporção não é igual

$H_0$ : Não há diferenças (Valor-P > 0,05)

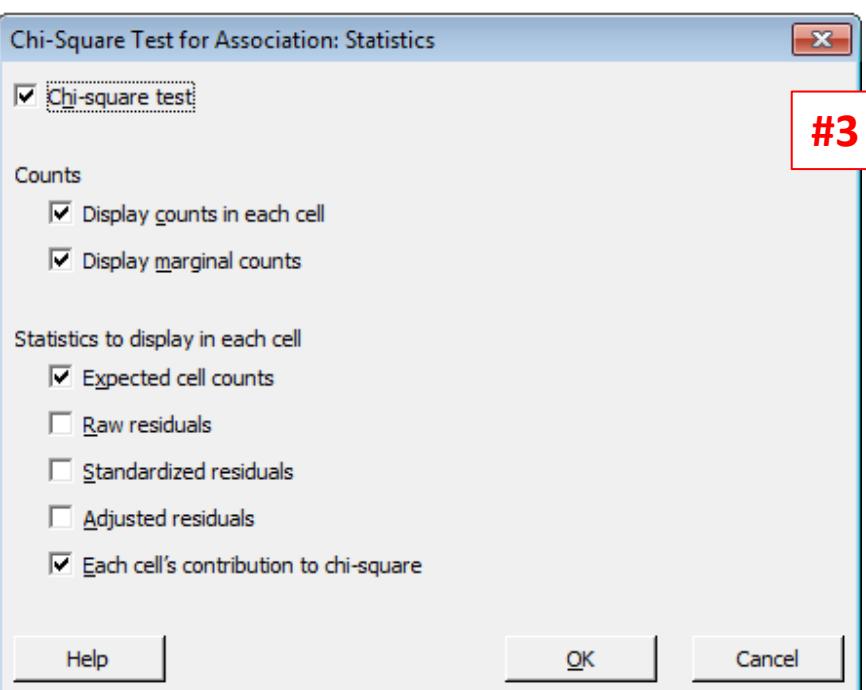
$H_a$ : Há diferenças (Valor-P ≤ 0,05)

# Exercício de aplicação do Chi-Square Test

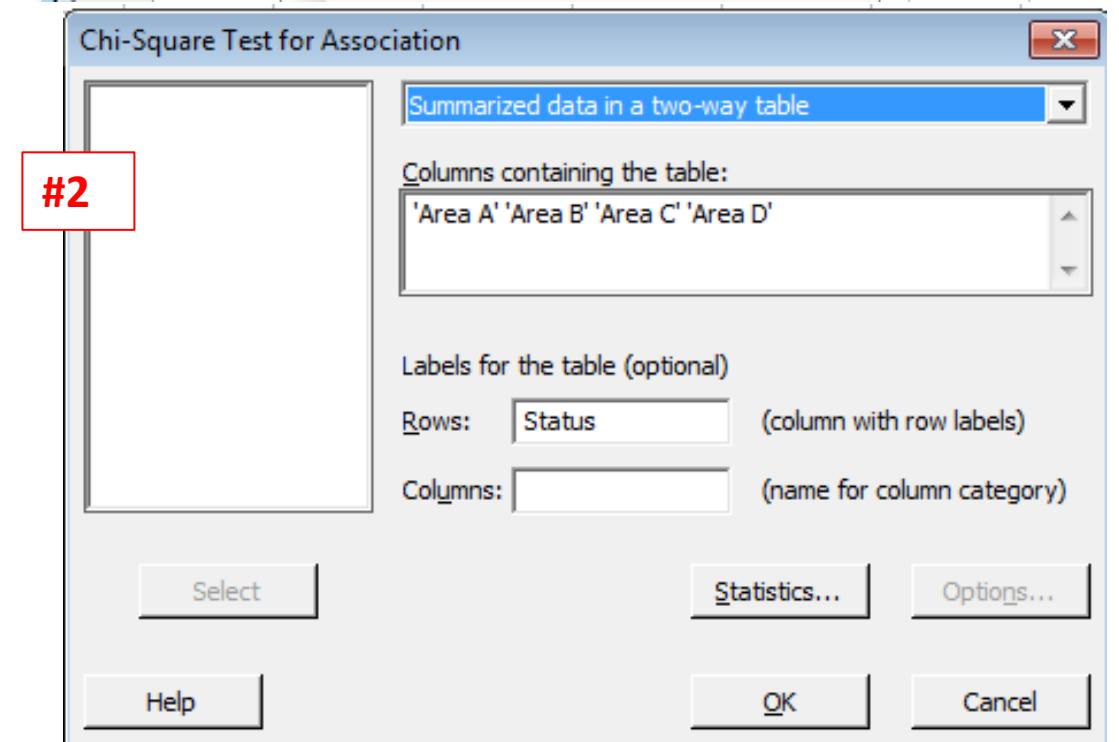
- **Questão a responder com o Chi-Square-Test:**  
Há uma diferença significativa no nível de satisfação dos clientes por Hotel?
- **Hipótese nula:**  
 $H_0$ : A proporção de satisfação é a mesma para os Hotéis
- **Teste:**  
Vamos rodar um Chi-Square Test com um software estatístico
- **Resultado:**  
O Valor-P é menor que 0,05?
- **Conclusão:**  
A equipe rejeitou a hipótese nula e concluiu que há uma diferença significativa entre o nível de satisfação por Hotel?

# Chi-Square Test

Chi-Square.xlsx



The Minitab interface is shown with the 'Stat' menu open. The 'Tables' option is highlighted. The 'Chi-Square Test for Association...' command is selected, which is highlighted in blue. A tooltip for this command is displayed, stating: 'Chi-Square Test for Association Determine whether two categorical variables are associated.'



# Chi-Square Test

## Chi-Square Test for Association: Status, Worksheet columns

Rows: Status    Columns: Worksheet columns

	Area A	Area B	Area C	Area D	All
Satisfied	72	79	35	9	195
	65.00	81.25	40.63	8.13	
	0.7538	0.0623	0.7788	0.0942	
Not Satisfied	8	21	15	1	45
	15.00	18.75	9.38	1.88	
	3.2667	0.2700	3.3750	0.4083	
All	80	100	50	10	240

Cell Contents:

Count  
Expected count  
Contribution to Chi-square

Pearson Chi-Square = 9.009, DF = 3, P-Value = 0.029

Likelihood Ratio Chi-Square = 9.245, DF = 3, P-Value = 0.026

\* NOTE \* 1 cells with expected counts less than 5



# Chi-Square Test – Usando o Assistente

Chi-Square.xlsx

#2

Teste Qui-Quadrado para Associação

Dados amostrais:

Como você inserirá seus dados?

Mudar linhas e colunas da tabela

Resultados estão nas colunas  
 Resultados estão nas linhas

Nome de X: Satisfaction

Nome de Y: Areas

Número de resultados: 2

Número de valores de X: 4

Areas	Satisfied	Not Satisfied
A	72	8
B	79	21
C	35	15
D	9	1

Configuração de teste

Qual nível de risco você pretende aceitar ao conduzir que existe uma associação quando não houver nenhuma?

Nível alfa: 0,05

OK Cancelar

#1

## Escolha um Teste de Hipóteses

Qual é o seu objetivo?

Comparar uma amostra a um alvo

Ajude-me a Escolher

REALIZAR UM TESTE

$\mu$  Teste t para 1 amostra

$\sigma$  Teste de desvio padrão para 1 amostra

$p$  Teste de % Defeituosos para 1 amostra

Qualidade de Ajuste Qui-Quadrado

Compare duas amostras entre si

Ajude-me a Escolher

REALIZAR UM TESTE

$\mu-\mu$  Teste t para 2 amostras

$\mu-\mu$  Teste t pareado

$\sigma$  Teste de desvio padrão para 2 amostras

$p$  Teste de % Defeituosos para 2 amostras

Teste Qui-Quadrado para Associação

Comparar mais de duas amostras

Ajude-me a Escolher

REALIZAR UM TESTE

$\mu$  ANOVA com um fator

$\sigma$  Teste de desvios padrão

$p$  Teste Qui-Quadrado para % de Defeituosos

Teste Qui-Quadrado para Associação

Fechar

# Chi-Square Test – Usando o Assistente

• • • • • • • • • • • • • • •

### Teste Qui-Quadrado para Associação: Satisfaction por Areas

#### Relatório Resumo

**Os perfis de porcentagem diferem?**

	0	0,05	0,1	> 0,5
<b>Sim</b>	P = 0,029			
<b>Não</b>				

As diferenças entre os perfis de porcentagem de resultados são significativas ( $p < 0,05$ ). Você pode concluir que existe uma associação entre Satisfaction e Areas.

**Carta de Perfis de Porcentagem**  
Compare os perfis.

Categoria	Satisfied (%)	Not Satisfied (%)
Média	82%	18%
A	90%	10%
B	79%	21%
C	70%	30%
D	90%	10%

**Comentários**

- Teste: você pode concluir que há diferenças entre os perfis de porcentagem de resultado no nível de significância de 0,05.
- Carta de Perfis de Porcentagem: use-as para comparar o perfil de cada valor de Areas e o perfil da média.
- Carta do % da Diferença: procure barras longas para identificar os resultados com o maior % de diferença entre as contagens observadas e esperadas.

**% da Diferença entre Contagens Observadas e Esperadas**

Categoria	Satisfied (%)	Not Satisfied (%)
A	10%	-40%
B	-5%	10%
C	-15%	60%
D	10%	-40%

Positivo: ocorre mais frequentemente do que esperado  
Negativo: ocorre menos frequentemente do que esperado

# Business Case Logística Álcool Gel - Exercício

Uma semana crítica sugerida pelo time, é a 2ª semana de Abril, onde houve um aumento na participação de 2 CD's : Recife e Piracicaba

Dados foram coletados que estão no arquivo CoolGel\_Banco de Dados\_Logística, Aba OTIF WK2 Abr2020 Pirac\_Recife

Responda às seguintes perguntas:

- A quantidade de caixas em cada pedido afeta o resultado do In Full?
- A quantidade de caixas em cada pedido afeta o resultado do On Time?

# Business Case Mercado – PDV Álcool Gel – Exercício

Com base nos seguintes dados faça uma análise de Chi Square e responda às seguintes perguntas:

- Há diferença no uso de cartaz entre Farmácias e Mercados?
- O número de SKU's por loja impacta a visualização do produto ou não?

Loja	Cartaz Loja - Sim	Cartaz Loja - Não
Farmácia	174	70
Mercado	307	141

#SKU/Loja	Viu produto - Sim	Viu produto - Não
1 SKU	45	117
2 SKU	82	158
3 SKU	169	264
4+ SKU	206	322

Esses dados também estão no arquivo Coolgel\_Banco de Dados\_PDV.xlsx na Aba Market Share Decline

# Régressão

# Matriz de Seleção de Ferramentas

**Discreto**

**X**

**Contínuo**

**Discreto** **Contínuo**

**Y**

- Diagrama de Dispersão
- Regressão Linear Simples

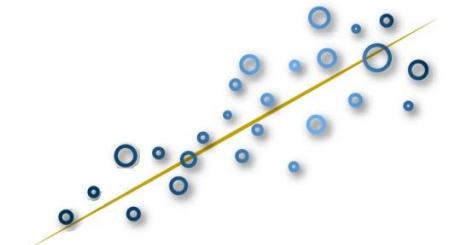
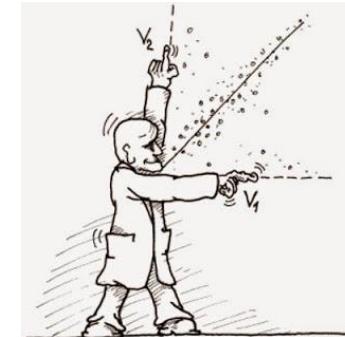
- Box Plot
- Teste de Hipóteses: Testes t, ANOVA, Mood's, Variância

- Regressão Logística

- Pareto
- Teste de Proporção
- Chi Quadrado

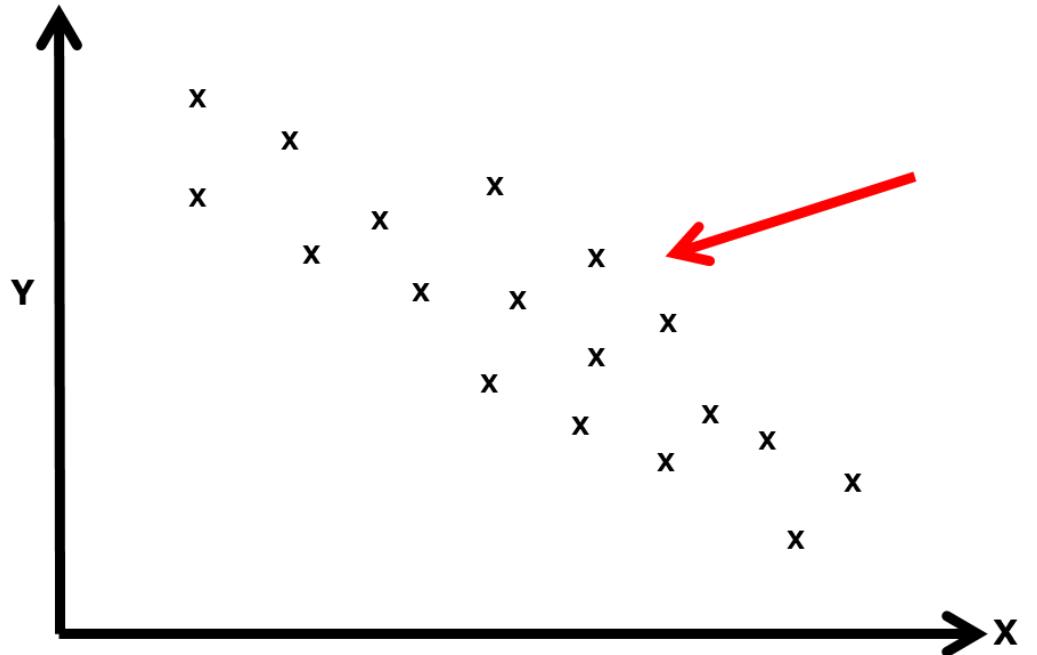
# Regressão

- Correlação
  - Regressão Linear



# Correlação

- Mede o grau de associação entre duas variáveis
- Mede a força linear da relação entre as variáveis Y e X
- Principal métrica: Coeficiente de Correlação de Pearson,  $r$  ( $r$  varia entre -1 e +1)
  - Relação positiva perfeita =>  $r = 1$
  - Sem relação =>  $r = 0$
  - Relação negativa perfeita =>  $r = -1$



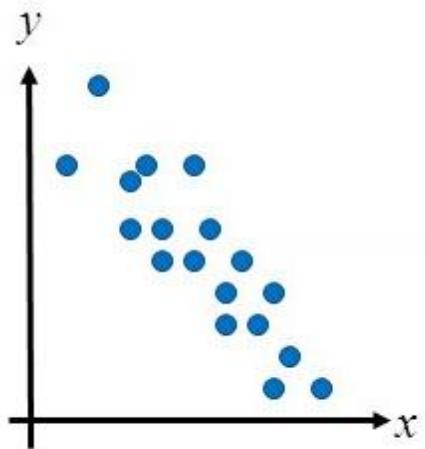
# Correlação

- Existe uma correlação quando duas (ou mais) variáveis apresentam uma tendência conjunta. Podemos observar esta tendência através de um gráfico de dispersão

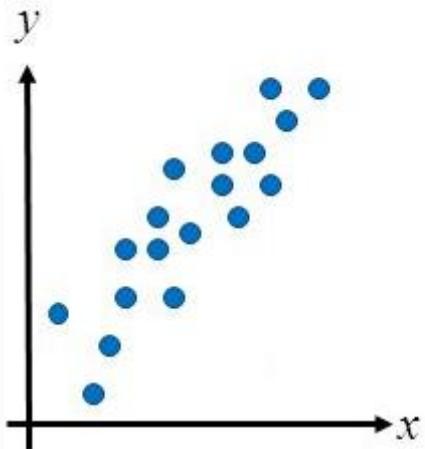
$$r = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

A medida da força de correlação entre as variáveis é chamada de coeficiente de correlação linear ( $r$ ).

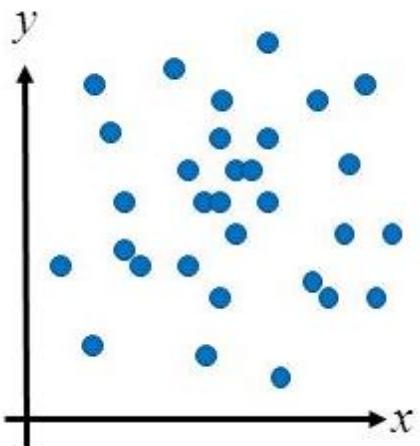
## Correlação Negativa



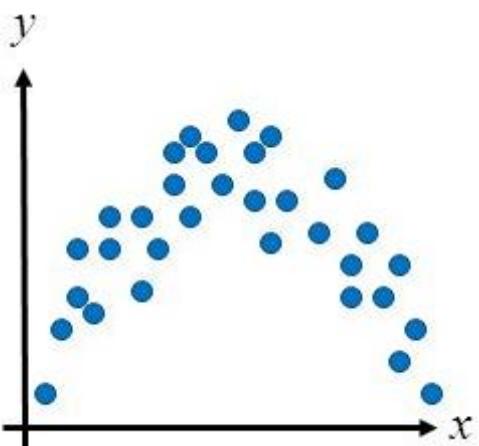
## Correlação Positiva



Coeficiente de Correlação -  $r$



## Correlação Nula



## Correlação Não Linear

- $r$  – coeficiente de correlação linear (varia entre -1 e +1)
- Mede o quanto próximos os pontos estão de uma reta

# Correlação – Consideração Importante!

**A existência de uma correlação entre duas (ou mais) variáveis não implica necessariamente em uma causalidade entre elas**

A causalidade pode ser comprovada quando ao mesmo tempo:

- Estatisticamente há evidências de uma correlação entre as variáveis.

Teste de Hipótese:

$H_0$ : não há correlação entre os fatores; variáveis independentes

$H_a$ : há correlação entre os fatores; variáveis dependentes

Pvalue <  $\alpha$  (normalmente 5%)

- Por observar o processo ou atuar sobre o mesmo (ex.: DOE) é possível verificar se relação causa-e-efeito pode ocorrer

# Correlação - Exemplo

- A Empresa “Bebidas Etílicas E Refrigerantes” (BEER) iniciou um Projeto para identificar as causas e propor ações corretivas com relação ao Market Share para marcas Premium em um dado mercado, que tem caído de forma constante nos últimos dois anos.
- Foram identificadas e coletados os dados das variáveis disponíveis para a análise (todas são relatadas como médias de loja para esse mercado).
- Tudo isso é baseado no volume que a BEER tem contra seus concorrentes, com base em sua participação de mercado em relação a seus concorrentes.
  - 100 é igualdade para todos os concorrentes
  - Maior que 100 favorece a BEER

# Correlação - Exemplo

- As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado):
  - PTC (Price to Customer) por caixa
  - SOV/SOM – Share of Voice/Share of Market: uma medida do investimento em mídia da BEER em relação aos concorrentes em comparação com a participação da BEER no mercado em relação aos mesmos concorrentes
  - Brand Health – Como os clientes favoráveis percebem a marca (neste caso, nossas marcas Premium) - quanto maior melhor - geralmente não é uma correlação de 1: 1, mas quanto maior for Brand Health maior volume de vendas
  - Espaço relativo (Relative space)
  - Distribuição Relativa (Relative distribution)
  - Exibições Relativas (Relative displays)

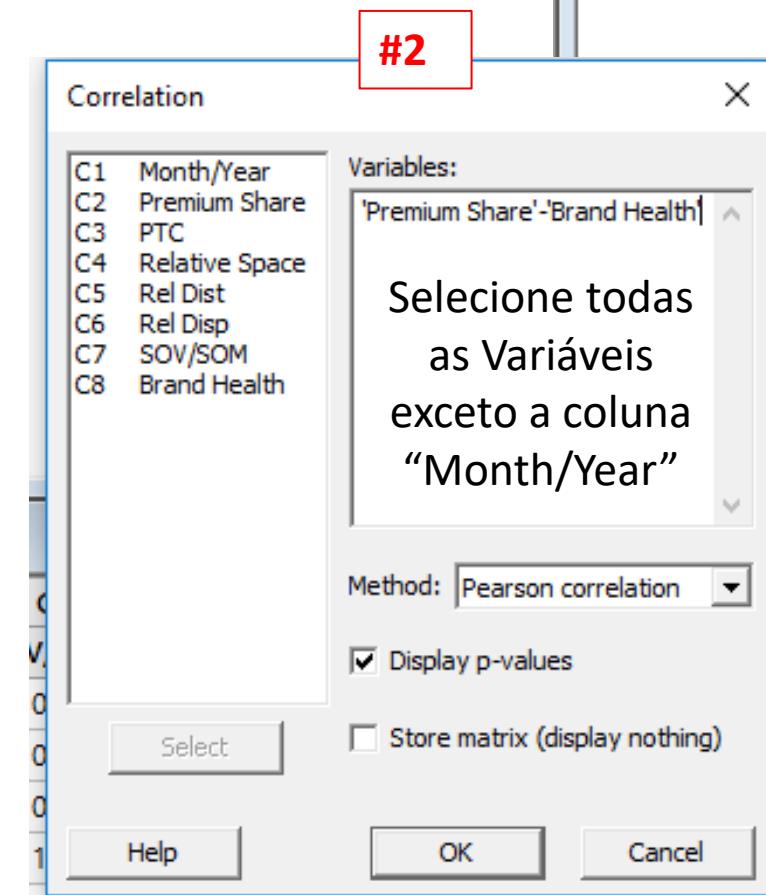
# Correlação - Exemplo

Os dados estão no arquivo BEER Data1.xlsx na aba “Market Share Decline”

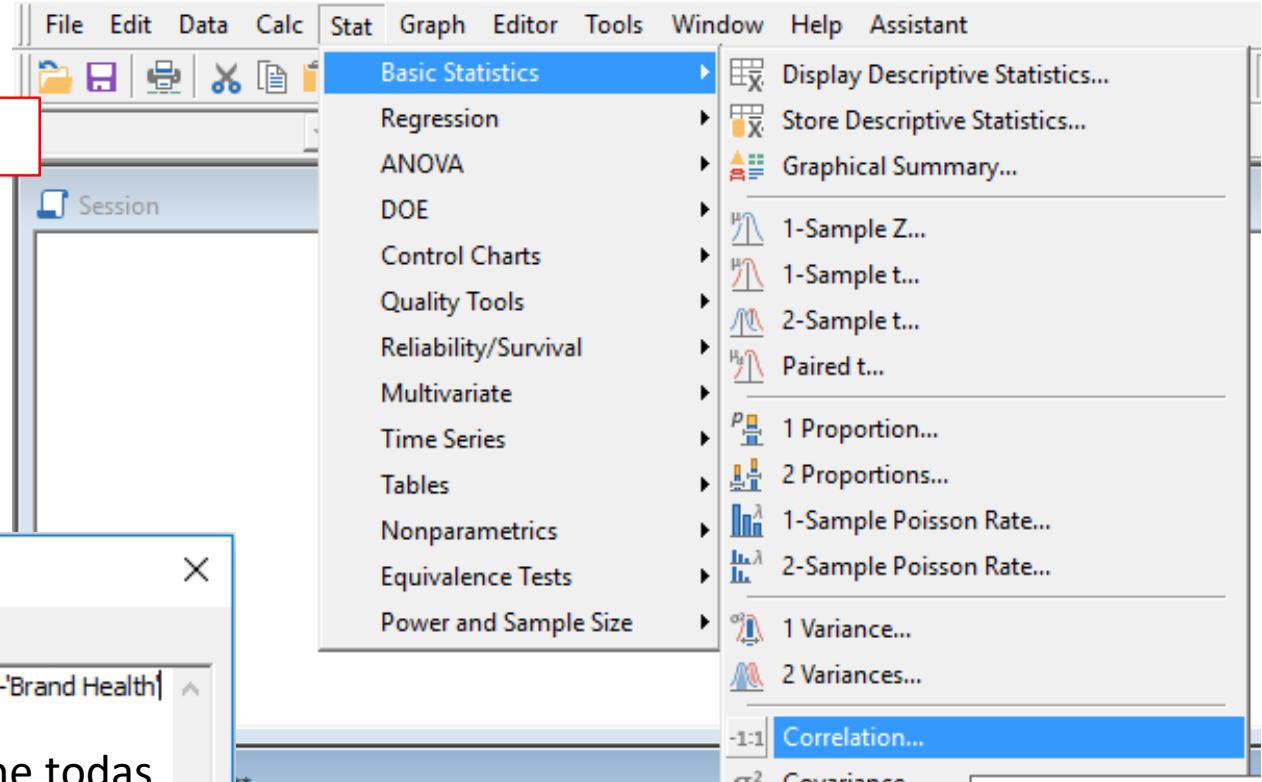
- Quais variáveis influenciam o market share?
- Existem variáveis de entrada correlacionadas entre si?
- Se sim, como isso muda sua resposta à pergunta anterior?

# Correlação

BEER Data1.xlsx



Minitab - Untitled



Display Descriptive Statistics...

Store Descriptive Statistics...

Graphical Summary...

1-Sample Z...

1-Sample t...

2-Sample t...

Paired t...

1 Proportion...

2 Proportions...

1-Sample Poisson Rate...

2-Sample Poisson Rate...

1 Variance...

2 Variances...

-1:1 Correlation...

$\sigma^2$  Covariance...

Correlation

Measure the strength and

# Correlação

- Encontre todos os P-values menores que 0,05 para Premium Share
- Share
  - Negativo com PTC
  - Positive com Rel Space, Distribution e Displays
- Tudo isso faz sentido

## Correlações de Pearson pareadas

Amostra 1	Amostra 2	N	Correlação	IC de 95% para p	Valor-p
PTC	Premium Share 24	24	-0,870	(-0,943; -0,719)	0,000
Relative Space	Premium Share 24	24	0,712	(0,434; 0,867)	0,000
Rel Dist	Premium Share 24	24	0,619	(0,287; 0,818)	0,001
Rel Disp	Premium Share 24	24	0,677	(0,376; 0,849)	0,000
SOV/SOM	Premium Share 24	24	-0,290	(-0,620; 0,129)	0,170
Brand Health	Premium Share 24	24	0,013	(-0,393; 0,414)	0,953

# Correlação

- Encontre todos os P-Values menores que 0,05 para as entradas (Xs)
- PTC, Space, Distribution e Displays estão todos correlacionados
- Quais deles estão realmente direcionando Market Share?
- Quais são confundidos?

*Não há resposta estatística para quais são os drivers e quais são confundidos. É aí que entra a expertise do analista. Talvez nem todos sejam drivers.*

## Correlações de Pearson pareadas

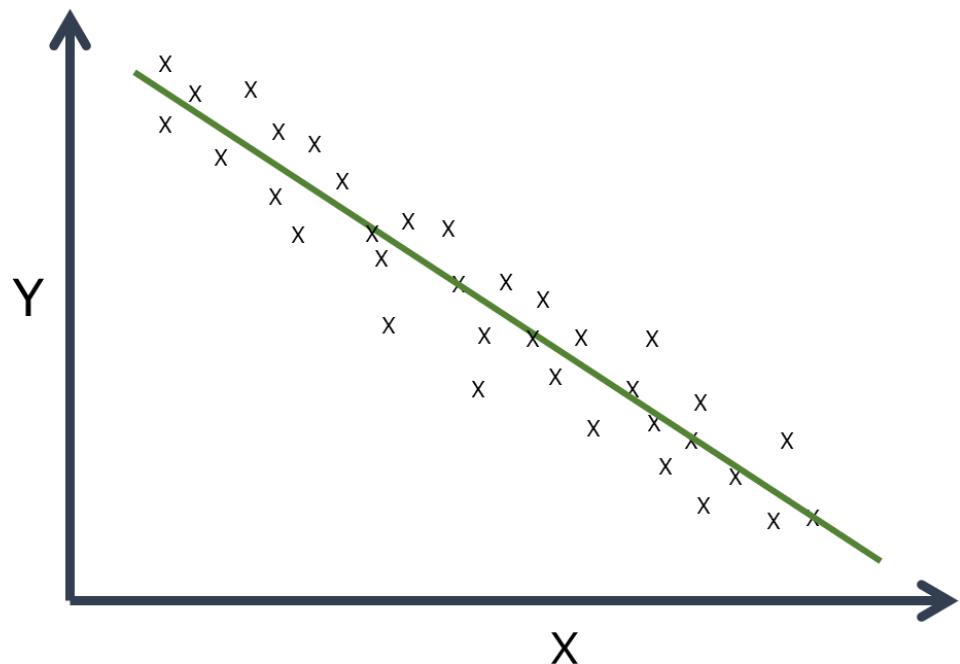
Amostra 1	Amostra 2	N	Correlação	IC de 95% para p	Valor-p
Relative Space	PTC	24	-0,577	(-0,795; -0,227)	0,003
Rel Dist	PTC	24	-0,469	(-0,734; -0,081)	0,021
Rel Disp	PTC	24	-0,473	(-0,736; -0,085)	0,020
SOV/SOM	PTC	24	0,242	(-0,178; 0,588)	0,254
Brand Health	PTC	24	0,153	(-0,267; 0,524)	0,475
Rel Dist	Relative Space	24	0,887	(0,754; 0,950)	0,000
Rel Disp	Relative Space	24	0,790	(0,567; 0,905)	0,000
SOV/SOM	Relative Space	24	-0,187	(-0,549; 0,234)	0,382
Brand Health	Relative Space	24	-0,197	(-0,556; 0,225)	0,357
Rel Disp	Rel Dist	24	0,735	(0,472; 0,878)	0,000
SOV/SOM	Rel Dist	24	-0,092	(-0,477; 0,324)	0,670
Brand Health	Rel Dist	24	-0,218	(-0,571; 0,203)	0,306
SOV/SOM	Rel Disp	24	-0,269	(-0,606; 0,151)	0,204
Brand Health	Rel Disp	24	-0,224	(-0,576; 0,197)	0,292
Brand Health	SOV/SOM	24	-0,010	(-0,411; 0,395)	0,965

# Regressão Linear

- É a função matemática que descreve a relação entre duas variáveis
- Quantifica a relação entre as variáveis Y e X ( $Y = b_0 + b_1X$ )
- Principal métrica:
  - Coeficiente de Determinação,  $R^2$  (varia entre “0,0” e “1,0”, ou de zero até 100%)
    - Nada da variação de Y é explicada pelo X =>  $R^2 = 0,0$
    - Toda a variação de Y é explicada pelo X =>  $R^2 = 1,0$

# Regressão Linear Simples

- Equação baseada no Método dos Mínimos Quadrados, no qual a reta ajustada é aquela que minimiza o quadrado das distâncias entre o valor real e o valor ideal encontrado pela reta



**MODELO LINEAR:**  $Y = b_0 + b_1 X$

Coeficiente  
Linear

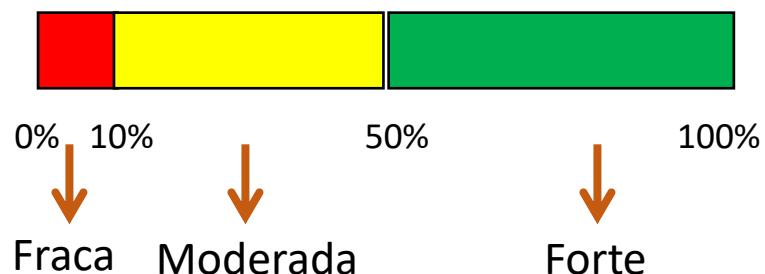
Coeficiente  
Angular

# Regressão Linear Simples

O Coeficiente de Determinação  $R^2$  indica:

- O quanto forte é a relação entre as variáveis

- Quanto o fator “X” afeta na variação de “Y”



$$R^2 = \frac{\text{Variação Explicada}}{\text{Variação Total}}$$

# Regressão - Exemplo

- A Empresa “Bebidas Etílicas E Refrigerantes” (BEER) iniciou um Projeto para identificar as causas e propor ações corretivas com relação ao Market Share para marcas Premium em um dado mercado, que tem caído de forma constante nos últimos dois anos.
- Foram identificadas e coletados os dados das variáveis disponíveis para a análise (todas são relatadas como médias de loja para esse mercado).
- Tudo isso é baseado no volume que a BEER tem contra seus concorrentes, com base em sua participação de mercado em relação a seus concorrentes.
  - 100 é igualdade para todos os concorrentes.
  - Maior que 100 favorece a BEER

# Regressão - Exemplo

- As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado):
  - PTC (Price to Customer) por caixa
  - SOV/SOM – Share of Voice/Share of Market: uma medida do investimento em mídia da BEER em relação aos concorrentes em comparação com a participação da BEER no mercado em relação aos mesmos concorrentes
  - Brand Health – Como os clientes favoráveis percebem a marca (neste caso, nossas marcas Premium) - quanto maior melhor - geralmente não é uma correlação de 1: 1, mas quanto maior for Brand Health maior volume de vendas
  - Espaço relativo (Relative space)
  - Distribuição Relativa (Relative distribution)
  - Exibições Relativas (Relative displays)

# Regressão - Exemplo

Os dados estão no arquivo BEER Data1.xlsx na aba “Market Share Decline”

- Durante o exercício de Correlação, você observou que a PTC era o fator mais forte
- Realize uma regressão simples do Market Share como uma função do PTC.

# Regressão

BEER Data1.xlsx

**#3**

**Regressão**

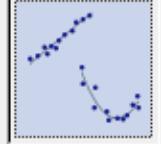
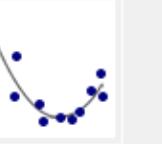
Variáveis contínuas

Coluna Y: Premium Share  
Coluna X: PTC

Os dados estão registrados na worksheet em ordem cronológica

Tipo de modelo de regressão

O Minitab pode selecionar o melhor modelo de ajuste ou você pode selecionar um modelo.

Selecionar para mim  Linear  Quadrático 

Configuração de teste

Qual risco você está disposto a aceitar para concluir que existe uma relação entre Y e X quando não houver uma?

Nível alfa: 0,05

Selecionar **OK** **Cancelar**

**#1**

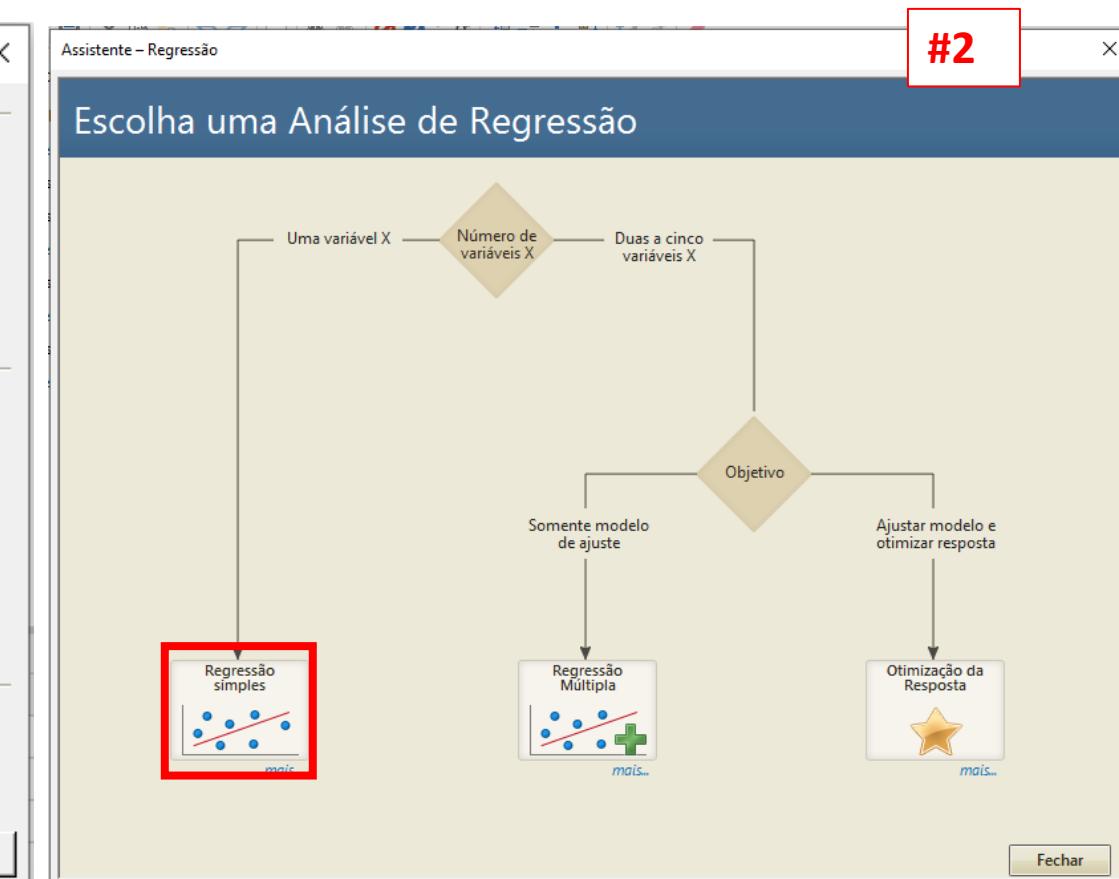
Assistente Ferramentas Adicionais

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
Testes de Hipótese...  
**Regressão...**

**Análise de Regressão**

Usar para modelar a relação entre uma ou mais entradas do processo (X) e a saída do processo (Y).

Cartas de Controle...



# Regressão - Exemplo

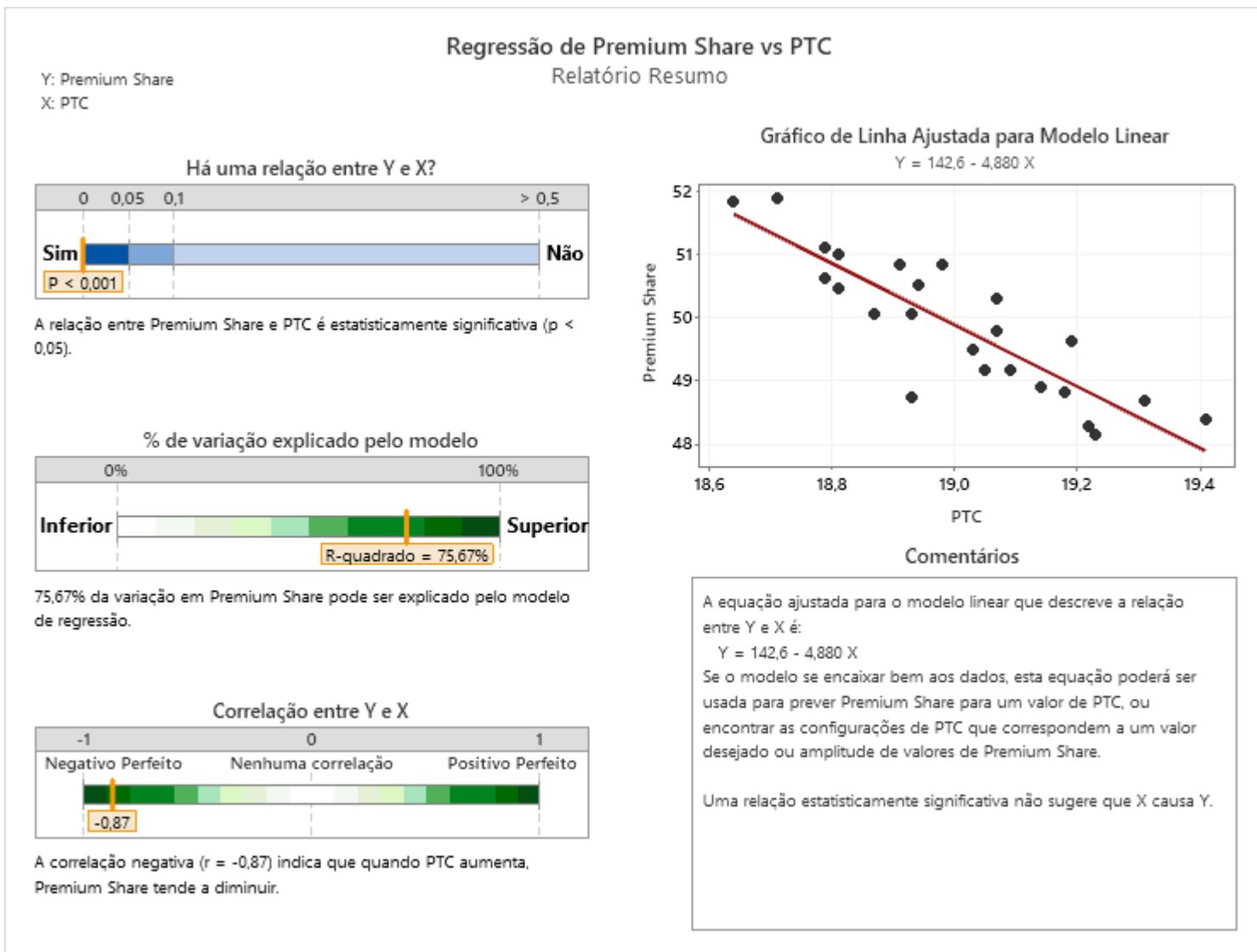
- Quantidade de dados é questionada
  - Geralmente, 15 ou mais conjuntos de pontos são suficientes. Quanto maior, melhor
- Um dos valores é destacado
  - será necessário investigar para ver seu impacto na inclinação do relacionamento
- Normalidade está OK

## Regressão de Premium Share vs PTC Cartão de Relatório

Verificar	Status	Descrição
Quantidade de Dados		O tamanho de sua amostra ( $n = 24$ ) não é grande o suficiente para fornecer uma estimativa precisa da força da relação. As medições da força da relação, como um R-Quadrado e R-Quadrado (ajustado), podem variar muito. Para obter uma estimativa mais precisa, devem ser usadas grandes quantidades de amostras (normalmente 40 ou mais).
Dados Atípicos		Um ponto de dados tem um resíduo grande e não está bem ajustado pela equação. Esse ponto está marcado em vermelho nos gráficos e está na linha 24 da worksheet. Como dados atípicos podem exercer uma forte influência sobre os resultados, Tente identificar a causa de sua natureza atípica. Corrija quaisquer erros de entrada de dados ou medições. Considere remover os dados que estão associados com causas especiais e refazer a análise.
Normalidade		Como você tem no mínimo 15 pontos de dados, a normalidade não é um problema. Se o número de pontos de dados for pequeno e os resíduos não forem normalmente distribuídos, o valor-p usado para determinar se há uma relação significativa entre X e Y pode não ser exato.
Ajuste do Modelo		Você deve avaliar os dados e o ajuste do modelo em termos de suas metas. Veja o gráfico de linha ajustada para ter certeza de que: <ul style="list-style-type: none"><li>• A amostra cobre adequadamente a amplitude de valores X.</li><li>• O modelo se ajusta bem a todas as curvaturas dos dados (evita o superajuste).</li><li>• A reta se ajusta bem nas áreas de interesse especial.</li></ul>

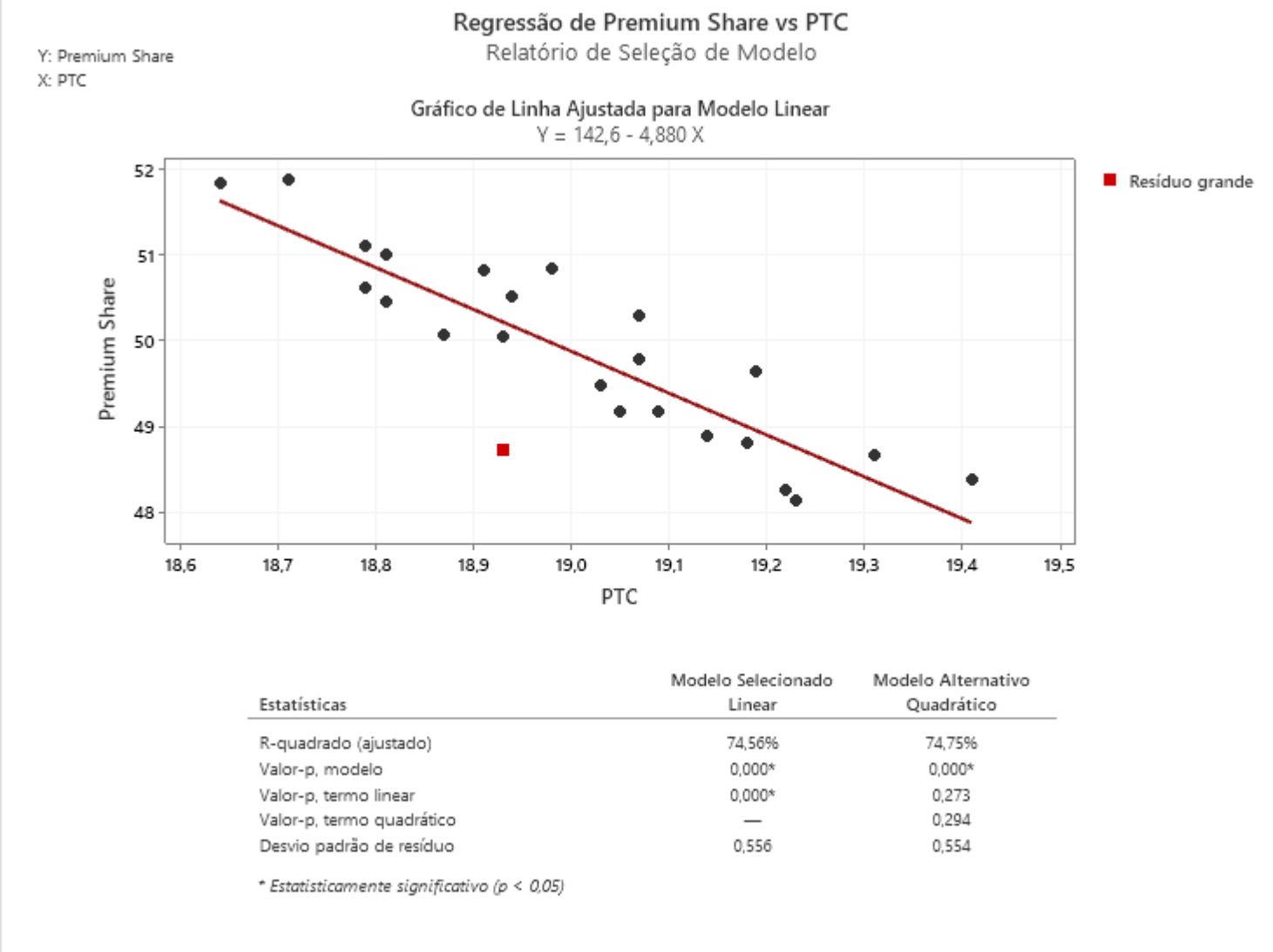
# Regressão - Exemplo

- O modelo é estatisticamente significativo:
  - (p-value < 0,001)
- Há uma correlação negativa:
  - (quanto maior PTC, menor é o share)
- $R^2 = 75,6\%$  - correlação forte



# Regressão - Exemplo

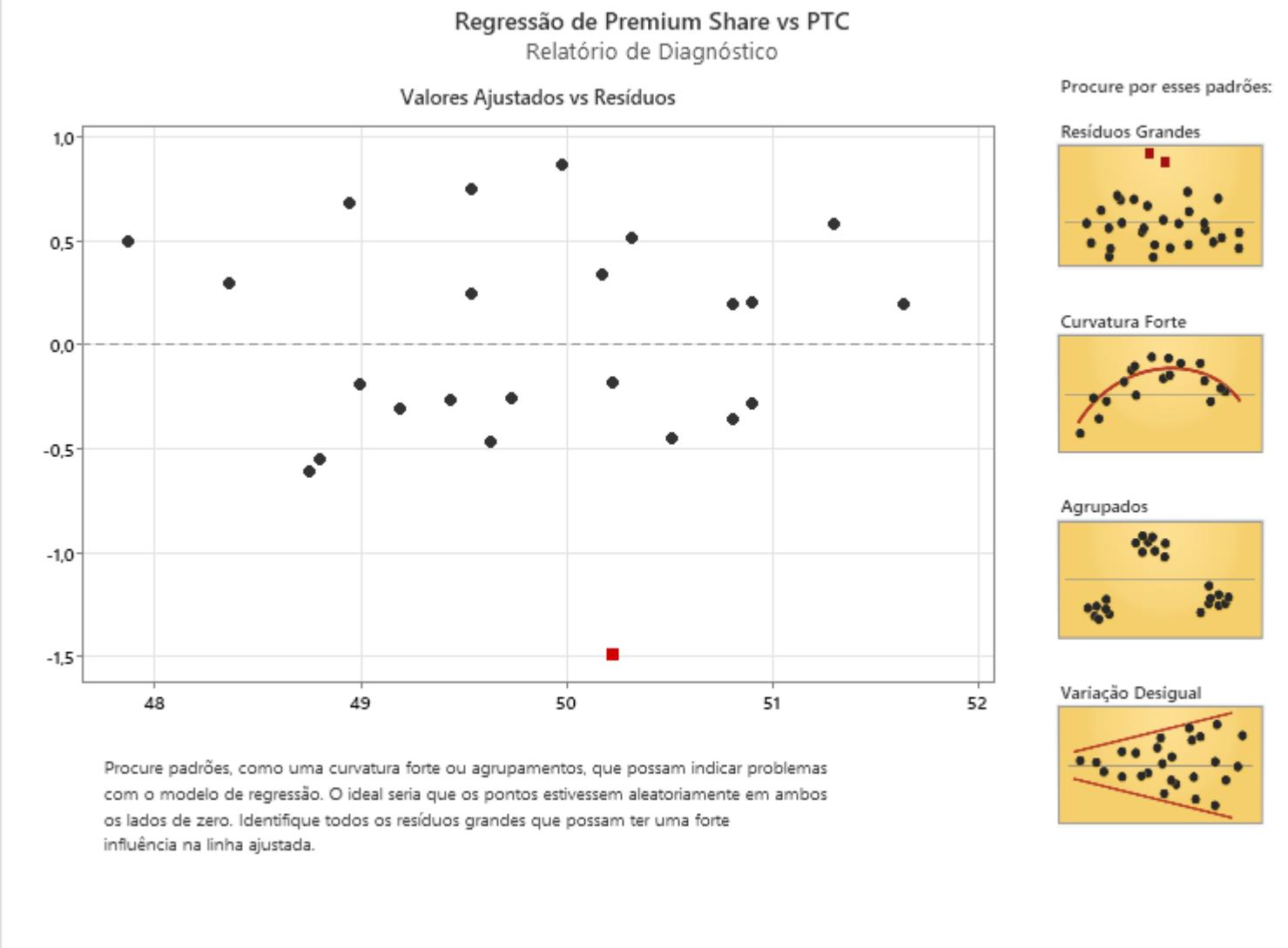
- Correlação linear para este intervalo de valores de PTC é apropriada
- Há um outlier no meio do intervalo PTC



## Regressão - Exemplo

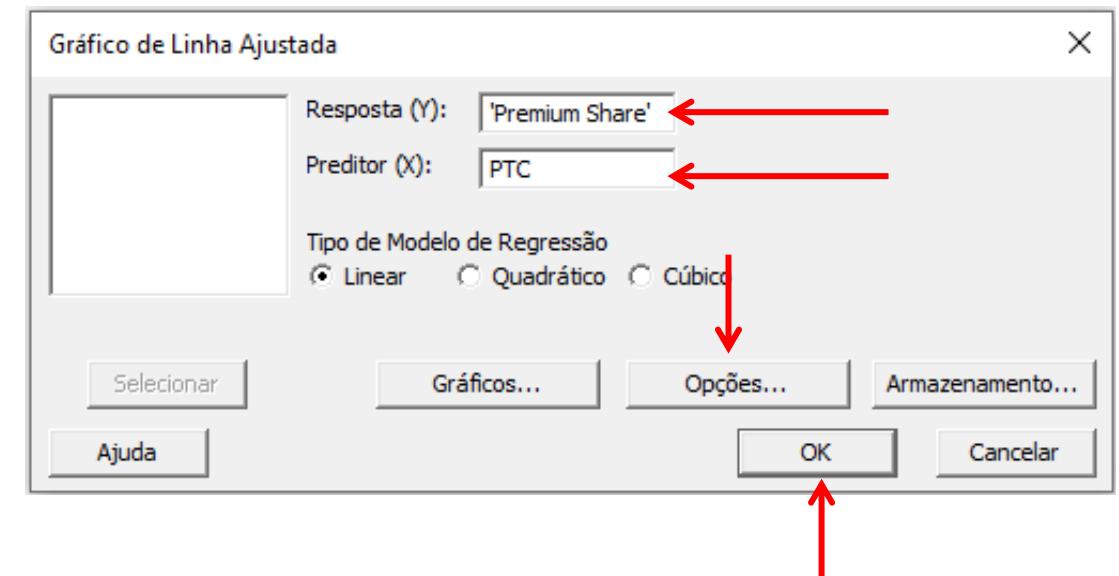
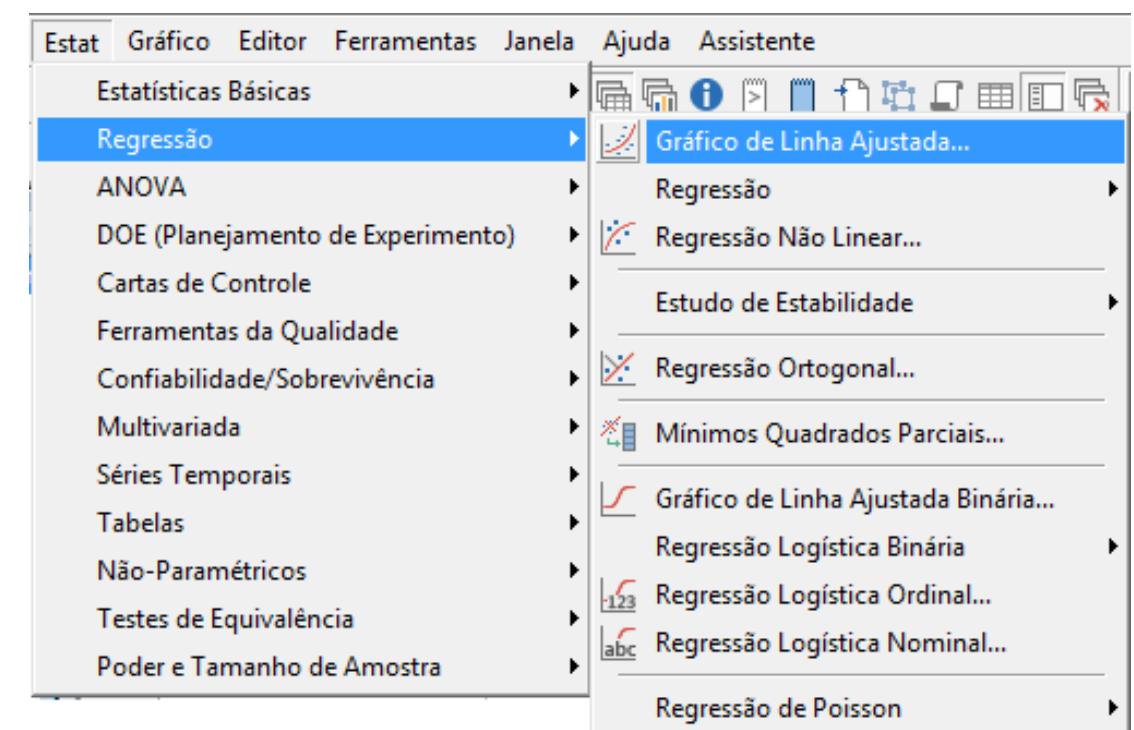
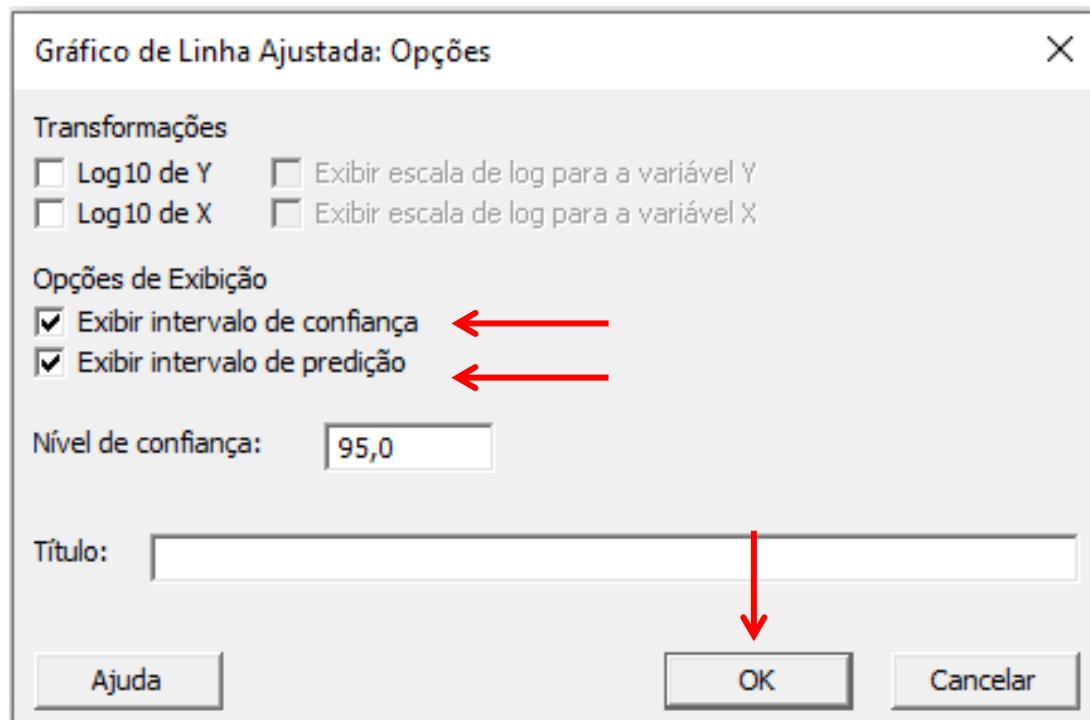
Há um outlier no meio do intervalo  
(não afeta muito a inclinação da  
equação e nem a validade do modelo)

Não há padrões observados no range  
avaliado





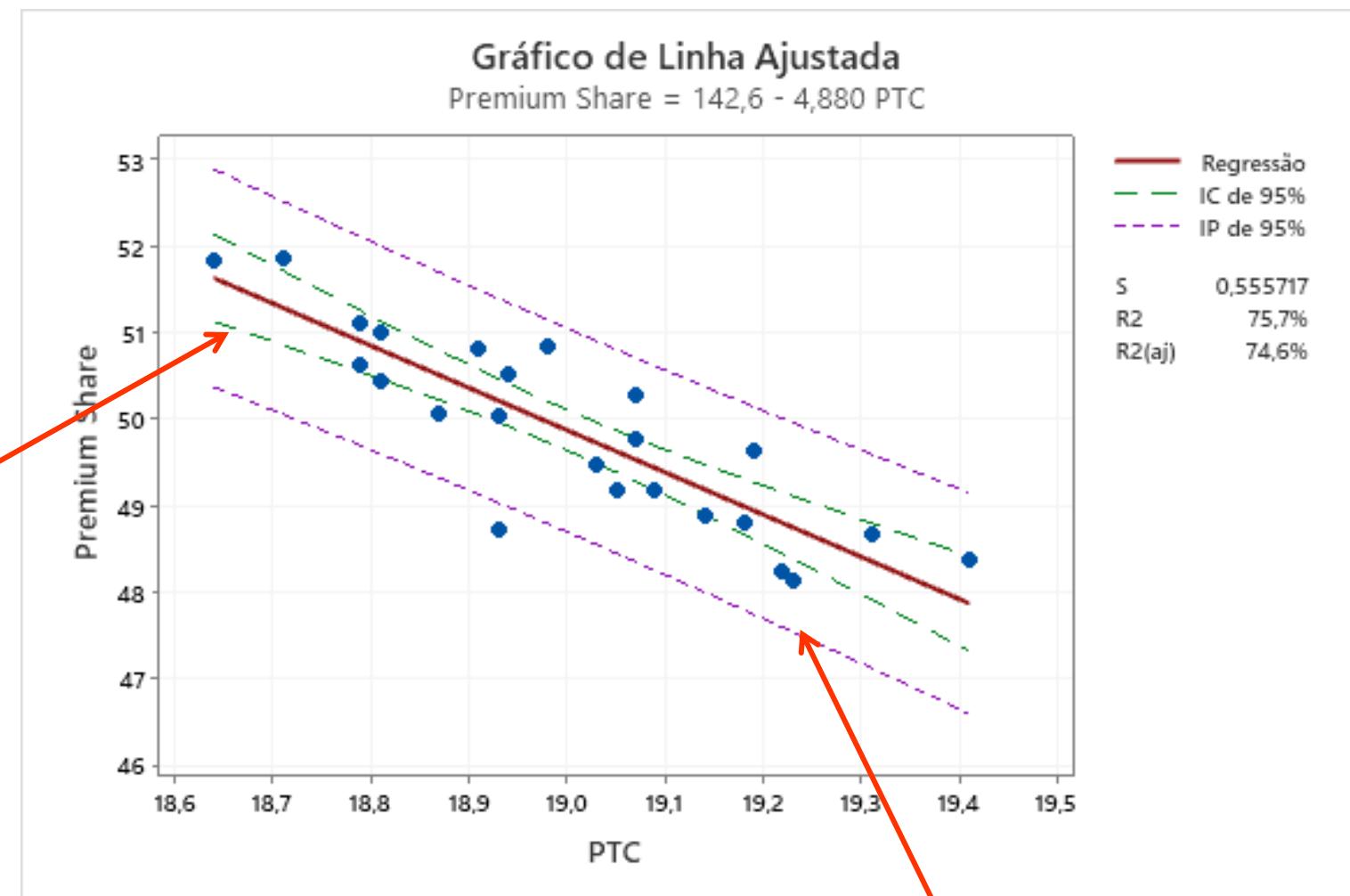
# Regressão Linear Simples – Exemplo Predição de Resultados



# Regressão Linear Simples – Exemplo

## Predição de Resultados

IC: amplitude de valores prováveis para as médias da população



IP: amplitude de valores prováveis para novas observações

# Business Case Produtividade Álcool Gel

Visando atuar no problema de qualidade relativo a rejeição de alguns lotes abaixo da especificação mínima de envase que é 500 ml (o máximo admitido é 505 ml), foram levantados dados de volume e tempo de envase.

Realize a seguinte análise:

- Regressão: para verificar qual a relação entre as variáveis de tempo e volume de enchimento
- Identifique sugestões de melhoria

Arquivo: Case CoolGel\_Dados\_Produtividade  
Aba: VolxTempo

# Business Case Mercado – PDV Álcool Gel

As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado no período de abril/2018 – março/2020):

- PTC (Price to Customer) por frasco
- Espaço relativo (Relative space)
- Distribuição Relativa (Relative distribution)
- Exibições Relativas (Relative displays)
- Evolução do Market Share para o mesmo período

# Business Case Mercado – PDV Álcool Gel – Exercício

Os dados estão no arquivo Coolgel\_Bando de Dados\_PDV.xlsx  
na aba “Market Share Decline”

Usem os dados para efetuar análises de Correlações e  
Regressões

- Tragam conclusões sobre Market Share e essas variáveis
- Indiquem potenciais melhorias
- Indique a estratégia de ação futura

# Regressão Múltipla

## Regressão Múltipla

- Para análise de Regressão Múltipla, se deseja entender a relação de 1 Y com mais do que 1 X;
- São diversos X's variando ao mesmo tempo, e sobre esta variação se pretende entender seus eventuais efeitos sobre a variável resposta.
- Há algumas formas distintas de se realizar esta análise e a forma sugerida neste material busca entender a correlação das variáveis X's, e escolher aquelas que trazem maior impacto sobre a variável Y.
- Com a seleção destas variáveis, pode-se encontrar o modelo  $Y = F(X)$ .

## Regressão Múltipla

- Começa-se por fazer uma análise exploratória no conjuntos dos dados visando:
  - Identificar graficamente correlações entre os X's e Y e entre os próprios X's;
  - Calcular os Coeficientes de Correlação Linear entre os X's e entre estes e o Y;
- Com isto já se terá uma noção das eventuais relações existentes entre os X's e Y e entre os próprios X's – o que poderá implicar em retirada de alguns deles do modelo final a ser construído;
- Este processo é interessante se não houver muitos X's a serem analisados, caso contrário pode-se partir direto para análise com as ferramentas analíticas.

# Regressão Múltipla - Exemplo

- A Empresa “Bebidas Etílicas E Refrigerantes” (BEER) iniciou um Projeto para identificar as causas e propor ações corretivas com relação ao Market Share para marcas Premium em um dado mercado, que tem caído de forma constante nos últimos dois anos.
- Foram identificadas e coletados os dados das variáveis disponíveis para a análise (todas são relatadas como médias de loja para esse mercado).
- Tudo isso é baseado no volume que a BEER tem contra seus concorrentes, com base em sua participação de mercado em relação a seus concorrentes.
  - 100 é igualdade para todos os concorrentes.
  - Maior que 100 favorece a BEER

# Regressão Múltipla - Exemplo

- As seguintes variáveis estavam disponíveis para análise (todas são relatadas como médias de loja para esse mercado):
  - PTC (Price to Customer) por caixa
  - SOV/SOM – Share of Voice/Share of Market: uma medida do investimento em mídia da empresa BEER em relação aos concorrentes em comparação com a participação da BEER no mercado em relação aos mesmos concorrentes
  - Brand Health – Como os clientes favoráveis percebem a marca (neste caso, nossas marcas Premium) - quanto maior melhor - geralmente não é uma correlação de 1: 1, mas quanto maior for Brand Health maior volume de vendas
  - Espaço relativo (Relative space)
  - Distribuição Relativa (Relative distribution)
  - Exibições Relativas (Relative displays)

# Regressão Múltipla - Exemplo

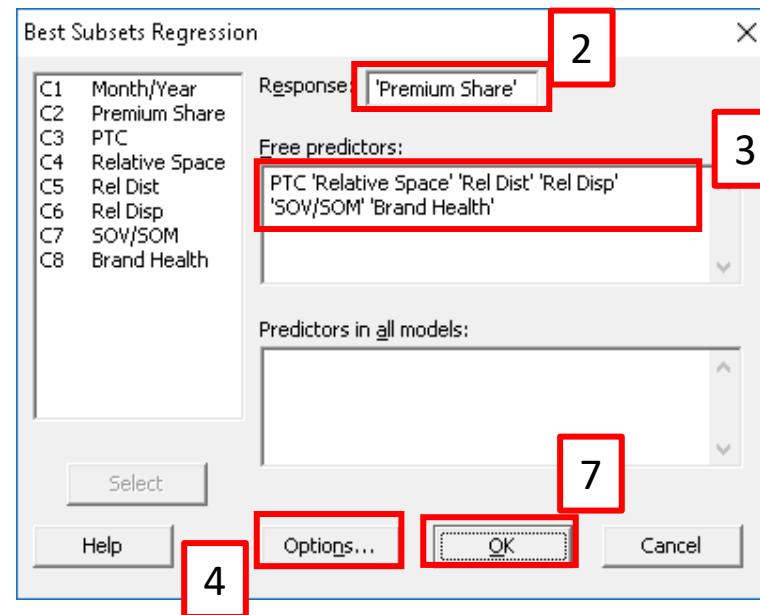
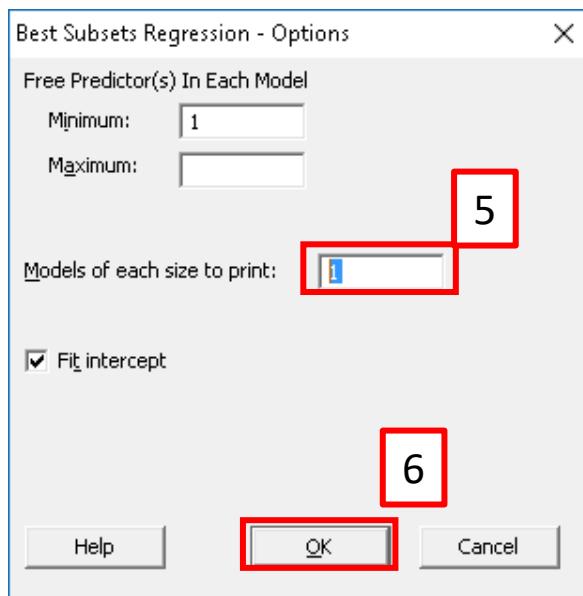
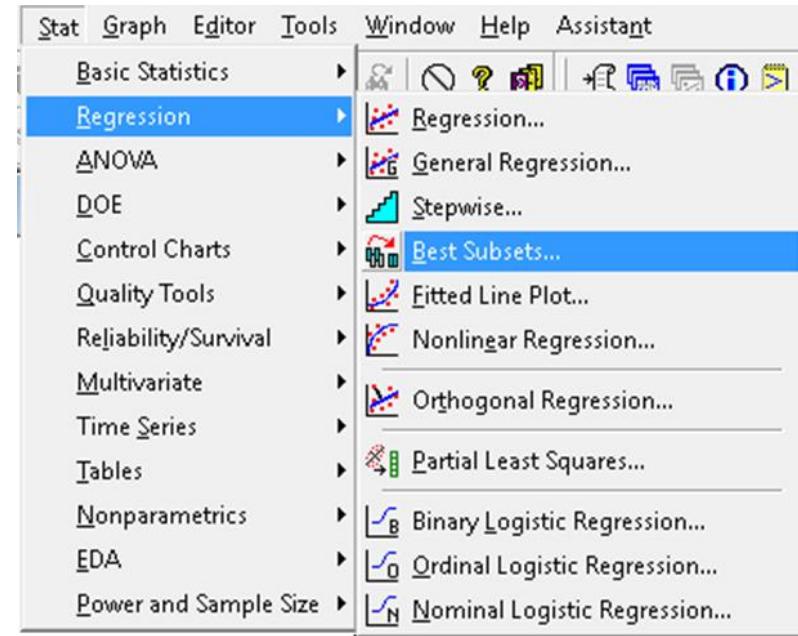
Os dados estão no  
arquivo BEER Data1.xlsx  
na aba “Market Share  
Decline”

- Quais fatores poderiam compor  
uma equação preditiva para  
Market Share?



# Regressão Múltipla – Exemplo

BEER Data1 .xlsx



1

4



# Regressão Múltipla – Exemplo

BEER Data1.xlsx

Estat Gráfico Visualizar Ajuda Assistente Ferramentas Adicionais

Estatísticas Básicas

Regressão

ANOVA

DOE (Planejamento de Experimento)

Cartas de Controle

Ferramentas da Qualidade

Confiabilidade/Sobrevivência

Análise preditiva

Multivariada

Séries Temporais

Tabelas

Não-Paramétricos

Testes de Equivalência

Poder e Tamanho de Amostra

Gráfico de Linha Ajustada...

Regressão

Regressão Não Linear...

Estudo de Estabilidade

Regressão Ortogonal...

Mínimos Quadrados Parciais...

Gráfico de Linha Ajustada Binária...

Regressão Logística Binária

Regressão Logística Ordinal...

Regressão Logística Nominal...

Regressão de Poisson

Ajuste de Modelo de Regressão...

Melhores Subconjuntos...

**Melhores Subconjuntos**

Comparar todos os modelos possíveis usando um conjunto especificado de preditores, e exibir os modelos com os melhores ajustes que contêm um preditor, dois preditores e assim por diante.

Oti

**Regressão dos Melhores Subconjuntos**

C1 Mês/Ano  
C2 Premium Share  
C3 PTC  
C4 Relative Space  
C5 Rel Dist  
C6 Rel Disp  
C7 SOV/SOM  
C8 Brand Health

Resposta: 'Premium Share' 2

Preditores livres:  
3  
PTC 'Relative Space' 'Rel Dist' 'Rel Disp'  
'SOV/SOM' 'Brand Health'

Preditores em todos os modelos:

Selecionar

Ajuda

Opções...

OK 4

Cancelar

1

2

3

4

# Regressão Múltipla - Exemplo

- Selecione o melhor subconjunto de Xs seguindo a ordem dos critérios
- Neste exemplo, o melhor subconjunto é o que apresenta apenas “PTC”, “Rel Disp” e “Brand Health”, pois só pelo primeiro critério (maior R<sup>2</sup> ajustado) já se pode identificar

Vars	Mallows		S
	R-Sq	R-Sq(adj)	
1	75,7	74,6	19,2 0,55572 X
2	84,8	83,3	6,5 0,44989 X X
3	88,9	87,3	1,9 0,39333 X X X
4	89,4	87,1	3,1 0,39532 X X X X
5	89,4	86,5	5,0 0,40511 X X X X X
6	89,4	85,7	7,0 0,41629 X X X X X X

## Critérios de escolha das variáveis:

### **1º) R<sup>2</sup> e R<sup>2</sup><sub>Ajustado</sub>**

Nr variáveis é igual a 1 – usar maior R<sup>2</sup>.

Nr variáveis é diferente de 1 – usar maior R<sup>2</sup><sub>Ajustado</sub>.

### **2º) Mallow's Cp**

Preferir modelos com Cp mais próximo do nr de variáveis e mais o intercepto, ou seja, nr de variáveis preditoras + 1.

### **3º) S**

Buscar o menor S possível.

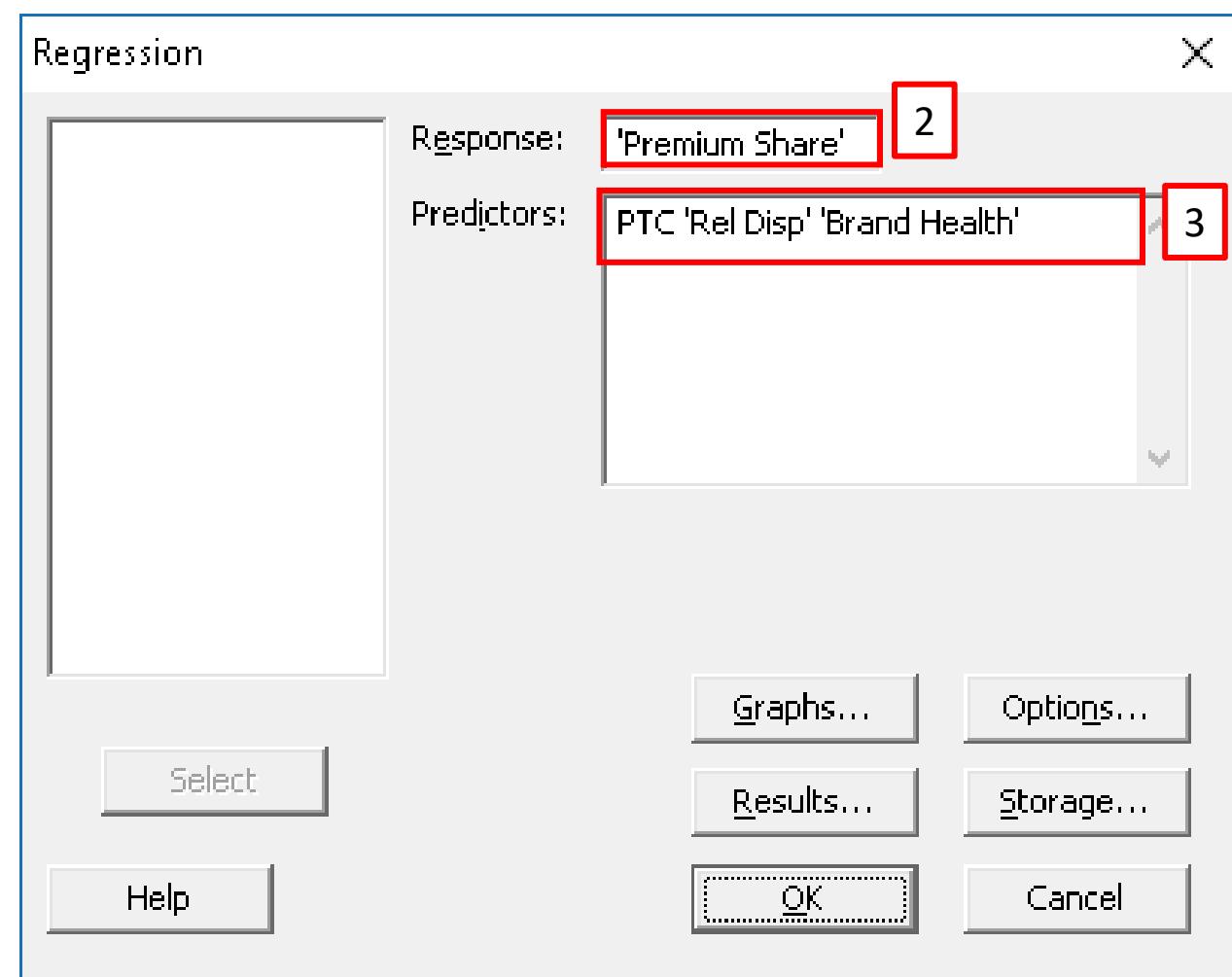
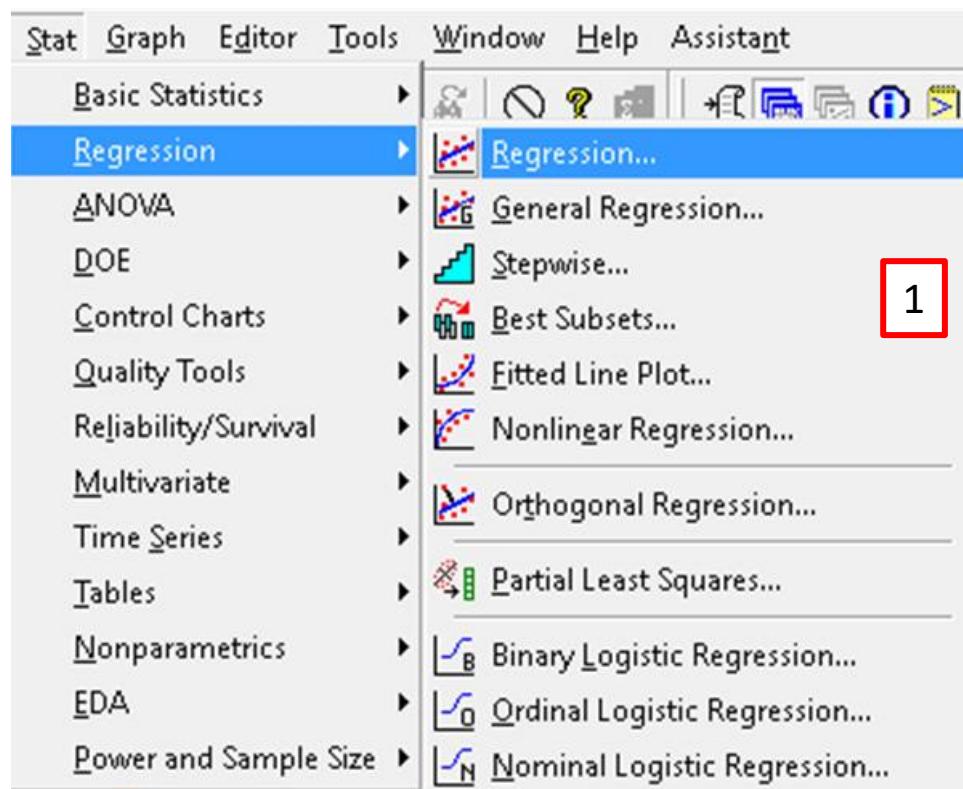
**Os critérios devem ser usados na ordem acima.**

Parâmetros de cada um dos modelos avaliados



# Regressão Múltipla – Exemplo

BEER Data1 .xlsx

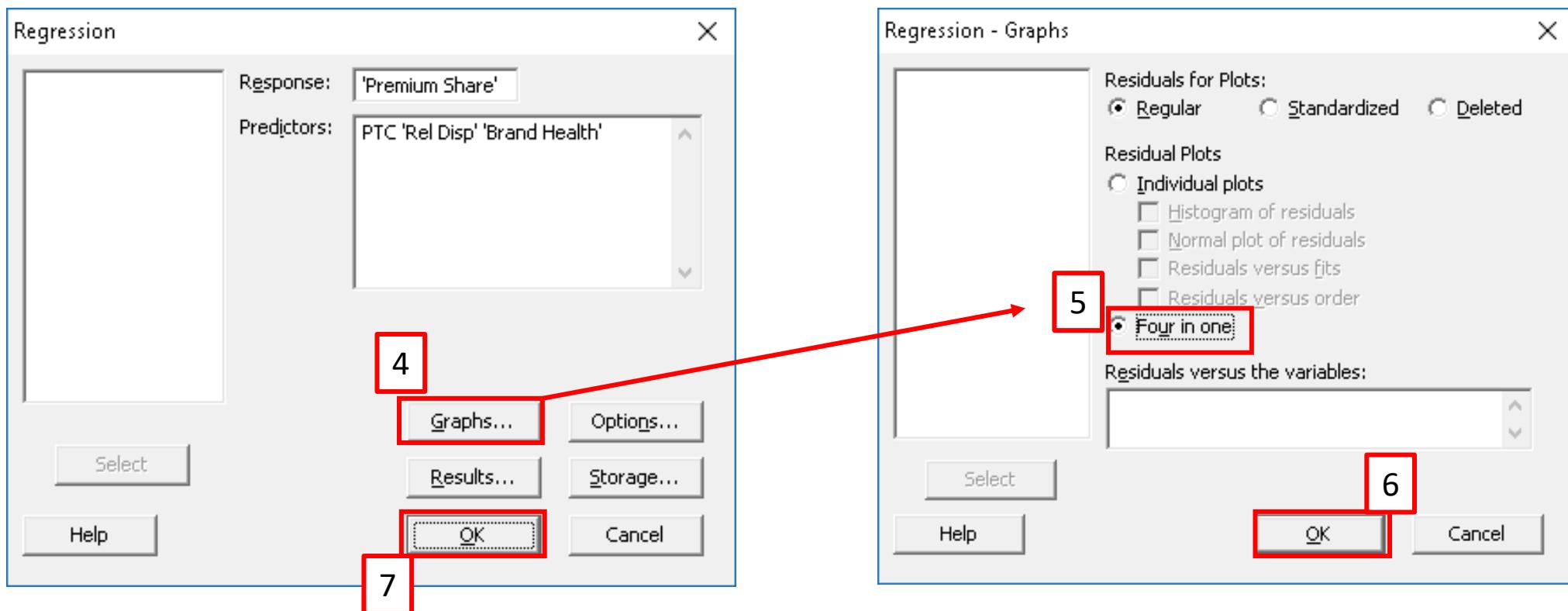


Uma vez selecionado o melhor subconjunto de Xs, efetuar a regressão múltipla apenas com as variáveis selecionadas em Best Subsets



# Regressão Múltipla – Exemplo

BEER Data1.xlsx





# Regressão Múltipla – Exemplo

BEER Data1.xlsx

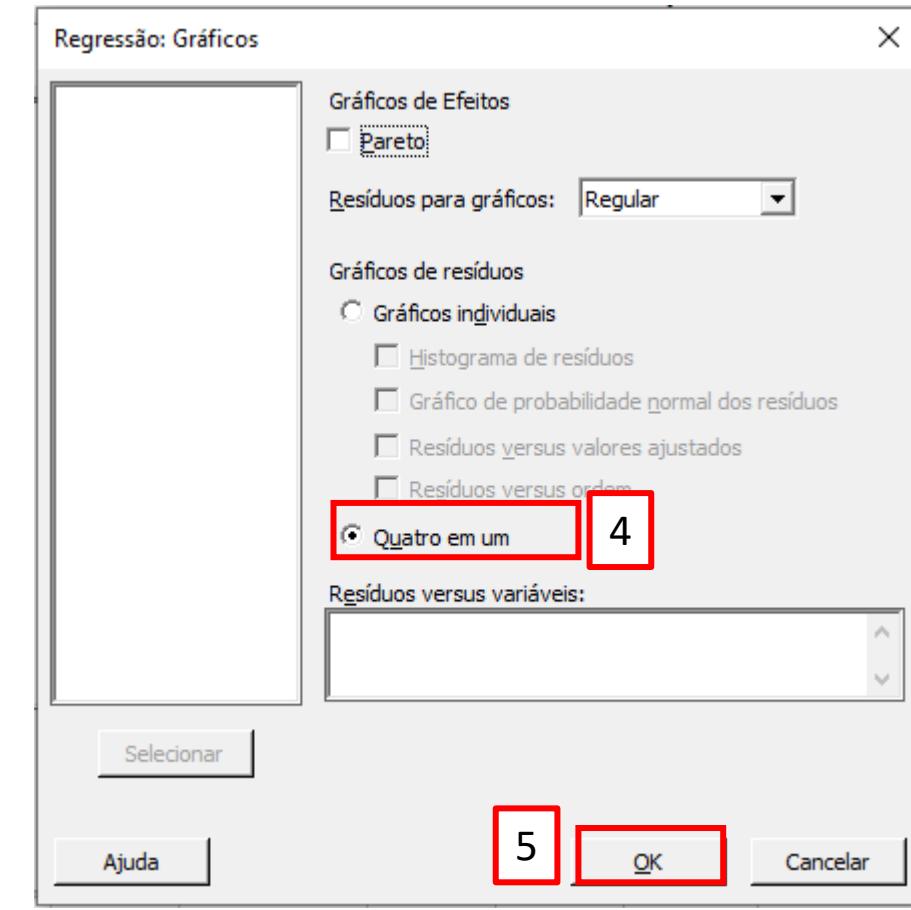
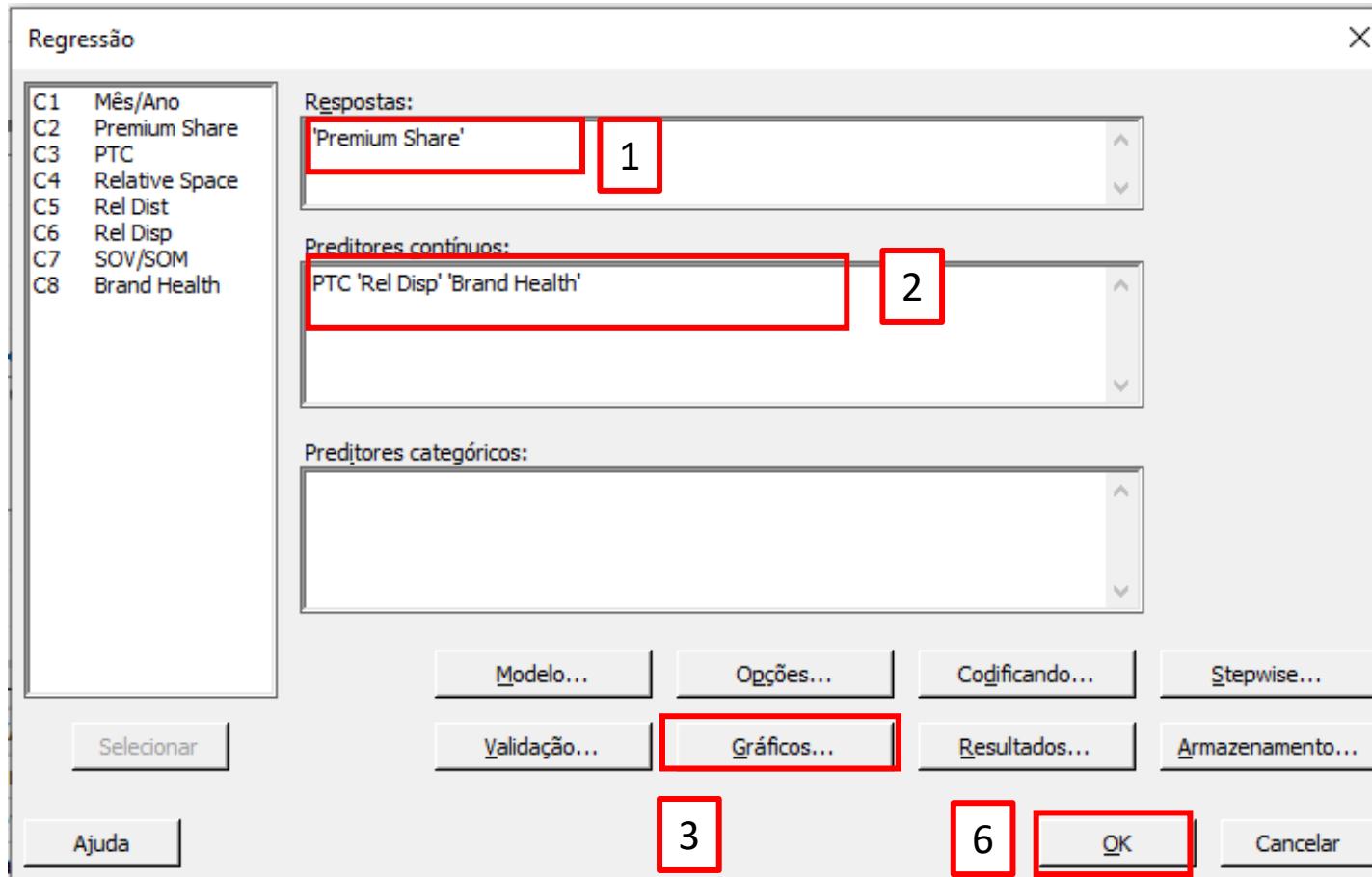
The screenshot shows the Minitab software interface with the 'Estat' tab selected in the menu bar. Under the 'Estat' menu, the 'Regressão' (Regression) option is highlighted. A sub-menu for 'Regressão' is open, listing various regression analysis options: Gráfico de Linha Ajustada..., Regressão, Regressão Não Linear..., Estudo de Estabilidade, Regressão Ortogonal..., Mínimos Quadrados Parciais..., Gráfico de Linha Ajustada Binária..., Regressão Logística Binária, Regressão Logística Ordinal..., Regressão Logística Nominal..., and Regressão de Poisson. The 'Ajuste de Modelo de Regressão...' (Model Fit) option is also visible in the main menu bar. A tooltip for 'Ajustar de Modelo de Regressão' (Fit Model) is displayed, describing it as a tool to model the relationship between categorical or continuous predictors and a response, including polynomial and interaction terms, or transforming the response if necessary.

Uma vez selecionado o melhor subconjunto de Xs, efetuar a regressão múltipla apenas com as variáveis selecionadas em Best Subsets



# Regressão Múltipla – Exemplo

BEER Data1.xlsx



Uma vez selecionado o melhor subconjunto de Xs, efetuar a regressão múltipla apenas com as variáveis selecionadas em Best Subsets

### Equação de Regressão

$$\text{Premium Share} = 117,63 - 4,043 \text{ PTC} + 0,02986 \text{ Rel Disp} + 0,1354 \text{ Brand Health}$$

Modelo/Equação que representa o Y em função dos Xs.

### Coeficientes

Termo	Coef	EP de Coef	Valor-T	Valor-P	VIF
Constante	117,63	9,54	12,33	0,000	
PTC	-4,043	0,475	-8,52	0,000	1,29
Rel Disp	0,02986	0,00668	4,47	0,000	1,33
Brand Health	0,1354	0,0495	2,73	0,013	1,06

### Avaliação do P-value

Para Regressões Múltiplas é avaliado não só se o modelo de uma forma geral é estatisticamente válido, mas também para cada variável é realizado o teste. Deve ser  $< 0,05$

### Sumário do Modelo

S	R2	R2(aj)	R2(pred)
0,393334	88,92%	87,26%	84,51%

### Avaliação VIF

Indica Multicolinearidade, que afeta a estimativa do coeficiente. O valor de VIF indica Multicolinearidade:

- sem: VIF = 1
- média:  $1 < \text{VIF} < 5$
- alta:  $\text{VIF} > 5$  (coeficientes mal estimados)

### Análise de Variância

Fonte	GL	SQ (A.j.)	QM (A.j.)	Valor F	Valor-P
Regressão	3	24,829	8,2763	53,50	0,000
PTC	1	11,233	11,2333	72,61	0,000
Rel Disp	1	3,090	3,0900	19,97	0,000
Brand Health	1	1,156	1,1561	7,47	0,013
Erro	20	3,094	0,1547		
Total	23	27,923			

### Parâmetros de Ajuste do Modelo ( $R^2$ , $R^2$ ajustado e S)

O  $R^2$  ajustado independe do número de fatores.

### Ajustados e Diagnósticos para Observações Atípicas

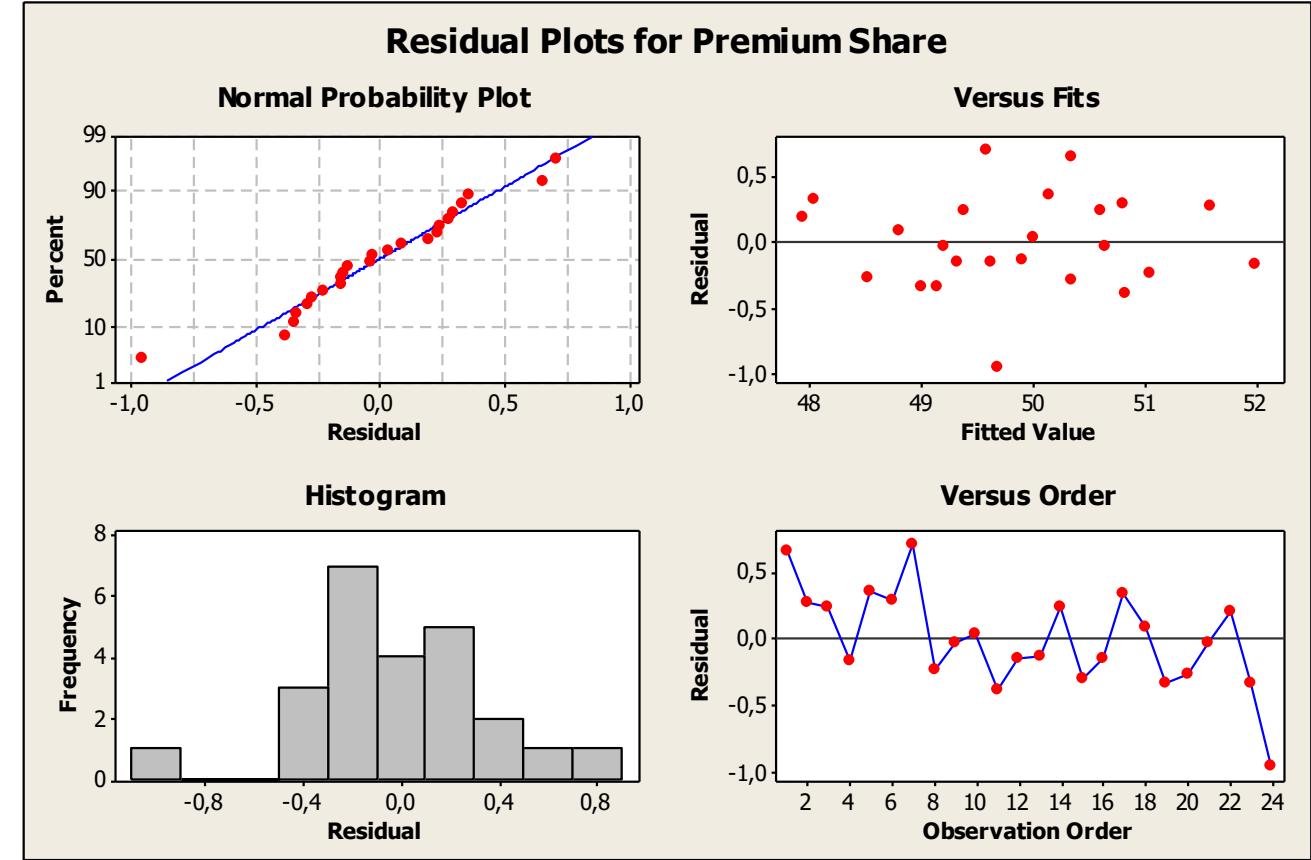
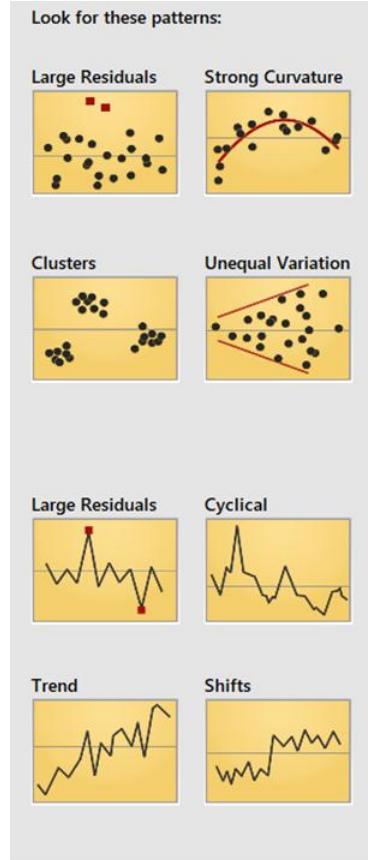
Obs.	Premium Share	Ajuste	Resíd	Resíd Pad
24	48,734	49,689	-0,955	-2,59 R

Resíduos que merecem alguma investigação. A Análise de Resíduos pode e deve ser combinada com a análise gráfica

# Regressão Múltipla – Exemplo

## Análise de Resíduos

Para a Análise de Resíduos, fique atento aos padrões:





# Regressão Múltipla – com Stepwise

BEER Data1.xlsx

Estat Gráfico Visualizar Ajuda Assistente Ferramentas Adicionais

Estatísticas Básicas

Regressão

ANOVA

DOE (Planejamento de Experimento)

Cartas de Controle

Ferramentas da Qualidade

Confiabilidade/Sobrevivência

Análise preditiva

Multivariada

Séries Temporais

Tabelas

Não-Paramétricos

Testes de Equivalência

Poder e Tamanho de Amostra

Gráfico de Linha Ajustada...

Regressão

Regressão Não Linear...

Estudo de Estabilidade

Regressão Ortogonal...

Mínimos Quadrados Parciais...

Gráfico de Linha Ajustada Binária...

Regressão Logística Binária

Regressão Logística Ordinal...

Regressão Logística Nominal...

Regressão de Poisson

Ajuste de Modelo de Regressão...

Melhores Subconjuntos...

**Ajustar de Modelo de Regressão**

Modelar a relação entre predores categóricos ou contínuos e uma resposta. Incluir facilmente termos polinomiais e de interação, ou transformar a resposta se necessário.

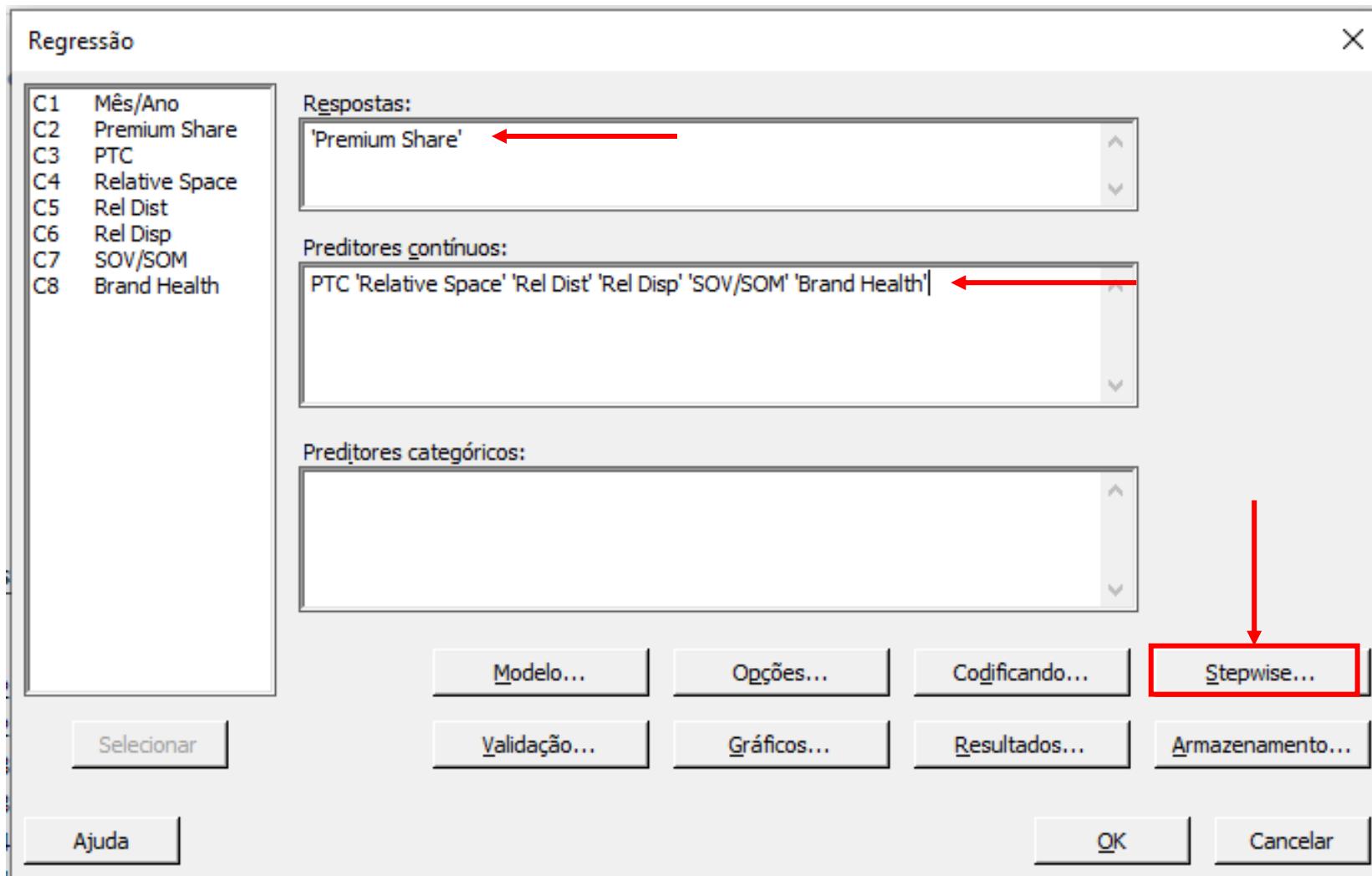
Gráfico de Contorno Sobreposto...

Otimizador de Resposta...



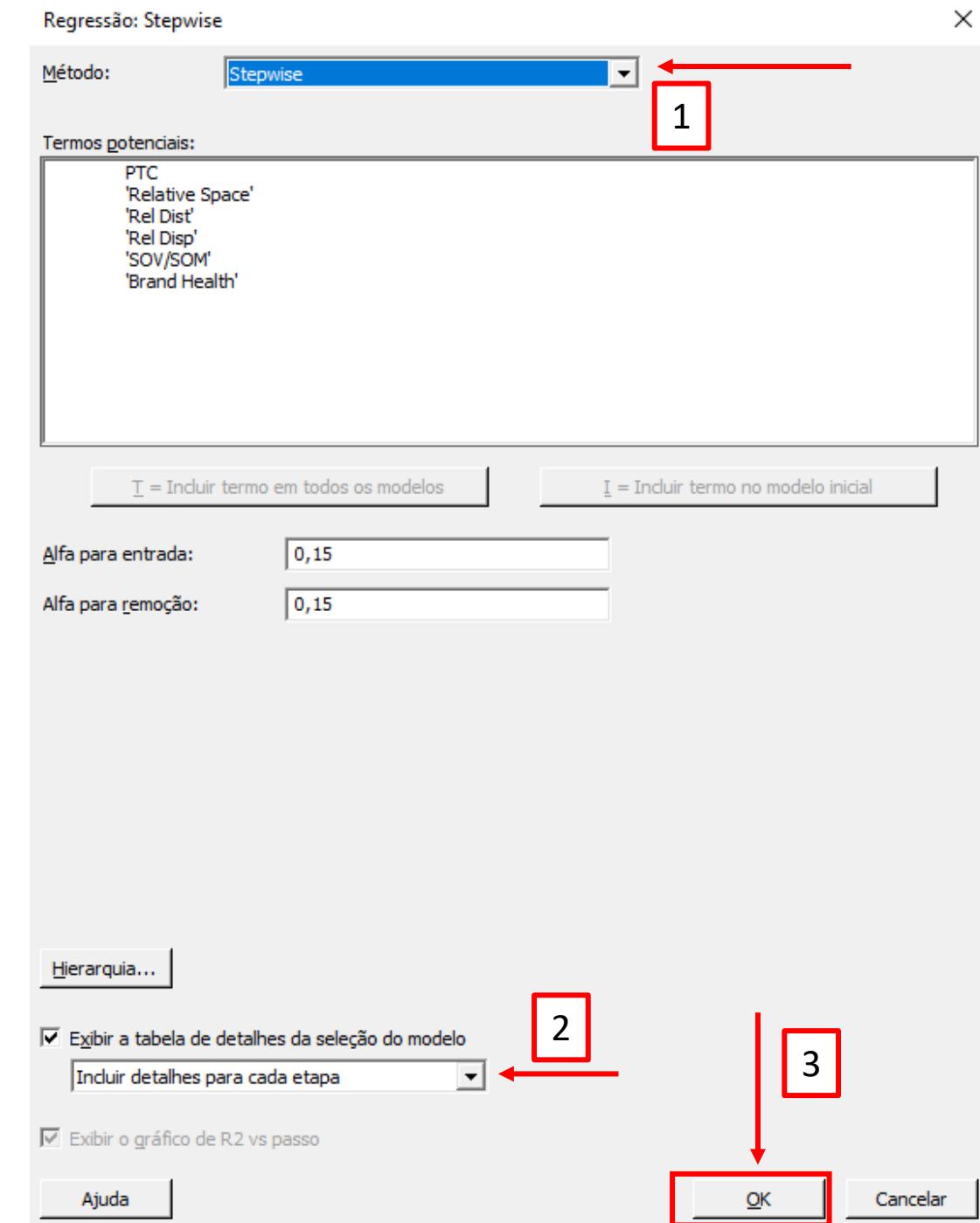
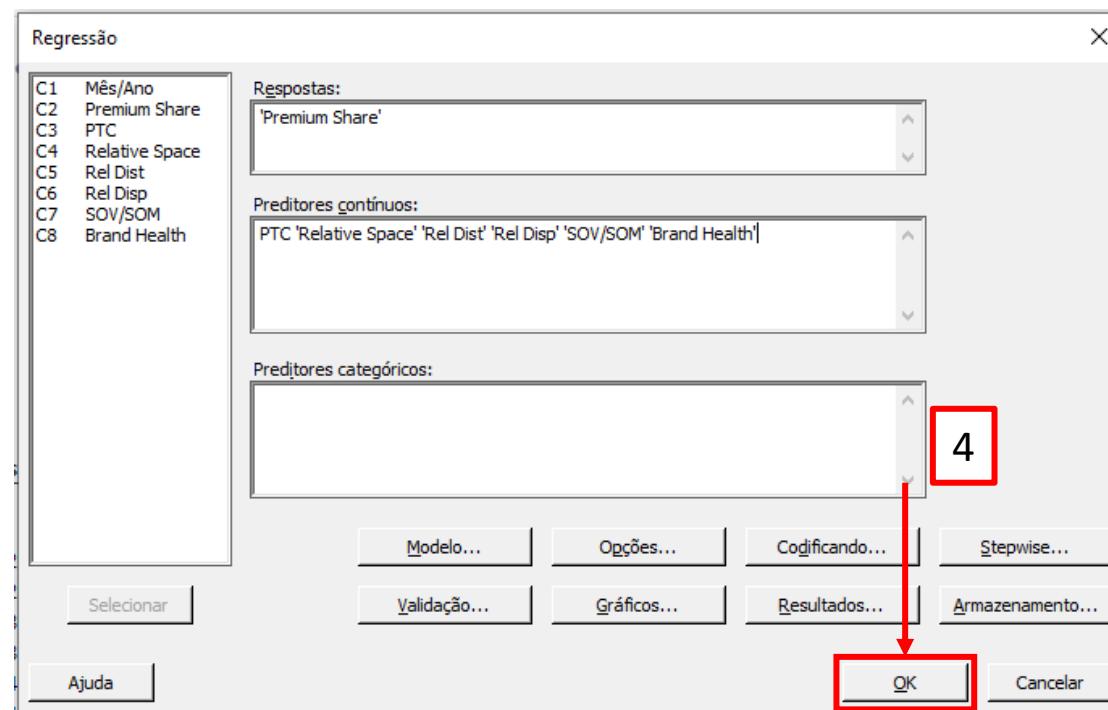
# Regressão Múltipla – com Stepwise

BEER Data1.xlsx



# Regressão Múltipla – com Stepwise

BEER Data1 .xlsx





# Regressão Múltipla – com Stepwise

BEER Data1 .xlsx

## Seleção Stepwise de Termos

Termos candidatos: PTC; Relative Space; Rel Dist; Rel Disp; SOV/SOM; Brand Health

	Passo 1		Passo 2		Passo 3	
	Coef	P	Coef	P	Coef	P
Constante	142,6		122,2		117,63	
PTC	-4,880	0,000	-3,972	0,000	-4,043	0,000
Rel Disp			0,02668	0,002	0,02986	0,000
Brand Health					0,1354	0,013
S	0,555717		0,449887		0,393334	
R2	75,67%		84,78%		88,92%	
R2(aj)	74,56%		83,33%		87,26%	
Cp de Mallows	19,20		6,53		1,86	
AICc	45,02		36,67		32,28	
BIC	47,36		39,28		34,83	

*a para entrada = 0,15; a para remoção = 0,15*

## Critérios de escolha dos modelos:

### 1º) P-Values

### 2º) R<sup>2</sup> e R<sup>2</sup> Ajustado

Nr variáveis é igual a 1 – usar maior R<sup>2</sup>.

Nr variáveis é diferente de 1 – usar maior R<sup>2</sup> Ajustado.

### 3º) Mallow's Cp

Preferir modelos com Cp mais próximo do nr de variáveis e mais o intercepto, ou seja, nr de variáveis preditoras + 1.

### 4º) S

Buscar o menor S possível.

### 5º) AICc e BIC

Os menores valores indicam melhores modelos

**Os critérios devem ser usados na ordem acima.**



# Regressão Múltipla – com Stepwise

BEER Data1.xlsx

## Equação de Regressão

Premium Share = 117,63 - 4,043 PTC + 0,02986 Rel Disp + 0,1354 Brand Health

Modelo/Equação que representa o Y em função dos Xs.

## Coefficientes

Termo	Coef	EP de Coef	Valor-T	Valor-P	VIF
Constante	117,63	9,54	12,33	0,000	
PTC	-4,043	0,475	-8,52	0,000	1,29
Rel Disp	0,02986	0,00668	4,47	0,000	1,33
Brand Health	0,1354	0,0495	2,73	0,013	1,06

## Sumário do Modelo

S	R2	R2(aj)	R2(pred)
0,393334	88,92%	87,26%	84,51%

## Análise de Variância

Fonte	GL	SQ (Aj.)	QM (Aj.)	Valor F	Valor-P
Regressão	3	24,829	8,2763	53,50	0,000
PTC	1	11,233	11,2333	72,61	0,000
Rel Disp	1	3,090	3,0900	19,97	0,000
Brand Health	1	1,156	1,1561	7,47	0,013
Erro	20	3,094	0,1547		
Total	23	27,923			

## Ajustados e Diagnósticos para Observações Atípicas

Obs.	Premium Share	Ajuste	Resíd	Resíd Pad
24	48,734	49,689	-0,955	-2,59 R

R Resíduo grande

A equação foi igual à obtida anteriormente!

O método Stepwise selecionou o melhor conjunto de preditores e este foi igual ao conjunto selecionado no método anterior (Best Subset).

## Benefícios:

- Stepwise: toma em conta o P-value de cada preditora
- Best Subsets: flexibilidade para selecionar o conjunto que melhor se adeque à realidade

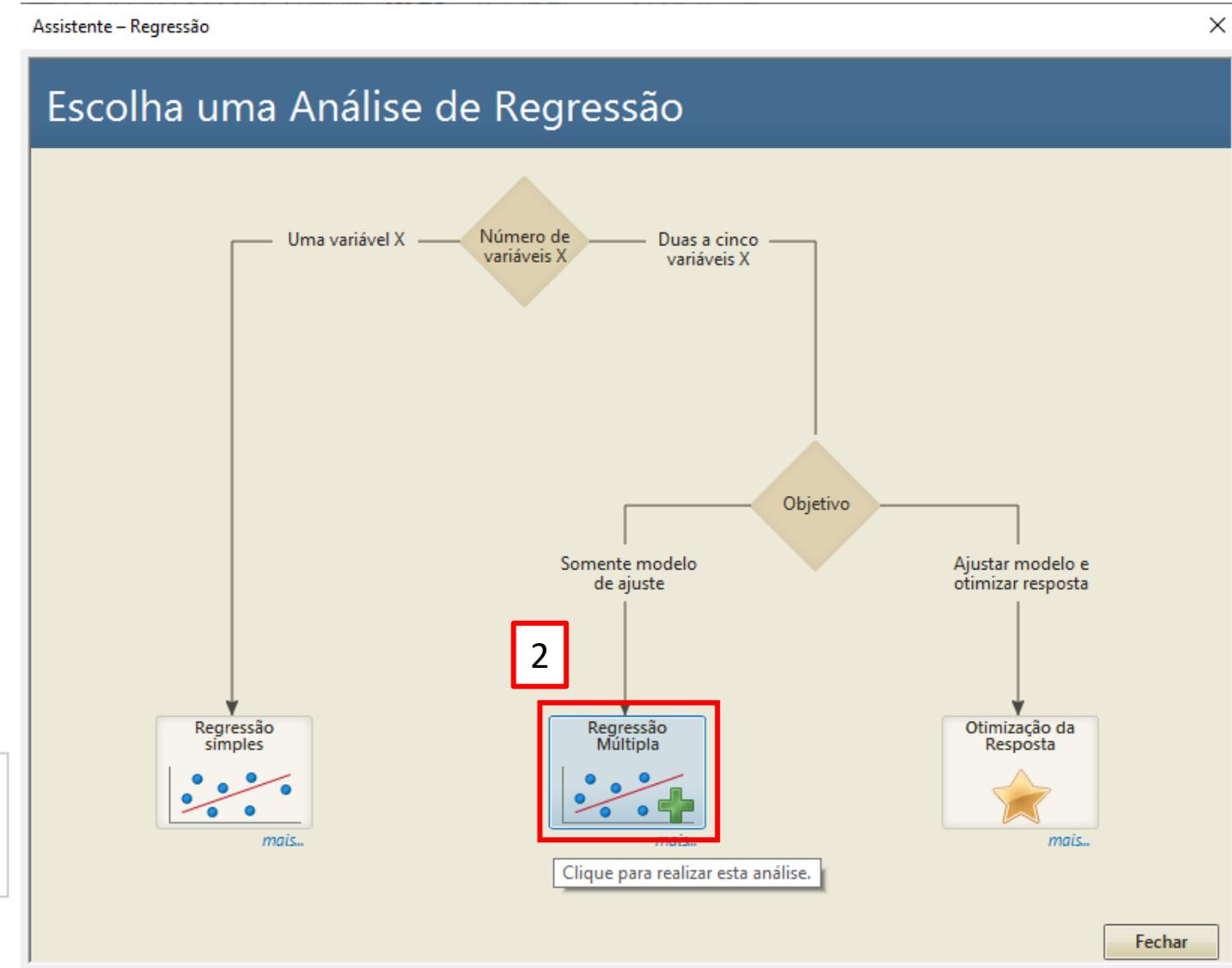
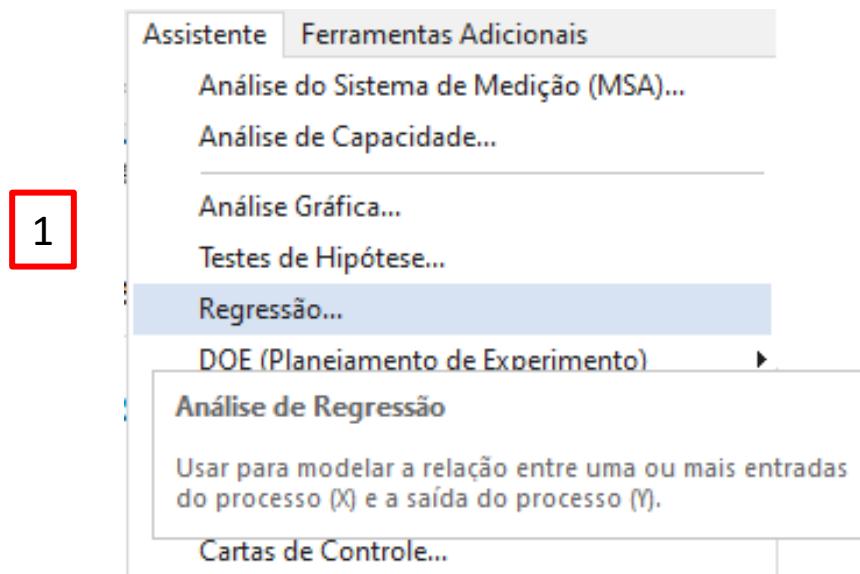
Parâmetros de Ajuste do Modelo ( $R^2$ ,  $R^2$  ajustado e  $R^2$  pred)

O  $R^2$  ajustado independe do número de fatores. O  $R^2$  pred serve para determinar o quanto bem o modelo prediz as respostas para novas observações.



# Regressão Múltipla – com uso do Assistente Exemplo 1

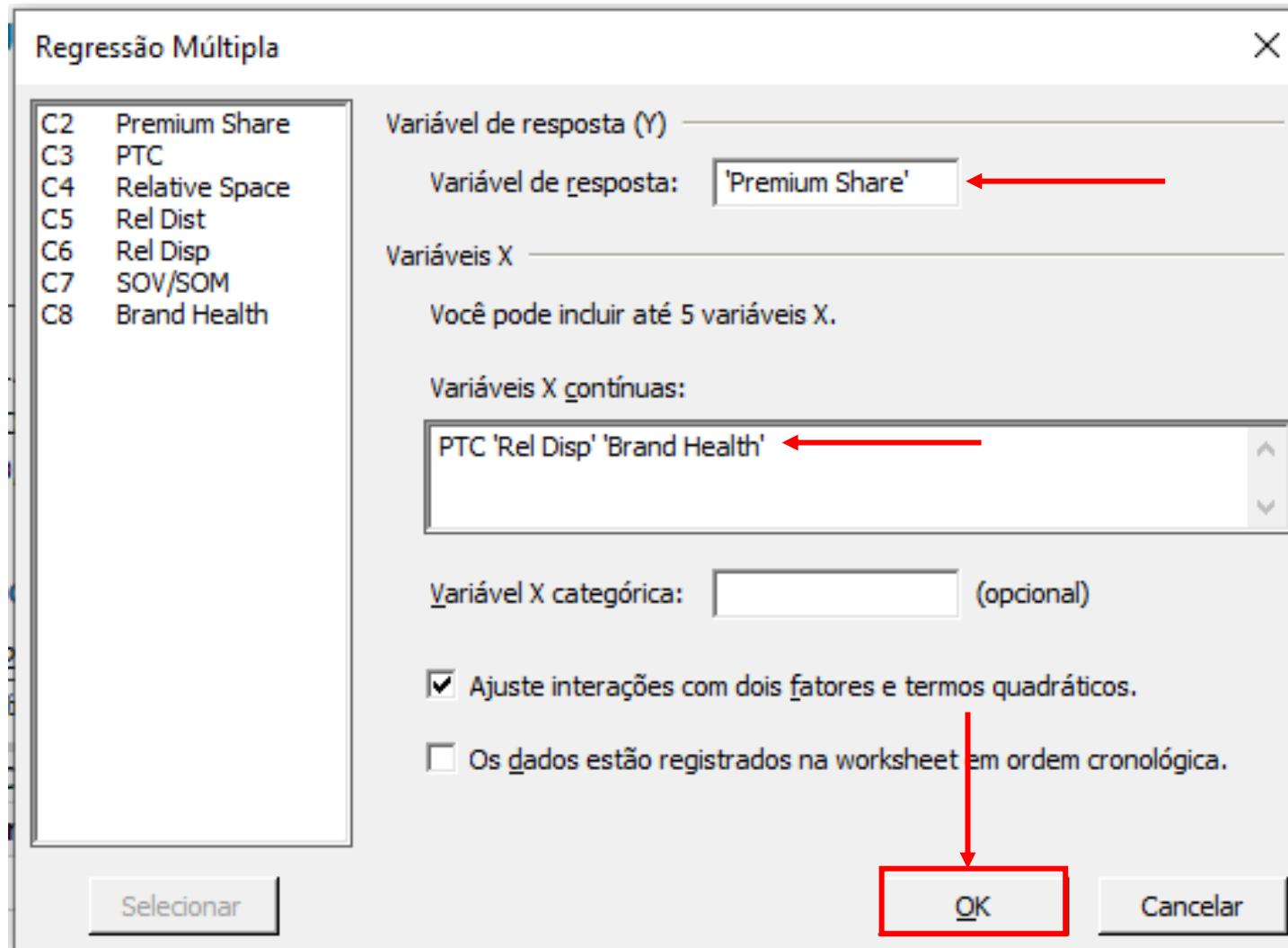
BEER Data1 .xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx

**Regressão Múltipla para Premium Shar**  
Relatório Resumo

Há uma relação entre Y e as variáveis X?

0	0,1	> 0,5
Sim		Não
P < 0,001		

A relação entre Y e as variáveis X no modelo é estatisticamente significativa ( $p < 0,10$ ).

Comentários

Os seguintes termos estão na equação ajustada que modela a relação entre Y e as variáveis X:

X1: PTC  
X2: Rel Disp  
X3: Brand Health

% de variação explicado pelo modelo

0%	100%
Inferior	Superior
R-quadrado = 88,92%	

88,92% da variação em Y pode ser explicado pelo modelo de regressão.

**Premium Shar versus Variáveis X**

PTC

Rel Disp

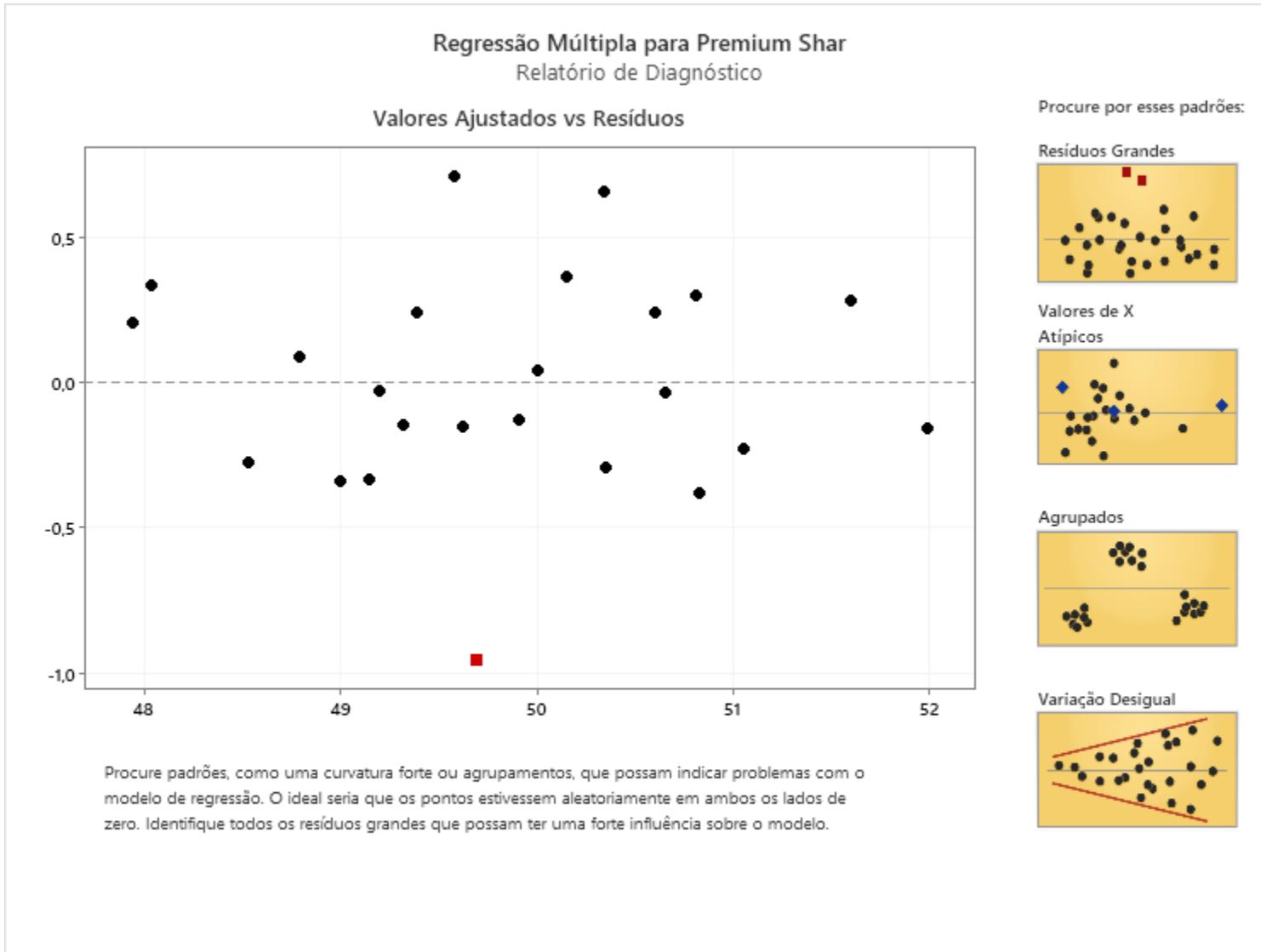
Brand Health

Um gráfico de fundo cinza representa uma variável X que não está no modelo.



# Regressão Múltipla – com uso do Assistente Exemplo

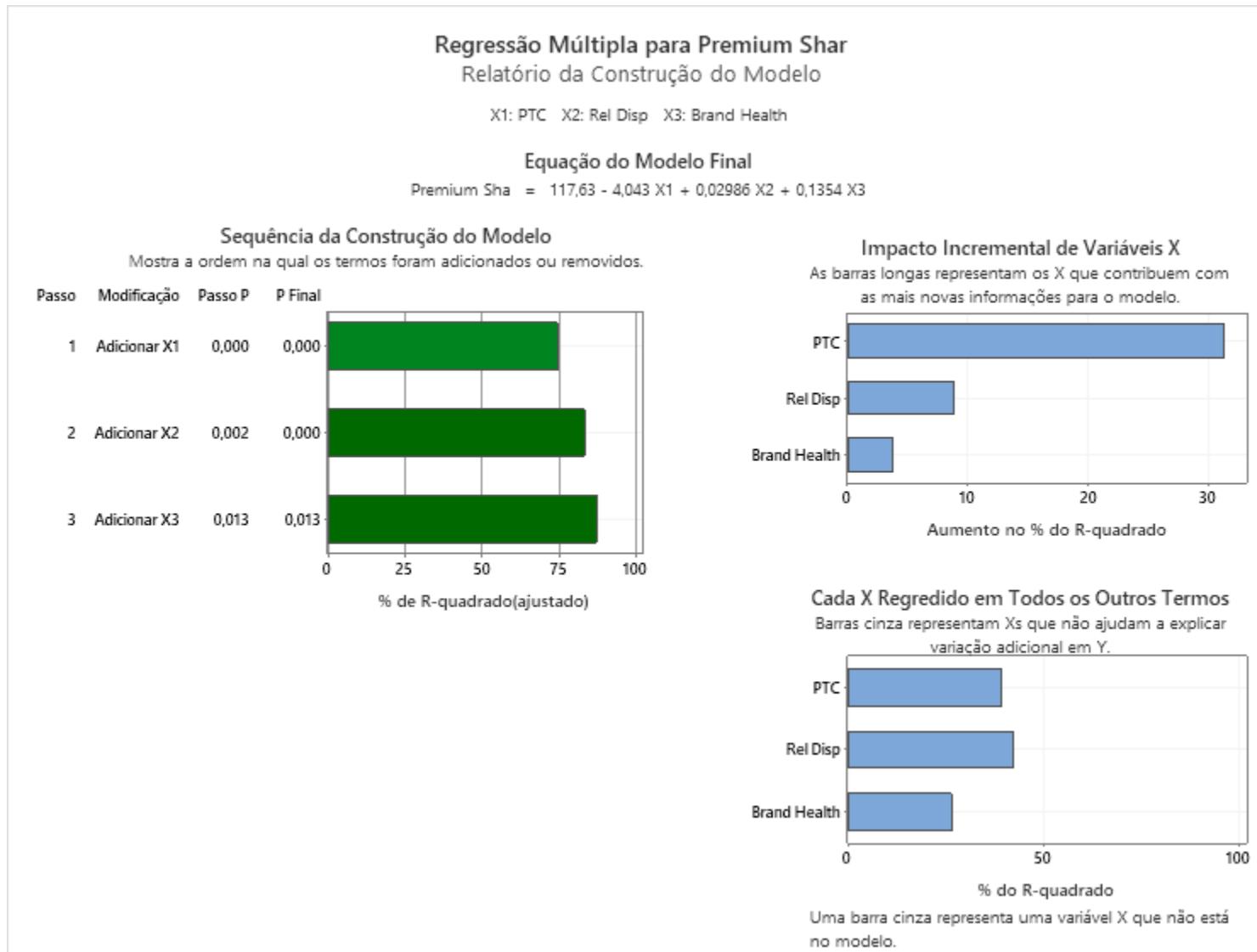
BEER Data1.xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1 .xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data.xlsx

Regressão Múltipla para Premium Shar Cartão de Relatório		
Verificar	Status	Descrição
Quantidade de Dados		O tamanho amostral ( $n = 24$ ) não é grande o bastante para fornecer uma estimativa muito precisa da força da relação. As medidas da força da relação, como R-Quadrado e R-Quadrado (ajustado), podem variar muito. Para obter uma estimativa precisa, amostras maiores (tipicamente 40 ou mais) devem ser usadas para um modelo desse tamanho.
Dados Atípicos		Um ponto de dados tem um resíduo grande e não está bem ajustado pela equação. Esse ponto está marcado em vermelho no Relatório de Diagnóstico e está na linha 24 da worksheet. Como dados atípicos podem exercer uma forte influência sobre os resultados, Tente identificar a causa de sua natureza atípica. Corija quaisquer erros de entrada de dados ou medições. Considere remover os dados que estão associados com causas especiais e refazer a análise.
Normalidade		Como você tem no mínimo 15 pontos de dados, a normalidade não é um problema. Se o número de pontos de dados for pequeno e os resíduos não forem normalmente distribuídos, os valores-p usados para determinar se há uma relação significativa entre X e Y podem não ser precisos.



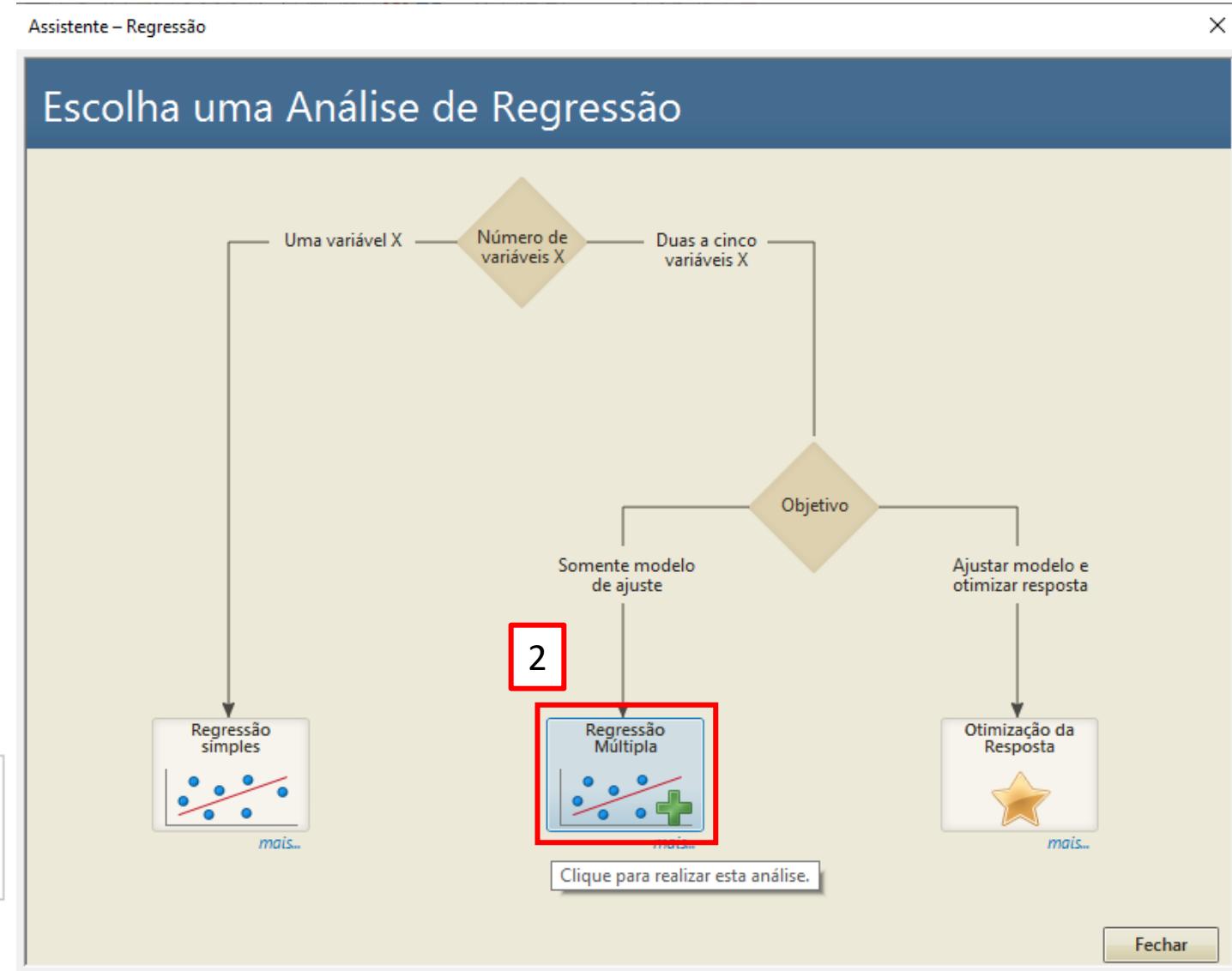
# Regressão Múltipla – com uso do Assistente Exemplo 2

BEER Data .xlsx

1

Análise do Sistema de Medição (MSA)...  
Análise de Capacidade...  
Análise Gráfica...  
Testes de Hipótese...  
Regressão...  
DOE (Planejamento de Experimento) ▶  
Análise de Regressão  
Usar para modelar a relação entre uma ou mais entradas do processo (X) e a saída do processo (Y).  
Cartas de Controle...

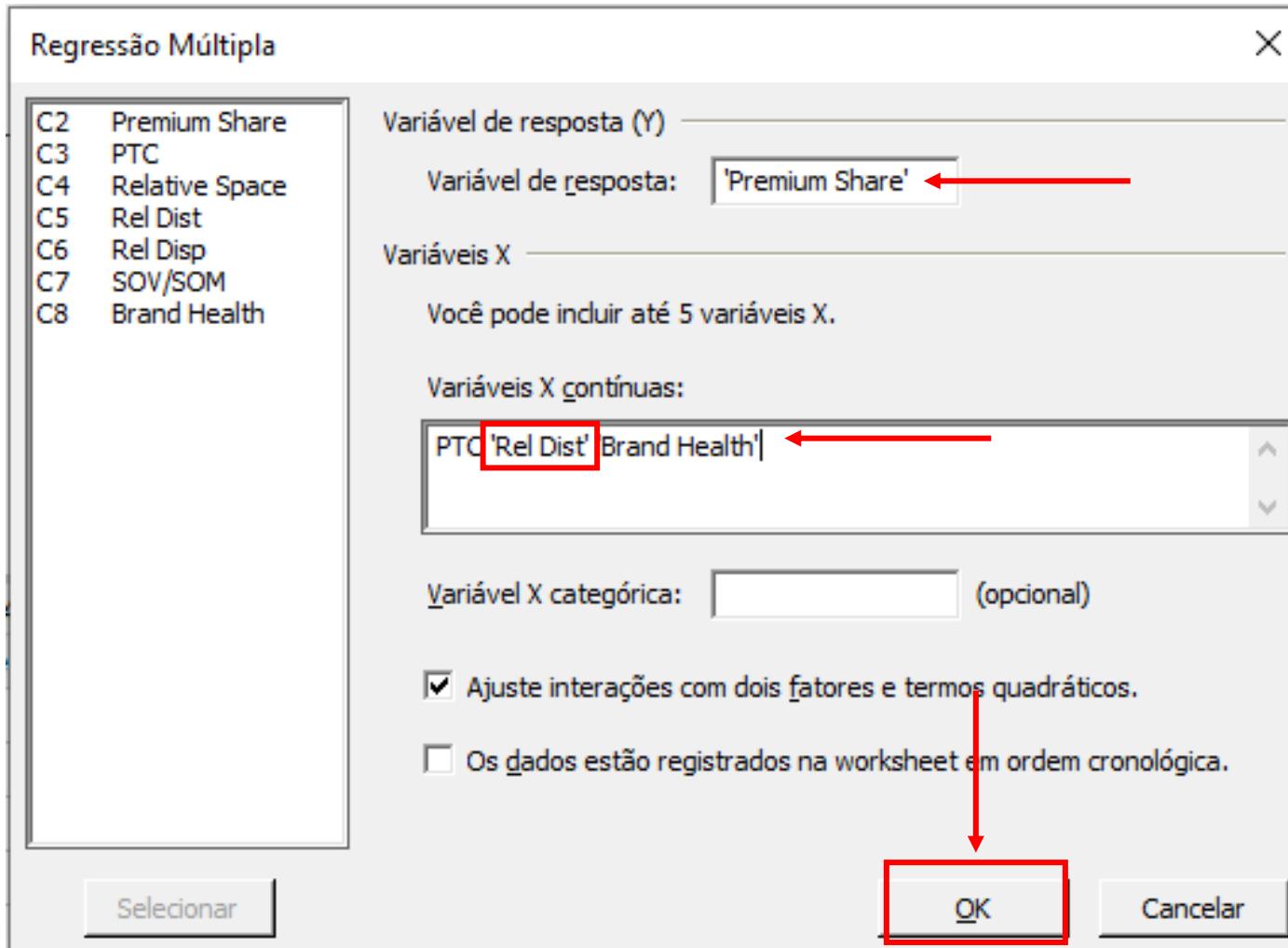
## Vamos utilizar outras variáveis preditoras!





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx



Alteramos as  
variáveis preditoras!



# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx

**Regressão Múltipla para Premium Shar**  
Relatório Resumo

Há uma relação entre Y e as variáveis X?

0	0,1	> 0,5
Sim		Não
P < 0,001		

A relação entre Y e as variáveis X no modelo é estatisticamente significativa ( $p < 0,10$ ).

Comentários

Os seguintes termos estão na equação ajustada que modela a relação entre Y e as variáveis X:

X1: PTC  
X2: Rel Dist  
X3: Brand Health  
X1\*X2

Se o modelo se ajusta bem aos dados, essa equação pode ser usada para prever Premium Shar para valores específicos das variáveis X, ou para encontrar as configurações para as variáveis X que correspondem a um valor desejado ou a uma amplitude de valores para Premium Shar.

% de variação explicado pelo modelo

0%	100%
Inferior	Superior
R-quadrado = 87,68%	

87,68% da variação em Y pode ser explicado pelo modelo de regressão.

**Premium Shar versus Variáveis X**

PTC

Rel Dist

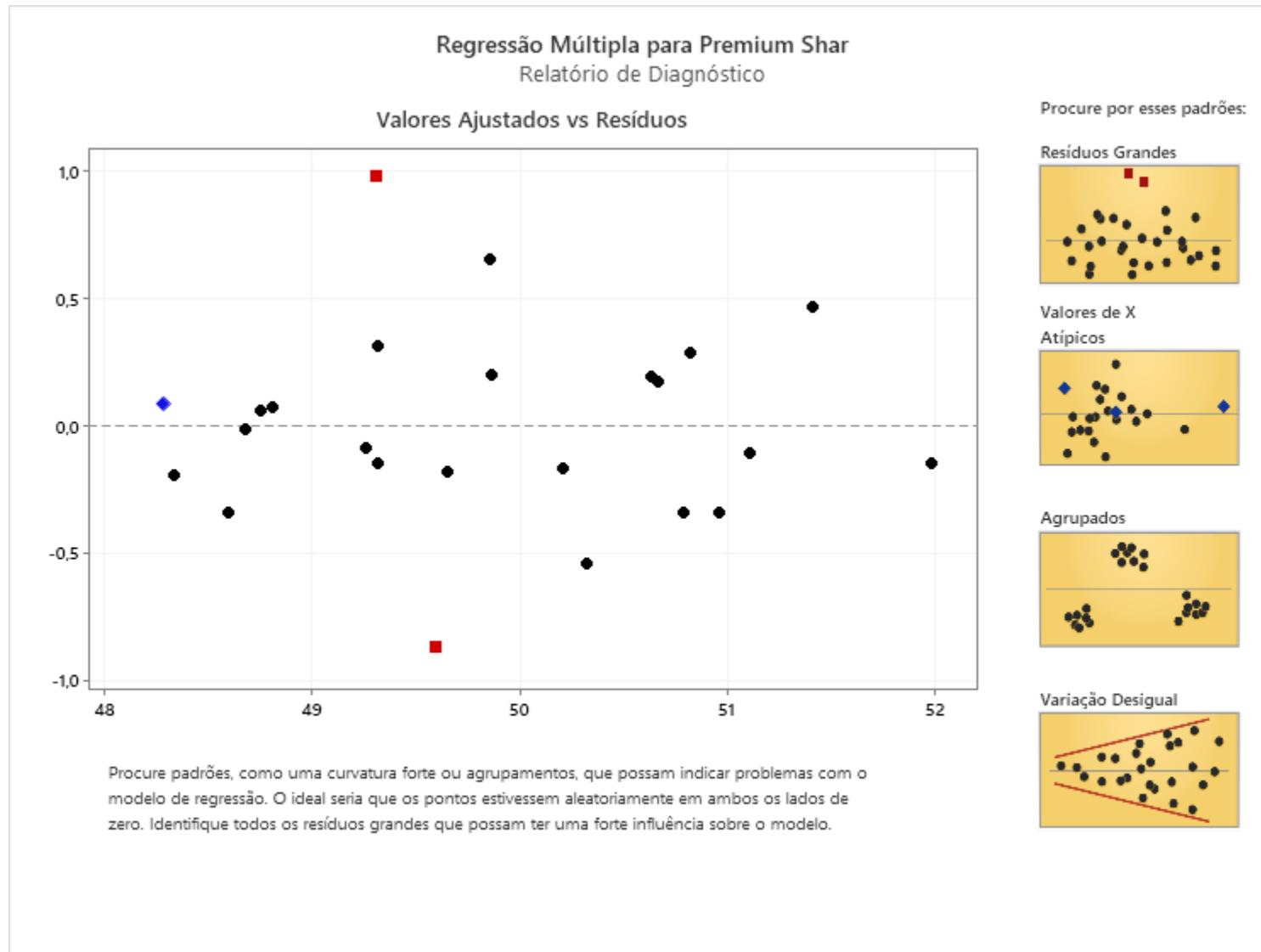
Brand Health

Um gráfico de fundo cinza representa uma variável X que não está no modelo.



# Regressão Múltipla – com uso do Assistente Exemplo

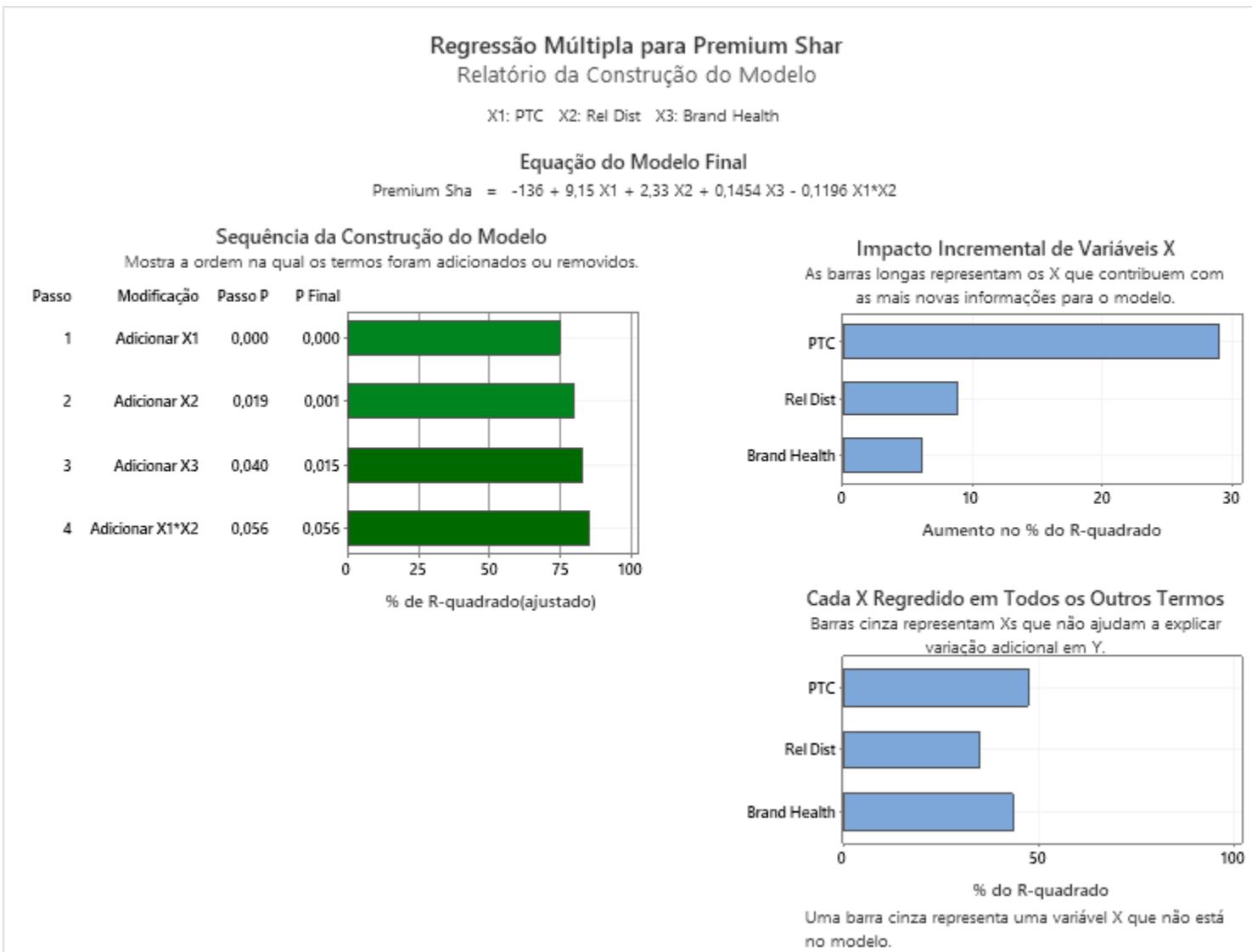
BEER Data1.xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

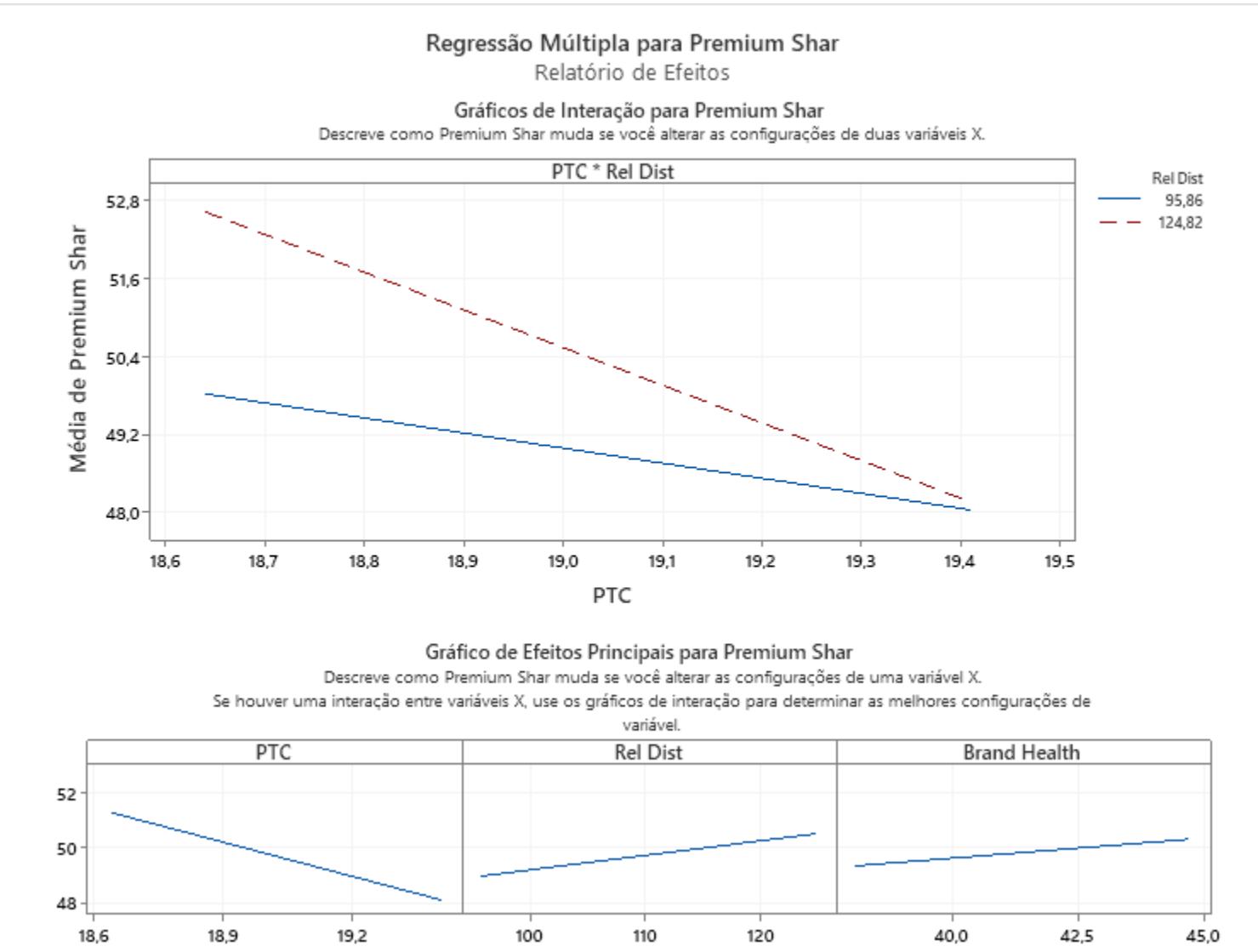
BEER Data1.xlsx





# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx





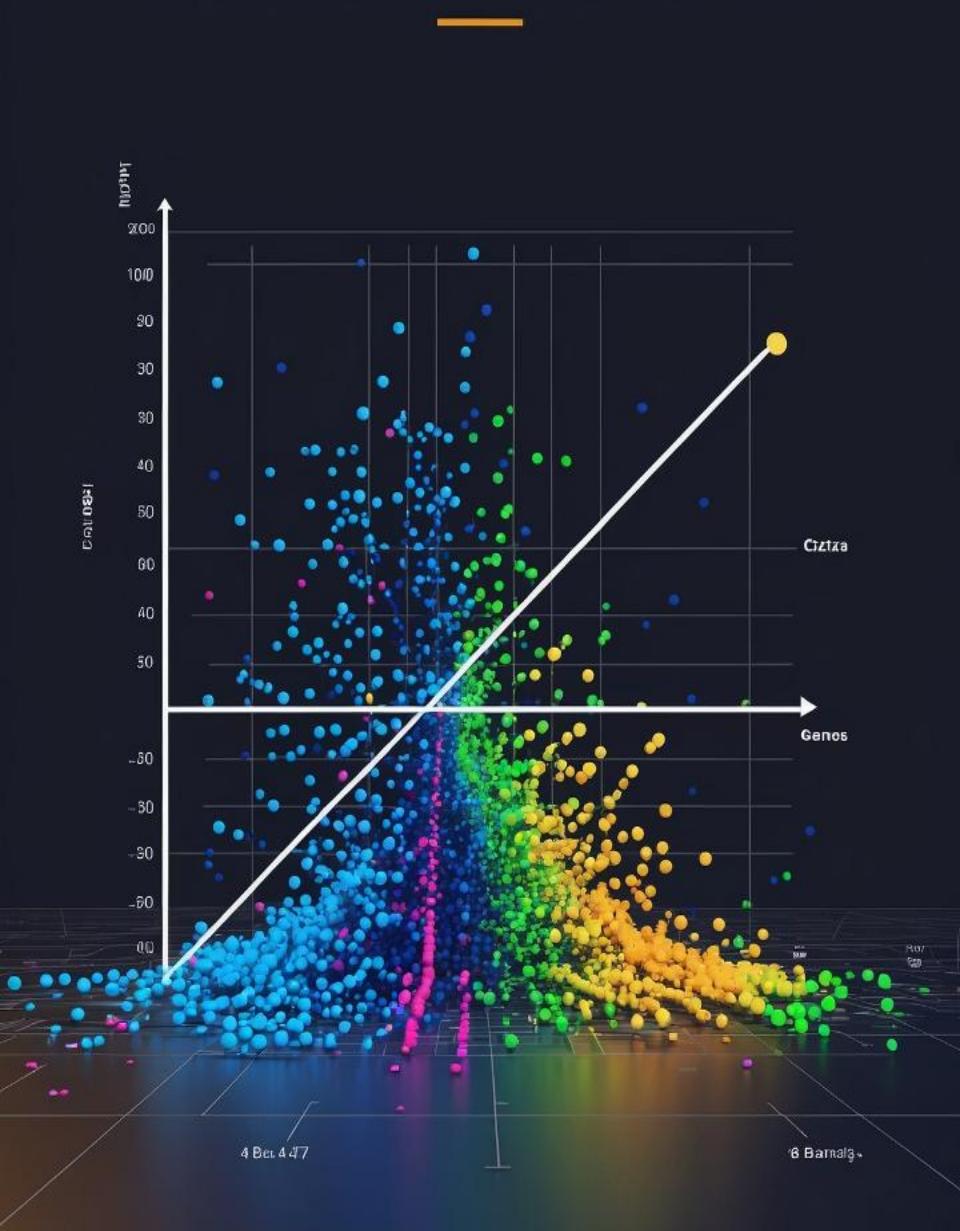
# Regressão Múltipla – com uso do Assistente Exemplo

BEER Data1.xlsx

Regressão Múltipla para Premium Shar Cartão de Relatório		
Verificar	Status	Descrição
Quantidade de Dados		O tamanho amostral ( $n = 24$ ) não é grande o bastante para fornecer uma estimativa muito precisa da força da relação. As medidas da força da relação, como R-Quadrado e R-Quadrado (ajustado), podem variar muito. Para obter uma estimativa precisa, amostras maiores (tipicamente 45 ou mais) devem ser usadas para um modelo desse tamanho.
Dados Atípicos		<ul style="list-style-type: none"><li>Resíduo grande: 2 pontos de dados têm grandes resíduos e não estão bem ajustados pela equação. Esses pontos estão marcados em vermelho no Relatório de Diagnóstico.</li><li>Valores atípicos de X: um ponto de dados tem um valor de X atípico, o que pode influenciar fortemente a equação do modelo. Esse ponto está marcado em azul no Relatório de Diagnóstico e está na linha 17 da worksheet. Você pode passar o cursor sobre um ponto ou usar o recurso da Função Brush do Minitab para identificar as linhas da worksheet. Como dados atípicos podem ter uma forte influência nos resultados, tente identificar a causa de sua natureza atípica. Corrija quaisquer erros de entrada de dados ou medições. Considere remover os dados que estão associados a causas especiais e refazer a análise.</li></ul>
Normalidade		Como você tem no mínimo 15 pontos de dados, a normalidade não é um problema. Se o número de pontos de dados for pequeno e os resíduos não forem normalmente distribuídos, os valores-p usados para determinar se há uma relação significativa entre X e Y podem não ser precisos.

**O Assistente:**

- **fornece modelos com interações de dois fatores e termos quadráticos e quando isso ocorre temos também gráficos de interação**
- **Não indica: colinearidade, P-Value dos fatores individuais,  $R^2$  pred e  $R^2$  ajustado total**



# Regressão Múltipla - Minitab

Bem-vindo à nossa apresentação sobre Regressão Múltipla utilizando o Minitab. Neste material, exploraremos conceitos fundamentais para compreender e interpretar os resultados de análises de regressão múltipla, com foco especial nos gráficos e métricas que o Minitab oferece para identificar multicolinearidade e avaliar o impacto de variáveis explicativas.

Abordaremos desde a interpretação básica até exemplos práticos que demonstram como utilizar essas ferramentas para tomar decisões mais fundamentadas em seus modelos estatísticos.

# Gráfico: "Impacto Incremental de Variáveis X" no Minitab

O Minitab oferece uma ferramenta poderosa para avaliar a contribuição de cada variável explicativa: o gráfico de "Impacto Incremental de Variáveis X".

Este gráfico é uma **análise incremental do R-quadrado ( $R^2$ )**. Ele mostra **quanto cada variável X contribui para o aumento do  $R^2$  do modelo, no momento em que ela entra na regressão**.

Essencialmente, responde à pergunta: "Ao adicionar essa variável ao modelo, quanto o  $R^2$  aumentou?"

Esta visualização permite identificar quais variáveis têm maior poder explicativo, auxiliando na seleção das mais relevantes para o modelo final.



# Como o gráfico é estruturado?

No eixo Y (vertical), são listadas as variáveis X do modelo. O eixo X (horizontal) mostra o **aumento percentual no R<sup>2</sup>** proporcionado por cada variável no momento da sua entrada. Cada linha representa uma variável X, e o valor mostrado é chamado de **Impacto Incremental no R<sup>2</sup>**.

Esta estrutura permite visualizar rapidamente quais variáveis trazem maior contribuição para o poder explicativo do modelo, facilitando decisões sobre quais variáveis manter ou remover.

## Variáveis X

Listadas no eixo vertical

## Aumento no R<sup>2</sup>

Mostrado no eixo horizontal

## Barras Horizontais

Representam o impacto de cada variável

## Comparação Visual

Facilita identificar as variáveis mais impactantes

# Como interpretar os valores

Um **impacto maior no R<sup>2</sup>** indica que essa variável **trouxe mais explicação** ao modelo. É, portanto, uma **variável mais relevante** para prever a variável resposta Y. Por outro lado, um **impacto pequeno ou nulo no R<sup>2</sup>** sugere que **essa variável contribui pouco ou quase nada** quando foi incluída, podendo ser irrelevante ou redundante com outras variáveis (possivelmente colinear).



## Impacto Alto (>10%)

Variável muito relevante, traz contribuição significativa para o modelo



## Impacto Médio (2-10%)

Variável moderadamente relevante, contribui para melhorar o modelo



## Impacto Baixo (<2%)

Variável pouco relevante, contribuição marginal ao poder explicativo



## Impacto Nulo ( $\approx 0\%$ )

Variável redundante ou irrelevante, candidata à remoção do modelo

# Por que a soma dos % de aumento do R-quadrado não dá 100%?

Porque **as contribuições das variáveis X não são independentes entre si**. Quando existe correlação entre as variáveis explicativas, suas contribuições para o  $R^2$  se sobrepõem, criando áreas de "explicação compartilhada" que não são atribuídas exclusivamente a nenhuma variável.

Considere um exemplo simples: suponha que  $X_1$  e  $X_2$  **ambas expliquem parte da mesma variabilidade** de  $Y$ . Se você **colocar  $X_1$  primeiro**, ela pode explicar, por exemplo, 20% do  $R^2$ . Depois, **ao adicionar  $X_2$** , ela só adiciona mais 5%, pois **parte da informação que ela explicaria já foi capturada por  $X_1$** .

Assim, o **incremento no  $R^2$  de cada variável é o valor marginal**, ou seja, o quanto ela ainda consegue explicar depois que as outras já estão no modelo.

# Exemplo numérico

Suponha que o  $R^2$  total do modelo seja 60%, indicando que 60% da variabilidade da variável resposta é explicada pelo conjunto de variáveis explicativas. No entanto, o gráfico de impacto incremental mostra:  $X_1 \rightarrow +20\%$ ,  $X_2 \rightarrow +15\%$ ,  $X_3 \rightarrow +10\%$ .

A soma desses valores é 45%, o que significa que os outros 15% estão "compartilhados" entre essas variáveis. Nenhuma variável sozinha levou esse pedaço, então ele não aparece no gráfico de impacto incremental.

## Decomposição do $R^2$ Total (60%)

- Contribuição única de  $X_1$ : 20%
- Contribuição única de  $X_2$ : 15%
- Contribuição única de  $X_3$ : 10%
- Contribuição compartilhada: 15%

A soma das contribuições únicas (45%) é menor que o  $R^2$  total (60%) devido à sobreposição de explicações entre as variáveis, que representa 15% da variabilidade explicada.

# Importante: A ordem das variáveis importa?

Sim! O impacto incremental depende da ordem de entrada das variáveis no modelo. Esta é uma consideração crucial ao interpretar o gráfico de "Impacto Incremental de Variáveis X" no Minitab.

O Minitab costuma utilizar diferentes abordagens para determinar esta ordem: **Stepwise** (Entrada por etapas), **Régressão Hierárquica** (você define a ordem), ou pode calcular o impacto mantendo todas variáveis no modelo, e mostrando a contribuição de cada uma mantendo as demais fixas.



# Conclusão simplificada

A soma dos aumentos de  $R^2$  no gráfico não dá 100% porque há sobreposição de explicações entre as variáveis. **O gráfico mostra apenas a contribuição exclusiva de cada variável, e não contabiliza a parte compartilhada com outras.**

Apesar desta limitação, o gráfico continua sendo uma ferramenta valiosa para identificar a importância relativa das variáveis explicativas e orientar decisões sobre quais variáveis manter ou remover do modelo.

## O que o gráfico mostra

Contribuição única e exclusiva de cada variável para o  $R^2$  do modelo, mantendo as outras variáveis constantes

## O que o gráfico não mostra

Contribuição compartilhada entre variáveis correlacionadas, que também faz parte do  $R^2$  total

## Implicação prática

A soma dos impactos incrementais será sempre menor ou igual ao  $R^2$  total do modelo, com a diferença representando a sobreposição de explicações

# "Cada X regredido em todos os outros termos"

O gráfico "Each X regressed on all the other Xs" é uma ferramenta poderosa do Minitab para identificar multicolinearidade. Este recurso visual permite analisar de forma sistemática as relações entre as variáveis explicativas do seu modelo de regressão múltipla.

Através deste gráfico, é possível visualizar rapidamente quais variáveis podem ser explicadas por outras, indicando potenciais redundâncias no modelo que poderiam comprometer a interpretação dos coeficientes e a estabilidade das estimativas.

Esta análise é fundamental para garantir que seu modelo de regressão múltipla seja robusto e que as interpretações derivadas dele sejam confiáveis e precisas.

# Gráfico: “Cada X regredido em todos os outros Termos”



## Regressão Individual

Cada variável X é tratada como dependente em uma regressão separada



## Análise de Correlação

Identifica o quanto cada X está correlacionado com os demais X's



## Valor de R<sup>2</sup>

Mostra a porcentagem da variabilidade de cada X explicada pelos outros X's



## Alerta de Multicolinearidade

Valores altos de R<sup>2</sup> indicam potencial problema no modelo

### R<sup>2</sup> Alto (>80%)

Forte multicolinearidade - a variável é quase totalmente explicada pelas outras

### R<sup>2</sup> Médio (40-80%)

Multicolinearidade moderada - a variável compartilha informação com outras

### R<sup>2</sup> Baixo (<40%)

Baixa multicolinearidade - a variável traz informação única ao modelo

## Decisão

Avaliar se é necessário remover ou combinar variáveis com alta multicolinearidade

# Exemplo prático

Imagine que você está analisando fatores que influenciam o salário de profissionais e tem as seguintes variáveis explicativas: X1: Idade, X2: Experiência, X3: Tempo de Empresa.

Se o gráfico indicar que o  $R^2$  de X2 (Experiência) regredido sobre X1 e X3 é 95%, isso significa que a variável Experiência pode ser quase totalmente explicada pelas outras variáveis. Este resultado sugere forte multicolinearidade.

Diante dessa situação, você deve considerar algumas alternativas: eliminar uma das variáveis redundantes, combinar variáveis (por exemplo, via Análise de Componentes Principais), ou apenas monitorar os efeitos se o modelo ainda for estável.



Análise de Dados

Identificação da multicolinearidade entre variáveis através do gráfico do Minitab



Visualização do  $R^2$

Gráfico mostrando que Experiência tem  $R^2$  de 95% quando regredida sobre Idade e Tempo de Empresa



Tomada de Decisão

Avaliação das alternativas para lidar com a multicolinearidade identificada

# Business Case Mercado – PDV Álcool Gel – Exercício

Os dados estão no arquivo Coolgel Banco de dados PDV.xlsx  
nas abas “Market Share Decline”

Usem os dados para efetuar análises uma Regressão  
Múltipla

- Tragam conclusões sobre Market Share e essas variáveis
- Indiquem potenciais melhorias
- Indique a estratégia de ação futura

# Régressão Logística

**Discreto**

**Contínuo**

**X**

**Discreto**

Matriz de  
Seleção de  
Ferramentas

- Diagrama de Dispersão
- Regressão Linear Simples

- Box Plot
- Teste de Hipóteses:  
Testes t, ANOVA,  
Mood's, Variância

- Regressão Logística

- Pareto
- Teste de Proporção
- Chi Quadrado

# Modelos de Regressão para dados discretos

Modelo	Resposta	Exemplo
Regressão Logística Binária	Duas categorias distintas	Compra / Não Compra; Passa / Falha
Regressão Logística Nominal	Três ou mais categorias não ordenadas	Vermelho / Amarelo / Verde; Doce / Amargo / Salgado
Regressão Logística Ordinal	Três ou mais categorias ordenadas	P / M / G / XG; 6-pack / 12-pack / 24-pack
Regressão de Poisson	Contagem de ocorrência	Defeitos em um pallet; Reclamações de clientes

# Regressão Logística Binária

## **Semelhanças com as demais Regressões:**

- Determina a relação entre uma variável de resposta e uma ou mais preditoras
- Variáveis preditoras podem ser contínuas ou discretas

## **Diferenças:**

- A regressão linear usa variáveis de resposta contínua
- A regressão logística binária usa variáveis de resposta discretas

# Regressão Logística Binária

**Variável resposta binária:**

- Se um evento ocorre, então a probabilidade é  $p$
- Se um evento não ocorre, então a probabilidade é  $(1-p)$

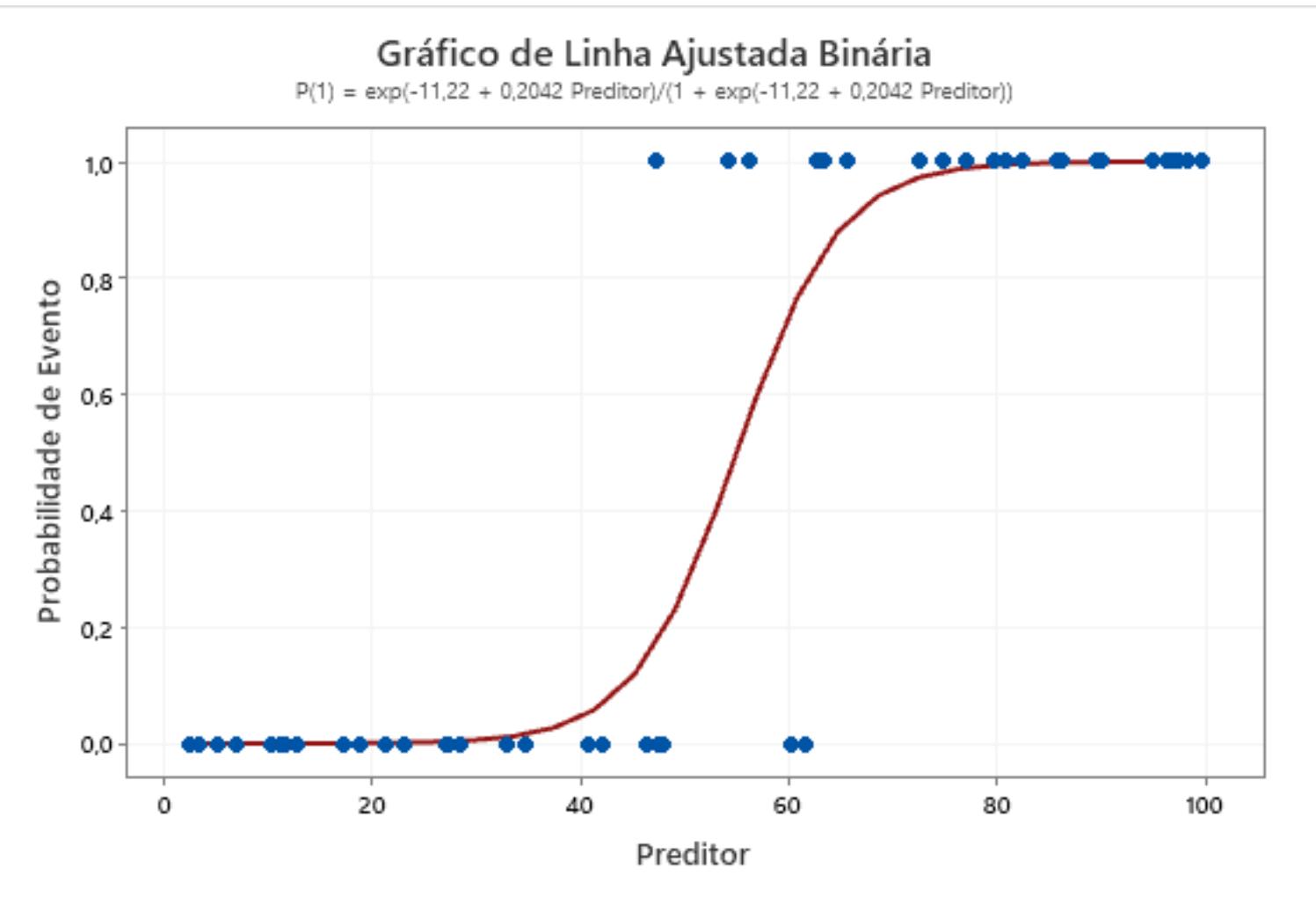
**A probabilidade de sucesso em relação ao fracasso é a razão de  $p / (1-p)$**

# Regressão Logística Binária

- Uma função de ligação transforma as probabilidades dos níveis de uma variável de resposta categórica em uma escala contínua
- O modelo de regressão logística é ajustado ao logaritmo natural das probabilidades
- O modelo estatístico para regressão logística é:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

# Regressão Logística Binária





# Regressão Logística - Exemplo

Foi efetuado uma amostragem em uma rede de supermercados para identificar se o valor do ticket de compra de uma pessoa, que comprou cerveja, impacta na compra de cerveja premium.

Arquivo: RL Ex1 Ex2.xlsx na aba Exemplo 1

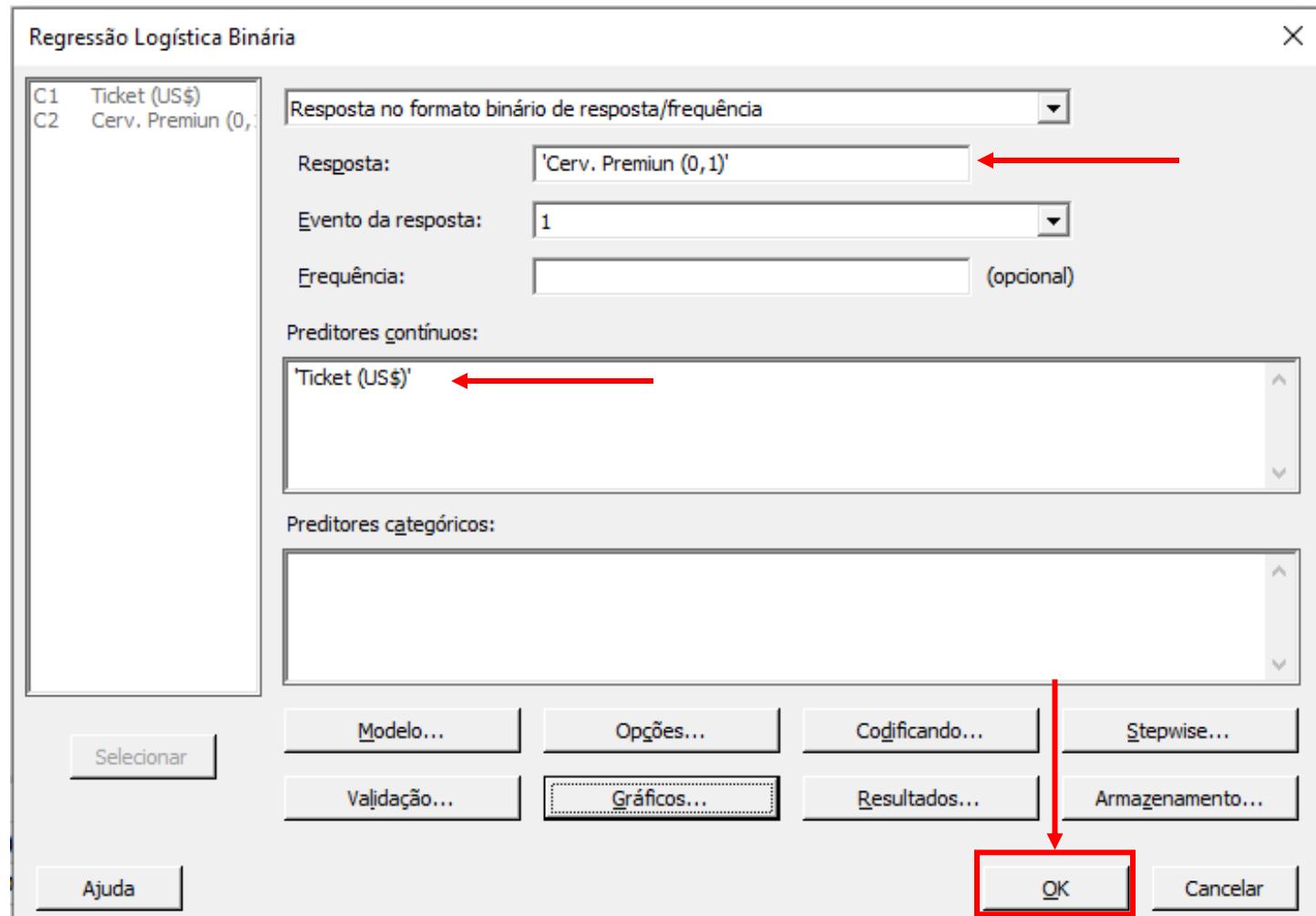
The screenshot shows the Minitab software interface with the 'Estat' menu open. The 'Regressão' option is selected. A sub-menu for 'Regressão Logística Binária' is displayed, with the first item, 'Ajuste de Modelo Logístico Binário...', highlighted. The sub-menu also includes options for 'Ajustar Regressão Logística Binária' and 'Realizar regressão logística em respostas binárias.'

- Estat
- Gráfico
- Visualizar
- Ajuda
- Assistente
- Ferramentas Adicionais
- Estatísticas Básicas
- Regressão
- ANOVA
- DOE (Planejamento de Experimento)
- Cartas de Controle
- Ferramentas da Qualidade
- Confiabilidade/Sobrevivência
- Análise preditiva
- Multivariada
- Séries Temporais
- Tabelas
- Não-Paramétricos
- Testes de Equivalência
- Poder e Tamanho de Amostra
- Gráfico de Linha Ajustada...
- Regressão
- Regressão Não Linear...
- Estudo de Estabilidade
- Regressão Ortogonal...
- Mínimos Quadrados Parciais...
- Gráfico de Linha Ajustada Binária...
- Regressão Logística Binária
- Regressão Logística Ordinal...
- Regressão Logística Nominal...
- Regressão de Poisson

# Regressão Logística – Exemplo 1



Foi efetuado uma amostragem em uma rede de supermercados para identificar se o valor do ticket de compra de uma pessoa, que comprou cerveja, impacta na compra de cerveja premium.  
Arquivo: RL Ex1 Ex2.xlsx na aba Exemplo 1





# Regressão Logística – Exemplo 1

## Método

Função de Ligação: Logito  
Linhas usadas: 50

## Sumário do Modelo

R2 Deviance	R2 (Ajustado)	Deviance	AIC	AICc	BIC	Área sob a curva ROC
42,51%	41,05%	43,43	43,69	47,26		0,8864

Varia de 0,5 a 1.  
1 indica que o modelo binário separa bem as classes

## Informações da Resposta

Variável	Valor	Contagem
Cerv. Premiun (0,1)	1	28 (Evento)
	0	22
Total		50

Para Eventos/Ensaios

Para eventos elevados

Para Resposta/Frequência

## Equação de Regressão

$$P(1) = \exp(Y') / (1 + \exp(Y'))$$

$$Y' = -9,36 + 0,347 \text{ Ticket (US$)}$$

## Coeficientes

Termo	Coef	EP de Coef	Valor-Z	Valor-P	VIF
Constante	-9,36	3,08	-3,04	0,002	
Ticket (US\$)	0,347	0,111	3,14	0,002	1,00

## Razões de Chances para Preditores Contínuos

	Razão de Chances	IC de 95%
Ticket (US\$)	1,4144	(1,1388; 1,7568)

Indica a probabilidade de um evento ocorrer, quanto maior mais provável de ocorrer o evento com o aumento da preditora. Probabilidade =  $e^{\beta}$

## Testes de Qualidade de Ajuste

Teste	GL	Qui-Quadrado	Valor-P
Deviance	48	39,43	0,806
Pearson	48	35,35	0,912
Hosmer-Lemeshow	8	2,77	0,948

> 0,05  
Indica se os dados seguem uma distribuição Binomial

## Análise de Variância

Fonte	GL	Teste de Wald	
		Qui-Quadrado	Valor-p
Regressão	1	9,83	0,002
Ticket (US\$)	1	9,83	0,002

## Ajustados e Diagnósticos para Observações Atípicas

Obs.	Probabilidade Observada	Ajuste			Resid Pad
		Resid	Resid Pad	R Resíduo grande	
6	0,0000	0,8666	-2,0073	-2,06 R	

Avaliar e atuar se necessário



# Regressão Logística – Exemplo 1

Estat Gráfico Visualizar Ajuda Assistente Ferramentas Adicionais

- Estatísticas Básicas
- Regressão**
- ANOVA
- DOE (Planejamento de Experimento)
- Cartas de Controle
- Ferramentas da Qualidade
- Confiabilidade/Sobrevivência
- Análise preditiva
- Multivariada
- Séries Temporais
- Tabelas
- Não-Paramétricos
- Testes de Equivalência
- Poder e Tamanho de Amostra

- Gráfico de Linha Ajustada...
- Regressão
- Regressão Não Linear...
- Estudo de Estabilidade
- Regressão Ortogonal...
- Mínimos Quadrados Parciais...
- Gráfico de Linha Ajustada Binária...**
- Gráfico de Linha Ajustada
- Régressao de Poisson

Gráfico de Linha Ajustada Binária

C1 Ticket (US\$)  
C2 Cerv. Premium (0, 1)

Resposta no formato binário de resposta/frequência

Resposta: 'Cerv. Premium (0, 1)' ←

Evento da resposta: 1

Frequência: (opcional)

Predictor: 'Ticket (US\$)' ←

Opcões... Gráficos...

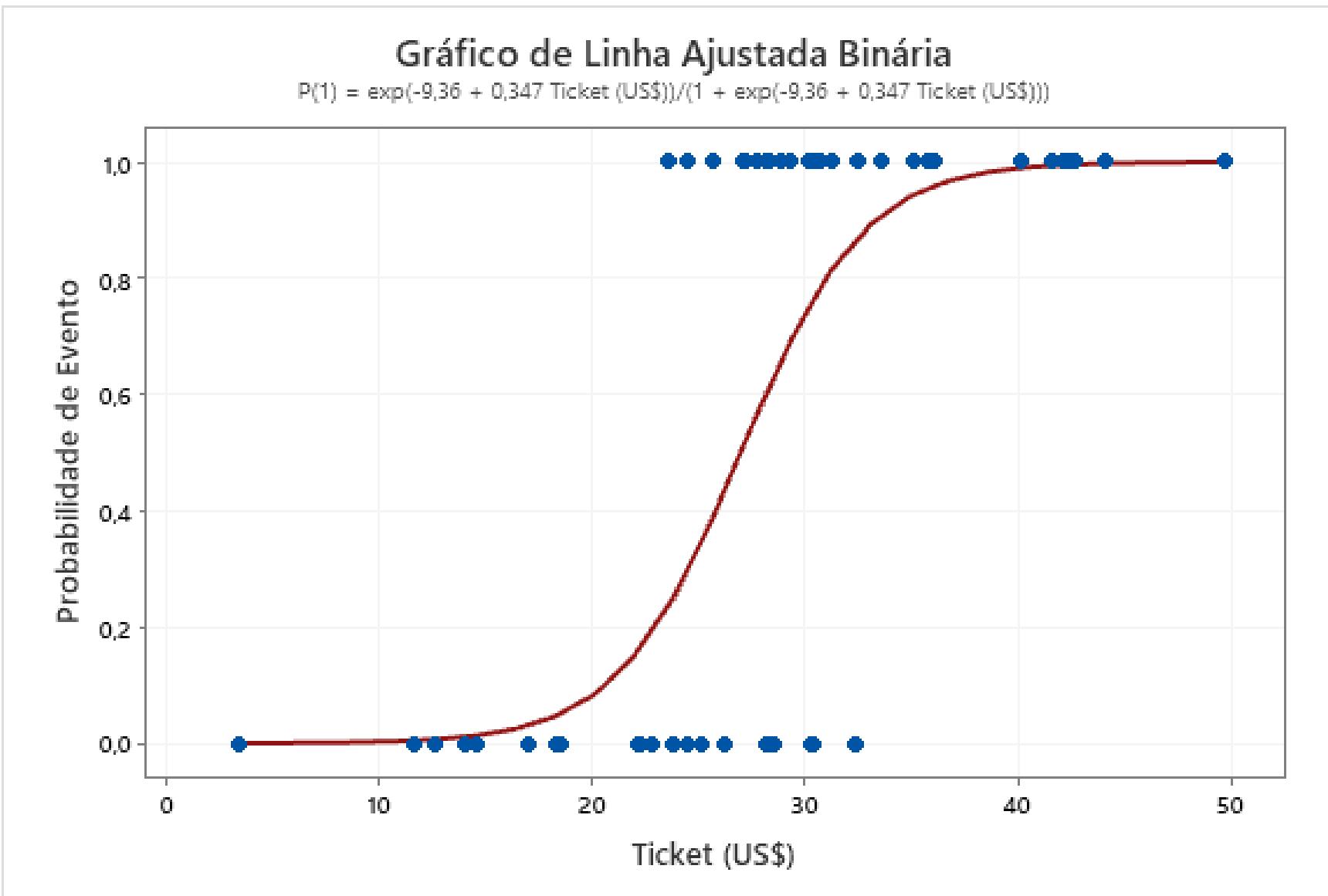
Selecionar Resultados... Armazenamento...

Ajuda OK Cancelar

Red arrows point to the Predictor field ('Ticket (US\$)') and the Response field ('Cerv. Premium (0, 1)'). A red box highlights the 'OK' button.



# Regressão Logística – Exemplo 1





# Regressão Logística – Exemplo 2

Foi efetuado uma amostragem em uma rede de supermercados para identificar se, para uma pessoa que comprou cerveja, o valor do ticket de compra, o tempo dela dentro do mercado e a existência de ponto extra impactam na compra de cerveja premium.

Arquivo: RL Ex1 Ex2.xlsx na aba Exemplo 2

The screenshot shows the Minitab software interface with the 'Estat' menu open. The 'Regressão' option is selected. A sub-menu for 'Regressão Logística Binária' is displayed, with the 'Ajuste de Modelo Logístico Binário...' option highlighted. The sub-menu also includes other options like 'Regressão Logística Ordinal...', 'Regressão Logística Nominal...', and 'Regressão de Poisson'.



# Regressão Logística – Exemplo 2

Foi efetuado uma amostragem em uma rede de supermercados para identificar se, para uma pessoa que comprou cerveja, o valor do ticket de compra, o tempo dela dentro do mercado e a existência de ponto extra impactam na compra de cerveja premium.

Arquivo: RL Ex1 Ex2.xlsx na aba Exemplo 2

Regressão Logística Binária

C1 Tempo no mercado  
C2 Ticket médio (US\$)  
C3 Ponto Extra  
C4 Compra (0, 1)

Resposta no formato binário de resposta/frequência

Resposta: 'Compra (0, 1)' ←

Evento da resposta: 1

Frequência: (opcional)

Preditores contínuos:

'Tempo no mercado (min)' 'Ticket médio (US\$)' ←

Preditores categóricos:

'Ponto Extra'

Selecionar

Modelo... Opções... Codificando... Stepwise...  
Validação... Gráficos... Resultados... Armazenamento...  
Ajuda OK Cancelar

The dialog box shows settings for a binary logistic regression analysis. The response variable is 'Compra (0, 1)'. The continuous predictors are 'Tempo no mercado (min)' and 'Ticket médio (US\$)'. The categorical predictor is 'Ponto Extra'. The 'Resultados...' button in the bottom toolbar is highlighted with a red box, indicating it should be selected before running the analysis.



# Regressão Logística – Exemplo 2

## Informações da Resposta

Variável	Valor	Contagem
Compra (0, 1)	1	23 (Evento)
	0	27
Total		50

## Equação de Regressão

$$P(1) = \exp(Y')/(1 + \exp(Y'))$$

### Ponto Extra

0       $Y' = -17,42 + 0,1310 \text{ Tempo no mercado (min)} + 0,2552 \text{ Ticket médio (US$)}$

1       $Y' = -12,75 + 0,1310 \text{ Tempo no mercado (min)} + 0,2552 \text{ Ticket médio (US$)}$

## Coeficientes

Termo	Coef	EP de Coef	Valor-Z	Valor-P	VIF
Constante	-17,42	6,60	-2,64	0,008	
Tempo no mercado (min)	0,131	0,120	1,09	0,275	1,20
Ticket médio (US\$)	0,2552	0,0874	2,92	0,003	1,80
Ponto Extra					
1	4,67	1,70	2,74	0,006	1,99

## Razões de Chances para Preditores Contínuos

	Razão de Chances	IC de 95%
Tempo no mercado (min)	1,1399 (0,9009; 1,4423)	
Ticket médio (US\$)	1,2908 (1,0875; 1,5320)	

## Razões de Chances para Preditores Categóricos

Nível A	Nível B	Razão de Chances	IC de 95%
Ponto Extra	1	0	106,3824 (3,7943; 2982,6648)

Razão de chances para o nível A em relação ao nível B

## Ajustados e Diagnósticos para Observações Atípicas

Obs.	Probabilidade Observada	Ajustado			Resíduo Pad
		Ajuste	Resíduo	X Atípicos X	
13	0,000	0,800	-1,795		-2,02 R
17	0,000	0,923	-2,266		-2,38 R
31	1,000	0,311	1,528		1,82 X
42	1,000	0,847	0,576		0,66 X

R Resíduo grande  
X Atípicos X

## Sumário do Modelo

R2 Deviance	R2 (Aj.)	Deviance	AIC	AICc	BIC	Área sob a curva ROC
73,18%	68,83%	68,83%	26,51	27,40	34,15	0,9758

## Testes de Qualidade de Ajuste

Teste	GL	Qui-Quadrado	Valor-P
Deviance	46	18,51	1,000
Pearson	46	23,10	0,998
Hosmer-Lemeshow	8	5,41	0,713

> 0,05

## Análise de Variância

Fonte	GL	Qui-Quadrado	Teste de Wald	Valor-p
Regressão	3	9,67		0,022
Tempo no mercado (min)	1	1,19		0,275
Ticket médio (US\$)	1	8,53		0,003
Ponto Extra	1	7,53		0,006

< 0,05

Remover Tempo no mercado do Modelo!



# Regressão Logística – Exemplo 2

Regressão Logística Binária

X

C1 Tempo no mercado  
C2 Ticket médio (US\$)  
C3 Ponto Extra  
C4 Compra (0, 1)

Resposta no formato binário de resposta/frequência

Resposta: 'Compra (0, 1)' ←

Evento da resposta: 1

Frequência: (opcional)

Preditores contínuos:

'Ticket médio (US\$)' ←

Preditores categóricos:

'Ponto Extra'

Selecionar

Modelo... Validação... Ajuda

Opções... Gráficos...

Codificando... Resultados...

Stepwise... Armazenamento...

OK Cancelar



# Regressão Logística – Exemplo 2

## Equação de Regressão

$$P(1) = \exp(Y)/(1 + \exp(Y))$$

## Ponto Extra

0  $Y' = -13,14 + 0,2519$  Ticket médio (US\$)

1  $Y' = -8,764 + 0,2519$  Ticket médio (US\$)

## Coeficientes

Termo	Coef	EP de Coef	Valor-Z	Valor-P	VIF
Constante	-13,14	4,44	-2,96	0,003	
Ticket médio (US\$)	0,2519	0,0866	2,91	0,004	1,97
Ponto Extra					
1	4,37	1,62	2,69	0,007	1,97

## Razões de Chances para Preditores Contínuos

	Razão de Chances	IC de 95%
Ticket médio (US\$)	1,2865 (1,0857; 1,5245)	

## Razões de Chances para Preditores Categóricos

Nível A	Nível B	Razão de Chances	IC de 95%
Ponto Extra	1 0	79,2701 (3,2877; 1911,2706)	

Razão de chances para o nível A em relação ao nível B

## Sumário do Modelo

R2 Deviance	R2 (Aj.) Deviance	AIC	AICc	BIC	Área sob a curva ROC
71,20%	68,30%	25,87	26,39	31,61	0,9710

## Testes de Qualidade de Ajuste

Teste	GL	Qui-Quadrado	Valor-P
Deviance	47	19,87	1,000
Pearson	47	32,78	0,943
Hosmer-Lemeshow	8	5,48	0,705

> 0,05

## Análise de Variância

Fonte	GL	Qui-Quadrado	Valor-p
Regressão	2	9,30	0,010
Ticket médio (US\$)	1	8,46	0,004
Ponto Extra	1	7,25	0,007

< 0,05

## Ajustados e Diagnósticos para Observações Atípicas

Obs.	Probabilidade				
	Observada	Ajuste	Resíd	Resíd Pad	
13	0,000	0,731	-1,620	-1,81	X
17	0,000	0,958	-2,520	-2,59	R
31	1,000	0,502	1,174	1,31	X
37	1,000	0,789	0,689	0,77	X
50	1,000	0,527	1,131	1,27	X

R Resíduo grande  
X Atípicos X

Podemos verificar o impacto individual:  
próximos slides!



# Regressão Logística – Exemplo 2

Estat Gráfico Visualizar Ajuda Assistente Ferramentas Adicionais

- Estatísticas Básicas
- Regressão**
- ANOVA
- DOE (Planejamento de Experimento)
- Cartas de Controle
- Ferramentas da Qualidade
- Confiabilidade/Sobrevivência
- Análise preditiva
- Multivariada
- Séries Temporais
- Tabelas
- Não-Paramétricos
- Testes de Equivalência
- Poder e Tamanho de Amostra

- Gráfico de Linha Ajustada...
- Regressão
- Regressão Não Linear...
- Estudo de Estabilidade
- Regressão Ortogonal...
- Mínimos Quadrados Parciais...
- Gráfico de Linha Ajustada Binária...**
- Gráfico de Linha Ajustada
- Criar linhas de regressão logística binária ajustadas com intervalos de confiança.
- Régressao de Poisson

Gráfico de Linha Ajustada Binária

C1 Tempo no mercado  
C2 Ticket médio (US\$)  
C3 Ponto Extra  
C4 Compra (0, 1)

Resposta no formato binário de resposta/frequência

Resposta: 'Compra (0, 1)' ←

Evento da resposta: 1

Frequência: (opcional)

Predictor: 'Ticket médio (US\$)' ←

Opções... Gráficos...

Selecionar Resultados... Armazenamento...

Ajuda OK Cancelar

A red arrow points to the Predictor field containing 'Ticket médio (US\$)'. Another red arrow points to the Response field containing 'Compra (0, 1)'.



# Regressão Logística – Exemplo 2

## Equação de Regressão

$$P(1) = \exp(-7,49 + 0,1765 \text{ Ticket médio (US$)}) / (1 + \exp(-7,49 + 0,1765 \text{ Ticket médio (US$)}))$$

## Coeficientes

Termo	Coef	EP de Coef	VIF
Constante	-7,49	2,18	
Ticket médio (US\$)	0,1765	0,0509	1,00

## Razões de Chances para Preditores Contínuos

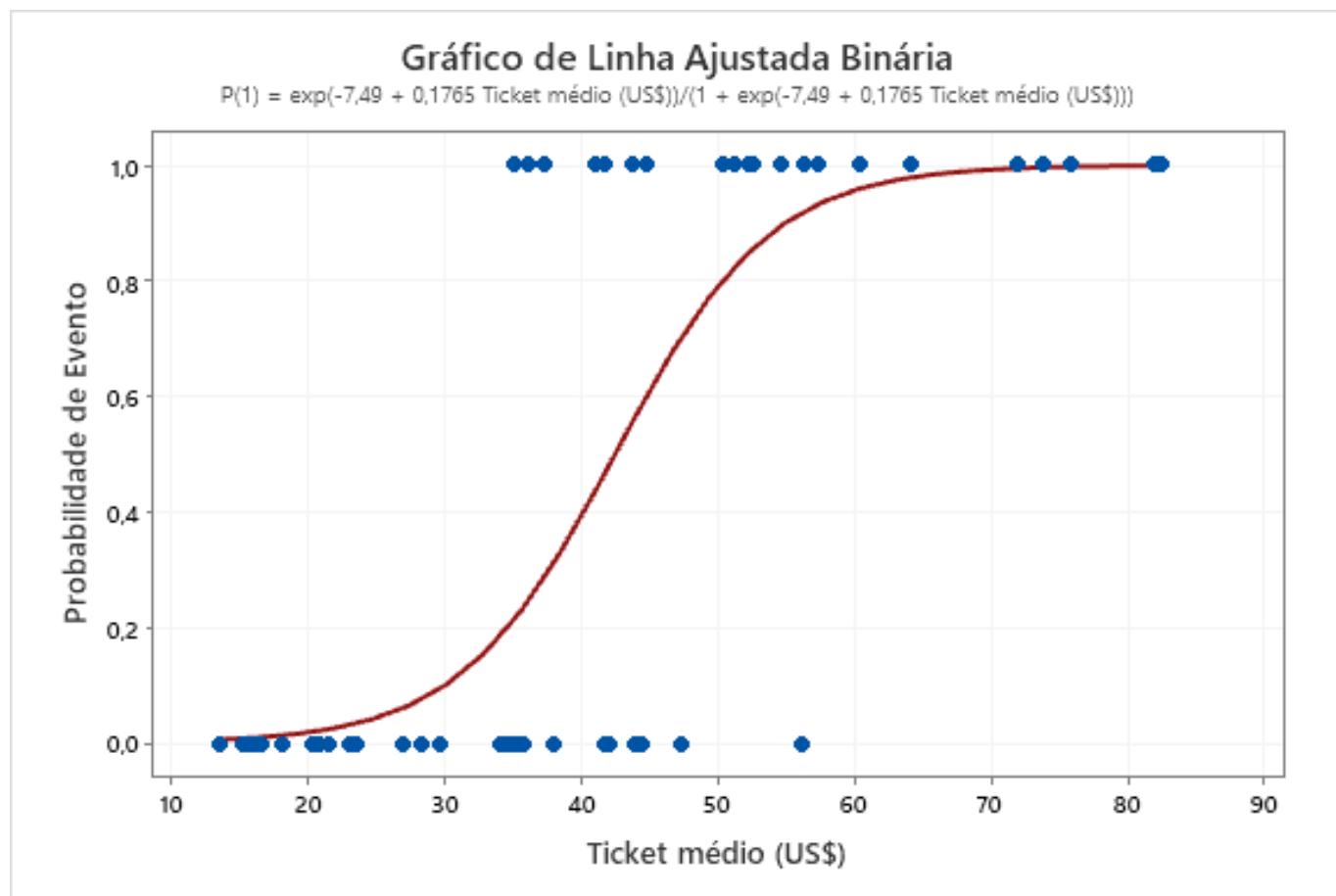
	Razão de Chances	IC de 95%
Ticket médio (US\$)	1,1930 (1,0797; 1,3183)	

## Sumário do Modelo

	R2 (Ajustado)			
R2 Deviance	Deviance	AIC	AICc	BIC
50,54%	49,09%	38,12	38,38	41,95

## Análise de Variância

Fonte	GL	Desv (Ajustado)	Média (Ajustada)	Qui-Quadrado	Valor-P
Regressão	1	34,87	34,8695	34,87	0,000
Ticket médio (US\$)	1	34,87	34,8695	34,87	0,000
Erro	48	34,12	0,7109		
Total	49	68,99			



## Ajustados e Diagnósticos para Observações Atípicas

Obs.	Probabilidade Observada	Probabilidade Ajustada		
		Resíd	Resíd Pad	
13	0,0000	0,9179	-2,2358	-2,30 R

R Resíduo grande



# Regressão Logística – Exemplo 2

Estat Gráfico Visualizar Ajuda Assistente Ferramentas Adicionais

- Estatísticas Básicas
- Regressão
- ANOVA
- DOE (Planejamento de Experimento)
- Cartas de Controle
- Ferramentas da Qualidade
- Confiabilidade/Sobrevivência
- Análise preditiva
- Multivariada
- Séries Temporais
- Tabelas
- Não-Paramétricos
- Testes de Equivalência
- Poder e Tamanho de Amostra

- Gráfico de Linha Ajustada...
- Regressão
- Regressão Não Linear...
- Estudo de Estabilidade
- Regressão Ortogonal...
- Mínimos Quadrados Parciais...
- Gráfico de Linha Ajustada Binária...
- Gráfico de Linha Ajustada
- Criar linhas de regressão logística binária ajustadas com intervalos de confiança.
- Regressão de Poisson

Gráfico de Linha Ajustada Binária

Resposta no formato binário de resposta/frequência

Resposta: 'Compra (0, 1)' ←

Evento da resposta: 1

Frequência: (opcional)

Predictor: 'Ponto Extra' ←

Selecionar

Opções... Gráficos...

Resultados... Armazenamento...

Ajuda

OK → Cancelar



# Regressão Logística – Exemplo 2

## Equação de Regressão

$$P(1) = \exp(-1,099 + 1,735 \text{ Ponto Extra}) / (1 + \exp(-1,099 + 1,735 \text{ Ponto Extra}))$$

## Coeficientes

Termo	Coef	EP de Coef	VIF
Constante	-1,099	0,471	
Ponto Extra	1,735	0,626	1,00

## Razões de Chances para Preditores Contínuos

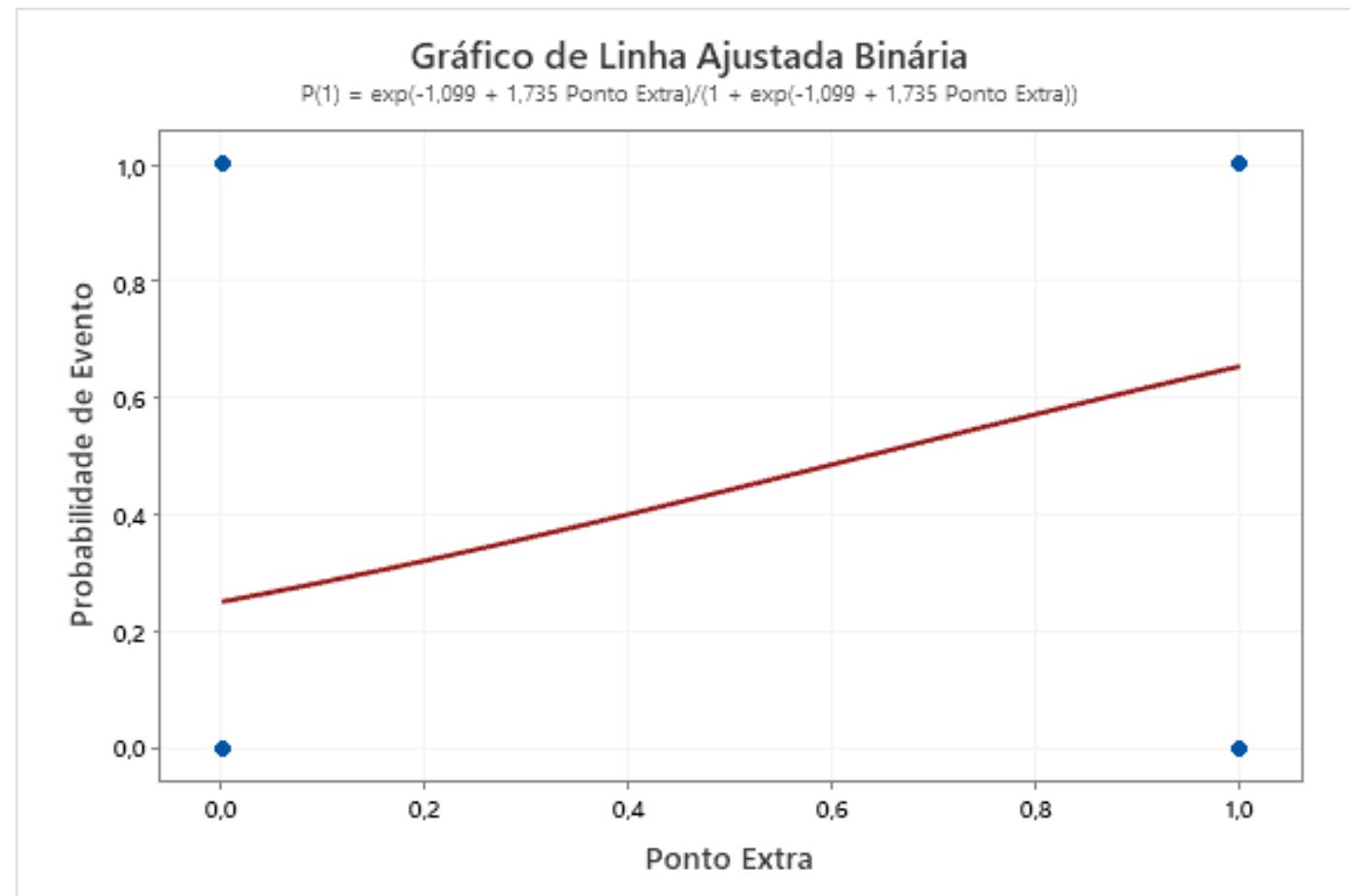
	Razão de Chances	IC de 95%
Ponto Extra	5,6667 (1,6607; 19,3362)	

## Sumário do Modelo

	R2 (Aj.)				
R2 Deviance	Deviance	AIC	AICc	BIC	
12,26%	10,81%	64,53	64,79	68,36	

## Análise de Variância

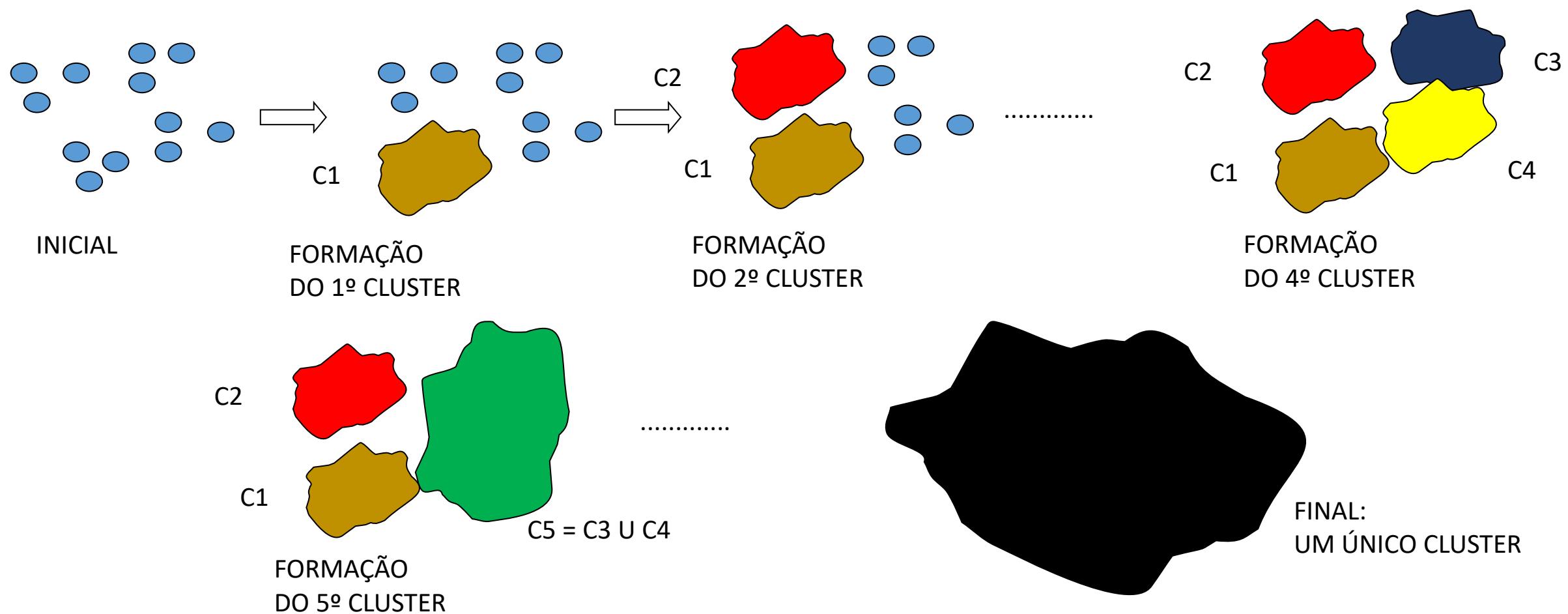
Fonte	GL	Desv (Aj.)	Média (Aj.)	Qui-Quadrado	Valor-P
Regressão	1	8,461	8,461	8,46	0,004
Ponto Extra	1	8,461	8,461	8,46	0,004
Erro	48	60,534	1,261		
Total	49	68,994			



# Análise de Clusters

# Clusters – Medidas de Similaridade

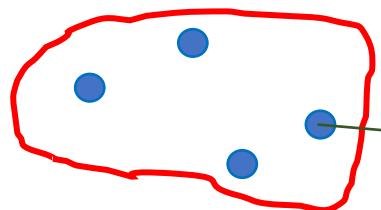
As medidas de similaridade serão utilizadas para formar os clusters como esquema a seguir:



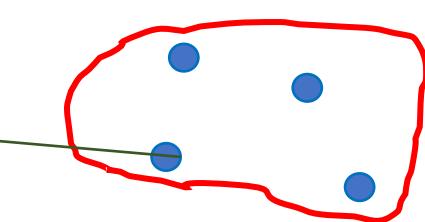
# Clusters – Métodos de Ligação

A questão principal é: como podemos ligar os clusters?

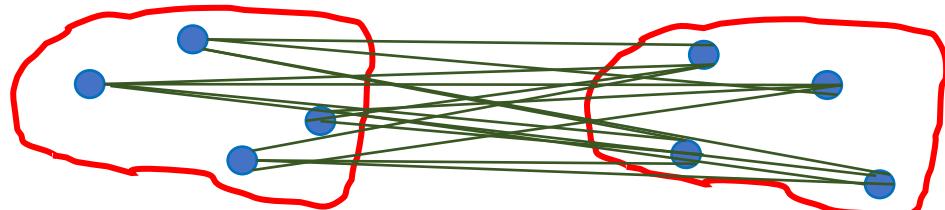
Os principais métodos são mostrados abaixo:



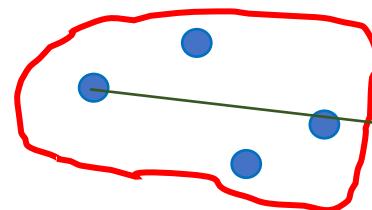
MÍNIMO



MÁXIMO



MÉDIA



CENTRÓIDE

Neste treinamento sempre utilizaremos a ligação pelo centróide.

# Clusters – Métodos de Agrupamento

Modelo	Objetivos & Quando utilizar	Exemplo
Agrupamento de Observações	<p>Para fazer grupos de observações que compartilham características comuns.</p> <p>Esta análise é apropriada quando você não tem nenhuma informação inicial sobre como formar os grupos.</p>	Fazer agrupamento de CDD's pelo Índice de Segurança de cada um deles
Agrupado de Variáveis	Se você deseja agrupar variáveis, em vez de observações	Desejo agrupar variáveis que tem impacto similar no Índice de Segurança
Agrupamento de K-Médias	Se você tiver informações suficientes para fornecer pontos de referência iniciais para os grupos, ou seja, alguns pontos já estão pré-classificados	Tenho 3 CDD's com Índices de Segurança conhecidos e classificados como Alto, Médio e Baixo e quero utilizar os valores desses 3 CDD's como referência para classificar os demais em torno deles
Análise Discriminante	Se você souber os grupos finais corretos para as observações e quiser classificar novas observações naqueles grupos	Tenho 3 grupos de CDD's identificados como sendo de Índice de Segurança Alto, Médio e Baixo e quero classificar alguns outros CDD's nesses grupos



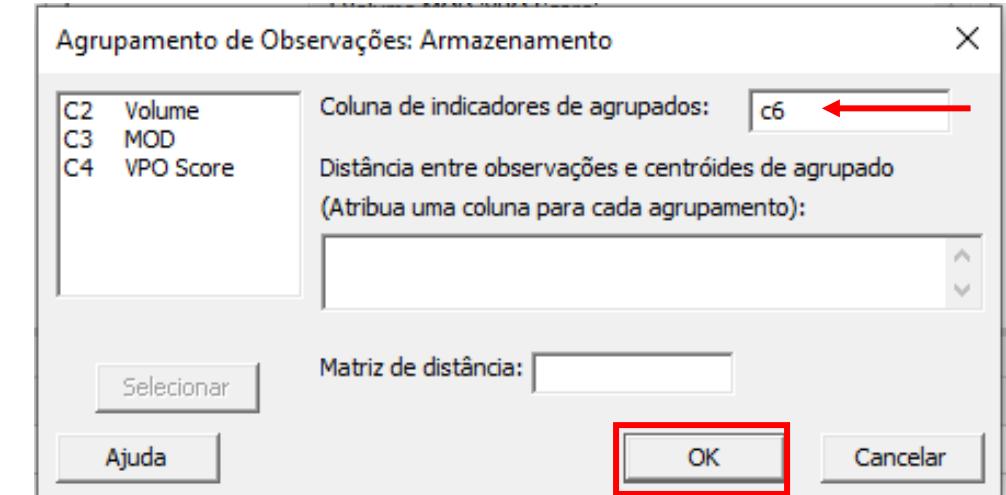
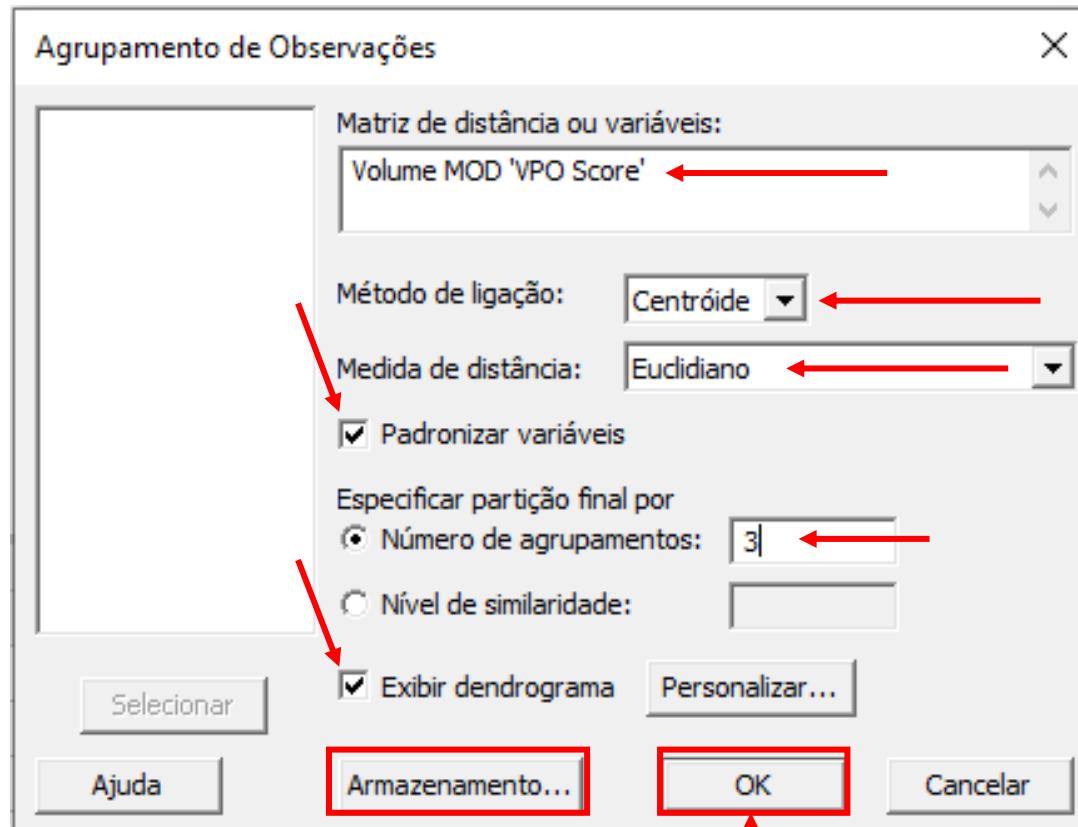
# Clusters – Agrupamento de Observações Exemplo

Foram obtidos dados de diversas plantas de produção com o objetivo de efetuar agrupamento entre elas para possibilitar uma melhor estratégia de atuação com os distintos grupos. Os dados estão no Arquivo: Clusters.xlsx na aba Plantas.

The screenshot shows the Minitab software interface with the 'Estat' menu open. The 'Multivariada' option is highlighted. A sub-menu for 'Agrupamento de Observações...' is displayed, with its description: 'Classificar observações em grupos com base em suas semelhanças. Usar quando você não tiver informações de agrupamento a priori para suas observações.'

Estat	Gráfico	Visualizar	Ajuda	Assistente
Estatísticas Básicas				
Regressão				
ANOVA				
DOE (Planejamento de Experimento)				
Cartas de Controle				
Ferramentas da Qualidade				
Confiabilidade/Sobrevivência				
Análise preditiva				
Multivariada				
Séries Temporais				
Tabelas				
Não-Paramétricos				
Testes de Equivalência				
Poder e Tamanho de Amostra				
Componentes Principais...				
Análise Fatorial...				
Análise de Item...				
Agrupamento de Observações...				
Agrupamento de Variáveis...				
Agrupamento de Observações				
Classificar observações em grupos com base em suas semelhanças. Usar quando você não tiver informações de agrupamento a priori para suas observações.				
Análise de Correspondência Múltipla...				

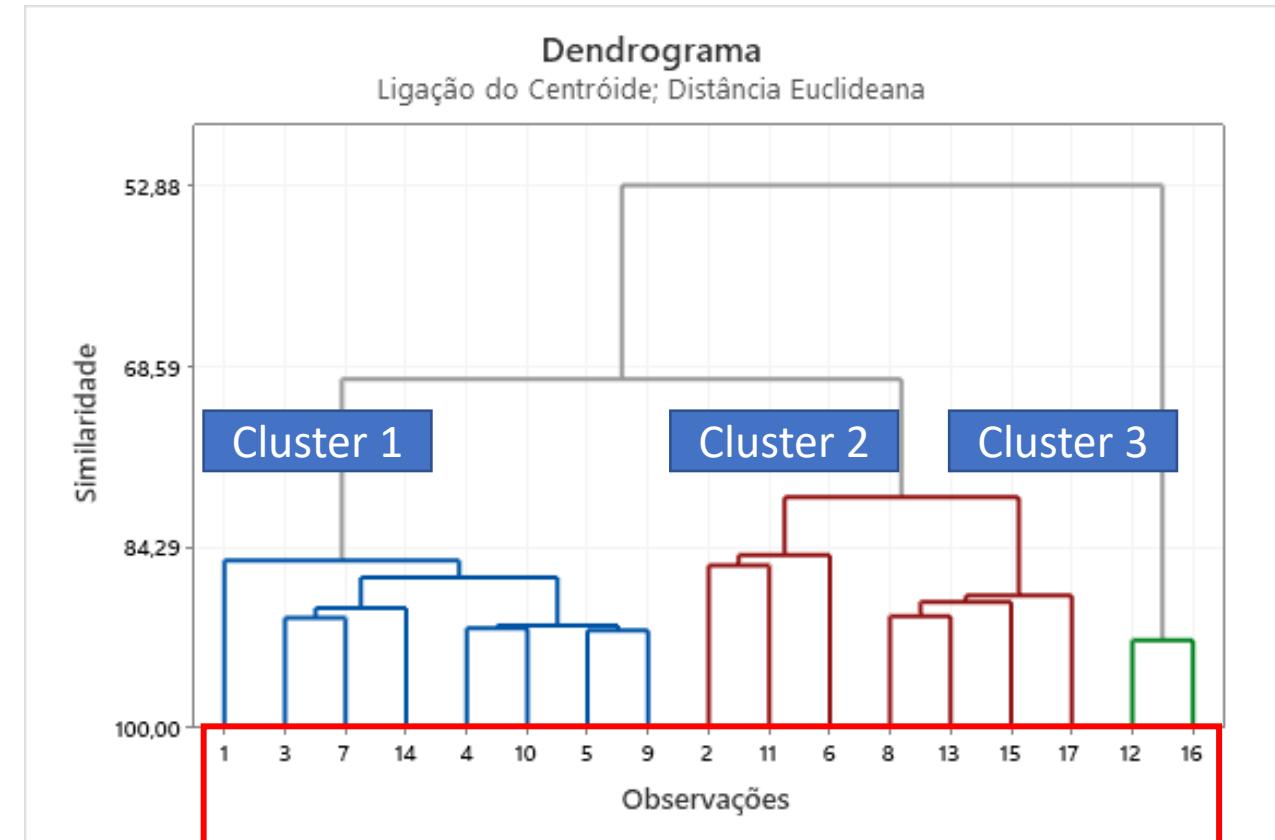
# Clusters – Agrupamento de Observações Exemplo





# Clusters – Agrupamento de Observações Exemplo

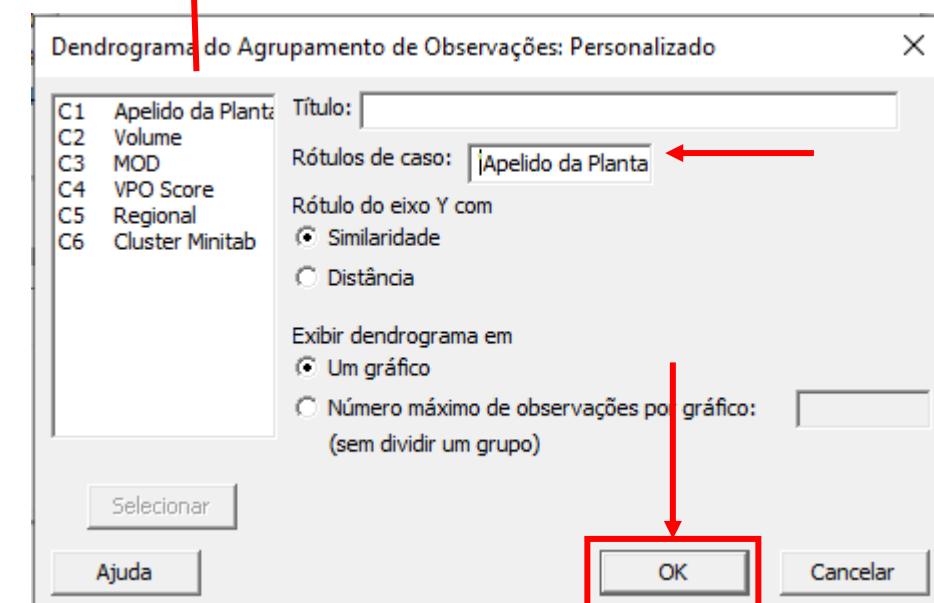
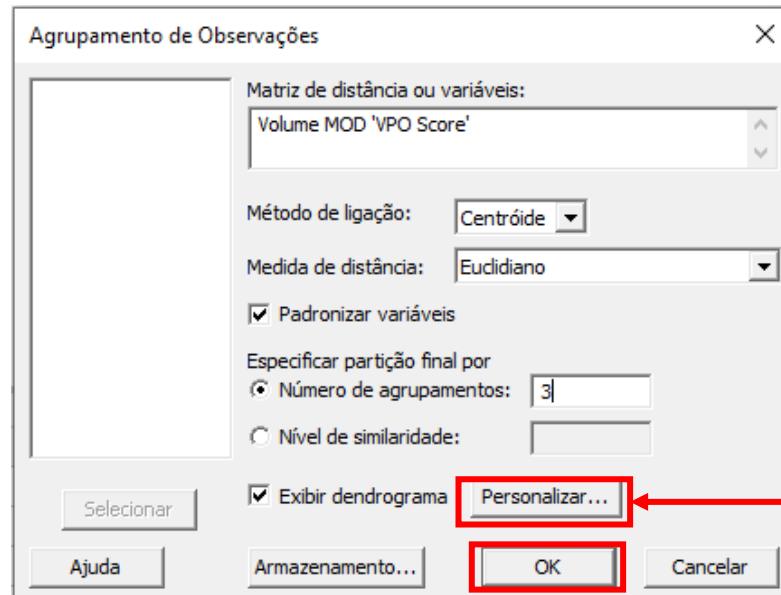
	C1-T	C2	C3	C4	C5-T	C6
	Apelido da Planta	Volume	MOD	VPO Score	Regional	Cluster Minitab
1	SHEILA	20000	500	80	A	1
2	MARCIA	30000	700	75	A	2
3	PAULA	25000	550	85	A	1
4	SHIRLEY	23000	400	85	A	1
5	DEBORA	18000	380	84	A	1
6	CAMILA	35000	730	78	B	2
7	OLIVIA	25000	510	83	B	1
8	CAROL	30000	682	80	B	2
9	THALITA	19000	400	86	B	1
10	AMANDA	22000	450	86	C	1
11	JOYCE	34500	708	73	C	2
12	DANIELA	42000	750	70	C	3
13	BRUNA	31500	650	82	C	2
14	CINTIA	20999	500	83	C	1
15	DENISE	32000	700	84	D	2
16	FERNANDA	41500	750	72	D	3
17	SILVIA	28345	650	85	D	2





# Clusters – Agrupamento de Observações Exemplo

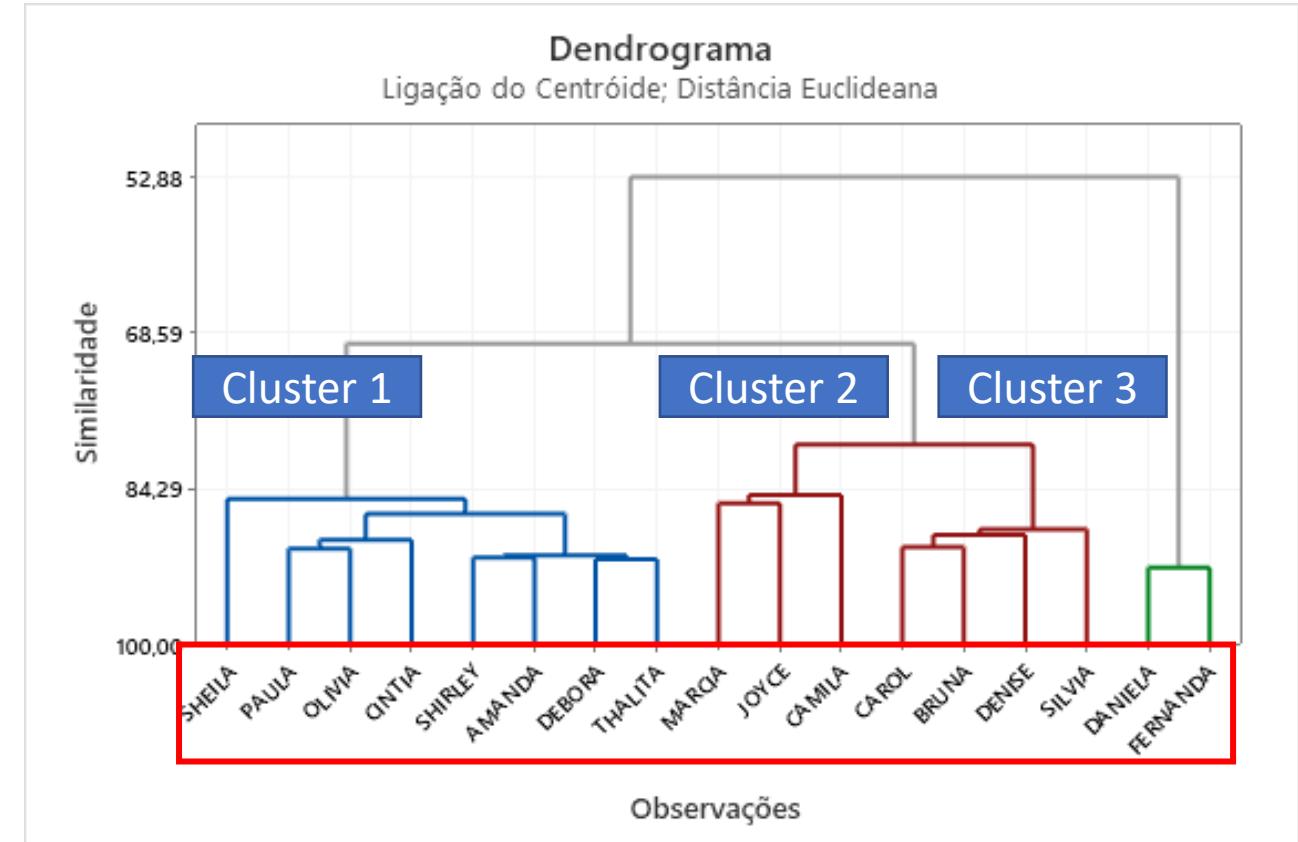
	C1-T	C2	C3	C4	C5-T	C6
	Apelido da Planta	Volume	MOD	VPO Score	Regional	Cluster Minitab
1	SHEILA	20000	500	80	A	1
2	MARCIA	30000	700	75	A	2
3	PAULA	25000	550	85	A	1
4	SHIRLEY	23000	400	85	A	1
5	DEBORA	18000	380	84	A	1
6	CAMILA	35000	730	78	B	2
7	OLIVIA	25000	510	83	B	1
8	CAROL	30000	682	80	B	2
9	THALITA	19000	400	86	B	1
10	AMANDA	22000	450	86	C	1
11	JOYCE	34500	708	73	C	2
12	DANIELA	42000	750	70	C	3
13	BRUNA	31500	650	82	C	2
14	CINTIA	20999	500	83	C	1
15	DENISE	32000	700	84	D	2
16	FERNANDA	41500	750	72	D	3
17	SILVIA	28345	650	85	D	2





# Clusters – Agrupamento de Observações Exemplo

	C1-T	C2	C3	C4	C5-T	C6
	Apelido da Planta	Volume	MOD	VPO Score	Regional	Cluster Minitab
1	SHEILA	20000	500	80	A	1
2	MARCIA	30000	700	75	A	2
3	PAULA	25000	550	85	A	1
4	SHIRLEY	23000	400	85	A	1
5	DEBORA	18000	380	84	A	1
6	CAMILA	35000	730	78	B	2
7	OLIVIA	25000	510	83	B	1
8	CAROL	30000	682	80	B	2
9	THALITA	19000	400	86	B	1
10	AMANDA	22000	450	86	C	1
11	JOYCE	34500	708	73	C	2
12	DANIELA	42000	750	70	C	3
13	BRUNA	31500	650	82	C	2
14	CINTIA	20999	500	83	C	1
15	DENISE	32000	700	84	D	2
16	FERNANDA	41500	750	72	D	3
17	SILVIA	28345	650	85	D	2

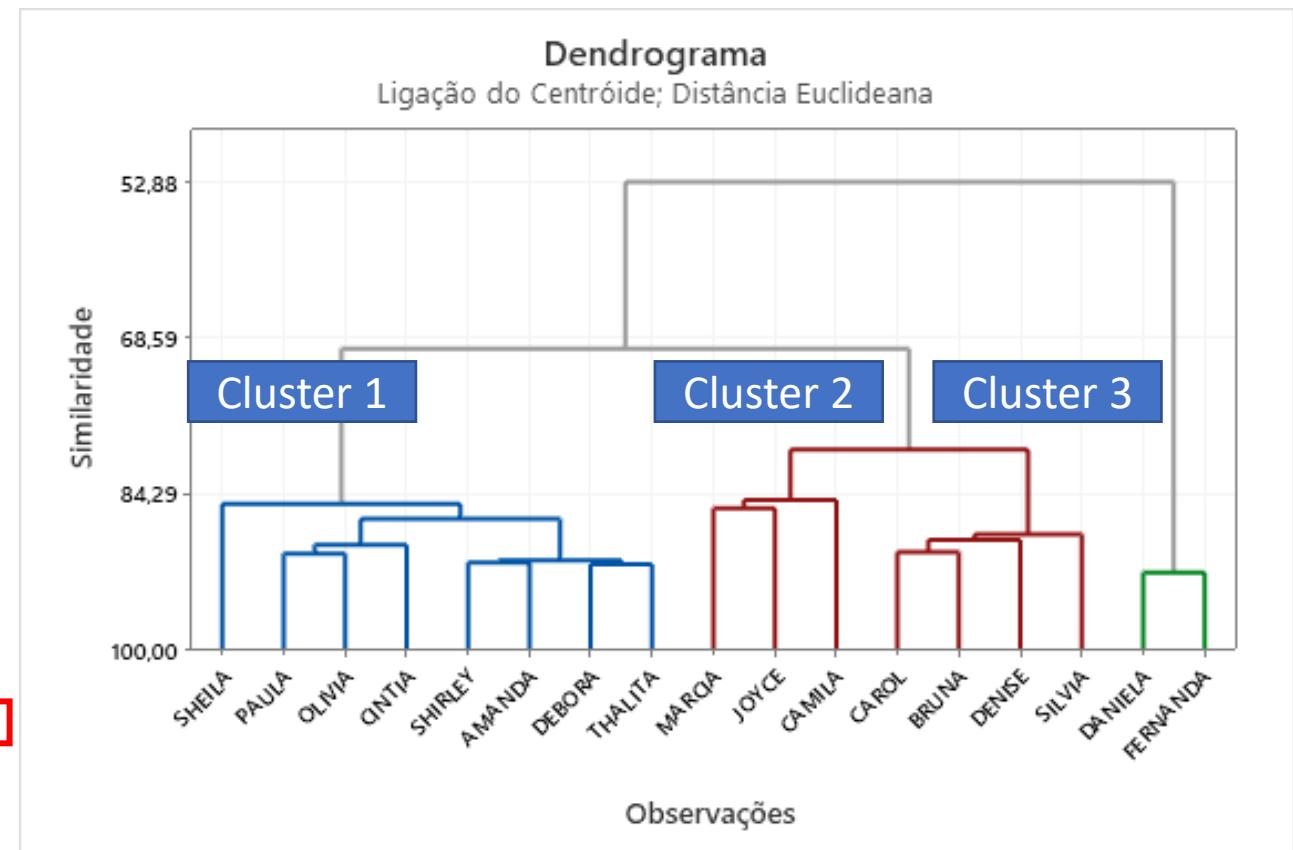




# Clusters – Agrupamento de Observações Exemplo

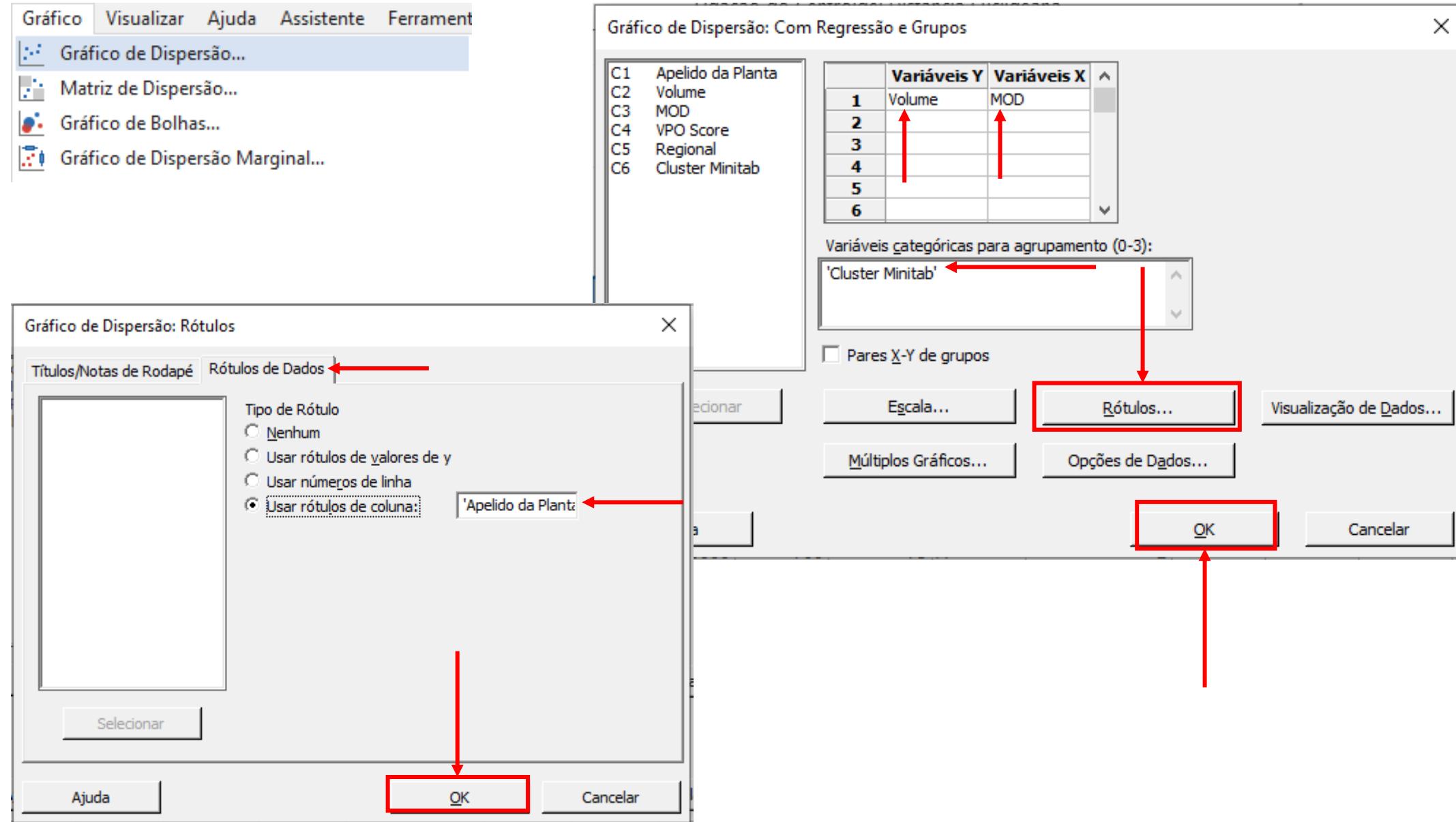
## Passos de Amalgamação

Passo	Número de agrupados	Número de similaridade	Nível de distância	Agrupados reunidos		Número de obs. no novo agrupado	
				12	16	12	2
1	16	92,3571	0,38908	12	16	12	2
2	15	91,4935	0,43304	5	9	5	2
3	14	91,3159	0,44208	4	10	4	2
4	13	91,1046	0,45283	4	5	4	4
5	12	90,4418	0,48658	3	7	3	2
6	11	90,2698	0,49533	8	13	8	2
7	10	89,5632	0,53130	3	14	3	3
8	9	89,0365	0,55812	8	15	8	3
9	8	88,4989	0,58548	8	17	8	4
10	7	86,9352	0,66508	3	4	3	7
11	6	85,8461	0,72053	2	11	2	2
12	5	85,4236	0,74203	1	3	1	8
13	4	84,9867	0,76428	2	6	2	3
14	3	79,9399	1,02119	2	8	2	7
15	2	69,7162	1,54165	1	2	1	15
16	1	52,8802	2,39871	1	12	1	17

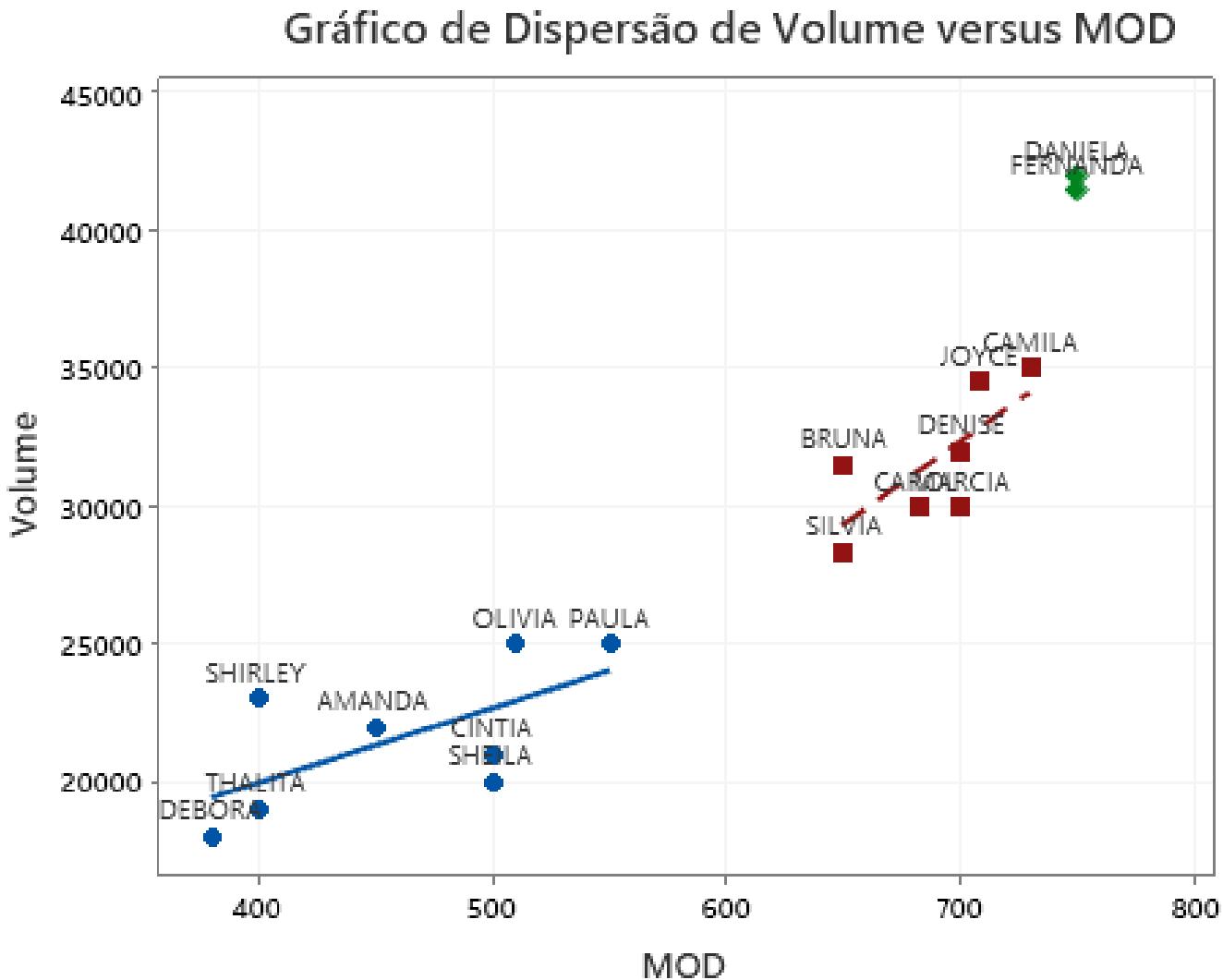


Auxilia a identificar um número mais adequado de clusters => quando a similaridade cai muito é por que estamos agrupando itens menos similares entre si.

# Clusters – Agrupamento de Observações Exemplo



# Clusters – Agrupamento de Observações Exemplo





# Clusters – Agrupamento de Observações Exemplo

Screenshot of Minitab software interface showing the process of creating a grouped boxplot.

The menu bar at the top includes: Gráfico, Visualizar, Ajuda, Assistente, and Ferramentas.

The "Gráfico" menu is expanded, showing various options like Gráfico de Dispersão..., Matriz de Dispersão..., Gráfico de Bolhas..., etc., with "Boxplot..." highlighted.

The "Boxplots" dialog box is open, showing two categories: "Um Y" and "Múltiplos Y's". Under "Um Y", the "Simples" option is selected, and under "Com Grupos", a boxplot with two groups (A and B) is displayed. This option is highlighted with a red box.

A red arrow points from the "OK" button in the "Boxplots" dialog to the "OK" button in the "Boxplot: Um Y, Com Grupos" sub-dialog.

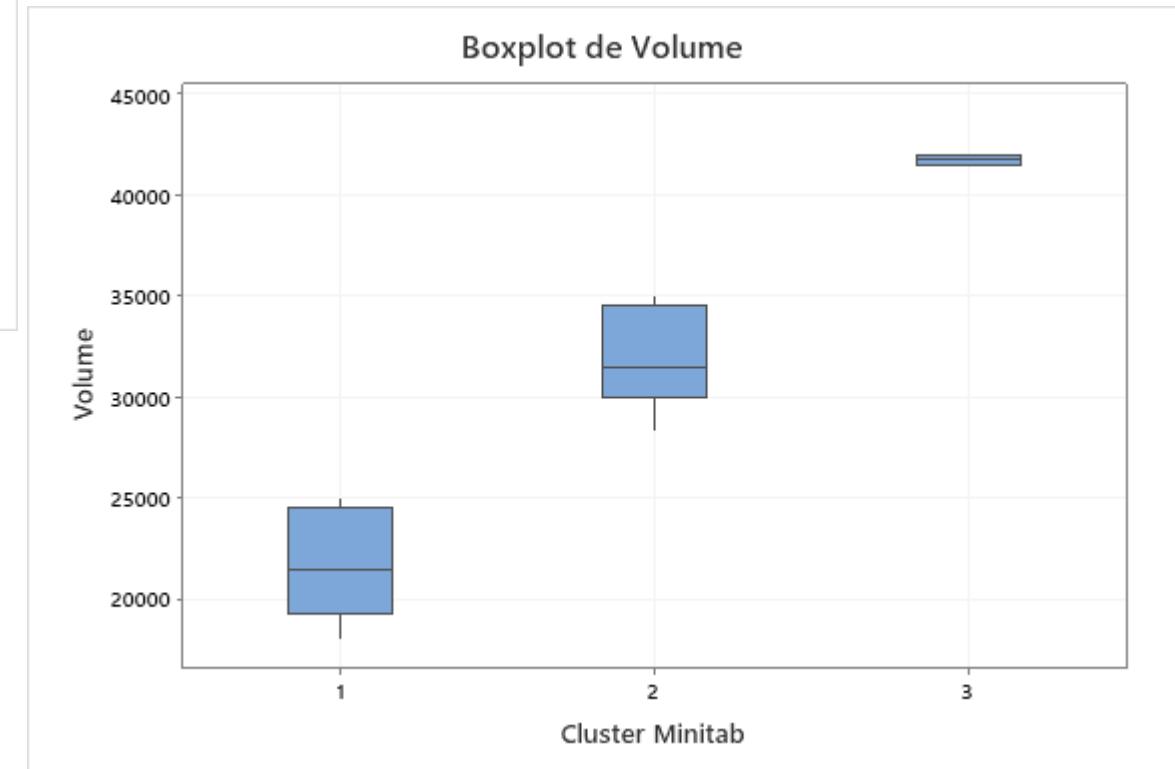
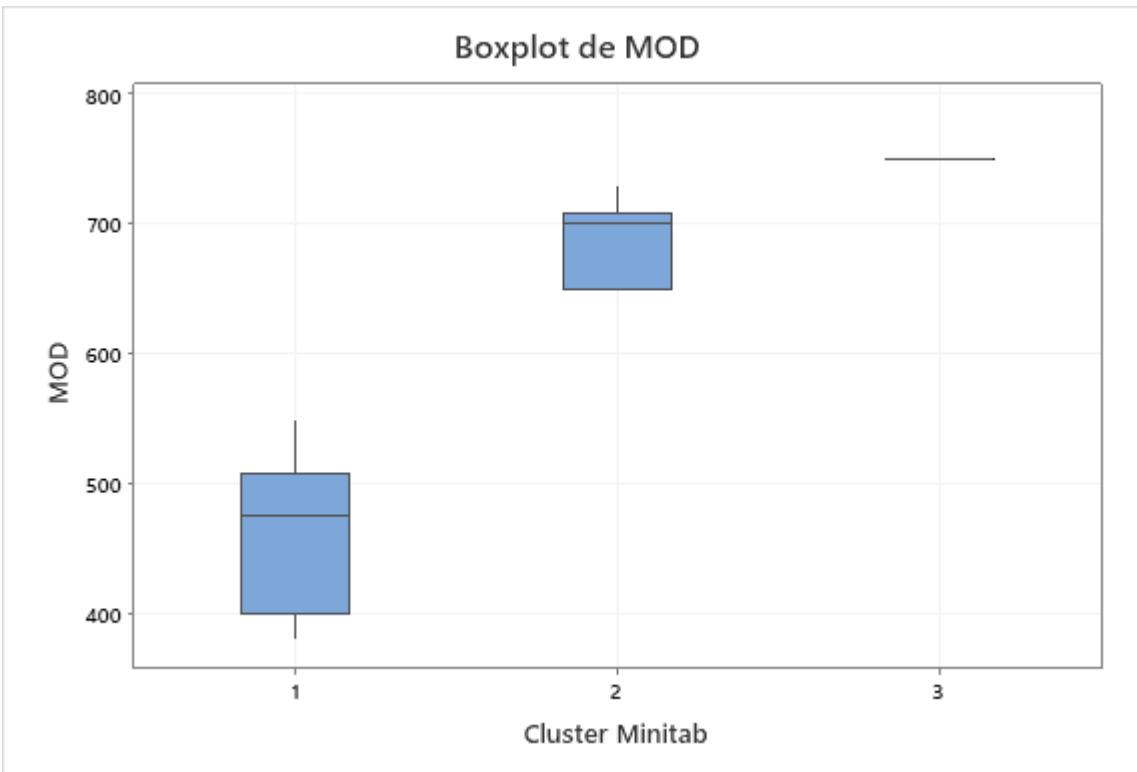
The "Boxplot: Um Y, Com Grupos" dialog box contains the following settings:

- Variáveis do gráfico: MOD Volume (highlighted with a red arrow).
- Variáveis categóricas para agrupamento (1-4, a mais externa primeiro): 'Cluster Minitab' (highlighted with a red arrow).
- Buttons: Escala..., Rótulos..., Visualização de Dados..., Selecionar, Múltiplos Gráficos..., Opções de Dados..., Ajuda, and Cancelar.

The "OK" button in the main "Boxplot: Um Y, Com Grupos" dialog is also highlighted with a red box.



# Clusters – Agrupamento de Observações Exemplo





# Clusters – Agrupamento de Variáveis Exemplo

Foram obtidos dados de diversas variáveis de CDD's distintos com o objetivo de reduzir o número de variáveis utilizadas em um dashboard de acompanhamento. O objetivo era ter no máximo 4 variáveis, seria isso adequado? Os dados estão no Arquivo: Clusters na aba CDDs.

The screenshot shows the Minitab software interface with the 'Estat' menu open. The 'Multivariada' option is selected, revealing a submenu with several statistical methods. The 'Agrupamento de Variáveis...' option is also highlighted. A detailed description of the 'Agrupamento de K-Médias...' method is provided at the bottom of the submenu.

Estat   Gráfico   Visualizar   Ajuda   Assistente

- Estatísticas Básicas
- Regressão
- ANOVA
- DOE (Planejamento de Experimento)
- Cartas de Controle
- Ferramentas da Qualidade
- Confiabilidade/Sobrevivência
- Análise preditiva
- Multivariada
- Séries Temporais
- Tabelas
- Não-Paramétricos
- Testes de Equivalência
- Poder e Tamanho de Amostra

Componentes Principais...

Análise Fatorial...

Análise de Item...

Agrupamento de Observações...

**Agrupamento de Variáveis...**

Agrupamento de K-Médias...

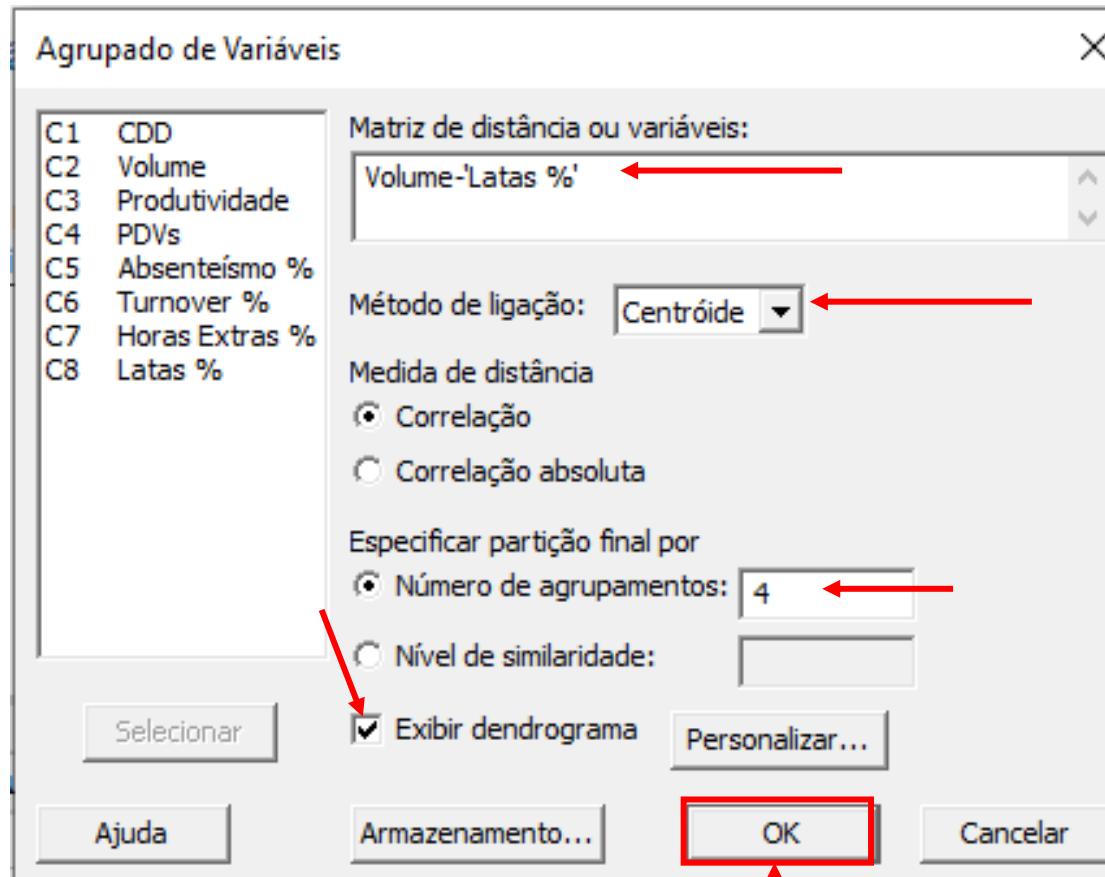
**Agrupamento de Variáveis**

Reducir o número de variáveis classificando-as em grupos com base em suas semelhanças.

Análise de Correspondência Múltipla...



# Clusters – Agrupamento de Variáveis Exemplo



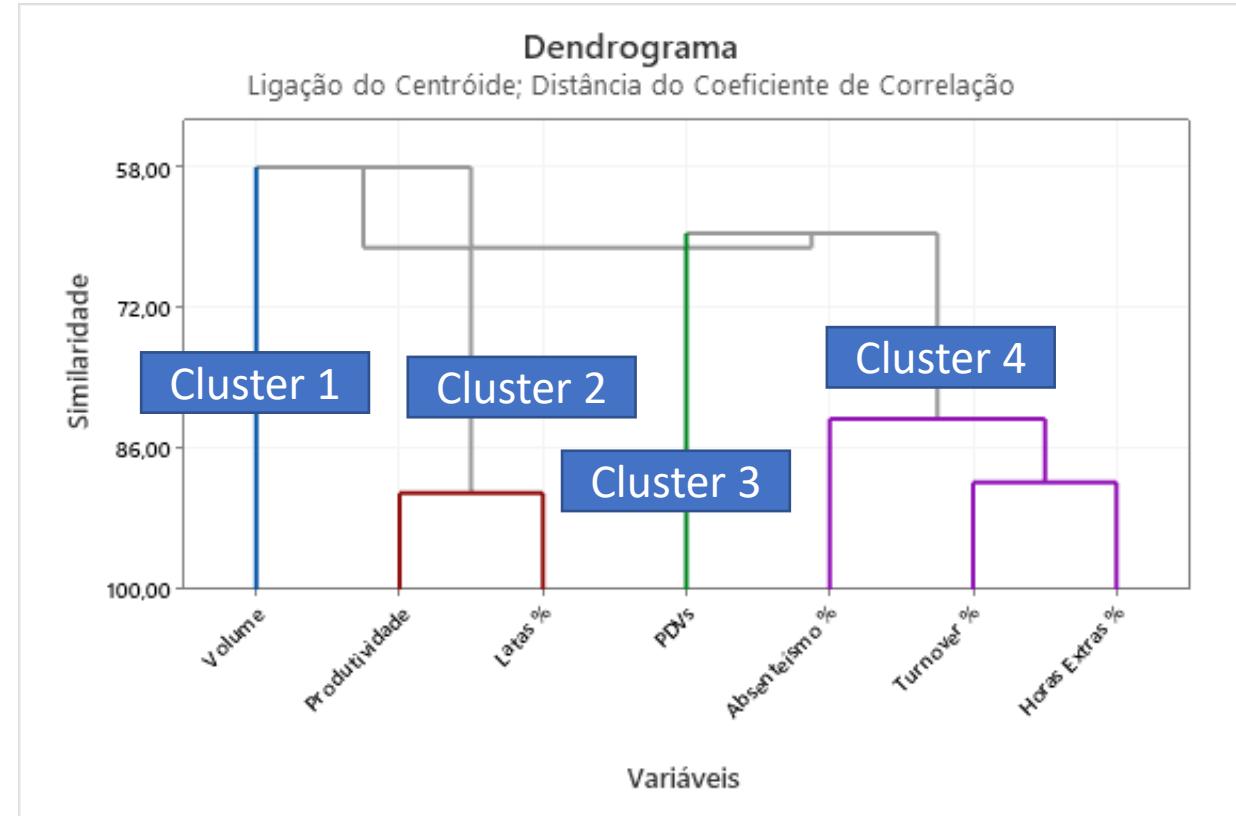


# Clusters – Agrupamento de Variáveis Exemplo

## Passos de Amalgamação

Passo	Número de agrupados	Nível de similaridade	Nível de distância	Agrupados reunidos	Novo agrupado	Número de obs. no novo agrupado
1	6	90,3799	0,192402	2	7	2
2	5	89,3283	0,213433	5	6	5
3	4	83,0400	0,339201	4	5	4
4	3	64,5778	0,708445	3	4	3
5	2	58,0039	0,839923	1	2	1
6	1	65,9971	0,680057	1	3	7

**Utilizar menos que 4 variáveis reduz muito a similaridade entre elas, porém podem ser selecionadas quatro variáveis (uma de cada cluster) para o acompanhamento no dashboard.**



# Business Case Mercado – PDV Álcool Gel – Exercício

Para avaliar se o acompanhamento e atuação de variáveis deveria ser separado por tipo de PDV (Mercado e Farmácia) faça análise de clusters (variáveis) tendo em conta:

- Market Share: MS Merc e MS Farm
- Distribuição Relativa: Distr Rel Merc e Distr Rel Farm
- Espaço Relativo: Esp. Relat. Merc e Esp. Relat. Farm.
- Exibições Relativas: Display Rel Merc e Display Rel Farm

Os dados estão no arquivo Coolgel\_Banco de Dados\_PDV.xlsx na aba “Market Share Decline”

## Passos de Amalgamação

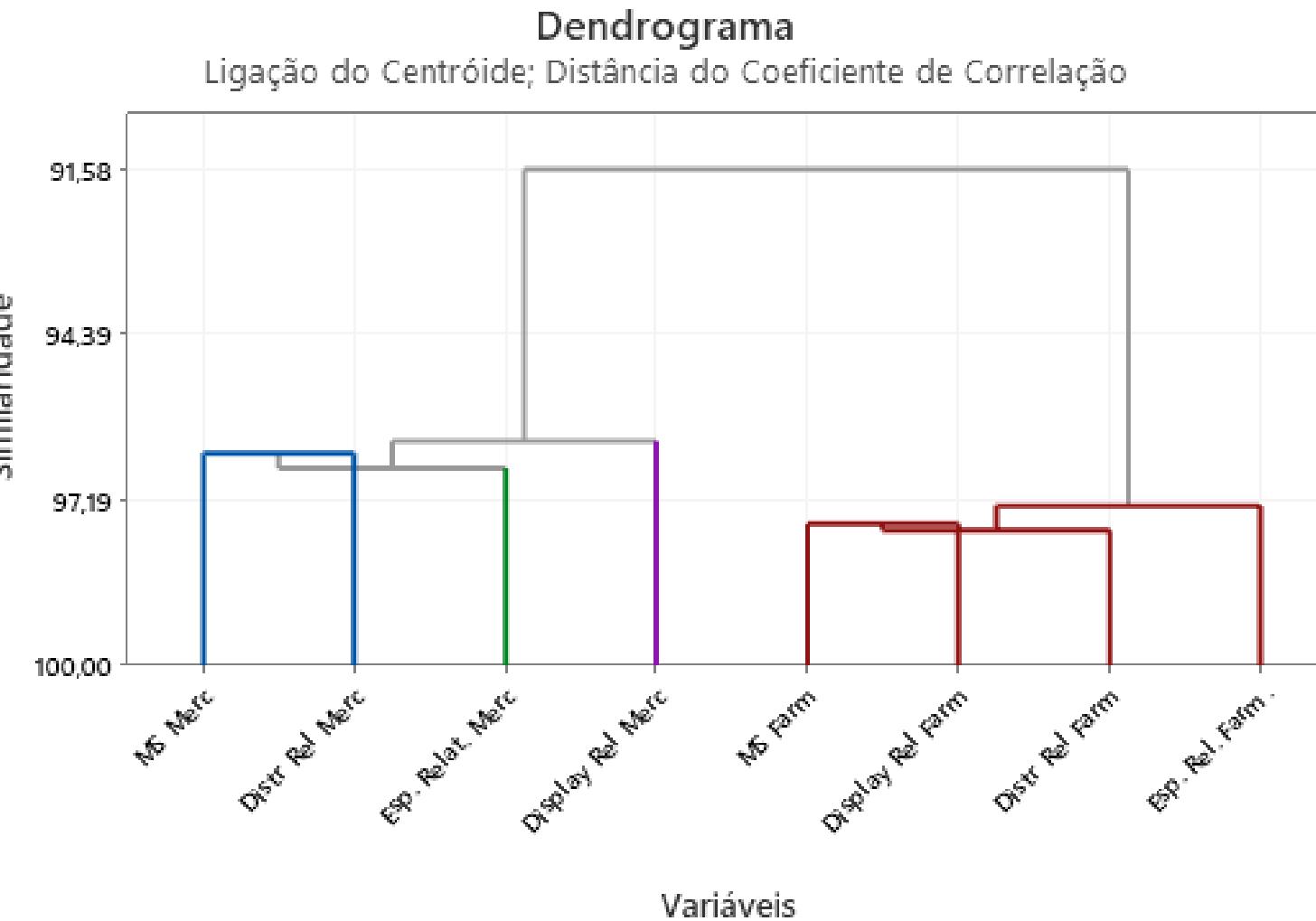
Passo	Número de agrupados	Nível de similaridade	Nível de distância	Agrupados reunidos		Novo agrupado	Número de obs. no novo agrupado
				reunidos	novos		
1	7	97,5980	0,048039	2	8	2	2
2	6	97,7024	0,045952	2	4	2	3
3	5	97,2955	0,054090	2	6	2	4
4	4	96,4006	0,071988	1	3	1	2
5	3	96,6525	0,066951	1	5	1	3
6	2	96,1933	0,076134	1	7	1	4
7	1	91,5833	0,168333	1	2	1	8

## Partição Final

### Variáveis

Agrupamento 1	MS	Merc	Distr	Rel	Merc						
Agrupamento 2	MS	Farm	Distr	Rel	Farm	Esp.	Rel.	Farm.	Display	Rel	Farm
Agrupamento 3	Esp.	Relat.	Merc								
Agrupamento 4	Display	Rel	Merc								

Business  
Case – PDV  
Análise de  
Cluster  
Variáveis



Business  
Case – PDV  
Análise de  
Cluster  
Variáveis

Os PDV's devem ser  
acompanhados e atuados de  
acordo com o tipo.

# Business Case Mercado – PDV Álcool Gel – Exercício

Para identificar o melhor conjunto mínimo de variáveis a utilizar para acompanhamento e atuação por tipo de PDV (Mercado e Farmácia) faça análise de clusters (variáveis) tendo em conta:

- Distribuição Relativa: Distr Rel Merc e Distr Rel Farm
- Espaço Relativo: Esp. Relat. Merc e Esp. Relat. Farm.
- Exibições Relativas: Display Rel Merc e Display Rel Farm

Os dados estão no arquivo Coolgel\_Banco de Dados\_PDV.xlsx na aba “Market Share Decline”

# Business Case – PDV Análise de Cluster Variáveis

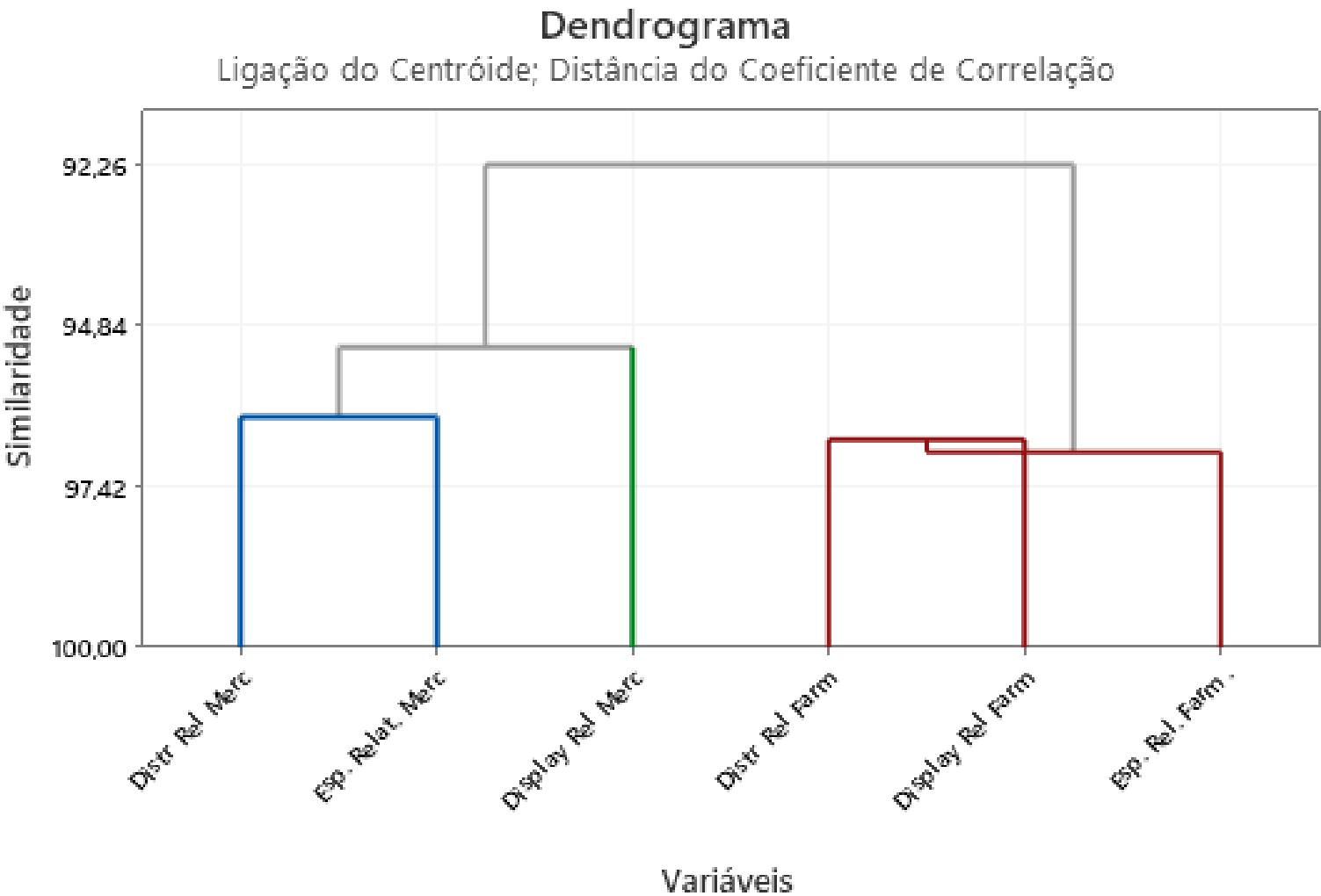
## Passos de Amalgamação

Passo	Número de agrupados	Nível de similaridade	Nível de distância	Agrupados reunidos	Novo agrupado	Número de obs. no novo agrupado
1	5	96,6663	0,066675	2 6	2	2
2	4	96,8708	0,062585	2 4	2	3
3	3	96,2957	0,074086	1 3	1	2
4	2	95,1885	0,096230	1 5	1	3
5	1	92,2599	0,154801	1 2	1	6

## Partição Final

### Variáveis

Agrupamento 1	Distr Rel Merc	Esp. Relat. Merc
Agrupamento 2	Distr Rel Farm	Esp. Rel. Farm.
Agrupamento 3	Display Rel	Merc



Business  
Case – PDV  
Análise de  
Cluster  
Variáveis

Pelo menos duas variáveis de  
acompanhamento para  
Mercado e uma para  
Farmácia.

# Business Case Mercado – PDV Álcool Gel – Exercício

Para identificar se existem agrupamentos em relação às observações mensais de dados e resultados de MS por tipo de PDV (Mercado e Farmácia) faça análise de clusters (observações) tendo em conta:

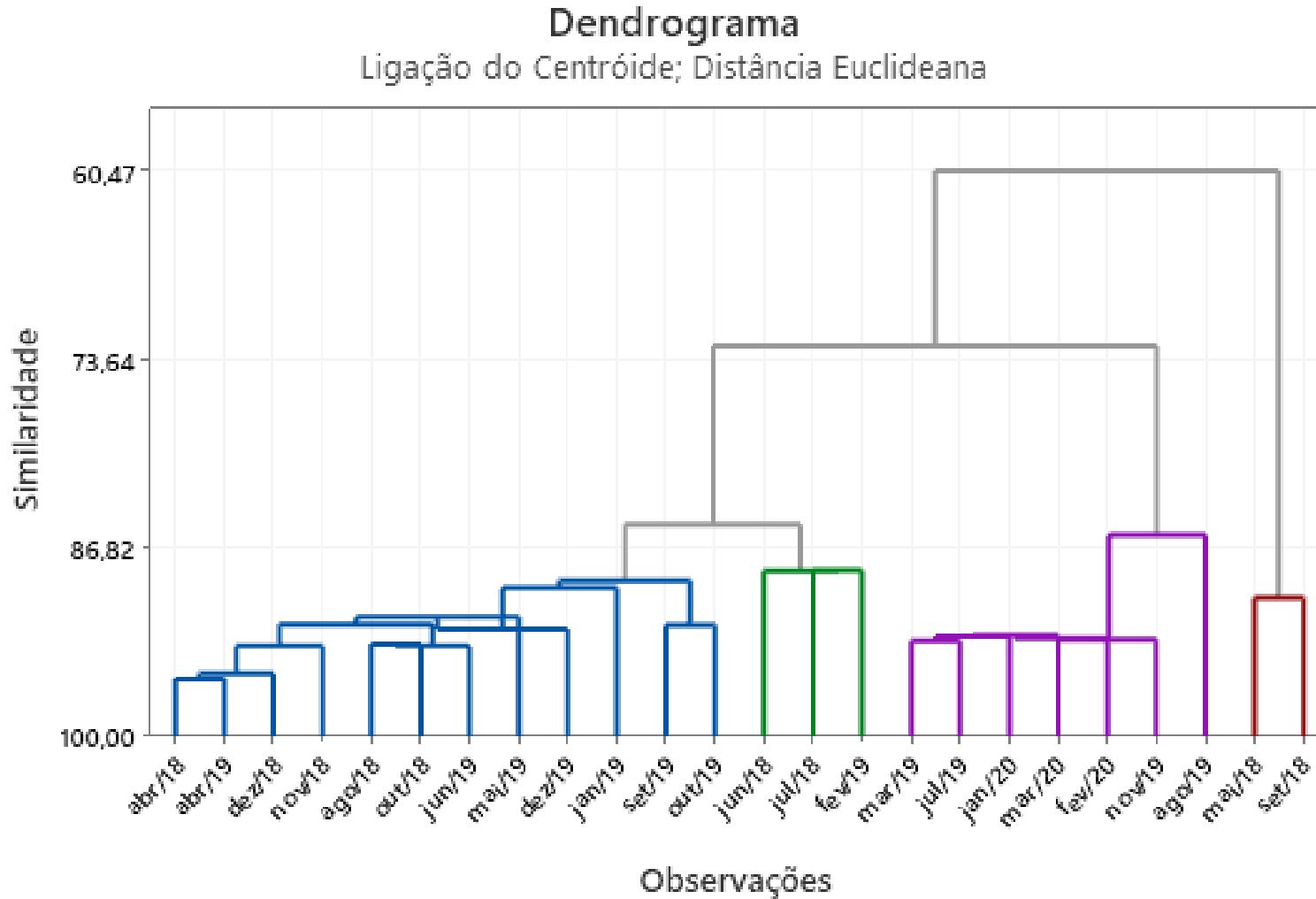
- Market Share: MS Merc e MS Farm
- Distribuição Relativa: Distr Rel Merc e Distr Rel Farm
- Espaço Relativo: Esp. Relat. Merc e Esp. Relat. Farm.
- Exibições Relativas: Display Rel Merc e Display Rel Farm

Os dados estão no arquivo Coolgel.xlsx na aba “Market Share Decline”

## Análise de Agrupamentos de Observações: Merc

Passo	Número de agrupados	Nível de similaridade	Nível de distância	Agrupados reunidos		Novo agrupado	Número de obs. no novo agrupado
1	23	95,9699	0,30331	1	13	1	2
2	22	95,5756	0,33299	1	9	1	3
3	21	93,6456	0,47824	1	8	1	4
4	20	93,5269	0,48717	5	7	5	2
5	19	93,6588	0,47724	5	15	5	3
6	18	93,3190	0,50282	12	16	12	2
7	17	92,9953	0,52718	12	22	12	3
8	16	92,9106	0,53356	12	24	12	4
9	15	93,1455	0,51588	12	23	12	5
10	14	93,2394	0,50881	12	20	12	6
11	13	92,1917	0,58766	18	19	18	2
12	12	92,1782	0,58867	1	5	1	7
13	11	91,6352	0,62955	1	14	1	8
14	10	92,4636	0,56719	1	21	1	9
15	9	90,3011	0,72995	2	6	2	2
16	8	89,5829	0,78400	1	10	1	10
17	7	89,1007	0,82029	1	18	1	12
18	6	88,3414	0,87744	4	11	4	2
19	5	88,4259	0,87108	3	4	3	3
20	4	85,8926	1,06174	12	17	12	7
21	3	85,1723	1,11595	1	3	1	15
22	2	72,6998	2,05464	1	12	1	22
23	1	60,4669	2,97530	1	2	1	24

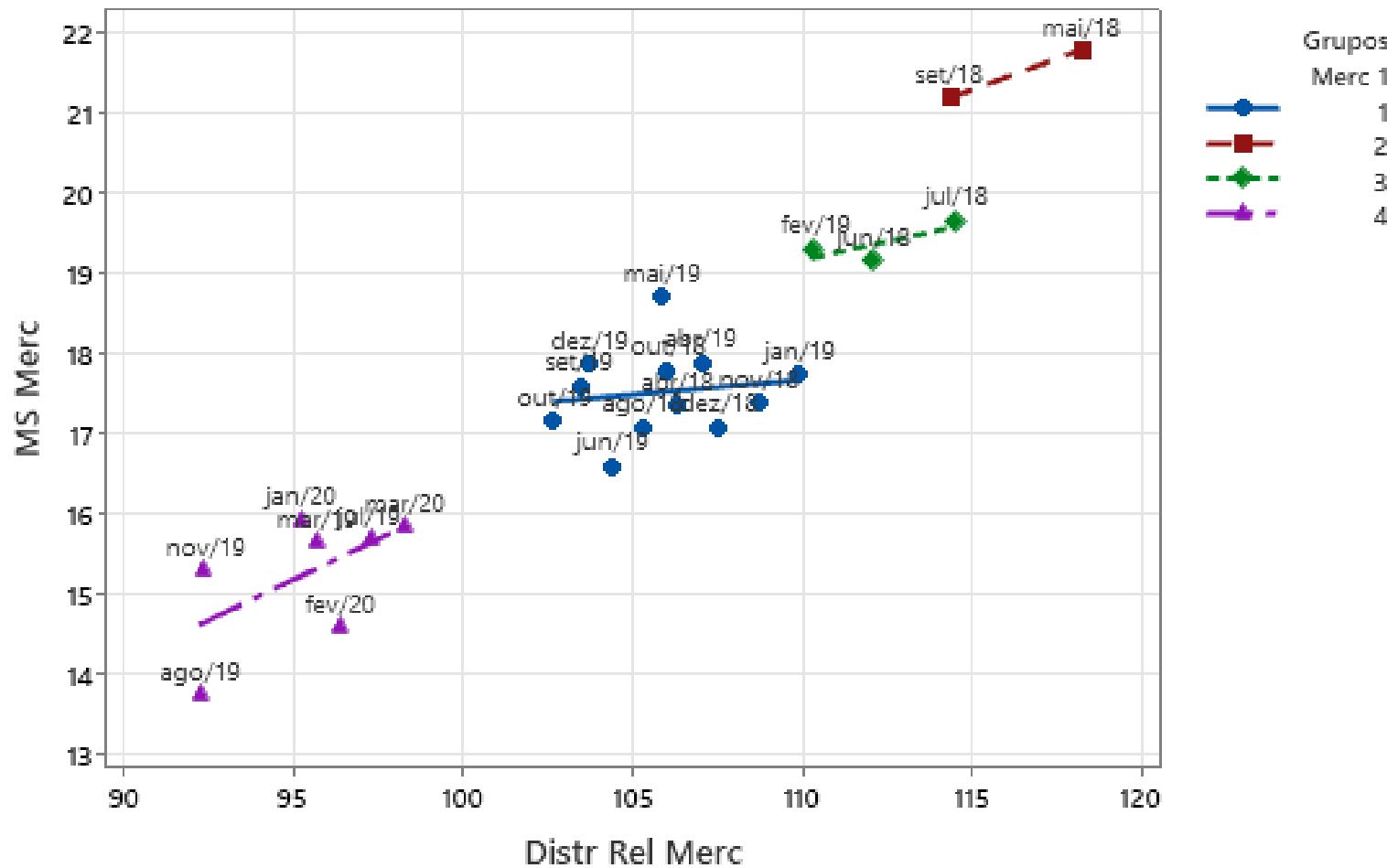
Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Mercado



Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Mercado

Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Mercado

Gráfico de Dispersão de MS Merc versus Distr Rel Merc

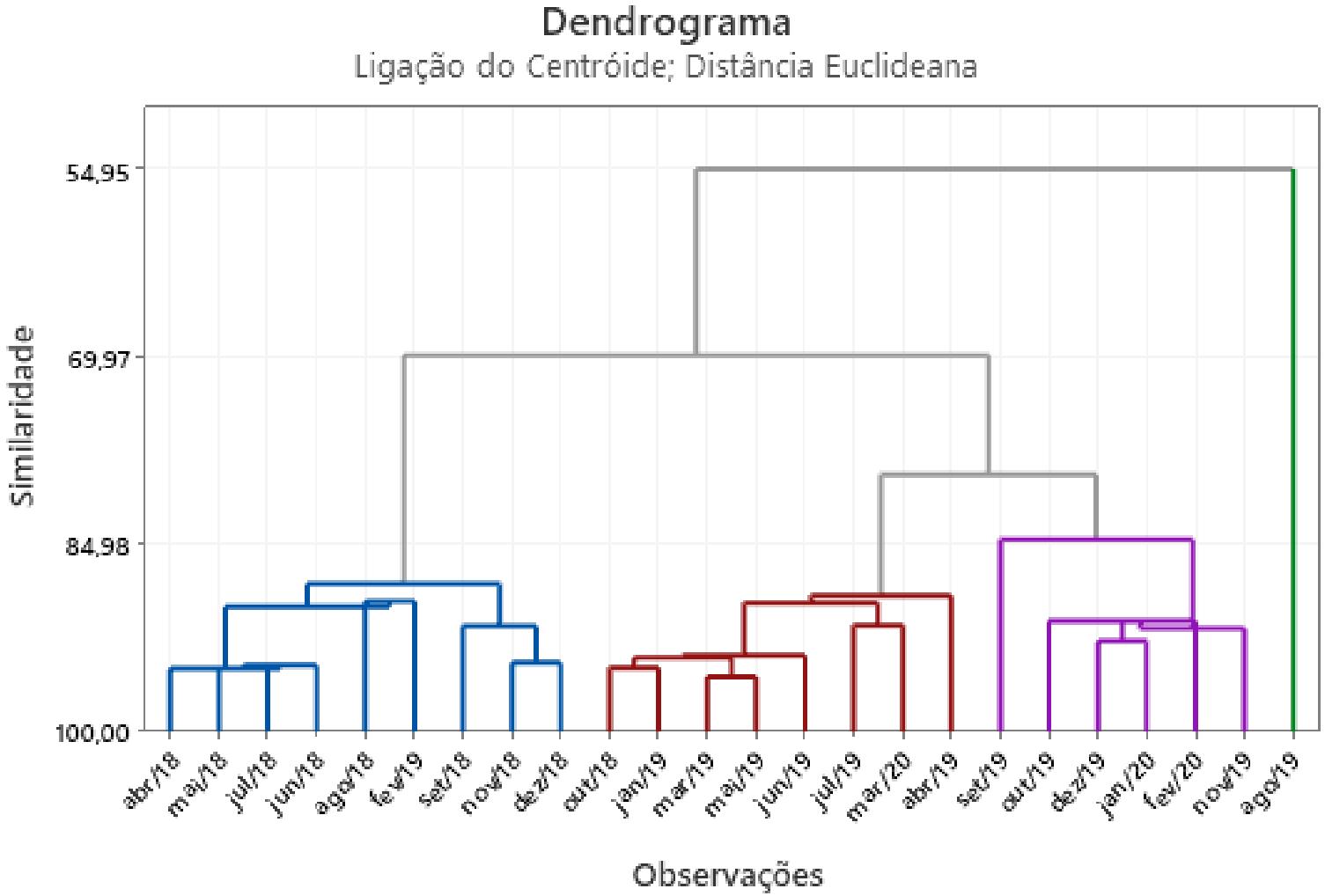


## Análise de Agrupamentos de Observações: Farm

### Passos de Amalgamação

Passo	Número de agrupados	Nível de similaridade	Número de distância	Agrupados reunidos		Novo agrupado	Número de obs. no novo agrupado
				reunidos	novos agrupados		
1	23	95,6003	0,29206	12	14	12	2
2	22	94,9154	0,33753	2	4	2	2
3	21	94,8750	0,34020	7	10	7	2
4	20	94,6756	0,35344	2	3	2	3
5	19	94,9069	0,33808	1	2	1	4
6	18	94,4492	0,36847	8	9	8	2
7	17	94,0811	0,39290	7	12	7	4
8	16	93,8910	0,40552	7	15	7	5
9	15	92,7333	0,48238	21	22	21	2
10	14	91,5256	0,56254	6	8	6	3
11	13	91,4760	0,56583	16	24	16	2
12	12	91,1309	0,58874	19	21	19	3
13	11	91,1280	0,58894	19	23	19	4
14	10	91,7082	0,55042	19	20	19	5
15	9	89,6895	0,68442	7	16	7	7
16	8	89,5138	0,69609	5	11	5	2
17	7	89,9803	0,66512	1	5	1	6
18	6	89,0894	0,72426	7	13	7	8
19	5	88,1272	0,78813	1	6	1	9
20	4	84,6152	1,02127	18	19	18	6
21	3	79,4236	1,36589	7	18	7	14
22	2	69,8344	2,00244	1	7	1	23
23	1	54,9526	2,99031	1	17	1	24

Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Farmácia



Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Farmácia

Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Farmácia

Gráfico de Dispersão de MS Farm versus Distr Rel Farm

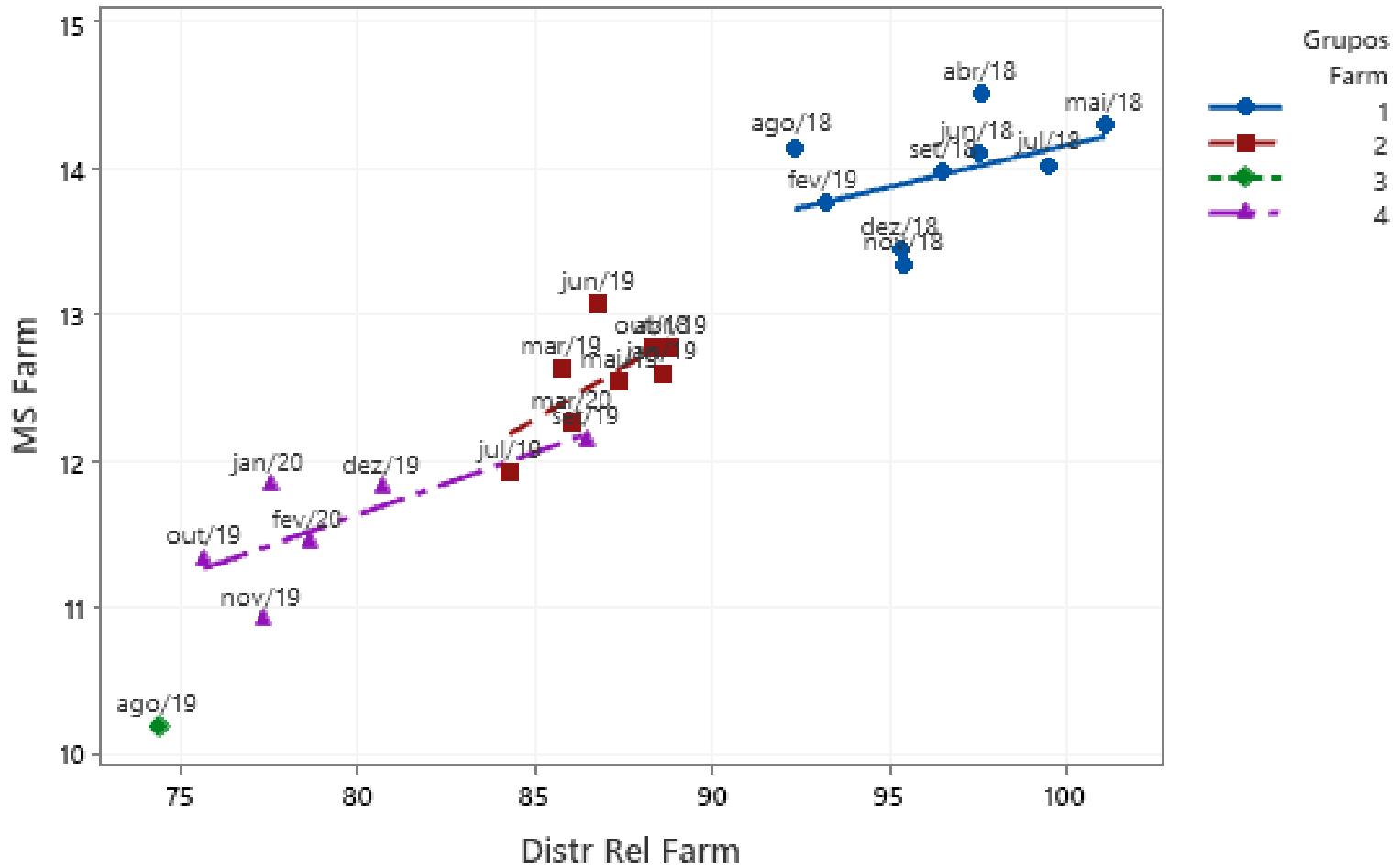


Gráfico de Dispersão de MS Merc versus Distr Rel Merc

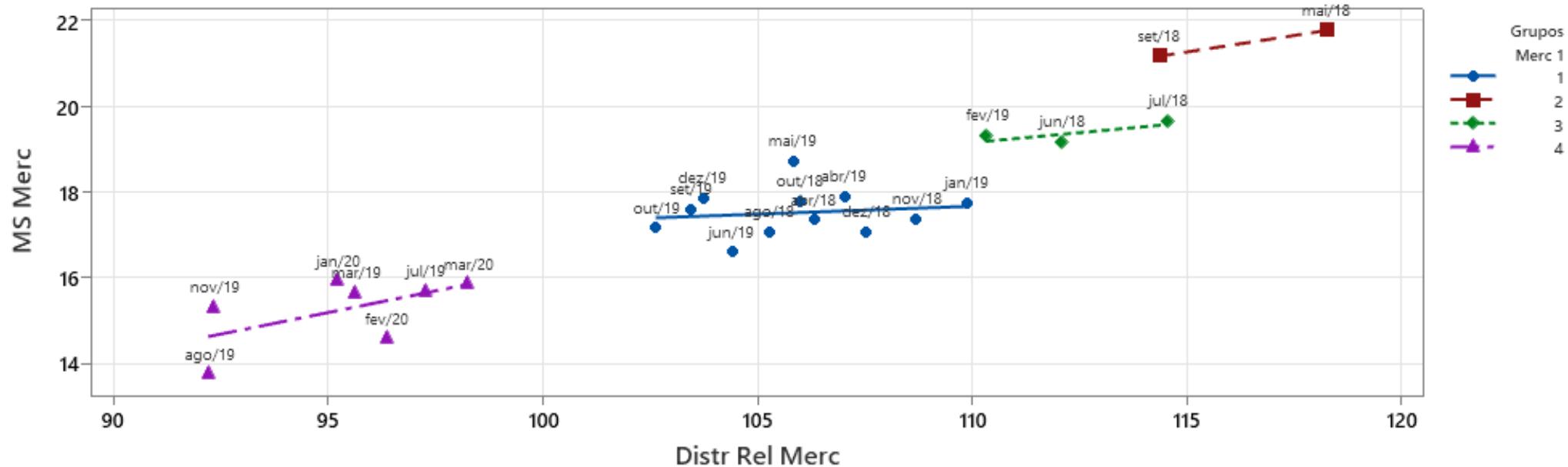
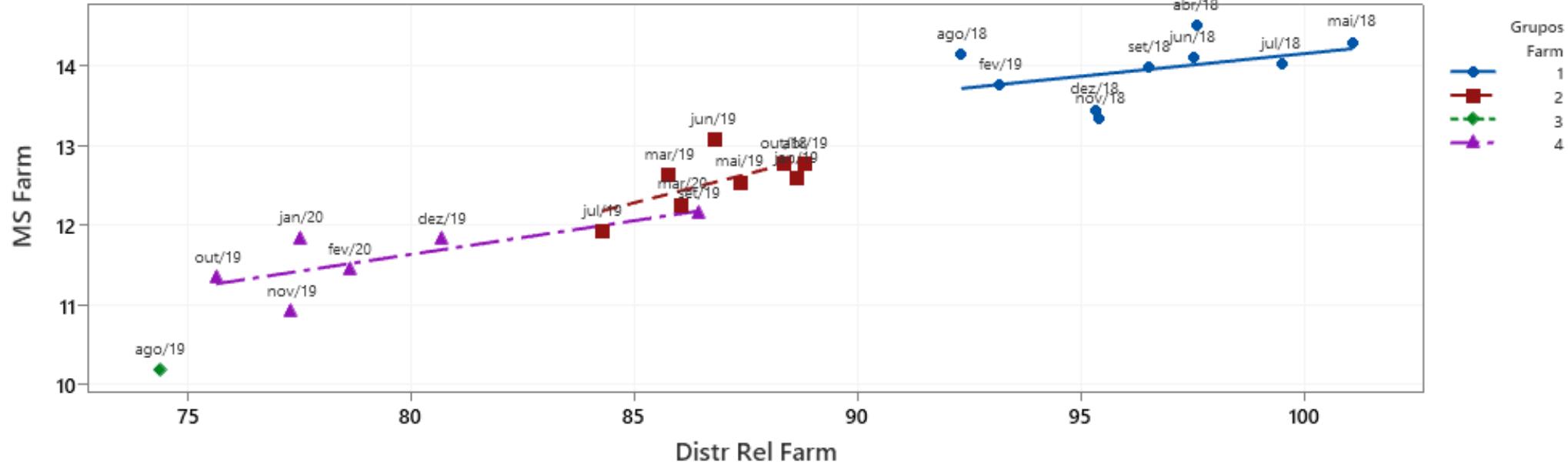
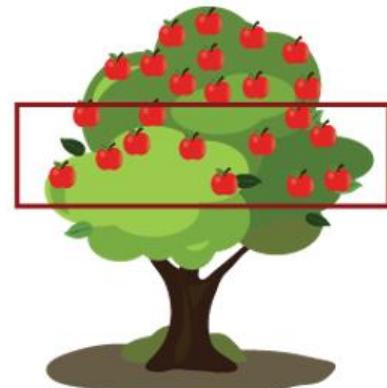


Gráfico de Dispersão de MS Farm versus Distr Rel Farm



Business  
Case – PDV  
Análise de  
Cluster  
Observações  
Mercado &  
Farmácia

# Atividade A.1 – Ciclo Iterativo Analyze



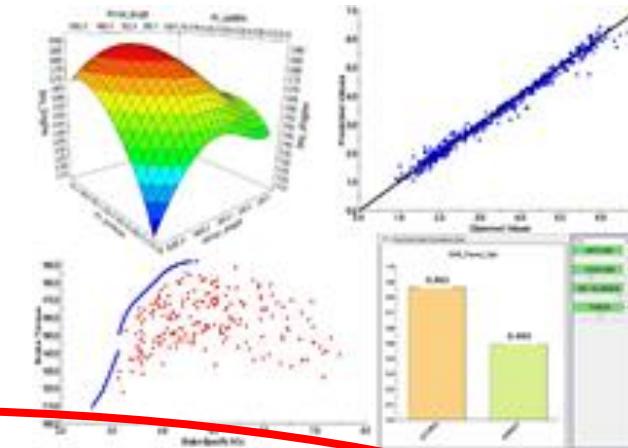
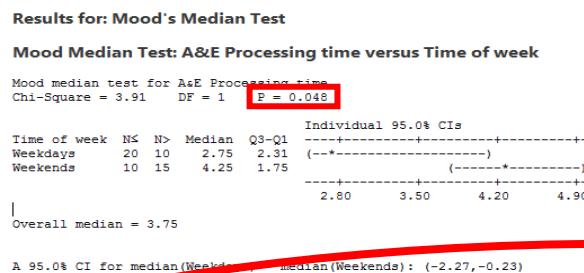
## Analyze Híbrido



## Identificar

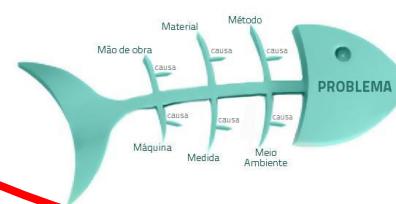
## A) Fatores Críticos & Impacto

### Ferramentas Gráficas e Analíticas



## **B) Riscos Potencias e Causas Raízes de Falhas**

### Ferramentas Qualitativas



5 Por quês?

## C) Melhorias Relacionadas com a consolidação das Variáveis Críticas

### Matriz Causa e Efeito

Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida
	CTQ 1	CTQ 2			
	Peso	Peso			
X1					
X2					
X3					
X4					
X5					
X6					
XN					

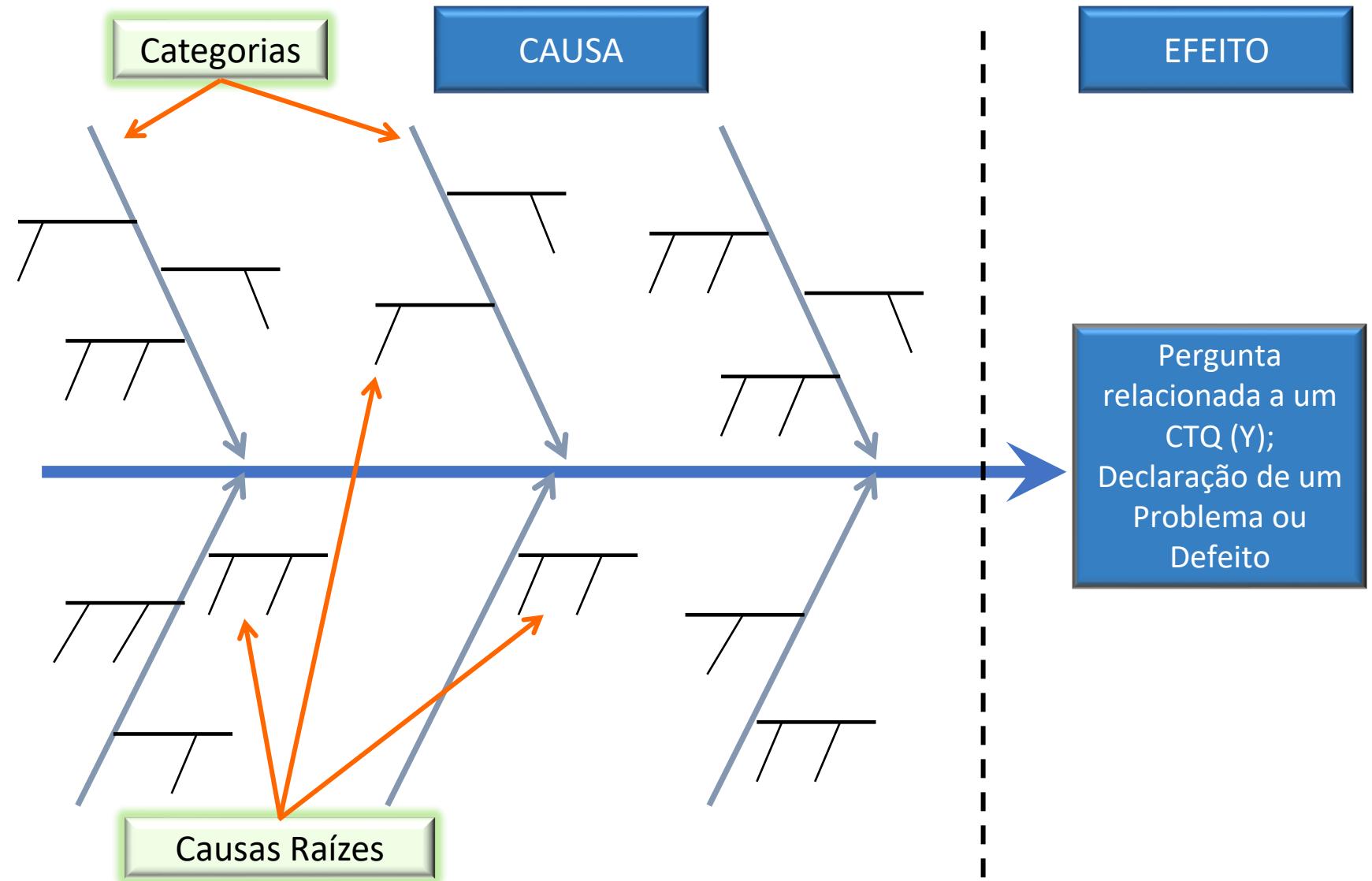
# Espinha de Peixe – Gráfico de Ishikawa ou de Causa & Efeito

# Espinha de Peixe

- Também conhecido como Diagrama de causa-efeito ou Diagrama de Ishikawa
- Permite o mapeamento do fatores (X's) que influenciam em um resultado (Y) na etapa de Measure
- Sugere possíveis causas de um problema para que sejam posteriormente confrontadas com os dados coletados na etapa de Analyze



# Espinha de Peixe

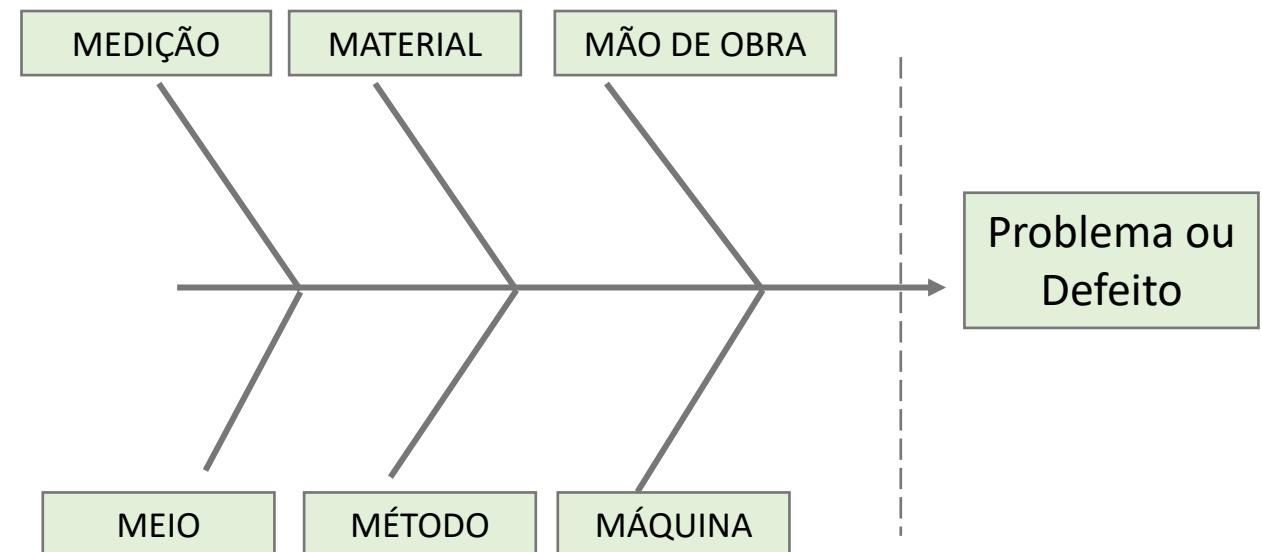


# Espinha de Peixe

As categorias descritas ao lado são conhecidas como 6 M's e servem como uma guia para a equipe pensar sobre as possíveis causas raízes que impactam o problema ou o defeito em questão. As categorias podem ter outros nomes

## Etapas na elaboração de um Diagrama de Causa e Efeito:

1. Desenhe um diagrama em branco sobre um quadro
2. Escreva o efeito (problema, defeito ou oportunidade) que se deseja investigar
3. Escreva os nomes selecionados para os eixos secundários (categorias que mais se adequem ao que se deseja investigar)

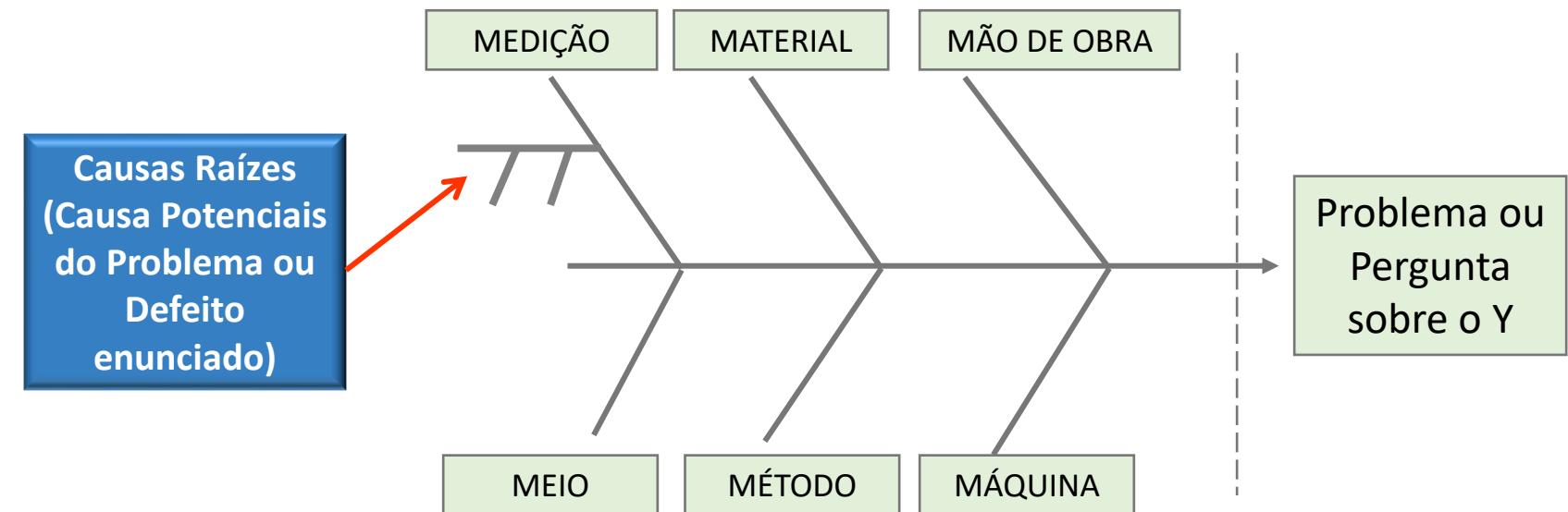


# Espinha de Peixe

Essencial: Fazer o Trabalho  
em Equipe!!

## Etapas na elaboração de um Diagrama de Causa e Efeito:

4. Faça um Brainstorming silencioso, com o uso de post-its, para identificar possíveis causas raízes e colocá-las junto às categorias apropriadas.
5. Permita um esclarecimento sobre as causas levantadas e agrupe as semelhantes
6. Identifique as mais prováveis e obtenha dados para permitir uma análise



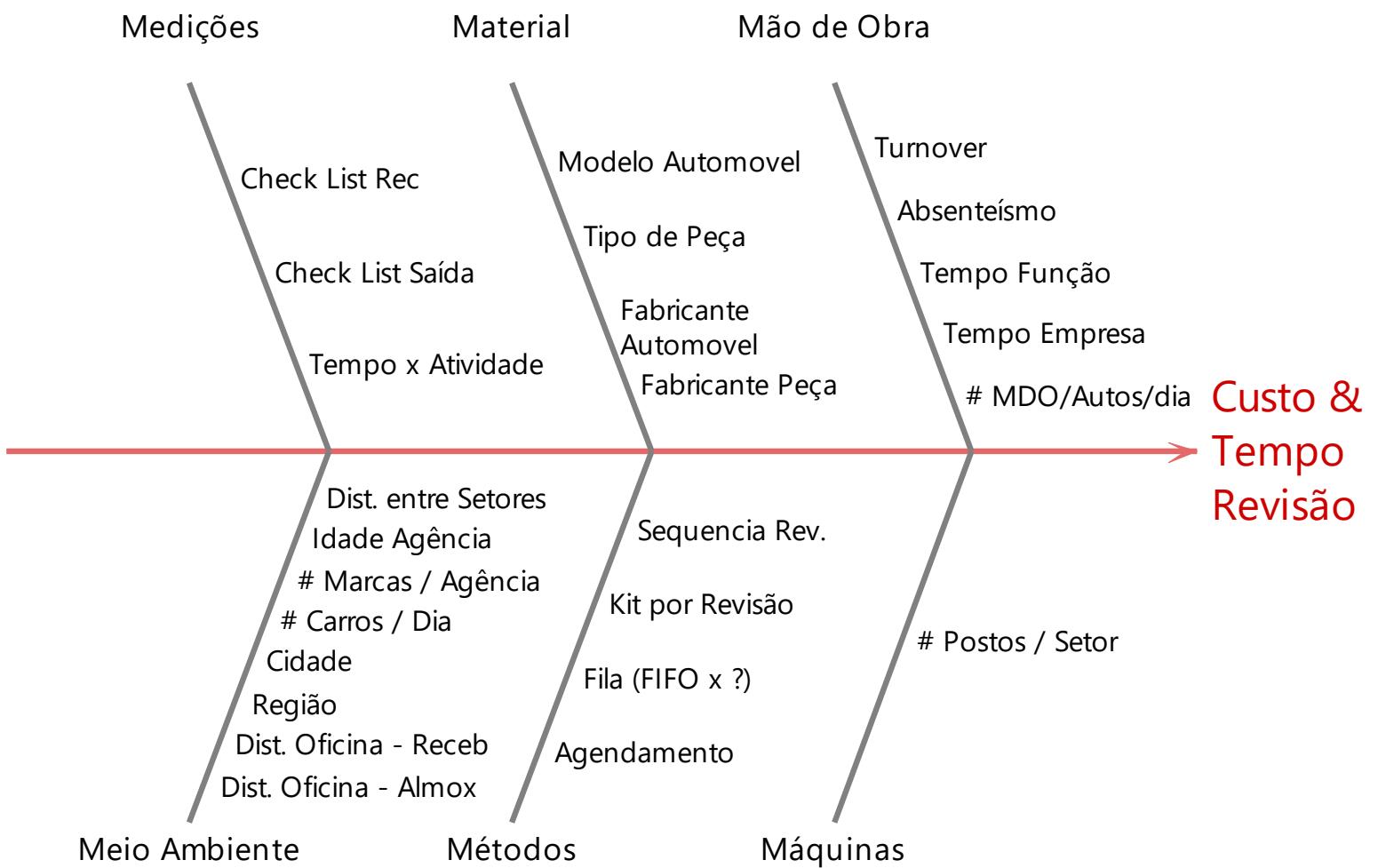
# Espinha de Peixe

## Measure

### Exemplo

Objetivo: identificar fatores (X's) que influenciam o resultado (Y – KPI do projeto)

### Diagrama de Causa e Efeito



# Espinha de Peixe Analyze

## Exemplo

Objetivo: identificar causas raízes de problemas que prejudicam o resultado (Y – KPI do projeto)

## Diagrama de Causa e Efeito

Medições                    Material                    Pessoal

Check List Rec mal preenchido

Check List Saída errado

Tempo x Atividade não controlado

Antes de férias e feriados

Oficina longe do Receb

Oficina longe do Almox

Material

Modelo Automovel errado

Peça com defeito

Peça errada

Falta de Peça

SOP não utilizado

Sem FIFO

Agendamento não efetuadp

Pessoal

Turnover elevado

Absenteísmo alto

Treinamento inadequado

Elevadores quebrados

Falta de ferramentas

Ambiente

Métodos

Máquinas

Custo e  
Tempo de  
Revisão  
Elevados

# Espinha de Peixe Analyze

Minitab

Arquivo: Espinha de  
Peixe.mpx

The screenshot shows the Minitab software interface with the 'Analyze' menu open. The 'Estat' tab is selected. The 'Ferramentas da Qualidade' option is highlighted in blue. A sub-menu for 'Causa e Efeito...' is displayed, also highlighted in blue. A tooltip for this option is shown in a callout box, explaining that it creates a fishbone diagram to identify potential causes and group them into categories.

Estatísticas Básicas	Gráficos de Ensaios...
Regressão	Gráfico de Pareto...
ANOVA	Causa e Efeito...
DOE (Planejamento de Experimento)	
Cartas de Controle	
Ferramentas da Qualidade	
Confiabilidade/Sobrevivência	
Análise preditiva	
Multivariada	
Séries Temporais	
Tabelas	
Não-Paramétricos	

# Espinha de Peixe Analyze

Minitab

Diagrama de Causa e Efeito

Galho	Causas	Rótulo	
1	da coluna	'Mão de Obra'	Pessoal
2	da coluna	Máquinas	Máquinas
3	da coluna	Material	Material
4	da coluna	Método	Métodos
5	da coluna	Medições	Medições
6	da coluna	'Meio Ambiente'	Ambiente
7			Sub...
8			Sub...
9			Sub...
10			Sub...

Efeito: Custo e Tempo de Revisão Elevados

Título:

Omitir o rótulo dos galhos

Omitir galhos vazios

Selecionar      Ajuda      OK      Cancelar

	C1-T	C2-T	C3-T	C4-T	C5-T	C6-T
	Medições	Material	Mão de Obra	Método	Máquinas	Meio Ambiente
1	Check List Rec mal preenchido	Modelo Automovel errado	Turnover elevado	Agendamento não efetuadp	Falta de ferramentas	Oficina longe do Almox
2	Check List Saída errado	Peça com defeito	Absenteísmo alto	Sem FIFO	Elevadores quebrados	Oficina longe do Receb
3	Tempo x Atividade não controlado	Peça errada	Treinamento inadequado	SOP não utilizado		Antes de férias e feriados
4		Falta de Peça				

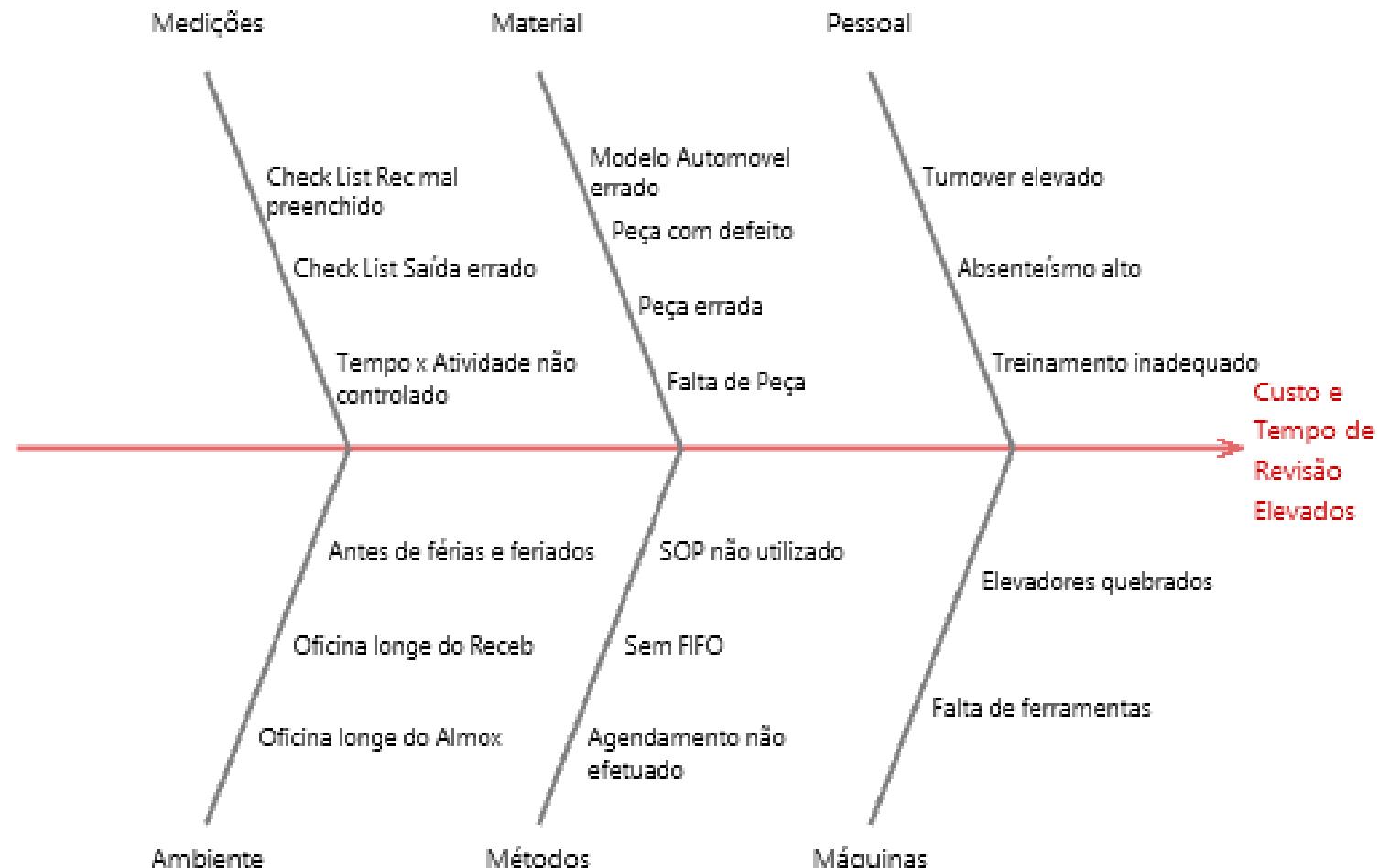
Arquivo: Espinha de  
Peixe.mpx

# Espinha de Peixe Analyze

Minitab

Arquivo: Espinha de  
Peixe.mpx

## Diagrama de Causa e Efeito



Worksheet: Analyze

# TO DO LIST

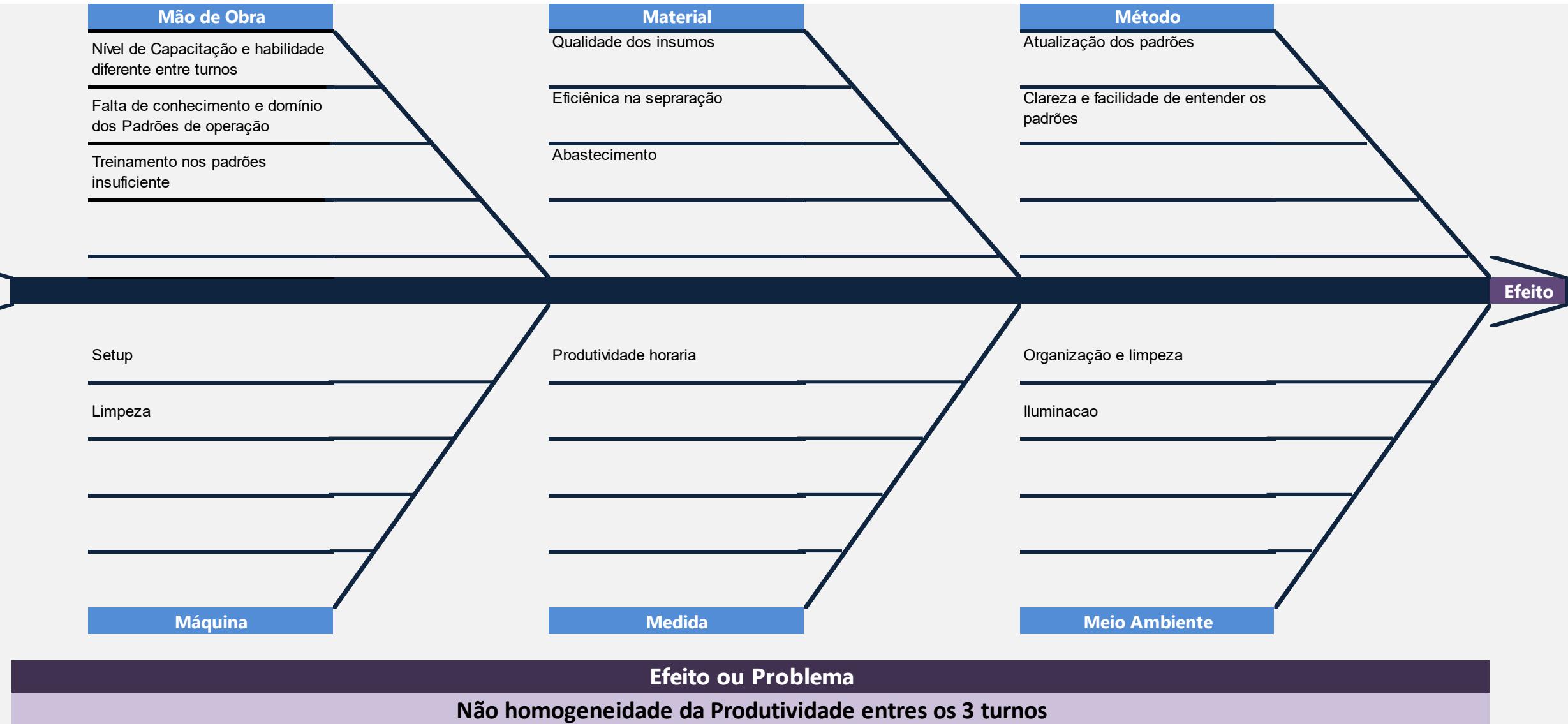
- 
- 
- 
- 
- 

Tema Aquecimento Global

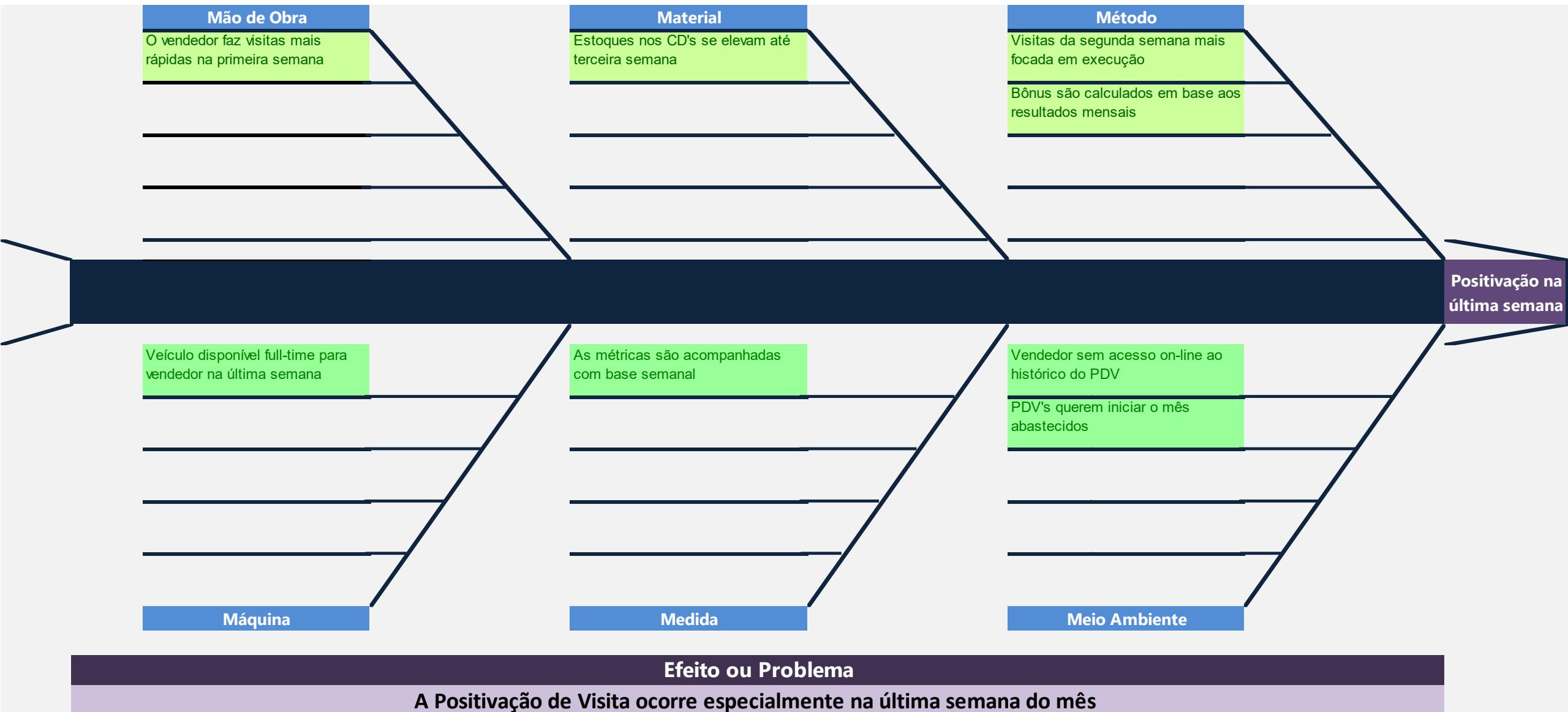
Faça Espinha de Peixe – Measure (X's)

Faça Espinha de Peixe – Analyze  
(causas raízes do problema)

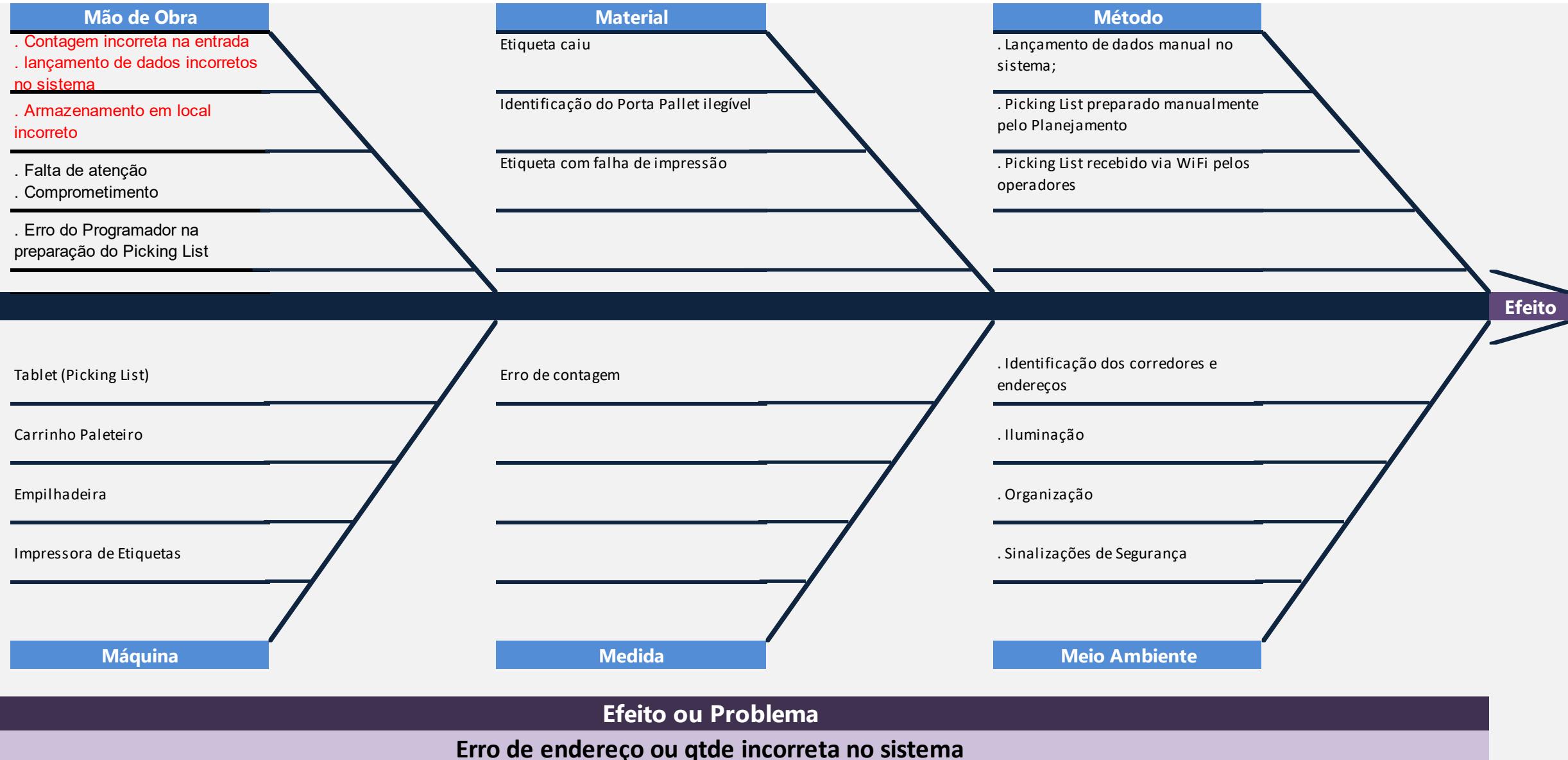
# Espinha de Peixe – Projeto Produtividade



# Espinha de Peixe – Projeto Mercado PDV



# Espinha de Peixe – Projeto Logística



# 5 Por quês?

# 5 Por quês?

Normalmente utilizado para buscar causas raízes de um problema

## Exemplo

### **1. Por que o carro não ligou ?**

Porque a bateria ficou sem carga

### **2. Por que a bateria descarregou?**

Porque o carro ficou muito tempo desligado

### **3. Por que o carro ficou muito tempo desligado?**

Porque na Pandemia não precisou ligar o carro.

### **4. Por que não precisou ligar o carro na Pandemia?**

Por que o carro só é ligado para sair

### **5. Por que o carro só é ligado para sair?**

Porque não tem rotina de ligar o carro periodicamente

# 5 Por quês?

Exemplo: queda de produtividade nos turnos

Sintoma?

Queda de Produtividade

Por Quê?

Queda das horas Trabalhadas nas Semanas

Por Quê?

Aumento no numero de funcionários que chegam atrasados

Por Quê?

A quantidade de ônibus urbanos foi reduzida na pandemia

# 5 Por quês? – Projeto Produtividade

**Problema:** *Nível de Capacitação e habilidade diferente entre turnos*

**1º Porque?** Por que o nível de capacitação e habilidade está diferente entre turnos?

**Resp.:** Devido à baixa fixação dos conceitos e técnicas do treinamento

**2º Porque?** Por que a baixa fixação d a baixa fixação dos conceitos?

**Resp.:** devido ao treinamento ter sido muito teórico

**3º Porque?** Por que o treinamento foi muito teórico?

**Resp.:** Porque o treinamento foi elaborado nos padrões e procedimentos atuais.

**4º Porque?** Por que o treinamento foi elaborado nesses padrões?

**Resp.:** devido à falta de um modelo para treinamento prático.

**5º Porque?** Por que não temos um modelo de treinamento prático?

**Resp.:** Pela falta de uma didática focada em andragogia.

**Ação:** A4. Desenvolver um treinamento moldado nos conceitos de Andragogia e reciclar os operadores.

A5. Implementar Matriz de Versatilidade para a gestão da capacitação do pessoal

# 5 Por quês? – Projeto Mercado PDV

**Problema:** O vendedor faz visitas mais rápidas na primeira semana

- 1 Por que? Ele necessita voltar para o escritório antes do final do dia
- 2 Por que? Para acessar o sistema de vendas
- 3 Por que? Para efetuar os relatórios mensais
- 4 Por que? Os gráficos e relatórios não são automatizados
- 5 Por que? Os dados são dispersos em várias telas

**Ação:** Automatizar elaboração de gráficos e relatórios e disponibilizar o sistema para os vendedores

# 5 Por quês? – Projeto Logística

**Problema:** *Contagem incorreta na entrada*

**1º Por que?** Por que acontece a contagem incorreta na entrada?

**Resp.:** Devido a um acúmulo de material no setor o que dificulta uma boa contagem.

**2º Por que?** Por que está tendo um acúmulo de material no setor?

**Resp.:** Porque estamos descarregando vários veículos simultaneamente.

**3º Por que?** Por que estamos descarregando vários veículos ao mesmo tempo.

**Resp.:** Estamos tendo horários de pico com a chegada de vários veículos (picos e vales)

**4º Por que?** Por que estamos tendo horários de pico?

**Resp.:** Porque estamos falhando na disciplina do seguimento da janela de entrega.

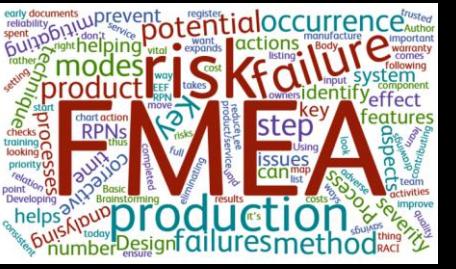
**5º Por que?** Por que Porque não estamos seguindo a janela de entregas?

**Resp.:** Porque perdemos o controle com o aumento da demanda

**Ação:** A8. Retomar a Janela de Entrega sendo mandatório o seu seguimento, sem exceções.

A9. 5S - Demarcação e organização do setor.

# FMEA – Análise de Efeito e Modos de Falha



# FMEA



# Edward Murphy



John Stapp

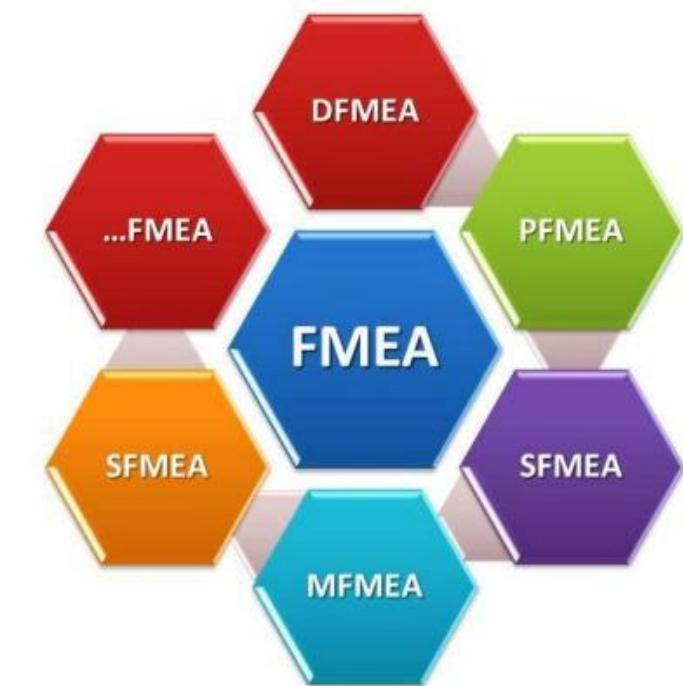


SE ALGUMA COISA PODE  
DAR ERRADO, DARÁ. E  
MAIS, DARÁ ERRADO DA  
PIOR MANEIRA, NO PIOR  
MOMENTO, E DE MODO  
QUE CAUSE O MAIOR  
DANO POSSÍVEL.



# FMEA

Failure Mode Effect Analysis

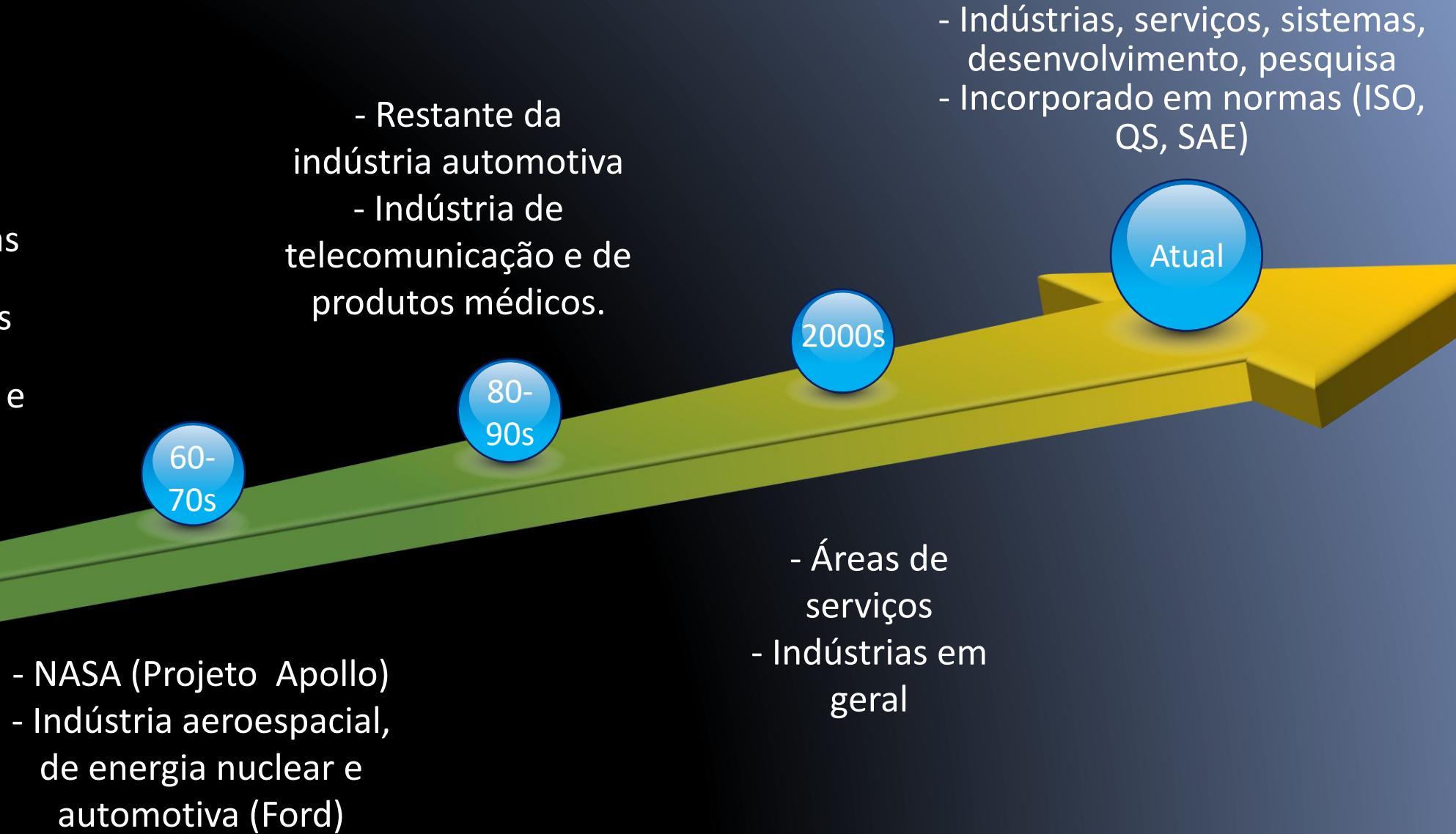


Risk



# FMEA – Histórico (Desenvolvimento & Uso)

FMEA foi desenvolvido pelas Forças Armadas Norte Americanas para minimizar perdas de pessoal e equipamentos



# FMEA

## Objetivos do uso:

- Identificar, quantificar e priorizar os riscos potenciais de maneira preventiva
- Identificar os modos de falhas, as severidades de seus efeitos, as suas causas e as probabilidades das ocorrências destas
- Identificar os atuais meios para detectar e/ou evitar a ocorrência destas causas e seus efeitos (falhas)
- Identificar e priorizar ações para minimizar os riscos por quantificar sua eficiência na redução dos riscos

# FMEA

## Benefícios:

- Foco na prevenção e não na correção
- Melhoria de qualidade, confiabilidade, segurança , produtividade e custos de produtos, processos, serviços, sistemas, desenvolvimento
- Documentação e rastreabilidade das ações recomendadas para mitigação do risco
- Gestão do conhecimento e do risco, com informações que podem ser utilizadas no futuro

# FMEA - Planilha

## Mapa de Processo



X1  
X2  
X3

- Dica: inicie por
- X's críticos qualitativos
  - X's críticos quantitativos
  - Etapas Críticas

Risk Probability Number

$$RPN = SEV * OCC * DET$$

Quanto mais elevado, maior o risco,  
maior a prioridade de ação!

Etapa Proc	Entradas (X's)	Modo de Falha	Efeito da Falha	SEV	Causas	OCC	Controles	use max SEV
Pode ser uma etapa do processo, atividade geral, área ou outro agrupamento racional.	Quais são as Entradas (X) do Processo ?	Como a Entrada (X) pode falhar? Que tipos de falhas podem ocorrer?	Se a falha ocorre qual será o seu efeito?	S E V	Quais são as causas que levam à esta falha?	O C C	Quais são os controles existentes para prevenir a ocorrência da causa ou para detectá-la?	D E T R P N

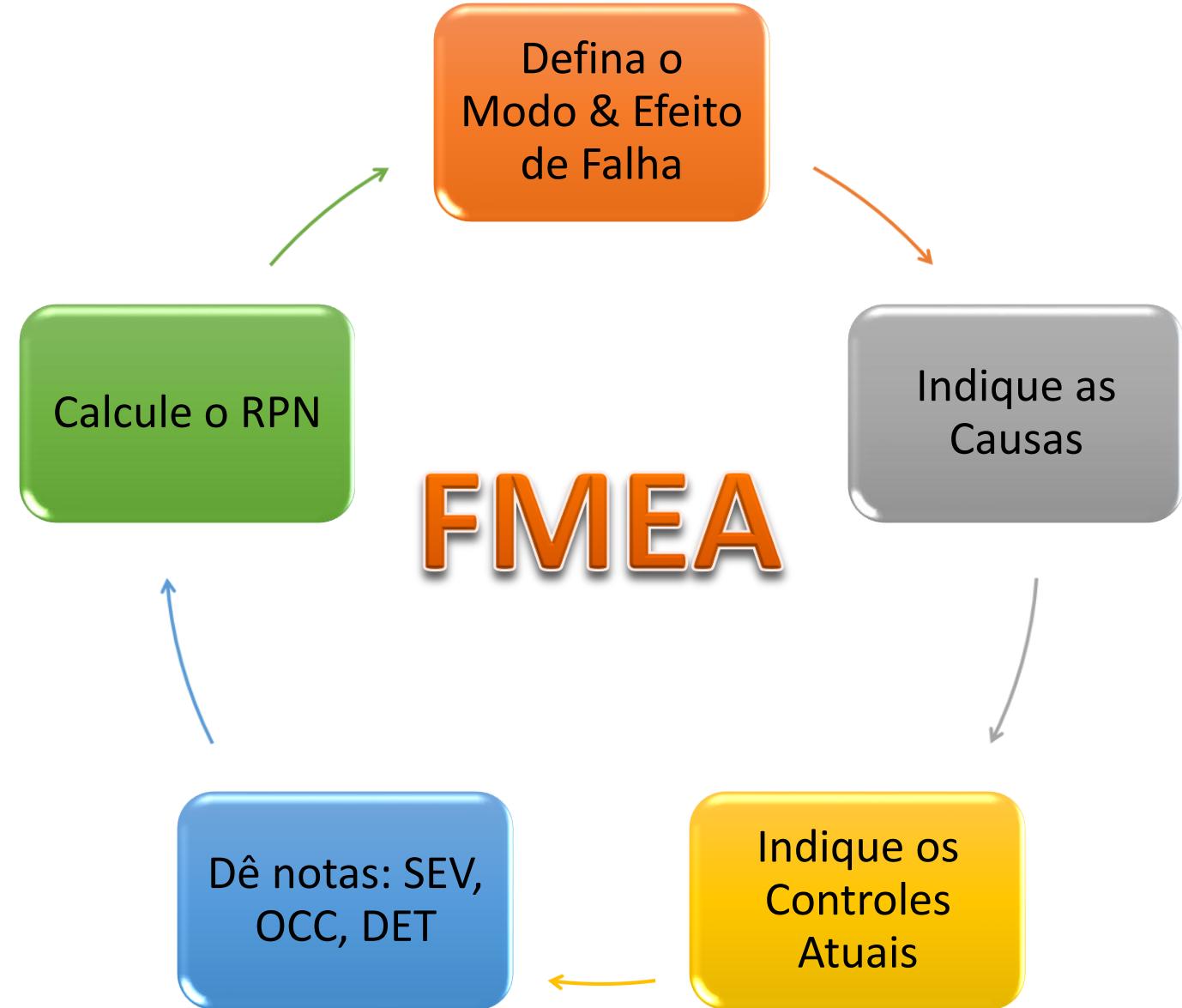
Severidade = quão sério é o efeito da causa para um cliente ou o KPI do projeto

Ocorrência = quão provável é que venha a ocorrer a causa que leva à falha

Detecção = qual a ineficiência do processo em detectar e/ou evitar a ocorrência da causa

# FMEA - Lógica

Sequência de ações para as Etapas e X's selecionados (críticos) para a Análise



## SEVERIDADE

Índice	Descrição	Referência
1	<ul style="list-style-type: none"><li>Uma oportunidade de melhorada é identificada, mas não é necessário tomar nenhuma ação. Consequência da Falha é negligenciável.</li></ul>	Sem danos detectados.
2	<ul style="list-style-type: none"><li>Consequência da falha é leve. O defeito provavelmente não será perceptível para o cliente.</li></ul>	Danos baixos para equipamentos e/ou meio ambiente.
3	<ul style="list-style-type: none"><li>Consequência da falha é moderada. Alguns clientes podem detectar e rejeitar o serviço/produto.</li></ul>	Danos moderados para equipamentos e/ou meio ambiente.
4	<ul style="list-style-type: none"><li>Produto, processo ou serviço é severamente degradado. O Cliente detectará e poderá trabalhar com o problema ou rejeitá-lo.</li></ul>	Danos significativos para equipamentos e/ou meio ambiente.
5	<ul style="list-style-type: none"><li>Consequência da falha é severa. O Produto ou serviço não funcionará ou não terá a performance necessitada pelo cliente. Dano mais elevado a algum KPI do Projeto.</li></ul>	Danos à pessoas, e/ou severos para equipamentos e/ou meio ambiente.

## FMEA – Pontuação

Pontuações:

- As tabelas apresentadas a seguir podem ser utilizadas
- Outras tabelas podem ser obtidas para casos específicos
- O próprio grupo do Projeto poderá definir seus critérios

## OCORRÊNCIA

Índice	Descrição	Referência
1	• Remota possibilidade de ocorrência. Improvável.	<= 1PPM
2	• Baixa taxa de ocorrência de falha (poucas falhas).	10 - 100 PPM
3	• Taxa de falha moderada (falha ocasional).	500 - 5000 PPM
4	• Taxa de falha frequente.	10000 - 20000 PPM
5	• Alta probabilidade de falha. É quase certo que a falha ocorrerá. Falha persistente.	> 50000 PPM

## FMEA – Pontuação

Pontuações:

- As tabelas apresentadas a seguir podem ser utilizadas
- Outras tabelas podem ser obtidas para casos específicos
- O próprio grupo do Projeto poderá definir seus critérios

# DETECÇÃO

Índice	Descrição	Método
1	<ul style="list-style-type: none"> <li>• Mais alta probabilidade do defeito ser detectado antes de chegar ao cliente.</li> <li>• Um controle que praticamente previne uma ocorrência da falha.</li> </ul>	
2	<ul style="list-style-type: none"> <li>• Chance muito alta do defeito ser detectado antes de chegar ao cliente.</li> <li>• Um controle que tipicamente previne a ocorrência do defeito</li> </ul>	
3	<ul style="list-style-type: none"> <li>• Moderada chance do defeito ser detectado antes de chegar ao cliente.</li> <li>• Um controle que detecta a causa após esta ter ocorrido.</li> </ul>	
4	<ul style="list-style-type: none"> <li>• Baixa chance de detectar o defeito antes de chegar ao cliente.</li> <li>• Poucos ou nenhum controle para prevenir a causa, alguns controles para detectar a causa</li> </ul>	
5	<ul style="list-style-type: none"> <li>• Mais baixa probabilidade do defeito ser detectado antes de chegar ao cliente. O cliente pode fazer uma reclamação contra o produto/serviço ou pode rejeitá-lo.</li> <li>• Sem controle consistente para prevenir ou detectar a causa</li> </ul>	

## FMEA – Pontuação

Pontuações:

- As tabelas apresentadas a seguir podem ser utilizadas
- Outras tabelas podem ser obtidas para casos específicos
- O próprio grupo do Projeto poderá definir seus critérios

# FMEA – Propostas Melhorias & PRPN

Pós Risk Probability Number

$$\text{PRPN} = \text{PSEV} * \text{POCC} * \text{PDET}$$

Quanto maior a diferença (RPN – PRPN),  
melhor será a ação proposta

use max SEV	Ação	Resp	Prazo				
R P N	Quais ações específicas são necessárias para melhorar reduzir o RPN?	Quem será encarregado da implementação da ação?	Quando a ação requerida ou planejada estará implementada?	P S	P O	P D	P R P N

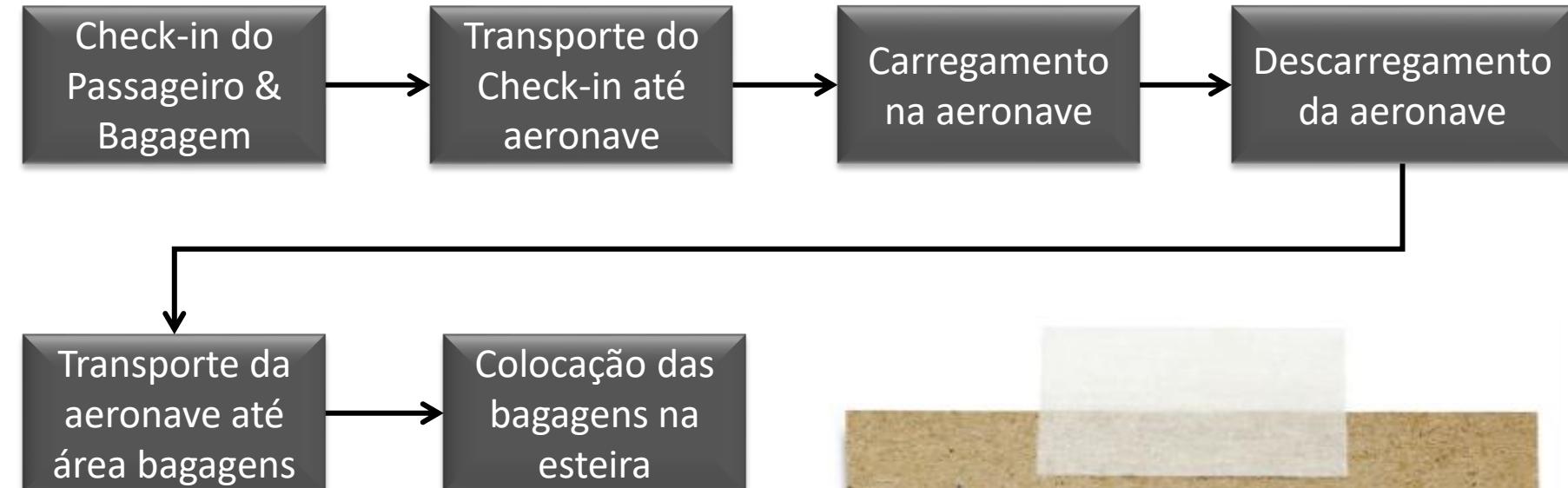
Pós Severidade = quanto sério será o efeito da causa para um cliente ou o KPI do projeto após a implantação da melhoria proposta

Pós Ocorrência = quanto provável será que venha a ocorrer a causa que leva à falha após a implantação da melhoria proposta

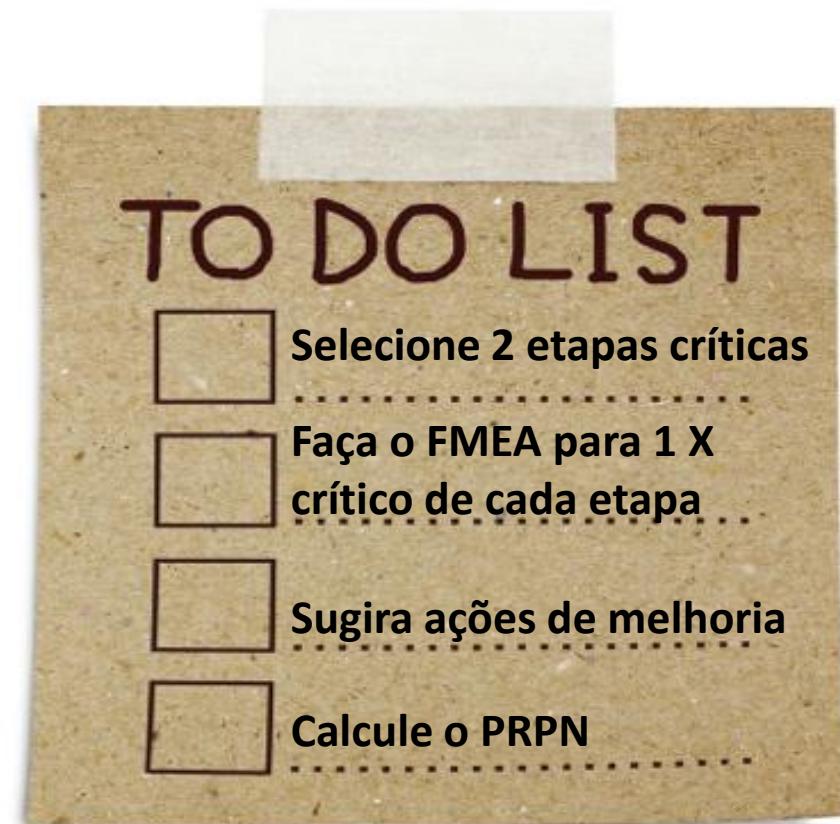
Pós Detecção = qual será a ineficiência do processo em detectar e/ou evitar a ocorrência da causa após a implantação da melhoria proposta

# FMEA - Exercício

Com o objetivo de reduzir o elevado nível de bagagens extraviadas de uma companhia aérea um GB decidiu utilizar o FMEA no processo ao lado



Arquivo:  
FMEA – Completo.xlsx



# FMEA – Projeto Produtividade

Etapa Proc	Entradas (X's)	Tipo de Falha	Efeito da Falha	S E V	Causas	O C C	Controles	use max SEV
Pode ser uma etapa do processo, atividade geral, área ou outro agrupamento racional.	Quais são as Entradas (X) do Processo ?	Como a Entrada (X) pode falhar? Que tipos de falhas podem ocorrer? Ou qual o resultado indesejado desta área?	Se a falha ocorre qual será o seu efeito?	S E V	Quais são as causas que levam à esta falha?	O C C	Quais são os controles existentes para prevenir a ocorrência da causa ou para detectá-la?	D E T R P N
Separação de Matéria Prima e garrafas plasticas 500ml	Eficiência no Abastecimento	Separação de insumos a menor	- Parada de linha e perda de produção - Reabastecimento.	7	- OP incorreta; - Falta de atenção; - Falta de condições adequadas para separação;	7	não tem	10 490
		Separação de insumos incorreta	- Parada de linha, perda de produção e reabastecimento; - Formulação Incorreta gerando retrabalho.	8	- OP incorreta; - Falta de atenção; - Falta de condições adequadas para separação; - Identificação dos materiais inadequada.	5	não tem	10 400
Pesagem e preparação da receita	Precisão na formulação	Formulação incorreta	- Perda de produção; - Correção com nova formulação e perda de material	8	Falta de aderencia aos padrões de formulação	4	Equipe treinada na IT de Formulação revisada (Sprint Define)	3 96
		Pesagem incorreta dos insumos	- Perda de produção; - Correção com nova formulação e perda de material	8	Balança descalibrada	4	Controle de Calibração	3 96
Envase	Envasamento fora do padrão	- Segregação e rejeição do lote fora de especificação; - Retrabalho para acerto do volume de envase.	7	Densidade fora de especificação	5	densímetro em laboratório a cada batch	3	105
				Ajuste de tempo de envase inconsistente	4	não tem	10	280
				Equipamento perdendo ajuste	3	não tem	10	210

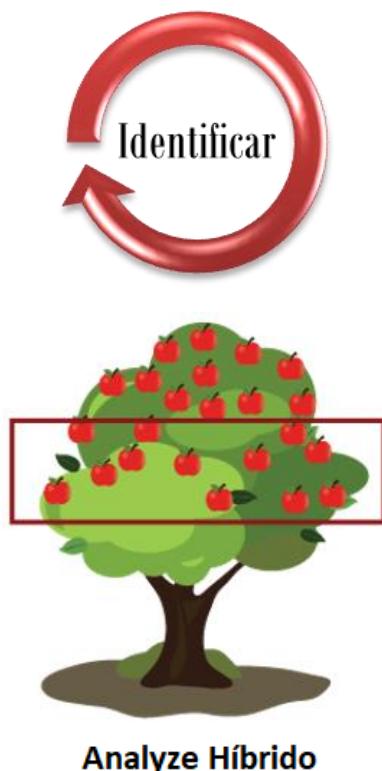
# FMEA – Projeto Mercado PDV

Etapa Proc	Entradas (X's)	Tipo de Falha	Efeito da Falha	S E V	Causas	O C C	Controles	use max SEV
Pode ser uma etapa do processo, atividade geral, área ou outro agrupamento racional.	Quais são as Entradas (X) do Processo ?	Como a Entrada (X) pode falhar? Que tipos de falhas podem ocorrer? Ou qual o resultado indesejado desta área?	Se a falha ocorre qual será o seu efeito?	S E V	Quais são as causas que levam à esta falha?	O C C	Quais são os controles existentes para prevenir a ocorrência da causa ou para detectá-la?	D E T  R P N
Visita Vendedor	Share de Prateleira	Não medir	Não corrigir o Share e perder vendas	10	Não ter tempo	4	Sem controle	10 400
		Medir errado	Não corrigir o Share e perder vendas	10	Não estar treinado	7	Formulário de Integração	6 420
Execução em Loja	Planograma	Não ter o Planograma	Não efetuar corretamente execução de Prateleira e perder vendas	10	Planograma não disponibilizado	8	Sem controle	10 800
		Não usar o Planograma	Não efetuar corretamente a execução de Prateleira e perder vendas	10	Não estar treinado	8	Sem controle	10 800
	Material de Execução	Não ter o Material	Não efetuar a execução do Material	8	Material acabou no estoque	8	Visual no sistema	6 384
		Utilizar mal o Material	Efetuar mal a execução do Material	7	Não estar treinado	8	Sem controle	10 560

# FMEA – Projeto Logística

Etapa Proc	Entradas (X's)	Tipo de Falha	Efeito da Falha	S E V	Causas	O C C	Controles	use max SEV
Pode ser uma etapa do processo, atividade geral, área ou outro agrupamento racional.	Quais são as Entradas (X) do Processo ?	Como a Entrada (X) pode falhar? Que tipos de falhas podem ocorrer? Ou qual o resultado indesejado desta área?	Se a falha ocorre qual será o seu efeito?	S E V	Quais são as causas que levam à esta falha?	O C C	Quais são os controles existentes para previnir a ocorrência da causa ou para detectá-la?	D E T R P N
Recebimento e Armazenamento	Nota Fiscal	Dados incorretos	Impossibilidade de receber o material	5	Preenchimento incorreto pela Fábrica	2	Inspeção manual na entrada	4 40
	Procedimento de Análise, Conferência e Movimentação	. Lançamento de dados incorreta no sistema	. Erro de inventário	5	. Não cumprir o procedimento; . Erro na execução.	3	Não tem	4 60
	Local de Armazenamento	Armazenar no endereço incorreto	. Vai gerar um erro no sistema, complicação durante picking	5	Desatenção do operador	3	Contagem Cíclica	3 45
		Endereço já ocupado por outro material	Impossibilidade de armazenar, atraso, retrabalho	5	Decorrente da falha anterior (Armazenar no endereço incorreto)	3	Contagem Cíclica	3 45
	Equipamentos	quebras, indisponibilidade	Tempo perdido - atrasos	4	Gestão de equipamentos - Preventivas não realizadas - utilização por outra unidade	3	Plano de Preventivas Não tem controle para utilização	3 36
Entrada e Análise dos Pedidos	Pedidos dos Clientes	indisponibilidade parcial ou completa dos itens do pedido	. Cliente não atendido; . Cliente busca outro fornecedor	5	. Falha no controle de inventário e programação	3	Controle sistemico do inventário	2 30
Roteirização	Software de Roteirização	Sistema do software não operante	Atraso no atendimento do pedido	4	Ausência de sistema	2	-----	5 40
	Padrão de Roteirização	Erro de endereço incorreto no sistema	. Impossibilidade de efetuar a entrega; . Retrabalho/atraso;	5	. Erro de digitação na entrada de dados no sistema; . Endereço desatualizado/não confere.	2	Experiência do motorista	4 40
		Não seguimento do roteiro	Excesso de movimentação e atraso na entrega	3	Desobediência do Motorista.	2	Controle de Kilometragem	2 12

# Atividade A.1 – Ciclo Iterativo Analyze



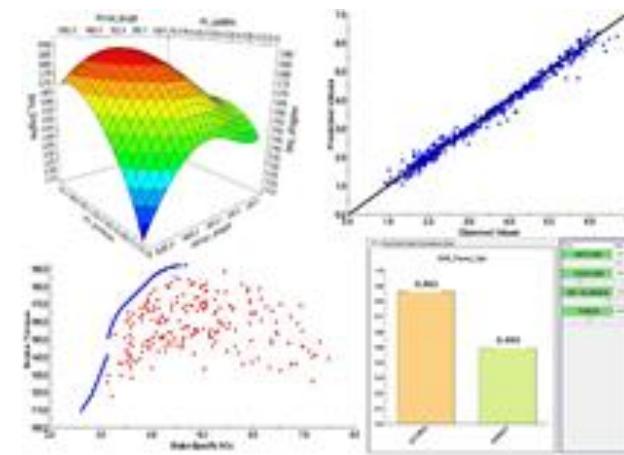
## A) Fatores Críticos & Impacto Ferramentas Gráficas e Analíticas

Results for: Mood's Median Test  
Mood Median Test: A&E Processing time versus Time of week  
Mood median test for A&E Processing time  
Chi-Square = 3.91 DF = 1 **P = 0.048**

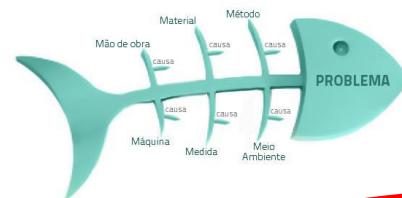
Individual 95.0% CIs					
Time of week	N<	N>	Median	Q3-Q1	(--*)-----
Weekdays	20	10	2.75	2.31	(--*)-----
Weekends	10	15	4.25	1.75	(---*---)

Overall median = 3.75

A 95.0% CI for median(Weekdays) - median(Weekends): (-2.27, -0.23)



## B) Riscos Potencias e Causas Raízes de Falhas Ferramentas Qualitativas



5 Por quês?

Etapa / entrada do Processo	Modo de Falha Potencial	Efeito da Falha	S m	Causas Potenciais	O cc	Controles atuais	D E T	R P N
Qual a etapa / entrada do processo sobre investigação?	Em quais caminhos a entrada prejudica?	Qual é o impacto sobre as variáveis de saída (Exigências do Cliente) ou exigências internas?	Qual grave é o efeito para o cliente?	Quais as causas para a entrada atuar prejudicialmente?	Com que frequência as causas ocorrem?	Quais são os controles e procedimentos existentes (Inspeção, teste, etc.) que podem detectar a Causa ou o Modo de Falha? Deveria incluir um número SOP.	Qual é bem você pode detectar a causa ou o Modo de Falha?	Ser/Outro/Dat 0

## C) Melhorias Relacionadas com a consolidação das Variáveis Críticas Matriz Causa e Efeito

Principais Entradas (X's)	Principais Saídas (Y's)			Efeto	Melhoria Sugerida
	CTQ 1	CTQ 2	CTQ N		
	Peso	Peso	Peso		
X1					
X2					
X3					
X4					
X5					
X6					
XN					

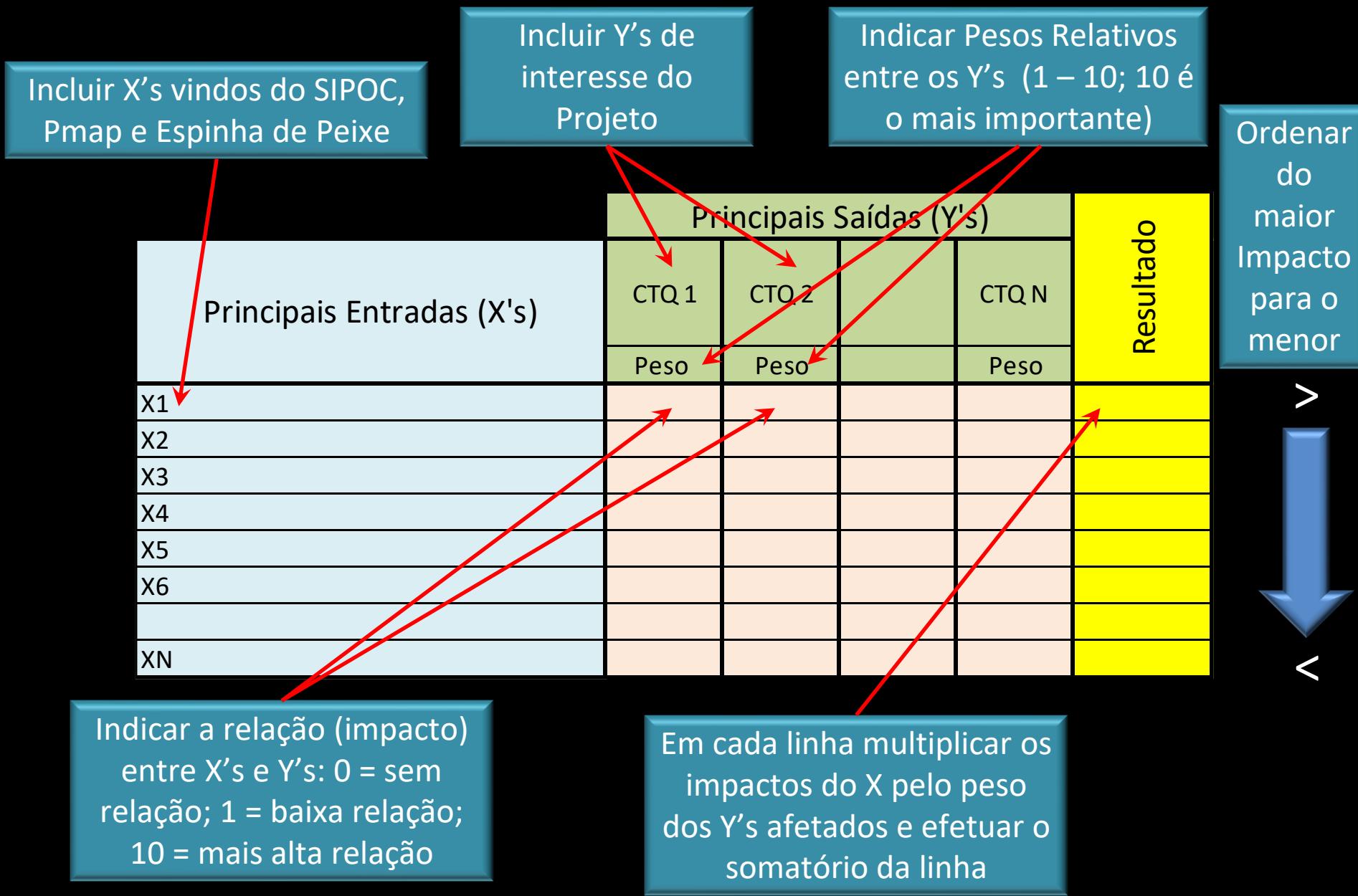
# MATRIZ CAUSA x EFEITO

Principais Entradas (X's)	Principais Saídas (Y's)				Resultado
	CTQ 1	CTQ 2		CTQ N	
	Peso	Peso		Peso	
X1					
X2					
X3					
X4					
X5					
X6					
XN					

# Matriz Causa e Efeito

- Seu objetivo é classificar os X's (entradas do processo) de acordo com o impacto (efeito) destes nos Y's (CTQ's ou KPI's do projeto)
- É um sumário das conclusões em relação aos X's (principais entradas do processo identificadas através do SIPOC, Mapa de Processo, Espinha de Peixe) através de ferramentas gráficas, analíticas e qualitativas
- É fonte de alimentação da Matriz Esforço x Impacto

# Matriz Causa e Efeito – Elaboração



# CICLO ITERATIVO – ANALYZE HÍBRIDO



## 1. Identificar

Fatores Críticos & Impacto

Riscos potenciais & Causas Raiz

Melhorias relacionadas com a consolidação  
das variáveis críticas



### Ideação – Técnicas e Ferramentas

#### Brainstorming

- Técnica para estimular a geração de um grande número de ideias em um curto espaço de tempo.

#### Workshop de Cocriação

- Encontro da equipe focado em realizar atividades em grupo com o objetivo de estimular a criatividade e a colaboração, fomentando a criação de soluções inovadoras.

#### Cardápio de Ideias

- Catálogo compilando a síntese de todas as sugestões e ideias geradas no processo de ideação. Pode incluir comentários relativos às ideias, eventuais desdobramentos e oportunidades de negócios.

#### Matriz de Posicionamento

- Ferramenta de análise estratégica das ideias geradas, utilizada na validação destas em relação aos Critérios Norteadores, bem como às necessidades das Personas criadas no projeto.



### IDEAÇÃO

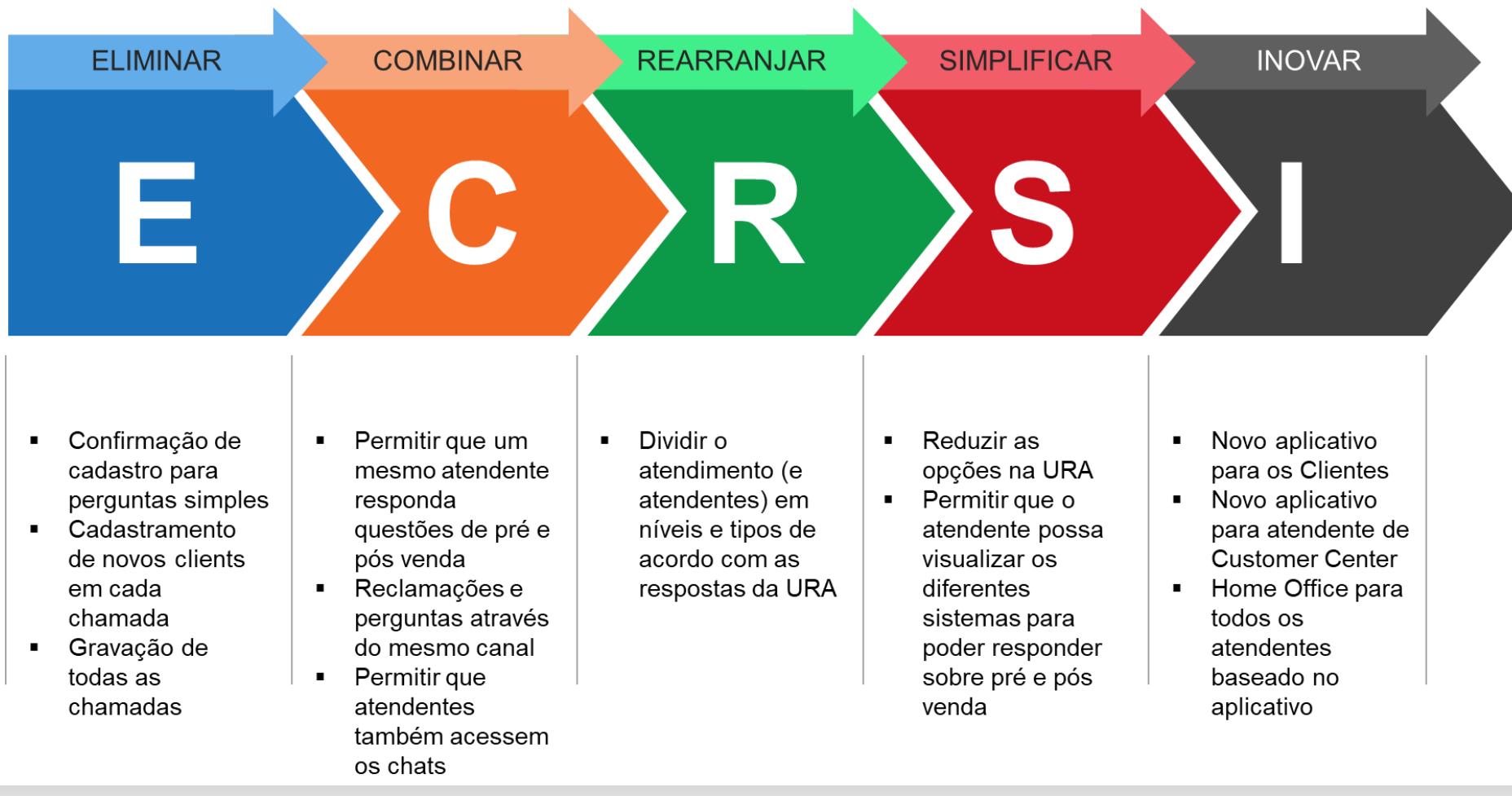


- Processo de formação de ideias e conceitos para resolver problemas específicos que os atores do processo (funcionários, fornecedores e clientes) estão enfrentando no momento
- Reunir uma equipe com diversidade criativa para geração de ideias inovadoras
- Reunir equipe multidisciplinar e multifuncional que combine talentos e criatividade para gerar novas ideias inovadoras
- Não existe uma solução ideal, o objetivo é gerar muitas ideias e Soluções possíveis que permita aprendizado e convergência para a solução final do cliente
- É a etapa que exigirá maior criatividade, pensamento disruptivo



# IDEAÇÃO

- Processo de formação de ideias e conceitos para resolver problemas específicos que os atores do processo (funcionários, fornecedores e clientes) estão enfrentando no momento
- Reunir uma equipe com diversidade criativa para geração de ideias inovadoras
- Reunir equipe multidisciplinar e multifuncional que combine talentos e criatividade para gerar novas ideias incríveis
- Não existe uma solução ideal, o objetivo é gerar muitas ideias e Soluções possíveis que permita aprendizado e convergência para a solução final do cliente
- É a etapa que exigirá maior criatividade, pensamento disruptivo



## Matriz ECRSI

- Auxilia a identificar diferentes opções de melhorias
- Sumariza as melhorias selecionadas por diferentes tipos

# Ideação – Técnicas e Ferramentas

## Brainstorming

- Técnica para estimular a geração de um grande número de ideias em um curto espaço de tempo.

## Workshop de Cocriação

- Encontro da equipe focado em realizar atividades em grupo com o objetivo de estimular a criatividade e a colaboração, fomentando a criação de soluções inovadoras.

## Cardápio de Ideias

- Catálogo compilando a síntese de todas as sugestões e ideias geradas no processo de ideação. Pode incluir comentários relativos às ideias, eventuais desdobramentos e oportunidades de negócio.

## Matriz de Posicionamento

- Ferramenta de análise estratégica das ideias geradas, utilizada na validação destas em relação aos Critérios Norteadores, bem como às necessidades das Personas criadas no projeto.

## Matriz Causa x Efeito & Melhorias

Principais Entradas (X's)	Principais Saídas (Y's)				Efeito	Melhoria Sugerida
	CTQ 1	CTQ 2		CTQ N		
	Peso	Peso		Peso		
X1						
X2						
X3						
X4						
X5						
X6						
XN						

- Seu objetivo é classificar os X's (entradas do processo) de acordo com o impacto (efeito) destes nos Y's (CTQ's ou KPI's do projeto)
- É um sumário das conclusões em relação aos X's (principais entradas do processo identificadas através do SIPOC, Mapa de Processo, Espinha de Peixe) através de ferramentas gráficas, analíticas e qualitativas
- É fonte de alimentação da Matriz Esforço x Impacto

# CICLO ITERATIVO – ANALYZE HÍBRIDO

Mesma abordagem para  
as demais etapas!

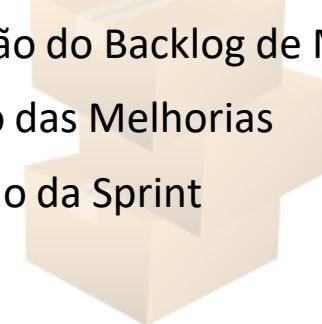


## 2. Priorizar

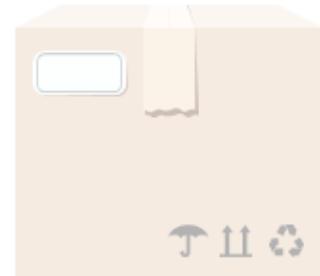
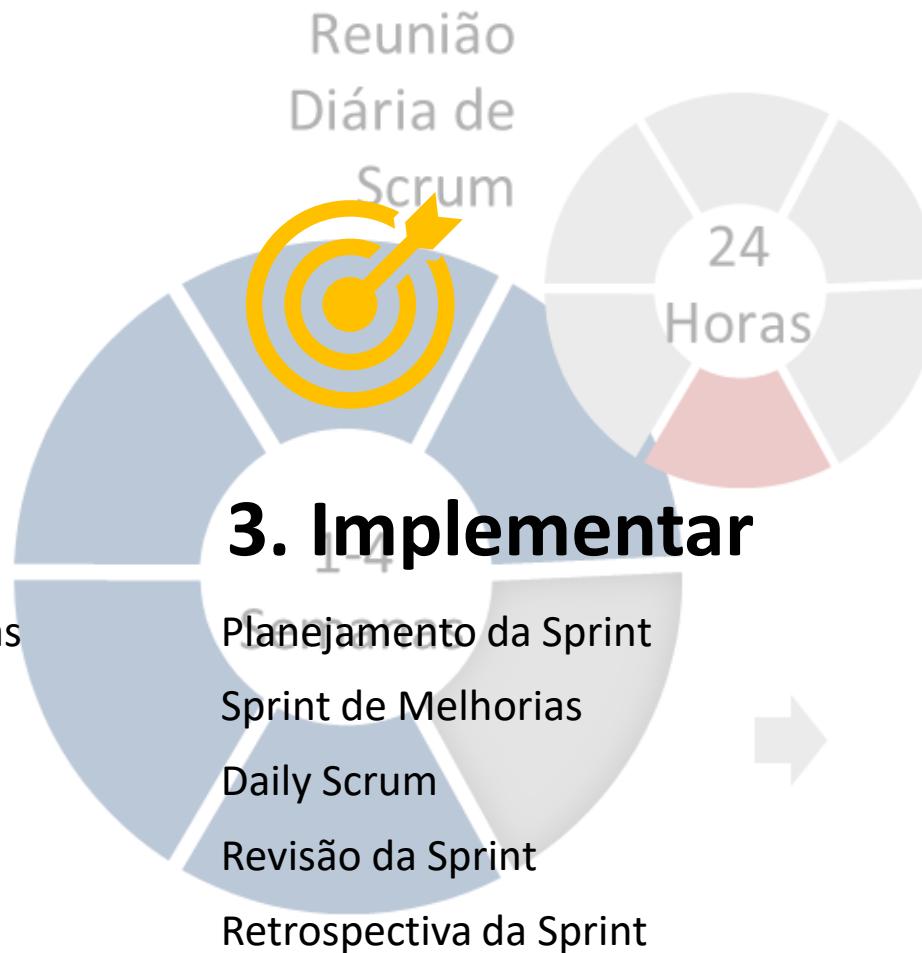
- Estruturação do Backlog de Melhorias
- Priorização das Melhorias
- Composição da Sprint



Backlog de  
Produtos



Backlog de  
Sprint



Novo  
Produto/Serviço

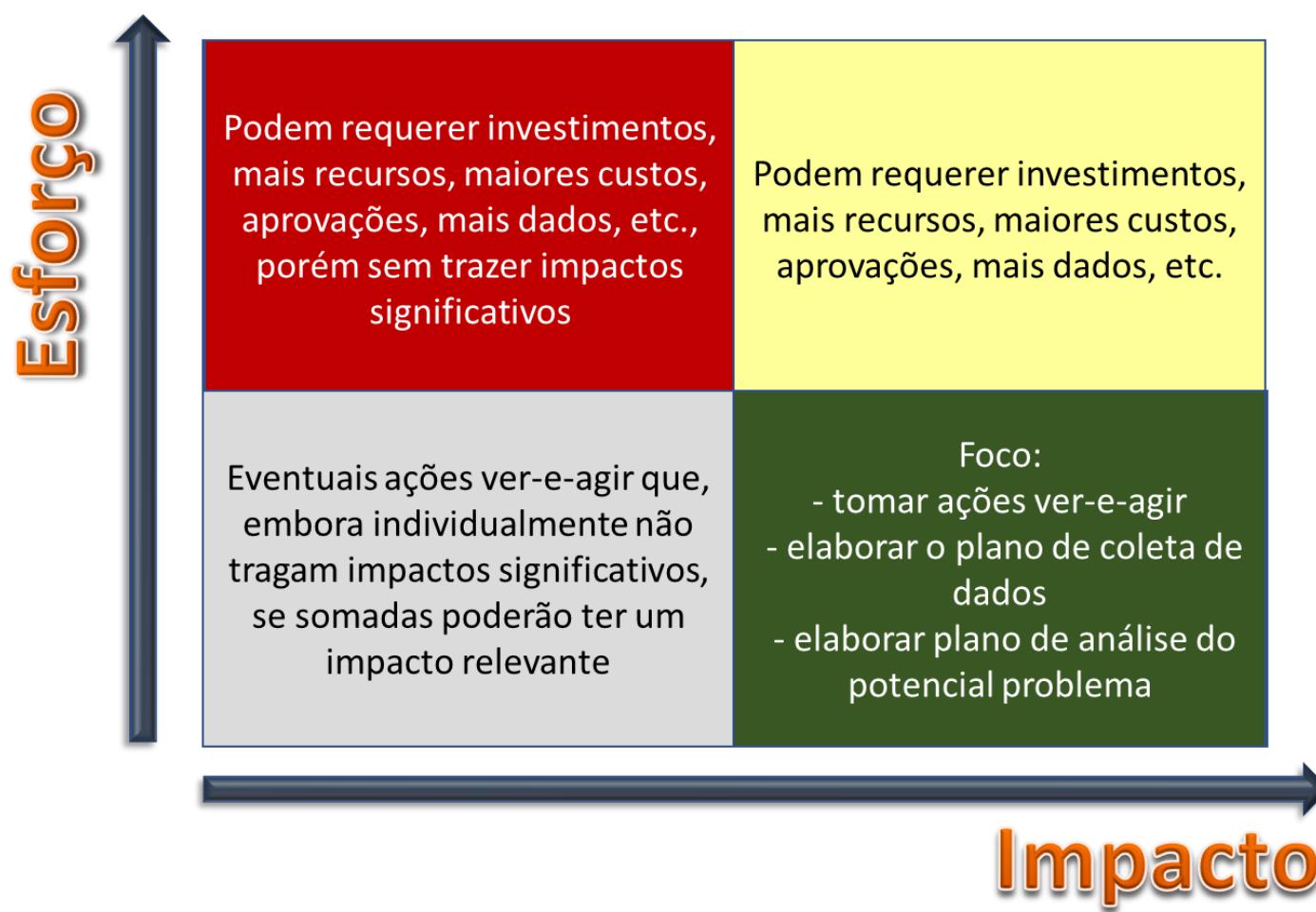
# MATRIZ ESFORÇO x IMPACTO

Matriz  
Causa x Efeito  
& Melhorias  
+  
Esforço

Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida	Esforço
	CTQ 1	CTQ 2	CTQ N			
	Peso	Peso	Peso			
X1						
X2						
X3						
X4						
X5						
X6						
XN						

- Seu objetivo é classificar os X's (entradas do processo) de acordo com o impacto (efeito) destes nos Y's (CTQ's ou KPI's do projeto)
- É um sumário das conclusões em relação aos X's (principais entradas do processo identificadas através do SIPOC, Mapa de Processo, Espinha de Peixe) através de ferramentas gráficas, analíticas e qualitativas
- Através de ações de Ideação encontramos as melhores sugestões de melhoria para cada X
- É fonte de alimentação da Matriz Esforço x Impacto

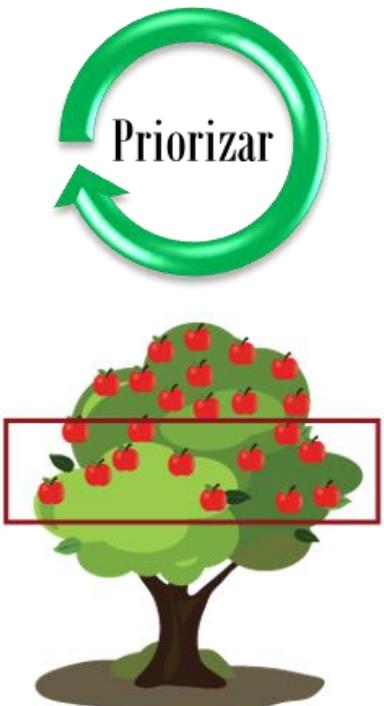
# Matriz Esforço x Impacto



Esforço em relação à: implantação das mudanças, coleta de dados, tempo para análise, aprovações, custos etc.

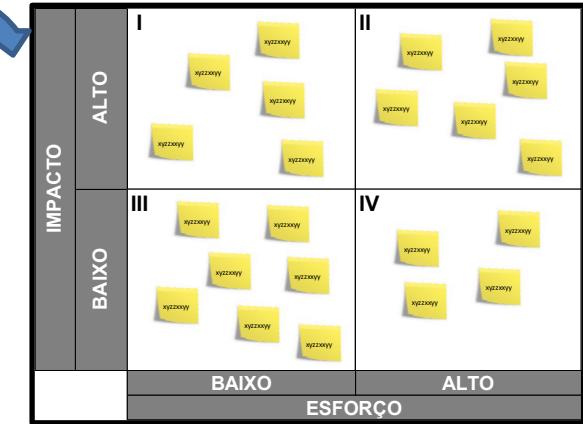
Impacto em relação às principais métricas de sucesso do projeto

# Atividade A.2 – Ciclo Iterativo Analyze



## A) Matriz Esforço/Impacto: Estruturar Backlog de Melhorias

Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida	Esforço
	CTQ 1	CTQ 2	CTQ N			
	Peso	Peso	Peso			
X1						
X2						
X3						
X4						
X5						
X6						
XN						



## B) Matriz Esforço/Impacto: Priorização das Melhorias

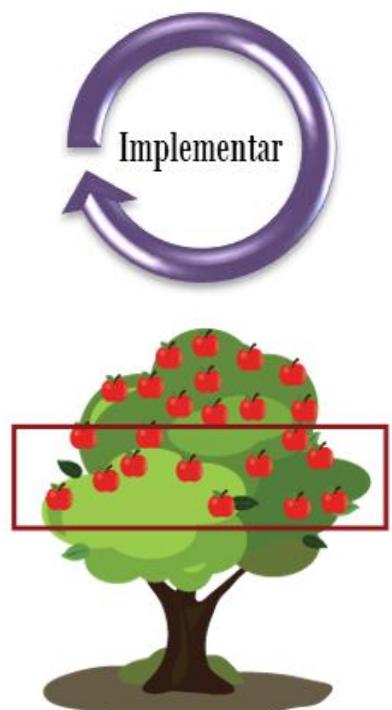


## C) Sprint Board: Composição dos Sprints

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D					
M					
A					
I					
C					

Analyze Híbrido

# Atividade A.3 – Ciclo Iterativo Analyze



Analyze Híbrido

## A) Planejamento do Sprint: Detalhamento das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D	5 yellow sticky notes				12 blue squares (3x4 grid)
M	4 yellow sticky notes				10 blue squares (2x5 grid)
A	3 yellow sticky notes	6 blue squares (2x3 grid)			
I	2 yellow sticky notes				
C	2 yellow sticky notes				

## B) Sprint Analyze: Implantação das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
A	2 yellow sticky notes	3 blue squares (1x3 grid)	2 blue squares (1x2 grid)	2 blue squares (1x2 grid)	3 blue squares (1x3 grid)

## C) Daily Scrum: Reuniões Diárias

## D) Revisão da Sprint



## E) Retrospectiva da Sprint

Com análise de Objetivos,  
Metas e Gestão à Vista

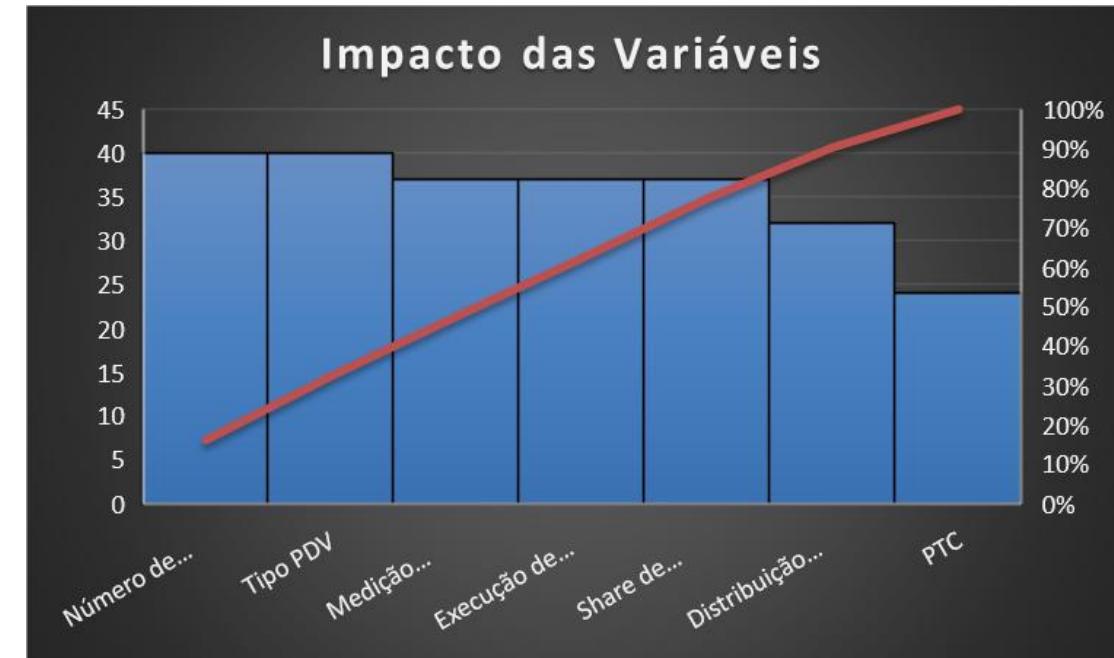
# Projeto Mercado - PDV – Ciclo A.1

## Analyze Identificar





# Matriz Causa & Efeito – Projeto Mercado – PDV – Analyze



Variável de Entrada (X)	Etapa do Processo	Peso do Requisito	Variáveis de Saída (Y)								Total	Sugestões de Melhoria
			Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8		
Visita	Visita	Número de SKU's	5	5							40	A1 - Aumento de SKU - fase 2
		Medição Execução	5	4							37	A2 - Roll-out de Medição da Execução
		Tipo PDV	5	5							40	Prioridade: Farmácia
	Execução	PTC	3	3							24	
		Execução de Loja	5	4							37	A3 - Roll-out Guia de Execução & Treino
Comprar	Execução	Share de Prateleira	5	4							37	A4 - Roll-out Planograma & Treino A5- Sist. de medição com foto A6- Implantar Controle/Acompanhamento
		Distribuição Relativa	4	4							32	
	Comprar	Display Relativo	3	2							21	A7 - Incrementar Distr. Rel. em Farm.

# Projeto Mercado - PDV – Ciclo A.2

## Analyze Priorizar





			ESFORÇO
IMPACTO	ALTO	A1 - Aumento de SKU - fase 2 A2 - Roll-out Medição da Execução A3 - Roll-out Guia de Execução & Treino A4 - Roll-out Planograma & Treino	A5- Sist. de medição com foto A6- Implantar Controle/Acompanhamento D5 - Plano Bonificação A7 - Incrementar Distr. Rel. em Farm. D6 - APP de Vendas: localização, visitas, medições, proposta de pedido, bônus
	BAIXO	D4 - Índice de Visitas Real x Prev.	M1 - Plano Adesão de PTC
	BAIXO	ALTO	ESFORÇO

# Projeto Mercado - PDV – Ciclo A.3 Analyze Implementar



# Sprint Board – Projeto Mercado – PDV – Analyze

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
<b>M</b>	M2 - Aumento de SKU				M2.1 Elaborado Material p/ Vendedor M2.2 Treinados Vendedores p/ Piloto M2.3 Piloto executado
	M3 - Medição Execução				M3.1 Elaborado Procedimento Medição M3.2 Elaborado Treinamento M3.3 Vendedores Treinados p/ Piloto M3.4 Piloto MSA efetuado
	M4 - Roll-out Guia de Execução & Treino				M4.1 - Roll-out de Execução em 1 Região
	M5 - Roll-out Planograma & Treino				M5.1 - Roll-out de Planograma em 1 Região
<b>A</b>	A1- Aumento de SKU fase 2	A1.1 - Continuar com o Piloto A1.2 - Coletar dados A1.3 - Analisar dados A1.4 - Elaborar plano de roll-out			
	A2 - Roll-out Medição da execução	A2.1 - Elaborar Plano de Treinamento A2.2.- Efetuar Treinamento A2.3 - Implantar coleta de dados A2.4 - Efetuar Análises dos dados			
	A3 - Roll-out Execução	A3.1 - Elaborar plano de roll-out A3.2 - Implantar o Plano de roll-out A3.3 - Coletar dados (iniciar Região 1) A3.4 - Analisar Dados			
	A4 - Roll-out Planograma	A4.1 - Elaborar plano de roll-out A4.2 - Implantar o Plano de roll-out A4.3 - Coletar dados (iniciar Região 1) A4.4 - Analisar Dados			

# Business Case Mercado – PDV Álcool Gel – Exercício

Nas etapas D e M foi implantado um melhor método de medição da Execução de Loja e efetuados pilotos de implantação de Planograma e Execução de Loja. Para avaliar o impacto destas ações foram coletados as variações de volume das lojas que tiveram evolução no Índice de Execução, devido à essas ações, e daquelas que não sofreram essas ações.

- Faça Análises Gráficas e Analíticas: Box Plots, ANOVA, 2-sample-t, teste de variância
- Quais são as conclusões?
- As ações efetuadas impactaram o Volume? Como?

Os dados estão no arquivo Índice de Execução.xlsx  
Os volumes estão em base 100 em comparação com o resultado médio de cada PDV no último trimestre.

# Business Case Mercado – PDV Álcool Gel – Exercício

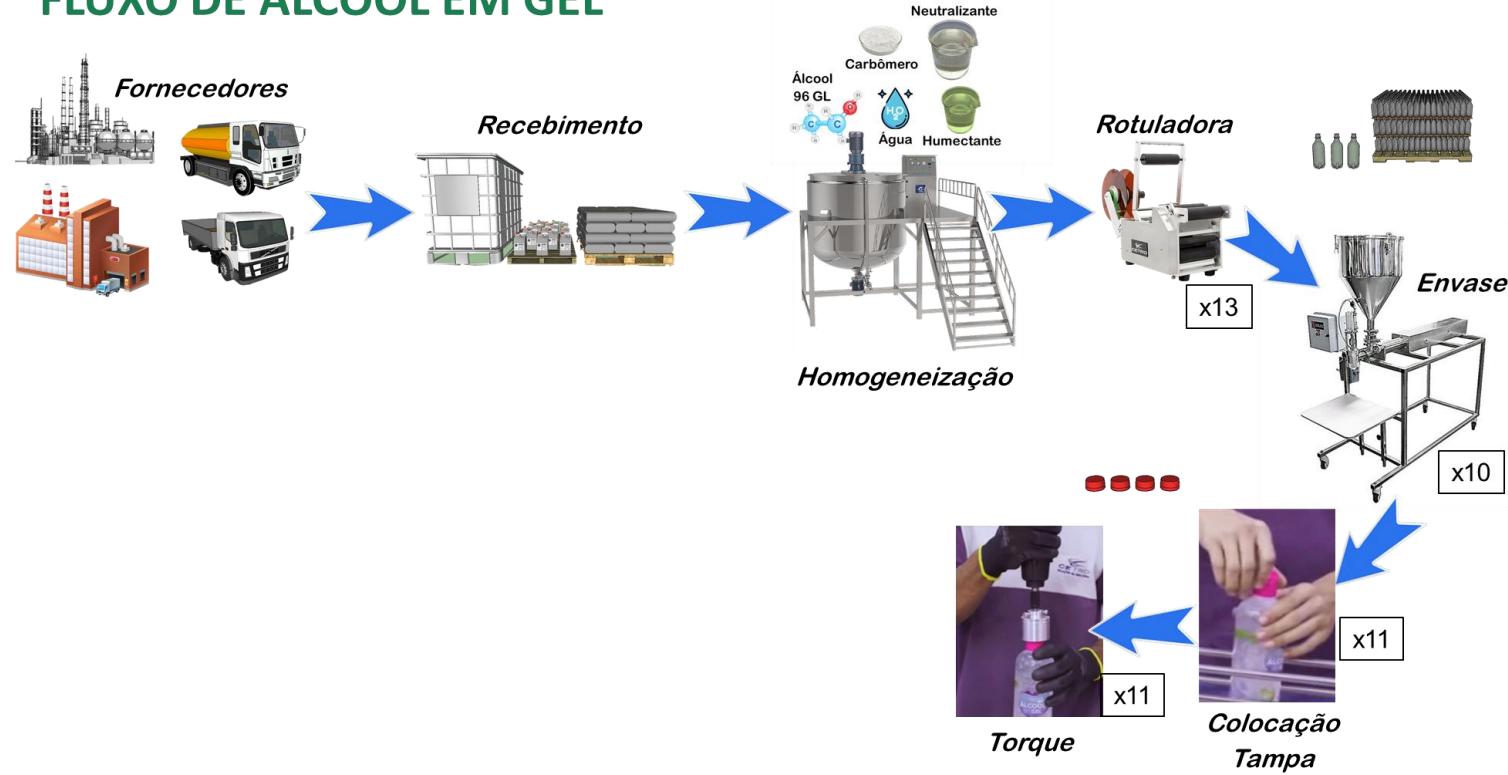
Nas etapas D e M foi efetuado um piloto de aumento de SKU's por loja. Para avaliar o impacto destas ações foram coletados as variações de volume das lojas que tiveram evolução no número de SKU's, devido à essas ações, e daquelas que não sofreram essas ações.

- Faça Análises Gráficas e Analíticas: Box Plots, ANOVA, 2-sample-t, teste de variância
- Quais são as conclusões?
- As ações efetuadas impactaram o Volume? Como?

Os dados estão no arquivo Aumento de SKU.xlsx  
Os volumes estão em base 100 em comparação com o resultado médio de cada PDV no último trimestre.

# Projeto Produtividade – Ciclo A.1 Analyze Identificar

## FLUXO DE ALCOOL EM GEL



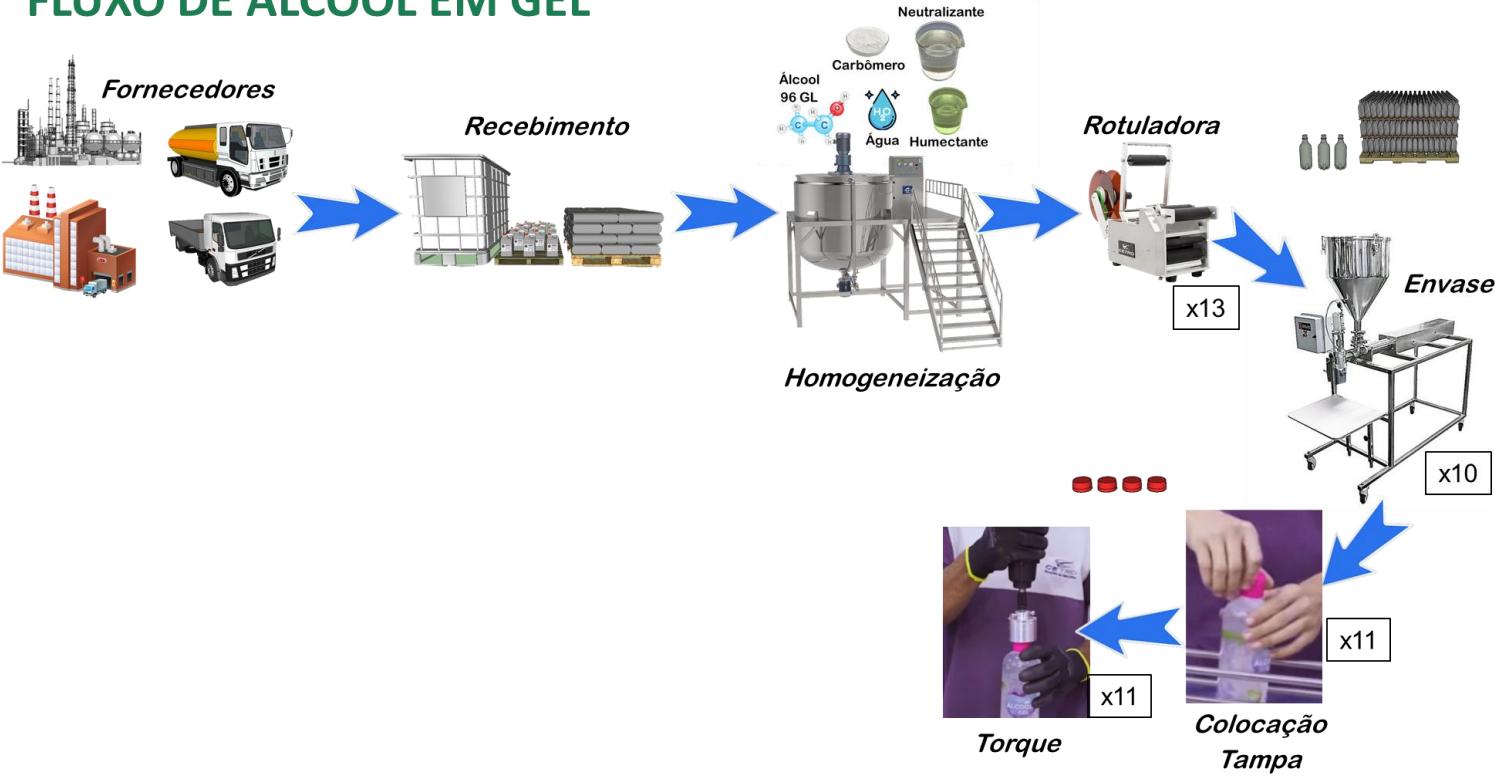
# Matriz Causa & Efeito – Projeto Produtividade – Analyze

Etapa do Processo		Variáveis de Saída (Y)					Total
		Eficiência de Produção 85%	Homogeneidade entre turnos	Qualidade do Alcool Gel	Custos de Produção		
		Peso	5	3	5	3	
Separação de Matéria Prima e garrafas plásticas 500ml	Peso de Insumos - Lote do dia	5	1	5	3	62	D2. Otimização de tamanho de lote
	Disponibilidade de Matéria prima	5	1	1	3	42	D1. Otimização de tamanho de lote junto aos fornecedores
	Organização e Identificação do Estoque	3	1	1	2	29	D15. Implantar Contagem Cíclica
	Qtd de Garrafas Plásticas	3	1	1	3	32	D3. Organização no setor de separação
	Identificação material inadequada	4	1	1	3	37	A2. Implementar Leitor de Código de Barras
Transporte MP, Rótulos e Garrafas para Produção	Eficiência no Abastecimento	5	5	1	3	54	D4. Otimização do processo de abastecimento para a Homogeneização
	Precisão na formulação	5	5	5	5	80	A1. Conferência 100% antes da entrega para produção
	Peso dos insumos - Batch	5	3	5	3	68	A2. Implementar Leitor de Código de Barras
Pesagem e preparação da receita	Não aderência ao procedimento de pesagem (tara, nívelamento)	3	3	5	2	55	D5. Implantar padrões robóticos para formulação
	Volume da água	3	3	5	2	55	M2. Padronizar Limpeza do Tanque
	Volume do Álcool	3	3	5	3	58	D6. Reduzir tempo de setup da Rotuladeira
Homogeneização	Limpeza do Tanque	5	3	3	1	52	D7. Eliminar causas de parada de posicionamento do rótulo
	Alinhamento para rotulação	5	5	1	2	51	D8. Reduzir e padronizar tempo de Envase
	Tempo de setup	4	3	1	2	40	A6. Revisar e definir especificação de tempo de envase com base na equação de regressão
Rotulação das Garrafas	Estabilidade da máquina - Rot	5	5	1	3	54	M1. Estabilizar Densidade do Álcool Gel
	Dosador	3	1	1	1	26	D9. Reduzir variabilidade no volume de envase
	Tempo de Setup	4	3	1	2	40	D6-2. Reduzir tempo de setup do envase
Envase	Densidade do Álcool Gel	5	3	5	3	68	M2. Padronizar Limpeza do Tanque
	Tempo de envase	5	3	3	3	58	D11. Melhorar padrão de rosqueamento
	Diam. Int - Tampas	5	1	1	3	42	D12. Otimizar ajustes da torqueadeira
Colocação Tampa e Torque	Apertad. com controle de torque	5	5	1	3	54	D11. Melhorar padrão de rosqueamento
	Qualidade do torque	2	2	1	3	30	D10. Combinar precolocação da tampa com torque
	Capacidade de Torque	3	3	1	1	32	A4. Desenvolver um treinamento moldado nos conceitos de Andragogia e reciclar os operadores.
	Fecham. Autom. de embalagem/lote	5	3	0	3	43	A5. Implementar Matriz de Versatilidade para a gestão da capacitação do pessoal
Embalagem	Organização da área	3	1	1	2	29	M3. Estabilizar Impressora de Etiquetas
	Impressora de etiquetas	5	5	0	3	49	D13. Melhorar organização na área de embalagem
	Nível de Capacitação e habilidade diferente entre turnos	5	5	3	4	67	D14. Equipamento automático para Stretch do lote
Geral						0	M4. Comprar equipamento para Fechamento Automático de Embalagens

# Projeto Produtividade – Ciclo A.2 Analyze Priorizar



## FLUXO DE ALCOOL EM GEL

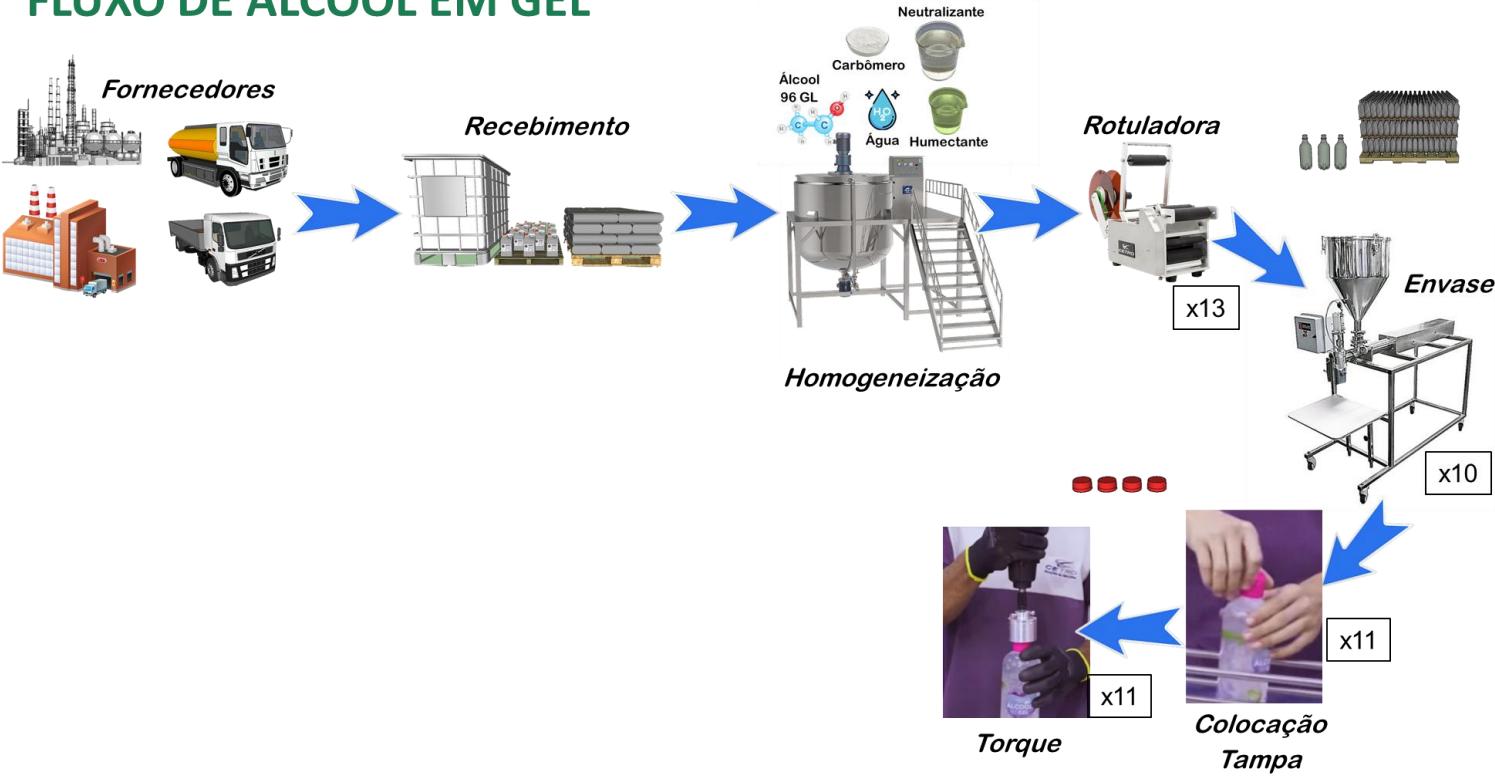


# Matriz Esforço & Impacto – Projeto Produtividade – Analyze

	A1. Conferência 100% antes da entrega para produção	A2. Implementar Leitor de Código de Barras
ALTO	A3. Revisão IT de Pesagem;	D2. Otimização de tamanho de lote interno
IMPACTO		D6-1. Reduzir tempo de setup da Rotuladeira
BAIXO		D6-2. Reduzir tempos de setup do Envase
BAIXO		D7. Eliminar causas de parada de posicionamento do rótulo
BAIXO		D8. Reduzir e padronizar tempo de Envase
BAIXO		D9. Reduzir variabilidade no volume de envase
BAIXO		D10. Combinar precolocação da tampa com torque;
BAIXO		D12. Otimizar ajustes da torqueadeira
BAIXO	<b>A4. Desenvolver um treinamento moldado nos conceitos de Andragogia e reciclar os operadores.</b>	
BAIXO	D13. Melhorar organização na área de embalagem	D11. Melhorar padrão de rosqueamento
BAIXO	M2. Padronizar Limpeza do Tanque	D14. Equipamento automático para Stretch do lote
BAIXO	M3. Estabilizar Impressora de Etiquetas	M3. Estabilidade Impressora de Etiquetas
BAIXO	<b>A5. Implementar Matriz de Versatilidade para a gestão da capacitação do pessoal</b>	D15. Implantar Contagem Cíclica
BAIXO	<b>A6. Revisar e definir especificação de tempo de envase com base na equação de regressão</b>	M4. Comprar equipamento para Fechamento Automático de Embalagens
BAIXO		
ESFORÇO	BAIXO	ALTO

# Projeto Produtividade – Ciclo A.3 Analyze Implementar

## FLUXO DE ALCOOL EM GEL



# Sprint Board – Projeto Produtividade – Analyze

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
<b>M</b>	D1. Otimização de tamanho de lote junto aos fornecedores				D1.4: Piloto Kanban nos fornecedores selecionados
	M1. Estabilizar Densidade do Alcool Gel				M1.1: Implementar Carta de Controle na Homogeneização
					M1.2: Segregar e corrigir densidade antes de enviar ao Envase
					M1.3: Auditoria diária pela Qualidade
	D10. Combinar precolocação da tampa com torque;				D10.1: Desenvolver dispositivo para rosqueamento e torque integrado
					D10.2: Testar e avaliar viabilidade
					D10.3: Caso OK implementar
<b>A</b>	D1. Otimização de tamanho de lote junto aos fornecedores	D1.5: Implantar e validar sistema kanban nos fornecedores selecionados			
	A1. Conferência 100% antes da entrega para produção	A1.1: Definir odêlo de conferência (Quem? Como?)			
		A1.2: Implementar Conferencia 100%			
		A3.1: Investigar inconsistencias da IT atual			
	A3. Revisão IT de Pesagem;	A3.2: Revisar IT de Pesagem			
		A3.3: Implementar			
		A4.1: Desenvolver Treinamento dentro do Conceito de Andragogia			
	A4. Desenvolver um treinamento moldado nos conceitos de Andragogia e reciclar os operadores.	A4.2: Rodar um piloto			
		A4.3: Preparar agenda de treinamento			
		A6.1: Analisar dados (Regressão)			
		A6.2: Validar processo utilizando parametros conforme equação de regressão			
	A6. Revisar e definir especificação de tempo de envase com base na equação de regressão	A6.3: Atualizar IT de Envase			
		A6.4: Treinar operadores de Envase dos 3 turnos			

# Acompanhamento de Resultados Projeto Produtividade – Analyze

	MELHORIA				
Baseline	D	M	A	I	C
Eficiência Geral	57%	66%	70%	75%	
Eficiência T1	51%	62%	68%	75%	
Eficiência T2	74%	76%	78%	81%	
Eficiência T3	47%	59%	65%	70%	
META	85%	85%	85%	85%	85%



# Projeto Logística – Ciclo A.1 Analyze Identificar



# Canvas Projeto Logística

Justificativas 	Produto 	Stakeholders 	Premissas 	Riscos 
Empresa tem perdido receitas e Market Share por não conseguir entregar pedidos após aumento de demanda, com OTIF de 82%. As entregas não estão sendo feitas no prazo (On Time 85%), e há ocorrências de entregas com erros diversos (In full 96%). Clientes tem mostrado insatisfação crescente e procurado outras alternativas, levando a perda de 15% da base de clientes, diminuindo receitas em 7%.	Processo robusto de Logística que garanta o objetivo SMART para entregas feitas "Sem Erro" e "On Time"	. Plant Manager . Ger. Projeto/Belt . Gerente Comercial . Gerente de Logística (SPONSOR)	. TI fornecerá suporte de 01 pessoa por 3 meses. A disponibilidade da rede será >80% do contratado. . 100% do volume será fornecido pela fábrica (SLA precisa ser >95%) . O estoque de produto fica nos CD's . A operação vai contar com 5 CD's (Piracicaba, Recife, Manaus, Joinville e Uberlândia)	. Fábrica não conseguir atender o aumento de demanda . Aumentar o grau de restrição por região devido à pandemia . Agentes de transporte com redução de pessoas devido às restrições definidas pelos órgãos públicos . Não conseguir homologar fornecedores logísticos para atender a demanda . Clientes definindo novas e diferentes rotinas e horários . Aumento do custo logístico devido à redução de mão de obra certificada
Objetivo SMART	Requisitos 	Equipe 	Grupo de Entregas 	Linha de Tempo 
Aumentar o OTIF (On Time in Full) para >=95% no final da etapa Improve (15/Jan/2021), e manter este na fase Control.	. Medição de entrega pelo indicador oficial OTIF . Inventário dentro do programado . Custo logístico dentro do Budget . Identificação, análise de causas de não atendimento e plano de ação. . Atendimento às 5 regiões do país . Processos padronizados para a área de Logística (Picking, recebimento, armazenamento, carregamento, roteirização) e Fornecedores . Fornecedores Logísticos homologados . Manual de operação e níveis de serviço para fornecedores logísticos	. Candidato Belt . Analista de PCP . Analista de Logística . Representante Comercial . Suporte de TI para o CD . Supervisor do CD . Sponsor	. Gestão da Rotina: dashboards, níveis de serviço, procedimentos, capacitação etc. . Plano de suporte à demanda por região (para as 5 regiões do país) . Plano de rastreamento de frota . Consolidação da operação com o novo Fluxo de Pedido	. Define: 23/Out/2020 . Measure: 13/Nov/2020 . Analyze: 27/Nov/2020 . Improve: Plano definido com Resp./ Data: 11/Dez/2020 Ações implementadas/verificadas: 15/Jan/2021 . Control: 19/Fev/2021
Benefícios			Restrições 	Custos 
. Satisfação e Fidelização de Clientes pela excelência nas entregas . Aumento de Market Share nas 5 regiões do país, em um cenário de aumento significativo da demanda . Retomada e Aumento de receita e rentabilidade	. Treinamento para equipe do CD e fornecedores logísticos . Sistema de rastreamento orçado/ aprovado . Novo Fluxo de pedido indo do PDV direto para o CD. . Dashboard / Power BI para gestão de indicadores	. Situação atual com variações das condições logísticas por estado durante a pandemia . Não aumentar o nível de segurança do inventário . Os pedidos não poderão seguir o Fluxo anterior indo para a fábrica.	. Alterações de Sistemas (a ser orçado) . Terminais de dados (R\$XXX) . Sistema de rastreamento (R\$YYY)	

Com alterações no Canvas

# Matriz Causa & Efeito – Projeto Logística – Analyze

Variável de Entrada (x)	Etapa do Processo	Peso do Requisito	Variáveis de Saída (Y)			Total
			On time	In full	OTIF	
			4	3	5	
Recebimento e Armazenamento	D11. Padronizar os dados de ERP e MRP		3	3	5	46
	A1. Análise antecipada da Nota Fiscal eletrônica		3	3	3	36
	A2. Contagem Dupla e Contagem Cega		3	3	4	41
	A3. Implementar Sistema de auditoria de localização		3	3	4	41
	A4. Implementar TPM e Controle de Utilização de Equipamentos		4	3	4	45
Entrada e Análise dos Pedidos	A5. Implantar sistema visual de controle de estoque com indicação de estoque mínimo		2	4	4	40
Carregamento do veículo	D4. Acuracidade físico x sistema		3	4	5	49
Roteirização	D3. Padronizar Roteirização		4	3	5	50
	Lista de pedidos		2	1	3	26
	Tempo teórico (objetivo) das entregas		1	2	2	20
Separação de Carga (Picking)	Lista de material (papel ou sistema)		1	2	2	20
	Romaneio: CD > PDV		1	2	2	20
	M4. Tempo de separação		5	2	4	46
	Manuseio no picking		1	4	3	31
	Custo Operacional		1	1	2	17
	Custo do Inventário		1	3	1	18
	A6. Programar o sistema para detectar não seguimento do roteiro de Picking		1	2	2	20
	M5. Sequência de entrega		4	2	4	42
Carregamento do veículo	Romaneio: CD > PDV		1	2	2	20
	D8. Tempo de carregamento		4	2	4	42
	Peso		1	1	1	12
Transporte e Entrega	Motorista do Agente Logístico		3	1	3	30
	NF e Romaneio		3	2	2	28
	M1. Embalagem e meios de manuseio (avarias)		3	4	4	44
	M6. Impacto financeiro de entregas não realizadas		3	3	3	36
	A7. Implementar sistema de inteligencia de emissão de Notas Fiscais e detecção de defeito (RPA)		4	3	4	45



# Projeto Logística – Ciclo A.2 Analyze Priorizar





## Matriz Esforço & Impacto – Projeto Logística – Analyze

		BAIXO	ALTO
		ESFORÇO	
IMPACTO	ALTO	A4. Implementar TPM e Controle de Utilização de Equipamentos A2. Contagem Dupla e Contagem Cega A3. Implementar Sistema de auditoria de localização A5. Implantar sistema visual de controle de estoque com indicação de estoque	D1. Inovar no sistema de distribuição M3. Disponibilizar canal de Venda Exclusivo D9. Contratação de novos agentes logísticos A7. Implementar sistema de inteligencia de emissão de Notas Fiscais e detecção
BAIXO	BAIXO	D10. Otimizar pedido	M7. Definir clientes VIP D12. Fazer in sourcing do sistema de transporte

# Projeto Logística – Ciclo A.3 Analyze Implementar

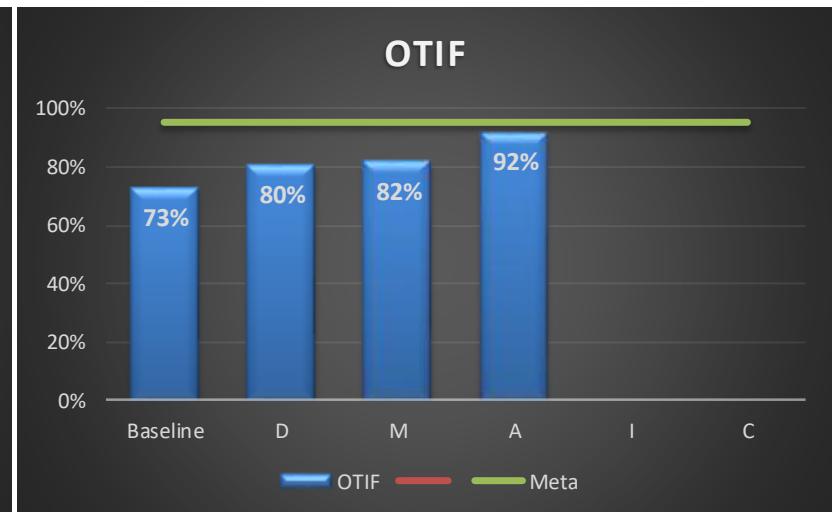
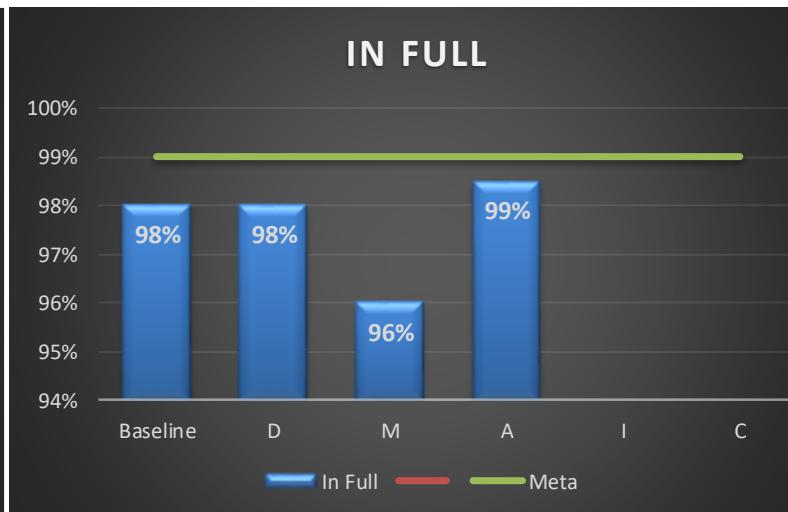
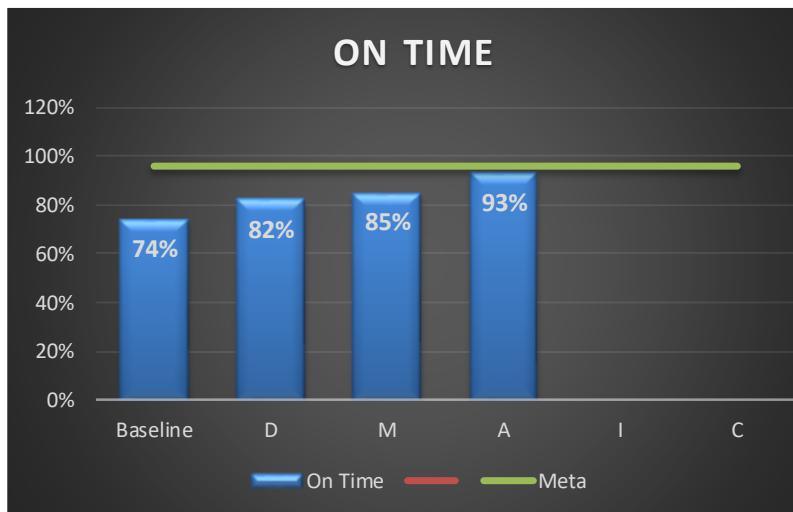


# Sprint Board – Projeto Logística – Analyze

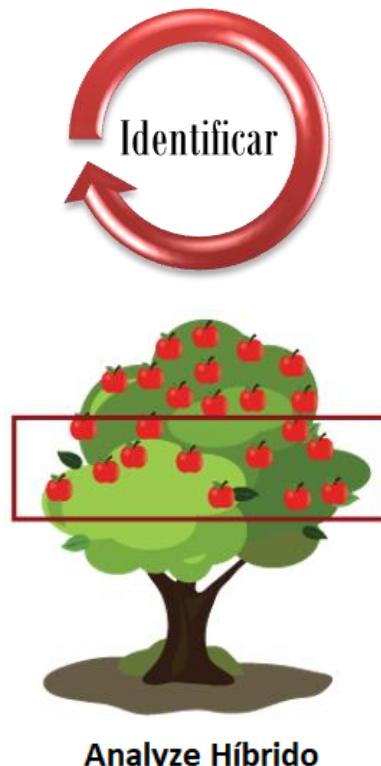
Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
M	D2. Otimizar roteiros				D2.1 Levantar PDV's, distribuidores e CD's D2.2 Kaizen de Otimização de roteiros
	D3. Padronizar Roteirização				D3.1 Definir e padronizar roteiros conforme necessidade levantada
	D7. Implementar Kanban				D7.1 Definir Kanban D7.2 Implementar Kanban
	D8. Otimizar Tempo de carregamento				D8.1 Levantar dados de carregamento de várias semanas D8.2 Comparar dados pré e pós aumento da demanda
	D11. Padronizar os dados de ERP/MRP				D11.1 Definir um padrão único para ERP/MRP
					D11.2 Treinar os times
					D11.3 Implementar
	M1. Padronizar embalagens e meios de manuseio				M1.1 Fazer estudo de embalagens para reduzir avarias M1.2 Levantar custos para estudo de caso
	M2. Poka Yoke Poka Yoke no processo abastecimento para melhorar acuracidade físicoX sistema				M2.1 implementar sistema de localização e controle de abastecimento e contagem
	M4.Kaizen para melhorar tempo de separação				M4. 1 Levantar tempos e direcionar Kaizen para redução do tempo de separação
A	M5.Padronizar sequência de entrega				M5. Levantar condições e sequenciar entrega ótima por região
	M6.Analisar impacto financeiro de entregas não realizadas				M6. 1 Levantar dados do impacto no pagamento dos clientes antes e depois do aumento da demanda
	D9. Contratação de novos agentes logísticos	D9.1 - Cotação com 3 Agentes Logísticos			
		D9.2 - Analise das cotações e separação da melhor proposta			
		D9.3 - Preparação de Proposta para avaliação da Diretoria			
	A4. Implementar TPM e Controle de Utilização de Equipamentos	A4.1 - Treinamento Pilares básicos do TPM			
		A4.2 - Kick-Off implementação			
		A4.3 - Preparar Controle de Utilização de Equipamentos e implementar			
	A2. Contagem Dupla e Contagem Cega	A2.1 - Escrever Procedimento de Contagem Dupla e Auditoria			
		A2.2 - Treinar Equipe - Definir Responsabilidades			
		A2.3 - Implementar			
	A3. Implementar Sistema de auditoria de localização	A3.1 - Escrever Procedimento de Auditoria			
		A3.2 - Treinar Equipe - Definir Responsabilidades			
		A3.3 - Implementar			
	A5. Implantar sistema visual de controle de estoque com indicação de estoque mínimo	A5.1 - Definir Estoques mínimos			
		A5.2 - Escrever Procedimento - Treinar Equipe e Implementar			

# Acompanhamento de Resultados Projeto Logística – Analyze

		MELHORIA				
Baseline		D	M	A	I	C
On Time	74%	82%	85%	93%		
Meta	96%	96%	96%	96%	96%	96%
In Full	98%	98%	96%	99%		
Meta	99%	99%	99%	99%	99%	99%
OTIF	73%	80%	82%	92%		
Meta	95%	95%	95%	95%	95%	95%

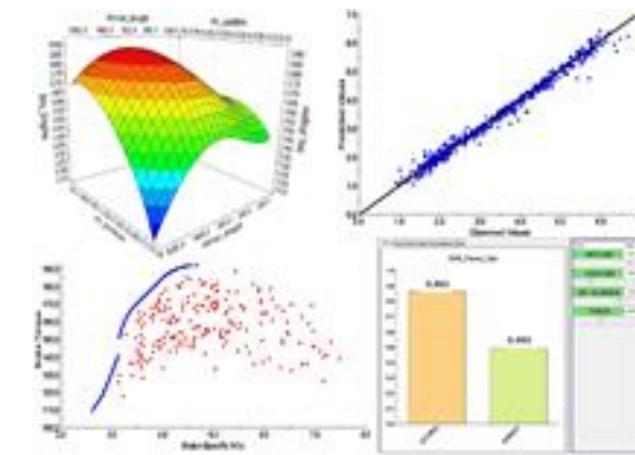
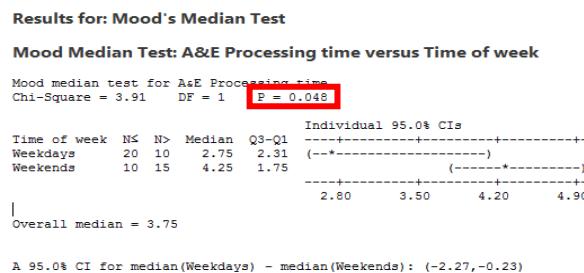


# Atividade A.1 – Ciclo Iterativo Analyze



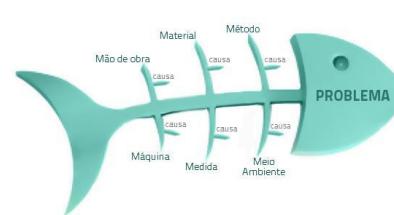
## A) Fatores Críticos & Impacto

### Ferramentas Gráficas e Analíticas



## **B) Riscos Potencias e Causas Raízes de Falhas**

### Ferramentas Qualitativas



5 Por quês?

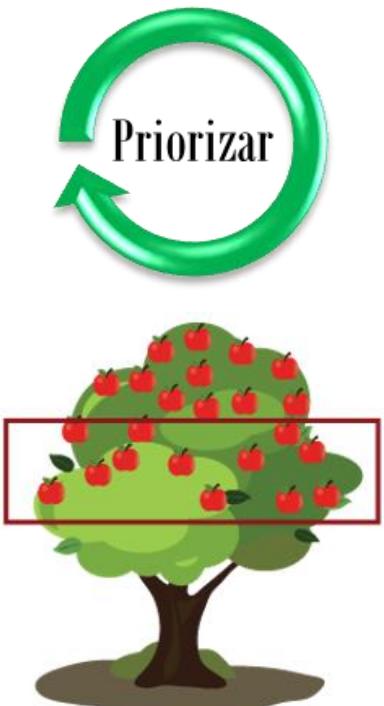
FMEA							
Etapa / entrada do Processo	Modo de Falha Potencial	Efeito da Falha	S E V	Causas Potenciais	O C C	Controles atuais	D E T R P N
Qual a etapa / entrada do processo sobre investigação?	Em quais caminhos a entrada prejudica?	Qual é o impacto sobre as variáveis de saída (Exigências do Cliente) ou exigências internas?	Qual o efeito para o cliente?	Quais as causas para a entrada atuar prejudicialmente?	Com que frequência as causas ou o EFM ocorrem?	Quais são os controles e procedimentos existentes (inspeção e teste) que já podem detectar a Causa ou o Modo de Falha? Deveria incluir um número SOP.	Qual é bom / ruim ou P/N? detetar a causa ou o FME?
							Sev*Occ*Det 0

## C) Melhorias Relacionadas com a consolidação das Variáveis Críticas

### Matriz Causa e Efeito

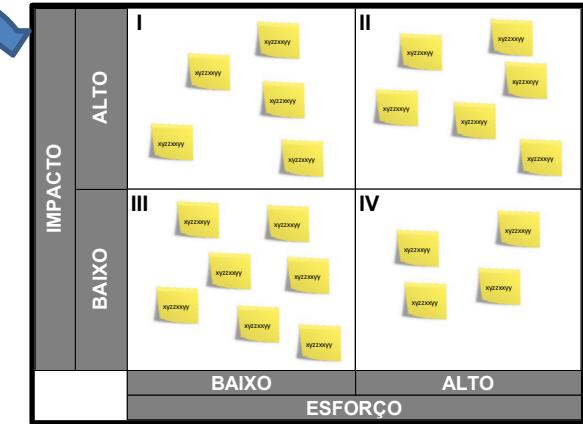
Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida
	CTQ 1	CTQ 2			
	Peso	Peso			
X1					
X2					
X3					
X4					
X5					
X6					
XN					

# Atividade A.2 – Ciclo Iterativo Analyze



## A) Matriz Esforço/Impacto: Estruturar Backlog de Melhorias

Principais Entradas (X's)	Principais Saídas (Y's)			Efeito	Melhoria Sugerida	Esforço
	CTQ 1	CTQ 2	CTQ N			
	Peso	Peso	Peso			
X1						
X2						
X3						
X4						
X5						
X6						
XN						



## B) Matriz Esforço/Impacto: Priorização das Melhorias

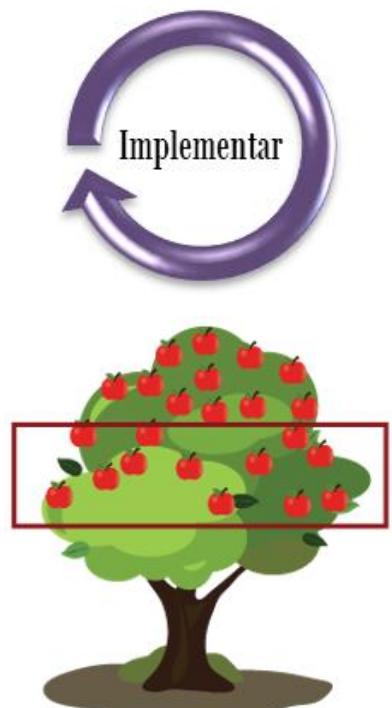


## C) Sprint Board: Composição dos Sprints

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D					
M					
A					
I					
C					

Analyze Híbrido

# Atividade A.3 – Ciclo Iterativo Analyze



Analyze Híbrido

## A) Planejamento do Sprint: Detalhamento das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
D	■■■■■				■■■■■
M	■■■■■				■■■■■
A	■■■■■	■■■■■			
I	■■■■■				
C	■■■■■				

## B) Sprint Analyze: Implantação das Melhorias

Sprint	Sprint Backlog	Ações de Melhoria			
		A fazer	Em Execução	Em Verificação	Realizado
A	■■	■■	■	■■	■■

## C) Daily Scrum: Reuniões Diárias

## D) Revisão da Sprint



## E) Retrospectiva da Sprint

Com análise de Objetivos,  
Metas e Gestão à Vista