Intervalos de confiança para pesquisas de opinião

Passo a passo da obtenção dos intervalos de confiança utilizando o Excel à configuração dos dashboards no Tableau e no Power BI

www.maiconcunha.com

consultoria@maiconcunha.com



Quando uma pesquisa é feita e uma amostra da população é utilizada é necessário saber o quão relevante é esse resultado com relação à população em geral.

Objetivos



Para determinar o resultado esperado na população utilizamos os intervalos de confiança.



Será apresentado um tutorial de como tratar os dados de uma pesquisa de opinião utilizando o Excel para o cálculo dos intervalos de confiança e tratamento dos dados e o Tableau e o Power BI para a construção de dashboards que contenham todas as informações necessárias de forma clara e simples.

Softwares e linguagens utilizadas nesse tutorial



Etapas



Obtenção dos dados



Transformação dos dados utilizando Excel



Cálculo dos intervalos de Confiança utilizando Excel



Preparação dos Dados para utilização no Tableau e no Power Bl



Construção do dashboard no Tableau



Construção do dashboard no Power BI

Os dados brutos da pesquisa se encontram em resultado pesquisa.xlsx.

No arquivo <u>dados.xlsx</u> são feitos os cálculos dos erros e das porcentagens. Nesse arquivo é feita a configuração do nível de confiança e do tamanho da população. O arquivo dados processados xisx é o arquivo que será utilizado no Power BI e no tableau e contém todas as informações necessárias para se construir os dashboards. Quando alguma alteração no arquivo dados.xisx é feita e salva, assim que o arquivo dados_processados.xisx é aberto novamente ele automaticamente atualiza os valores de erro.

O arquivo <u>dashboard.twbx</u> contém o dashboard obtido no Tableau.

você interage com o dashboard produzido no Tableau a partir do servidor, sem precisar ter o software instalado.

O arquivo <u>intervalo confianca.pbix</u> contém o dashboard obtido no Power BI.

Tour pelos arquivos

Dados Originais

Ao lado temos uma amostra do arquivo dados.xlsx de como os dados da pesquisa estão estruturados.

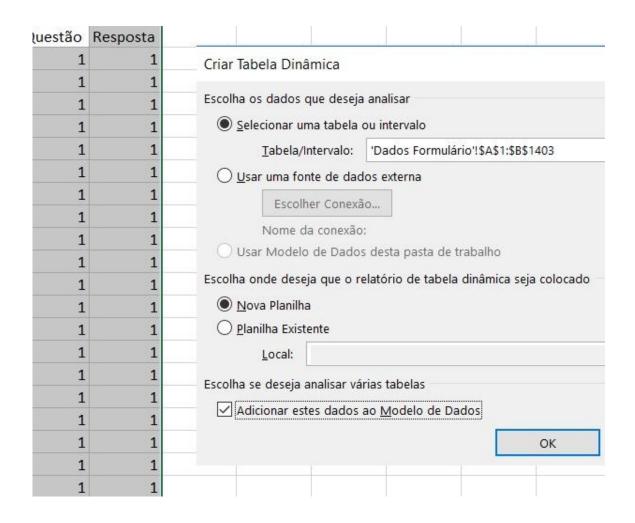
Cada questão tem sua resposta ao lado.

Não é possível identificar um número de formulário para cada pergunta e resposta.

Da forma como o dado está estruturado não é possível saber o número de perguntas e respostas para cada questão.

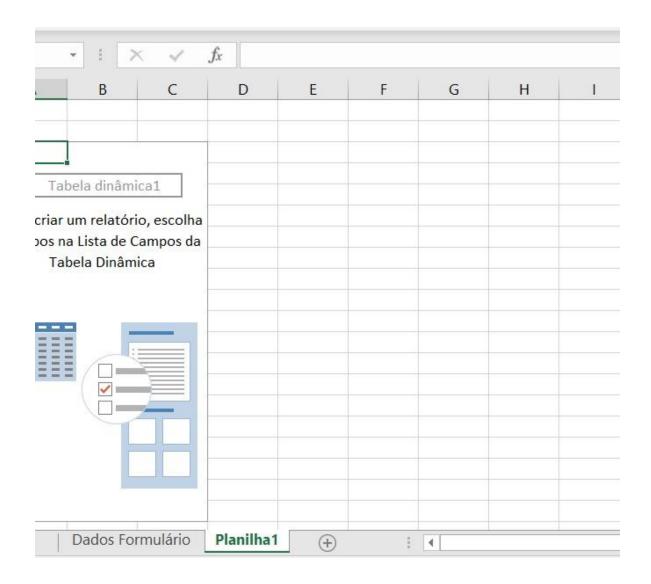
Para resolver isso devemos criar uma tabela dinâmica.

Questão	Resposta
1	1
1	2
2	2
3	2

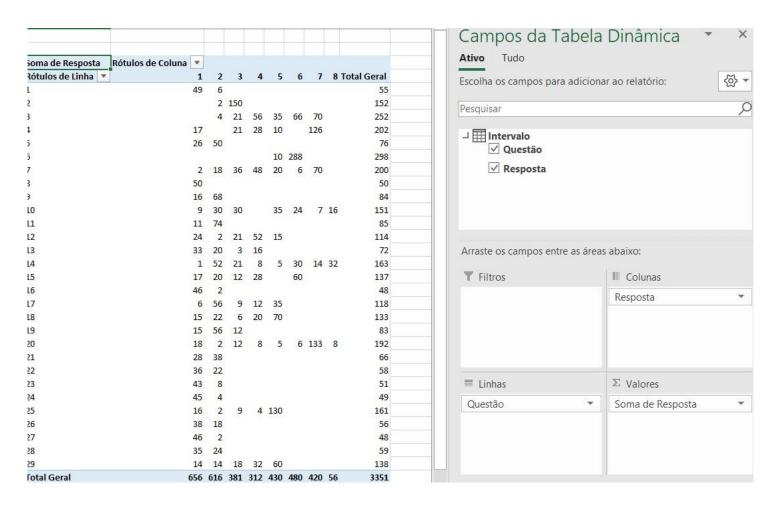


Criando a Tabela Dinâmica no Excel

- Clique na célula A1.
- Pressione "Ctrl + Shift + End" para selecionar todas as colunas com conteúdo.
- Com as duas primeiras colunas selecionadas, pressione "Ctrl + Shif + Pgdn" para selecionar todas as linhas.
- Com as células selecionadas clique em inserir tabela dinâmica e selecione e clique em OK.

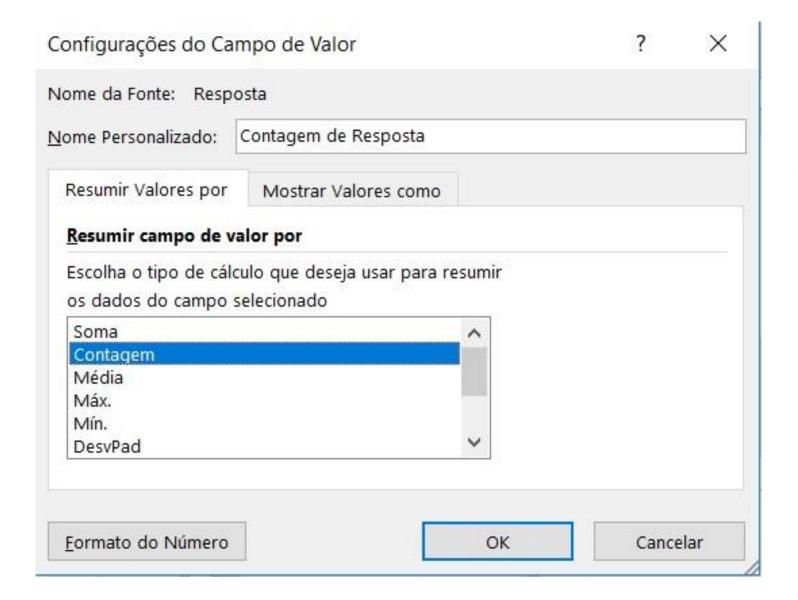


Resultado da Tabela Dinâmica



Obtendo os resultados da pesquisa

- Renomeie a planilha para Tabela Dinâmica.
- Dentro do campo Tabela Dinâmica, selecione Questão e arraste para o campo Linhas.
- Dentro do Campo Tabela
 Dinâmica, selecione Resposta e arraste para o campo Colunas.
- 4. Dentro do Campo Tabela Dinâmica, selecione Resposta e arraste para o campo Valores.



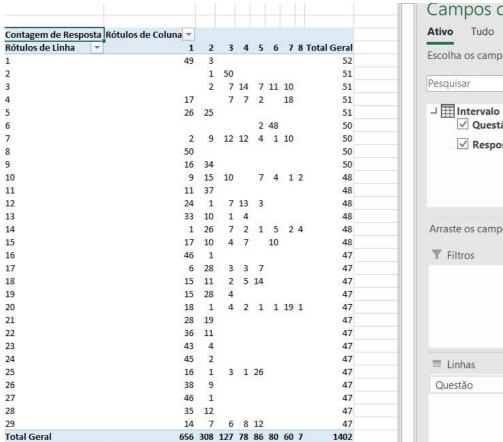
Observações

É possível observar que a Tabela Dinâmica faz um resumo dos dados originais.

Essa tabela ainda não é a desejada pois ela contém a soma das respostas, quando na verdade queremos contar o número de respostas distintas.

Para corrigir isso temos de alterar o campo Valores da tabela dinâmica.

Clique em Soma de Respostas -> Configuração de Campo de Valor e altere o campo para contagem.



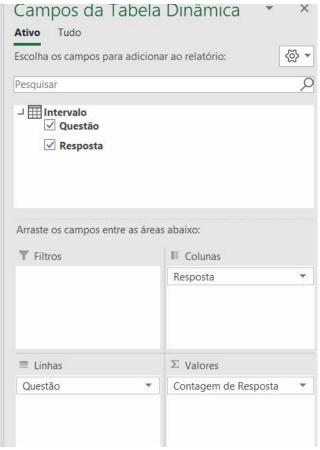


Tabela Dinâmica Correta

4	Α	В	С	D	Е	F	G	Н	1	J
1	Pergunta/Resposta	1	2	3	4	5	6	7	8	Total Geral
2	1	49	3							52
3	2		1	50						51
4	3		2	7	14	7	11	10		51
5	4	17		7	7	2		18		51
6	5	26	25							51
7	6					2	48			50
8	7	2	9	12	12	4	1	10		50
9	8	50								50
10	9	16	34							50
11	10	9	15	10		7	4	1	2	48
12	11	11	37							48
13	12	24	1	7	13	3				48
14	13	33	10	1	4					48
15	14	1	26	7	2	1	5	2	4	48
16	15	17	10	4	7		10			48
17	16	46	1							47
18	17	6	28	3	3	7				47
19	18	15	11	2	5	14				47
20	19	15	28	4						47
21	20	18	1	4	2	1	1	19	1	47
22	21	28	19							47
23	22	36	11							47
24	23	43	4							47
25	24	45	2							47
26	25	16	1	3	1	26				47
27	26	38	9							47
28	27	46	1							47
29	28	35	12							47
30	29	14	7	6	8	12				47

Copiando os dados da Tabela Dinâmica

Para Calcular os intervalos de confiança serão utilizados os dados obtidos com a tabela dinâmica.

Porém, vou copiar apenas os valores da tabela dinâmica para uma nova planilha já que muitas alterações serão feitas com esses dados.

Essa nova planilha será chamada Dados_Resumidos.

Fórmula dos intervalos de Confiança

Seja **p** a proporção do resultado da amostra.

Seja **N** o tamanho da população.

Seja <u>n</u> o tamanho da amostra.

O fator FPC é dado por FPC = (N-n)/(N-1)

O intervalo de confiança é dado por: ci = p $\pm Z_{\alpha/2}^* \sqrt{(1/n)^*p^*(1-p)^*FPC}$

Onde $Z_{\alpha/2}$ é o Valor crítico da distribuição normal em $_{\alpha/2}$

Configuração dos intervalos

Na planilha Dados_Resumidos, abaixo dos dados tabela dinâmica copiada vamos configurar o nível de confiança para o cálculo dos intervalos, além de definir o tamanho da população.

A princípio será definido uma confiança de 95% e uma população de 500 pessoas.

Próximo passo é definir o parâmetro Alpha.

В3	4 - 1	Χ .	/ .	f. (-INV.	NORM	P.N((1-B32)/	(2)
d	Α	P	C	D	E	F	G	Н	1
28	21	46	1						
29	28	35	12						
30	29	14	7	6	8	12			
31									
32	Confiança	95%							
22	População	500							
34	Alpha	1,96							

Cálculo do fator FPC

Clique na célula B33 da planilha Dados_Resumidos e altere o nome para "N". Esse passo é somente para facilitar o cálculo!

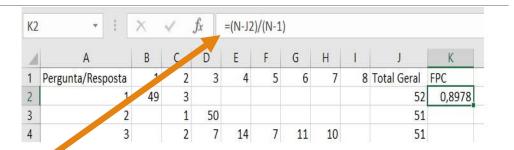
Clique na célula B34 da planilha Dados Resumidos e altere o nome para "Alpha".

Crie uma nova coluna com o nome FPC.

Na célula K2 insira a fórmula "=(N-J2)/(N-1)"

Arraste a célula selecionada K2 para obter o fator FPC das demais perguntas.

Observe que se a população for alterada o fator FPC é alterado automaticamente.



Pergunta/Resposta	1	2	3	4	5	6	7	8	Total Geral	FPC
1	49	3							52	0,8978
2		1	50						51	0,8998
3		2	7	14	7	11	10		51	0,8998
4	17		7	7	2		18		51	0,8998
5	26	25							51	0,8998
6					2	48			50	0,9018
7	2	9	12	12	4	1	10		50	0,9018

Proporção das Respostas

Os resultados dos intervalos de confiança serão baseados nas proporções das respostas, por isso vamos calcular esses valores:

Crie uma nova planilha com o nome Dados_Porcentagem, com o mesmo cabeçalho de perguntas e respostas da planilha Dados_Resumidos.

Na célula B2 da planilha Dados_Porcentagens entre com a fórmula "=Dados_Resumidos!B2/Dados_Resumidos!\$J 2" para calcular a porcentagem das respostas.

Arraste a fórmula para preencher os demais campos da tabela.

B2 + 1	× ✓	fx	-Dados	Recumi	dosIB2/I	ados R	esumidos	1\$12
DZ .	/\ Y	JA	-Dauos_	_ivesuiiii	u03:D2/1	Jauos_I	esumuos	2 د کې د
A	В	C	D	E	F	G	Н	- 1
1 Pergunta/Respos	ta	1	2	3	4 5	6	7	8
2	1 0,94	23						
3	2							
4	3							
5	4							
6	5							
Pergunta/Resposta	1	2	3	4	5	6	7	8
ergur ta/Resposta 1	1 0,942	2 0,06	3	4 0	5	6	7	8
		_					-	
1	0,942	0,06	0	0	0	0	0	C
1 2	0,942 0	0,06 0,02	0 0,98	0	0	0	0	0
1 2 3	0,942 0 0	0,06 0,02 0,04	0 0,98 0,14	0 0 0,27	0 0 0,14	0 0 0,22	0 0,2	0
1 2 3 4	0,942 0 0 0,333	0,06 0,02 0,04 0	0 0,98 0,14 0,14	0 0 0,27 0,14	0 0,14 0,04 0	0 0 0,22 0	0 0 0,2 0,35	C
1 2 3 4 5	0,942 0 0 0,333 0,51	0,06 0,02 0,04 0	0 0,98 0,14 0,14	0 0,27 0,14 0	0 0,14 0,04	0 0 0,22 0	0 0 0,2 0,35 0	(

Cálculo do Erro

Crie Uma planilha nova com o nome Erros.

Copie o cabeçalho com as perguntas e respostas da planilha Dados_Resumidos.

Na célula B2 coloque a seguinte fórmula "=Alpha*RAIZ((Dados_Porcentagem!B2)*(1-Dados_Porcentagem!B2)*(1/Dados_Resumidos!\$J2))" e arraste a célula para preencher os demais campos da tabela.

B2	· 1	×	~	fx	=Alp	ha*RA	IZ((Dad	los_Por	centag	e	em!B2)*(em!B2)*(1-Dados_	em!B2)*(1-Dados_Porcentage	em!B2)*(1-Dados_Porcentagem!B2)*(1/	em!B2)*(1-Dados_Porcentagem!B2)*(1/Dados_Res
A	Α		В	С	D	Е	F	G	Н		1	I J	I J K	I J K L	I J K L M
1	Pergunta/Respos	sta	1	2	3	4	5	6	7	1	8	8	8	8	8
2		1	0,063												
3		2													
Perg	unta/Resposta		1 2	3	4	5	6	7	8						
	1	0,06	3 0,063	0	0	0	0	0	0						
	2		0 0,038	0,038	0	0	0	0	0						
	3		0 0,053	0,094	0,122	0,094	0,113	0,109	0						
	4	0,12	9 0	0,094	0,094	0,053	0	0,131	0						

Observações sobre o cálculo do Erro



Os parâmetros tamanho da população e nível de confiança podem ser alterados na planilha Dados_Resumidos e automaticamente o Erro é calculado na planilha Erros;



Em geral o nível de confiança utilizado é 95% porém da maneira como foram construídas as planilhas qualquer nível de confiança pode ser obtido.

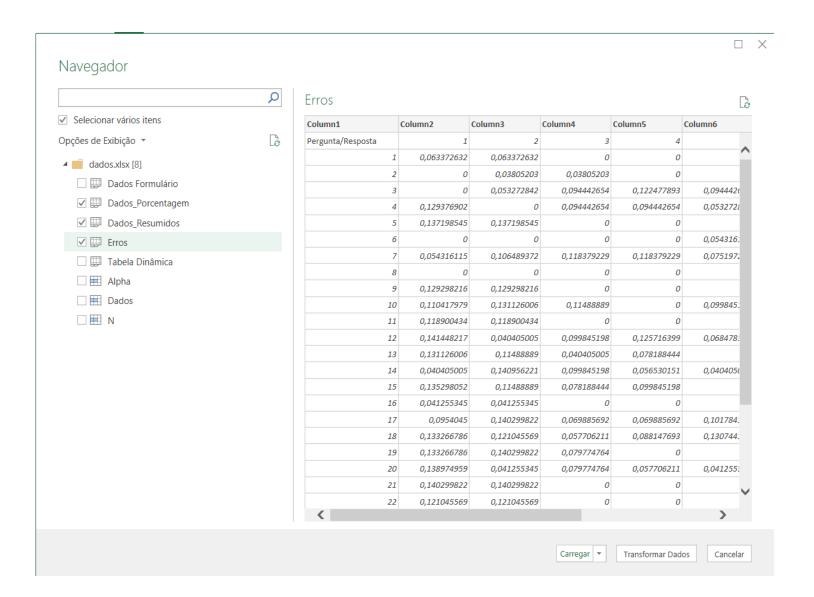


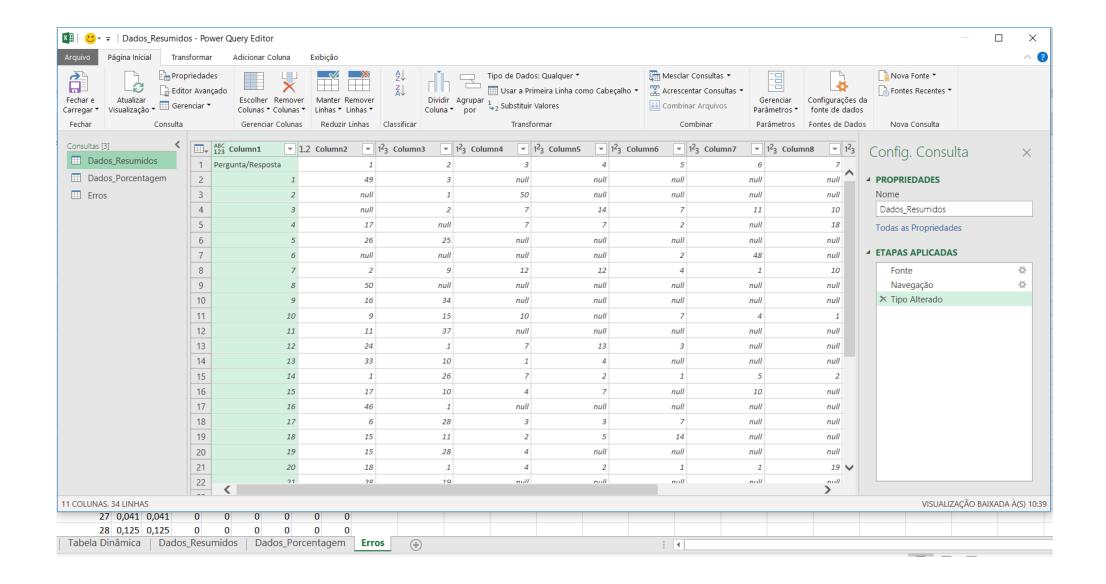
Esse cálculo não é um cálculo simples de intervalo de confiança (cuja função existe no Excel) pois leva em consideração o tamanho da população e calcula um fator de correção. Porém, para populações muito grandes esse fator é irrelevante. Você pode verificar isso alterando o tamanho da população para 1000000 por exemplo, nesse caso o fator fica próximo de 1 e é irrelevante.

Preparação dos Dados

Abra um novo arquivo em branco do Excel, clique em Obter Dados -> De arquivo -> Da pasta de Trabalho, selecione dados.xlsx e clique em importar;

Selecione
Dados_Porcentagem, Dados_Resumidos
e Erros e clique em transformar Dados





Observações



Do lado esquerdo do Power Query temos as três planilhas que importamos, todas elas serão utilizadas para obtenção de um arquivo final onde seja possível utilizar os dados para fazer os gráficos em qualquer software como o próprio Excel, Tableau ou o Power BI por exemplo.



Do lado direito estão os dados de cada uma das planilhas importadas de acordo com a seleção.

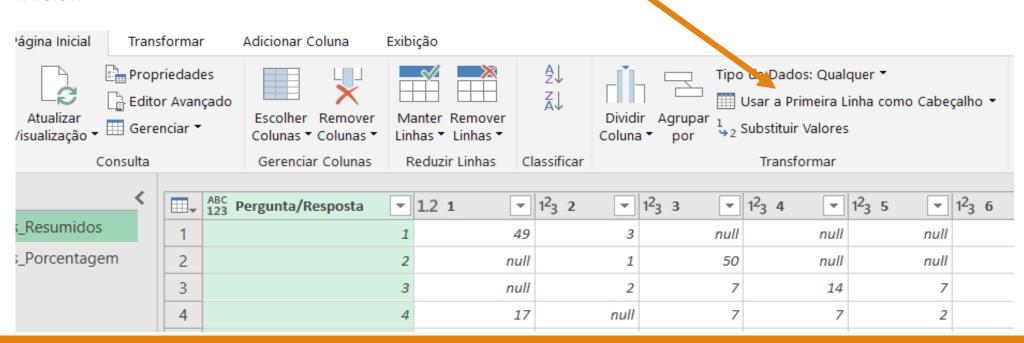


Com o power query é possível fazer várias transformações sobre esses dados, que estarão explicadas adiante.

Promovendo a primeira linha como cabeçalho.

Essa alteração deverá ser feita em todas as planilhas.

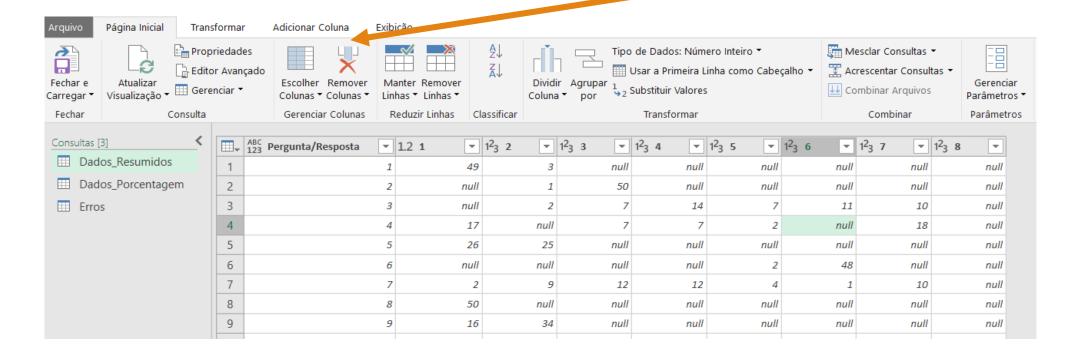
Basta selecionar a planilha e clicar em "Usar a primeira linha como cabeçalho" no menu Página Inicial.



Excluindo dados desnecessários (colunas).

Com a planilha Dados_Resumidos selecionada, selecione as colunas TotalGeral e FPC e clique em remover colunas.

Você deve obter uma tabela como abaixo:



Excluindo dados desnecessários (linhas).

Com a planilha Dados_Resumidos selecionada, vá em Remover linhas -> Remover Linhas inferiores.

Coloque 4 e clique em OK.

Você deve obter uma tabela como abaixo:

12	12	24	1	/	13	3	null	null	null
13	13	33	10	1	4	null	null	null	null
14	14	1	26	7	2	1	5	2	4
15	15	17	10	4	7	null	10	null	null
16	16	46	1	null	null	null	null	null	null
17	17	6	28	3	3	7	null	null	null
18	18	15	11	2	5	14	null	null	null
19	19	15	28	4	null	null	null	null	null
20	20	18	1	4	2	1	1	19	1
21	21	28	19	null	null	null	null	null	null
22	22	36	11	null	null	null	null	null	null
23	23	43	4	null	null	null	null	null	null
24	24	45	2	null	null	null	null	null	null
25	25	16	1	3	1	26	null	null	null
26	26	38	9	null	null	null	null	null	null
27	27	46	1	null	null	null	null	null	null
28	28	35	12	null	null	null	null	null	null
29	29	14	7	6	8	12	null	null	null

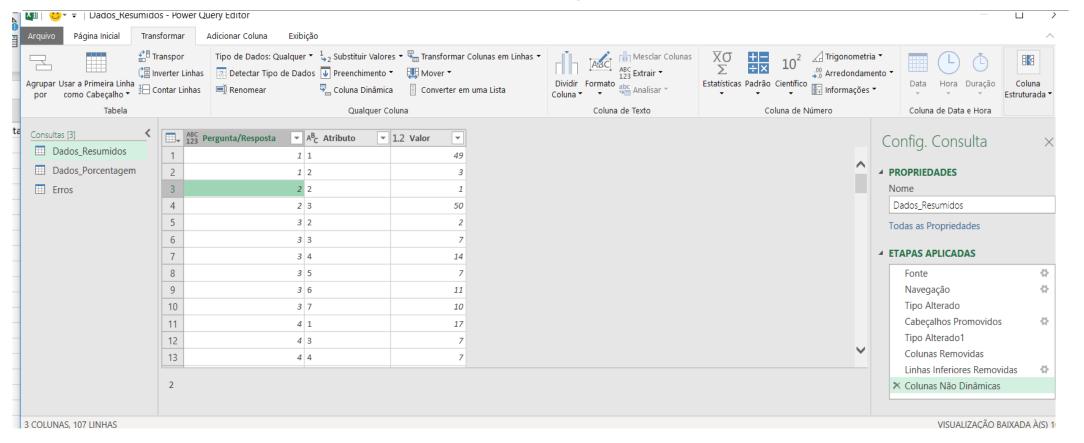
Transformando Colunas em Linhas;

Essa transformação será realizada nas três planilhas;

Por exemplo, na planilha Dados Resumidos selecione todas as colunas a partir da segunda e clique em Transformar -> Transforma colunas em linhas:



Resultado Transformação 4



Aplicando Transformação 4

O mesmo procedimento de Transformar colunas em linhas deverá ser realizado nas demais planilhas;

Observe que a planilha Dados_Resumidos possui 107 linhas e 3 colunas;

Observe que as planilhas Dados_Porcentagem e Erros possuem 3 colunas e 232 linhas;

Essa diferença ocorre pois a planilha Dados_Resumidos possui espaços sem preenchimento de valores. Essa diferença será de extrema importância quando for realizada a operação "JOIN" (mesclar) entre essas planilhas.





Aplicar nas 3 planilhas



Na planilha Dados_Resumidos as colunas foram alteradas para: Pergunta, Resposta e Contagem



Na planilha Dados_Porcentagem as colunas foram alteradas para: Pergunta, Resposta e Porcentagem



Na planilha Erros as colunas foram alteradas para: Pergunta, Resposta e Erro



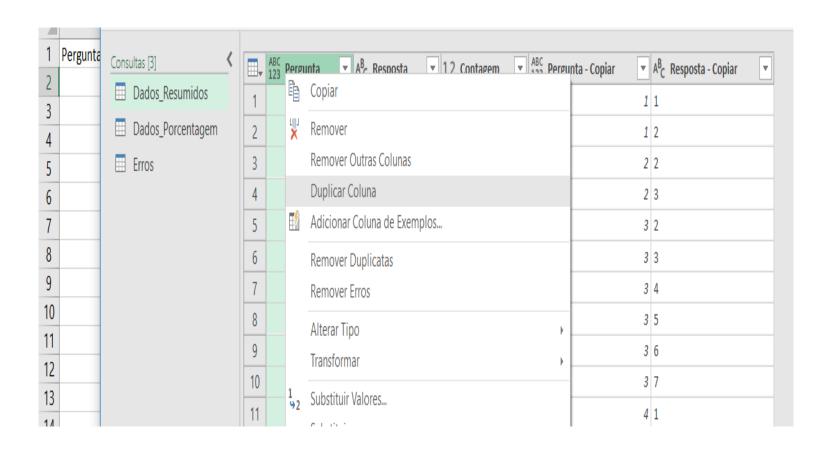
Para alterar o nome de uma coluna basta clicar em cima e reescrever

Criando uma chave:

A transformação será aplicada nas três planilhas.

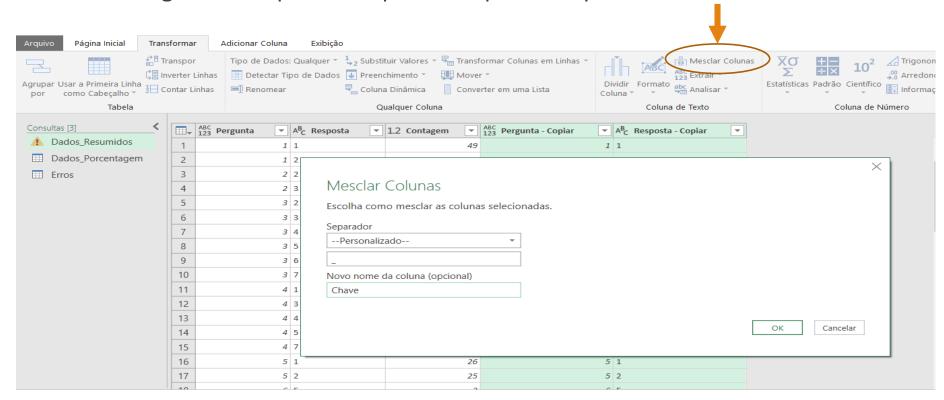
Selecione Pergunta e clique em Duplicar Colunas.

Selecione Resposta e clique em Duplicar Colunas.



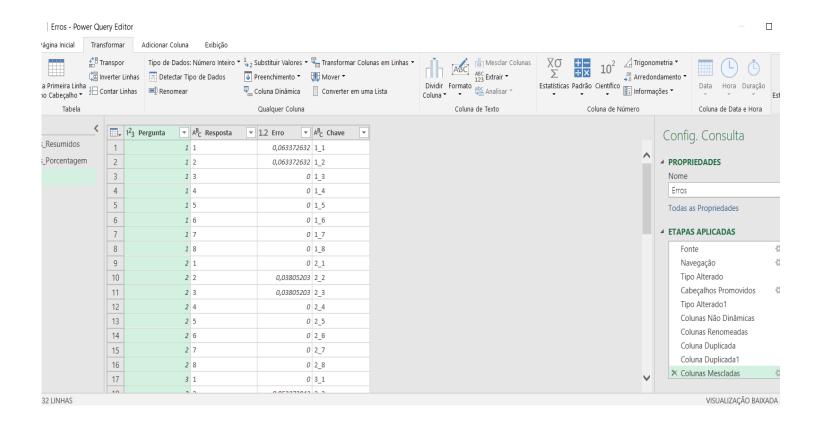
Criando uma chave:

Selecione Pergunta – Copiar e Resposta- Copiar e clique em Mesclar Colunas



Agora você tem uma chave que associa uma pergunta a uma resposta.

Repita o procedimento nas demais planilhas.

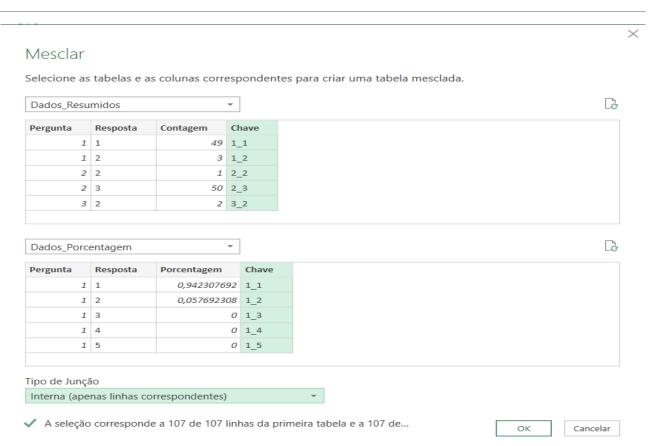


Mesclando Consultas (Join):

Clique em Mesclar Consultas -> Mesclar Consultas como novas.

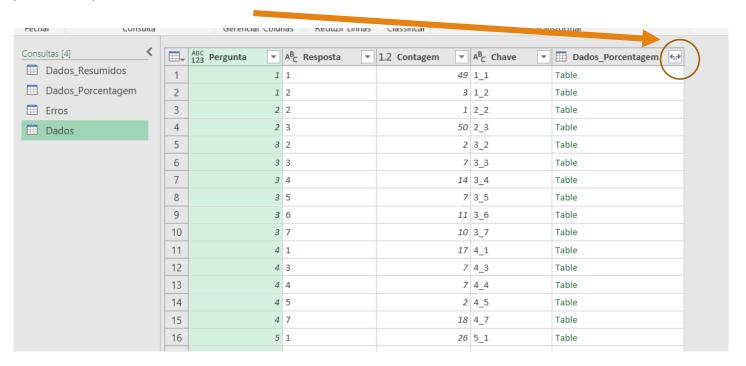
Primeiro vamos mesclar Dados_Resumidos com Dados_Porcentagem.

Selecione as opções como ao lado e clique em OK.

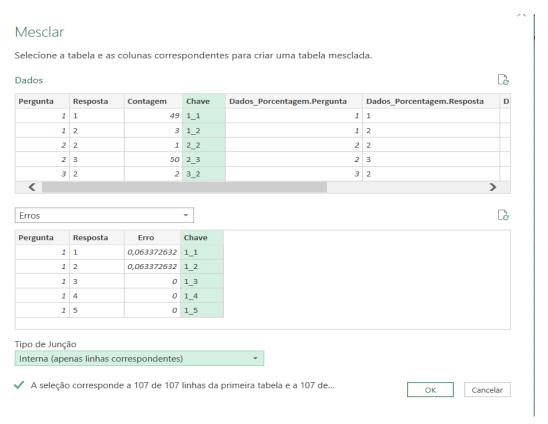


Observe que apareceu uma nova planilha com o nome de Merge1, vamos alterar esse nome para Dados

E clicar aqui para expandir os dados:



Com a planilha Dados, vamos agora clicar em mesclar consulta ("Join") e mesclar os Dados com Erros:



Exclusão das colunas repetidas.

Clique em expandir na planilha Dados e elimine todas as colunas que não serão utilizadas, restando apenas as seguintes colunas:

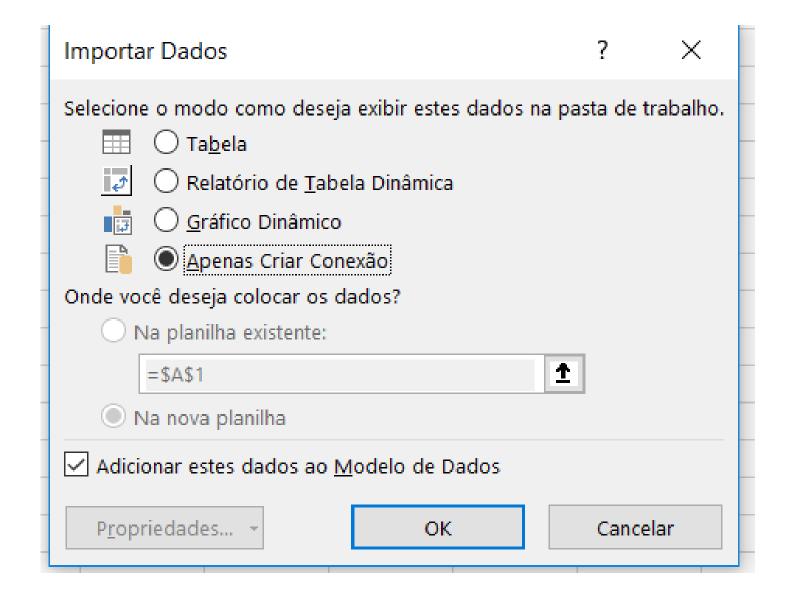
-	ABC Pergunta 💌	AE	B _C Resposta ▼	1.2 Contagem ▼	A	^B _C Chave ▼	1.2 Dados_Porcentagem.Porcenta	1.2 Erros.Erro
1	1	1		49	1	_1	0,942307692	0,0633726
2	1	2		3	1	_2	0,057692308	0,0633726
3	2	2		1	2	2_2	0,019607843	0,0380520
4	2	3		50	2	2_3	0,980392157	0,0380520
5	3	2		2	3	3_2	0,039215686	0,05327284
6	3	3		7	3	3_3	0,137254902	0,0944426
7	3	4		14	3	3_4	0,274509804	0,1224778
8	3	5		7	3	3_5	0,137254902	0,0944426
9	3	6		11	3	3_6	0,215686275	0,1128805
10	3	7		10	3	3_7	0,196078431	0,1089644
11	4	1		17	4	_1	0,333333333	0,12937690
12	4	3		7	4	_3	0,137254902	0,0944426
13	4	4		7	4	1_4	0,137254902	0,0944426
4.4		-				-	0.000045505	0.0500700

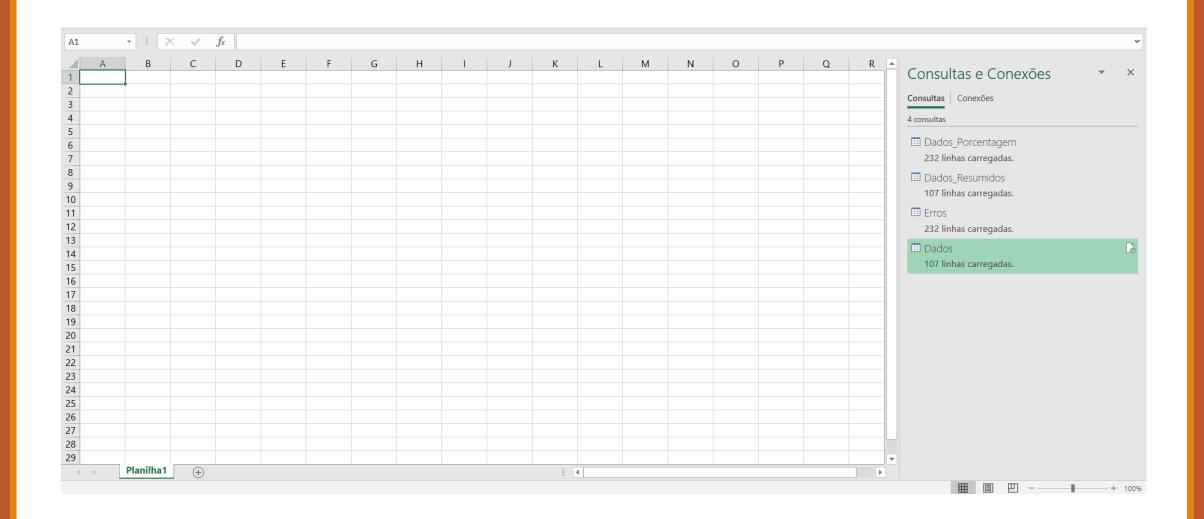
Carregando os Dados

Clique em Fechar e Carregar para.

Selecione Apenas Criar Conexão.

Selecione Adicionar estes dados ao modelo de Dados e clique em OK.



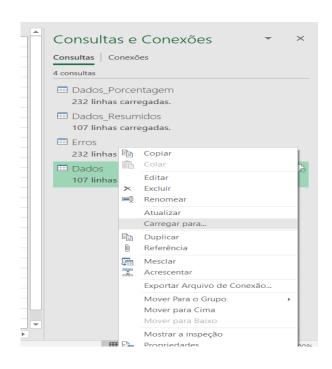


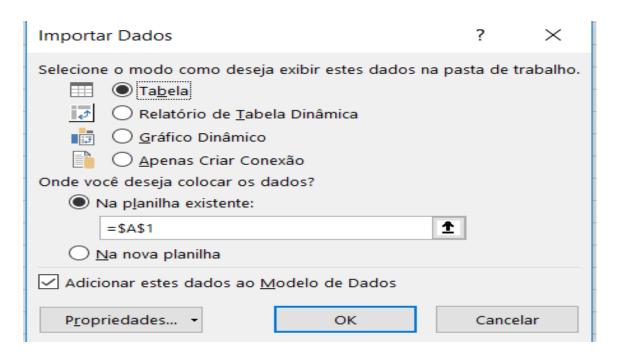
Carregamento de Dados

Aparentemente parece que não carregou nada, mas os dados estão no modelo de Dados.

Clique em Dados e Carregar para.

Selecione as opções e clique em OK:





4	Α	В	С	D	E	F	G
1	Pergunta 💌	Resposta 💌	Contagem 💌	Chave 💌	Dados_Porcentagem.Porcentagem 💌	Erros.Erro	
2	1	1	49	1_1	0,942307692	0,063372632	
3	1	2	3	1_2	0,057692308	0,063372632	
4	2	2	1	2_2	0,019607843	0,03805203	
5	2	3	50	2_3	0,980392157	0,03805203	
6	3	2	2	3_2	0,039215686	0,053272842	
7	3	3	7	3_3	0,137254902	0,094442654	
8	3	4	14	3_4	0,274509804	0,122477893	
9	3	5	7	3_5	0,137254902	0,094442654	
10	3	6	11	3_6	0,215686275	0,112880563	
11	3	7	10	3_7	0,196078431	0,108964431	
12	4	1	17	4_1	0,333333333	0,129376902	
13	4	3	7	4_3	0,137254902	0,094442654	
14	4	4	7	4_4	0,137254902	0,094442654	
15	4	5	2	4_5	0,039215686	0,053272842	
16	4	7	18	4_7	0,352941176	0,131155362	
17	5	1	26	5_1	0,509803922	0,137198545	
18	5	2	25	5_2	0,490196078	0,137198545	
19	6	5	2	6_5	0,04	0,054316115	
20	6	6	48	6_6	0,96	0,054316115	
21	7	1	2	7_1	0,04	0,054316115	
22	7	2	9	7_2	0,18	0,106489372	
23	7	3	12	7_3	0,24	0,118379229	
24	7	4	12	7_4	0,24	0,118379229	
25	7	5	4	7_5	0,08	0,075197256	
26	7	6	1	7_6	0,02	0,038805307	
27	7	7	10	7_7	0,2	0,110872306	
28	8	1	50	8_1	1	0	
29	9	1	16	9_1	0,32	0,129298216	
4	P	anilha1	+				: 4

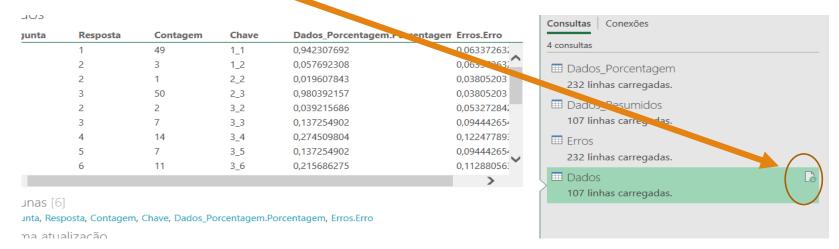
Observações

Altere o nome planilha1 para Dados.

No arquivo novo do Excel com nome dados_processados, estão todas as informações necessárias para se construir os dashboards com intervalos de confiança para as perguntas.

É importante notar que se for alterado a confiança e a população na planilha dados e o arquivo salvo, basta atualizar a planilha dados_processados para que os novos dados de erro sejam mostrados.

Por exemplo, vou alterar a confiança para 90% e atualizar dados_processados.



Observações:

Agora temos uma nova tabela de erros com 90% de confiança:

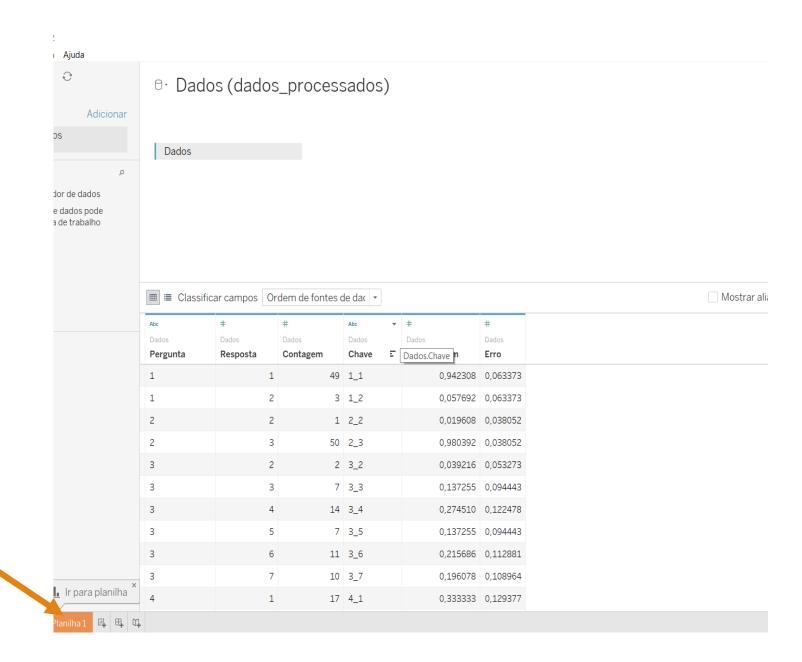
	Α	В	С	D	E	F	G			
1	Pergunta 💌	Resposta 💌	Contagem 💌	Chave 💌	Dados_Porcentagem.Porcentagem 🔻	Erros.Erro				
2	1	1	49	1_1	0,942307692	0,053183989				
3	1	2	3	1_2	0,057692308	0,053183989				
4	2	2	1	2_2	0,019607843	0,03193427				
5	2	3	50	2_3	0,980392157	0,03193427				
6	3	2	2	3_2	0,039215686	0,044707978				
7	3	3	7	3_3	0,137254902	0,079258774				
8	3	4	14	3_4	0,274509804	0,102786688				
9	3	5	7	3_5	0,137254902	0,079258774				
10	3	6	11	3_6	0,215686275	0,094732354				
11	3	7	10	3_7	0,196078431	0,091445833				
12	4	1	17	4_1	0,333333333	0,108576518				
13	4	3	7	4_3	0,137254902	0,079258774				
	4	4	7	4_4	0,137254902	0,079258774				
15	4	5	2	4_5	0,039215686	0,044707978				
16	4	7	18	4_7	0,352941176	0,110069049				
17	5	1	26	5_1	0,509803922	0,115140648				
18		2	25	5_2	0,490196078	0,115140648				
19	6	5	2	6_5	0,04	0,045583521				
20	6	6	48	6_6	0,96	0,045583521				
21	7	1	2	7_1	0,04	0,045583521				
22	7	2	9	7_2	0,18	0,089368699				
23	7	3	12	7_3	0,24	0,09934698				
24	7	4	12	7_4	0,24	0,09934698				
25	7	5	4	7_5	0,08	0,063107527				
26	7	6	1	7_6	0,02	0,03256644				
27	7	7		7_7	0,2	0,093046972				
28	8	1	50	8_1	1	0				
29	9	1		9_1	0,32	0,108510484				
Dados +										

Gráficos utilizando o Tableau

Importando os dados processados do arquivo dados_processados no Tableau.

Clique em conectar a um arquivo Excel e selecione o arquivo dados_processados.

Clique em Ir para planilha.

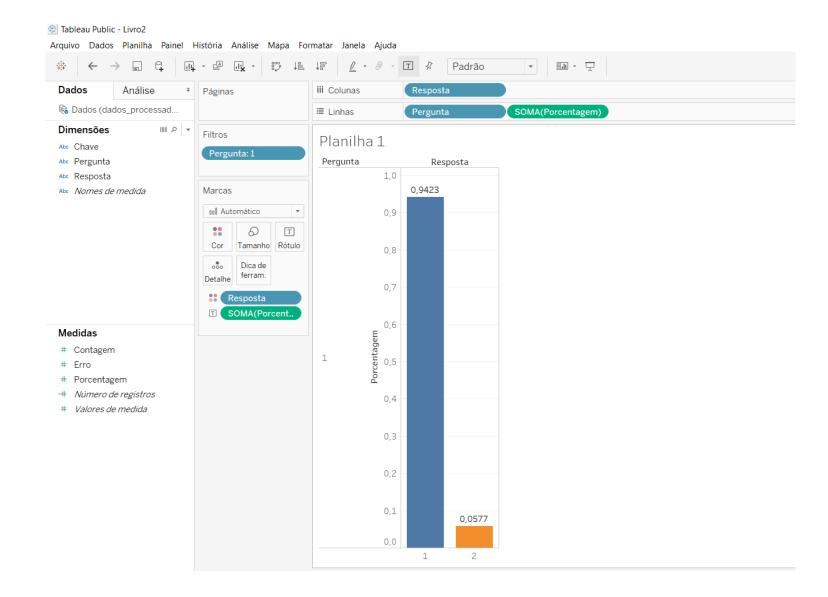


Organizando Dimensões e medidas

Altere o tipo de Dados da variável Resposta para cadeia de caracteres, clicando com o botão direito na medida.

Coloque a variável Resposta em Dimensões.

Organize as variáveis nas linhas e colunas como mostrado ao lado.



Criando Medidas de máximo, mínimo e intervalo

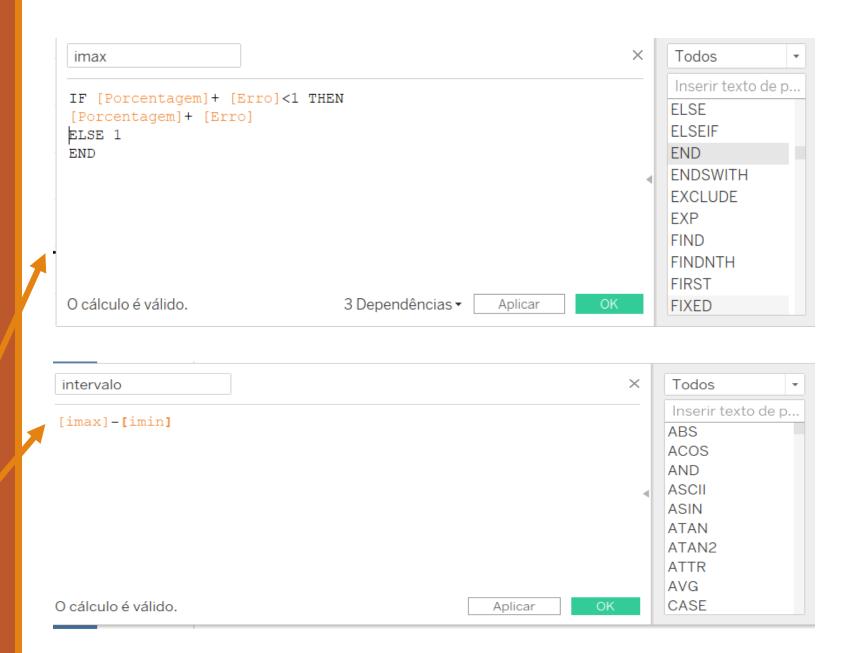
No campo medidas, clique em criar campo calculado.

Nomeie o campo como imin.

Imin = IF [Porcentagem]- [Erro]>0
THEN [Porcentagem]- [Erro] ELSE 0
END

Repita procedimento semelhante e crie imax.

Crie um novo campo calculado de nom intervalo = imax - imin

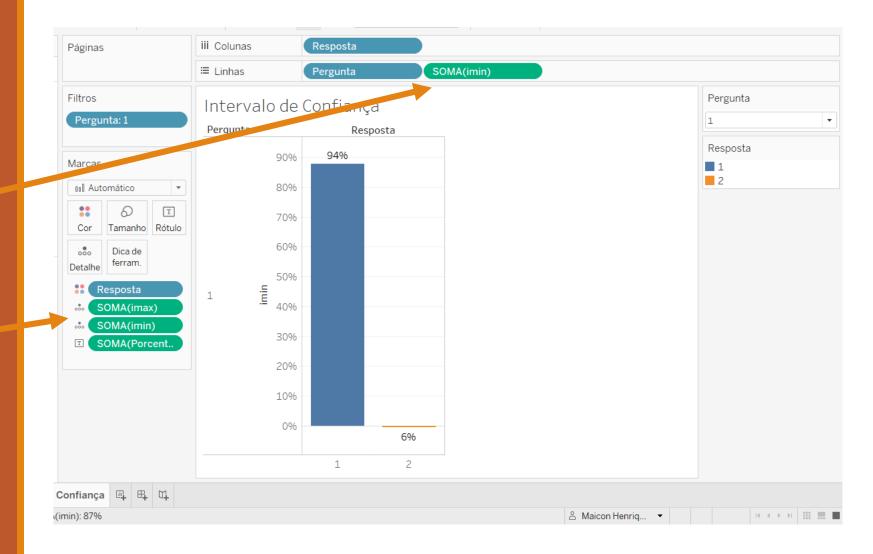


Criando as Barras de erro

Substitua a variável porcentagem por imin no gráfico.

Formate os eixos e os valores como desejar.

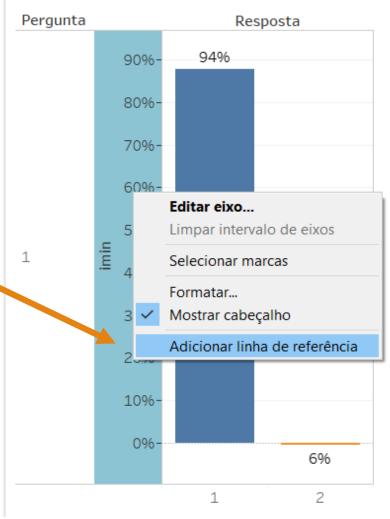
Adicione imax e imin em detalhes.



Criando as barras de Erros

Clique sobre o eixo do gráfico e em adicionar linha de referência:





Criando as Barras de Erros

Selecione Faixa por célula.

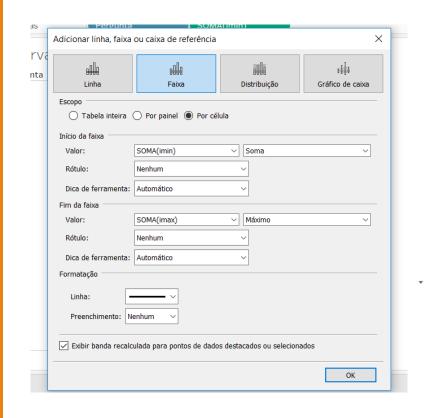
Identifique os valores de mínimo e máximo.

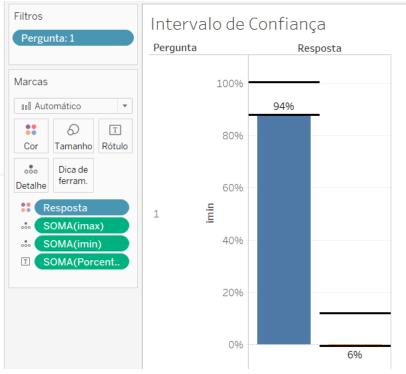
Altere a a expessura da linha.

Selecione preenchimento nenhum.

Clique em OK.

Resultado.

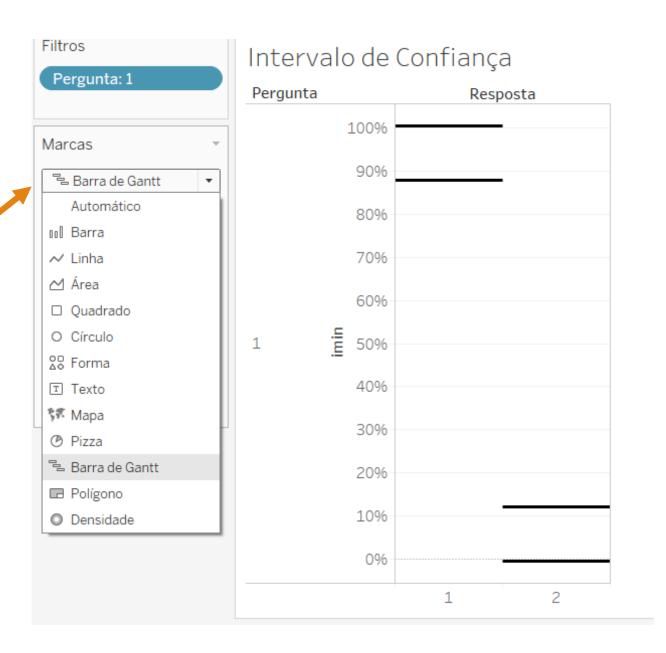




Alterando o Gráfico

Retire do texto a variável porcentagem.

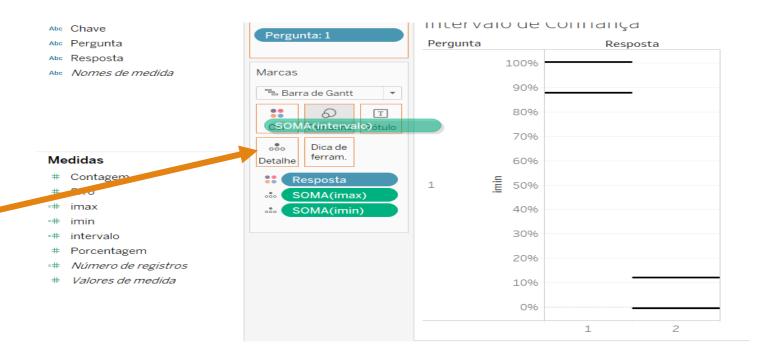
Altere o gráfico para Barra de Grant.

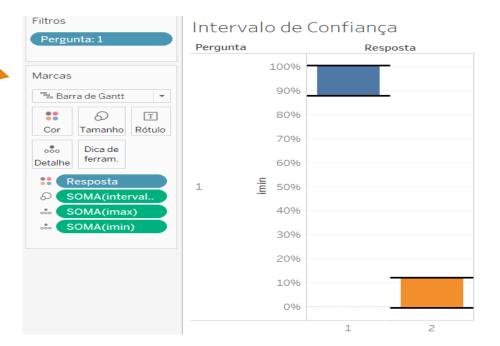


Tamanho da Barra

Arraste a variável Intervalo para tamanho.

O resultado deve ser o seguinte.



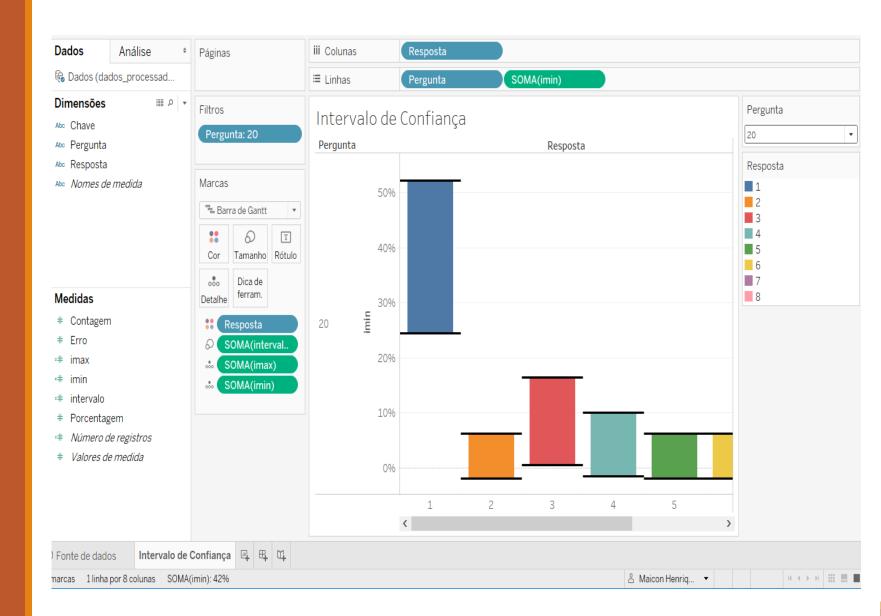


Pronto!

O dashboard com intervalo de confiança foi criado.

Basta selecionar uma pergunta e temos o gráfico.

Por exemplo a pergunta 20.



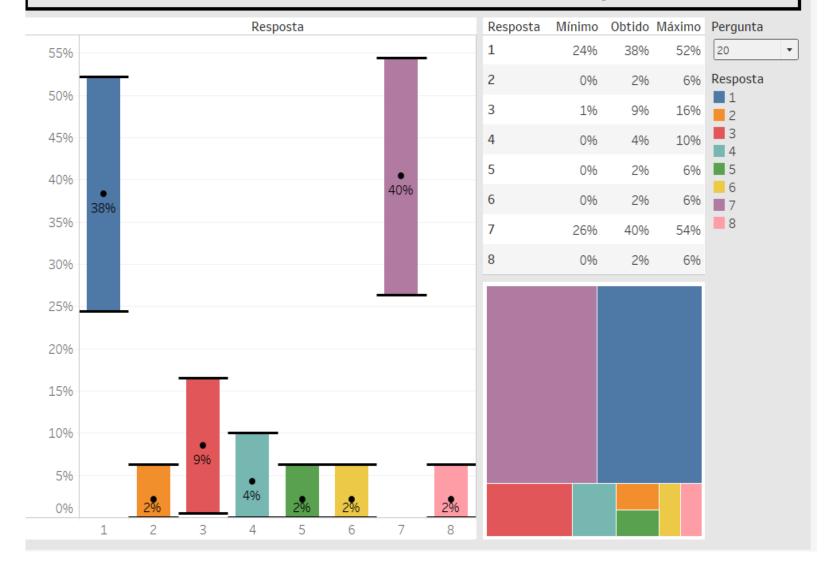
Melhorando o Dashboard

Algumas melhorias que podem ser feitas.

Link para o dashboard interativo aqui.

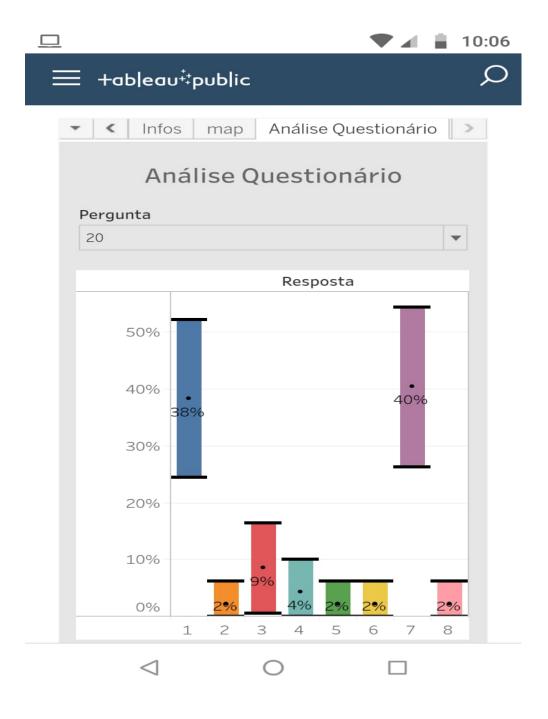
Visualizações alternativas para tablet e celular foram criadas de modo a manter o mesmo nível de interação com as diversas plataformas.

Análise Questionário - 95% Confiança



Visualizando com o celular

Versão simplificada é otimizada para celular.



Observações finais sobre o Dashboard no Tableau

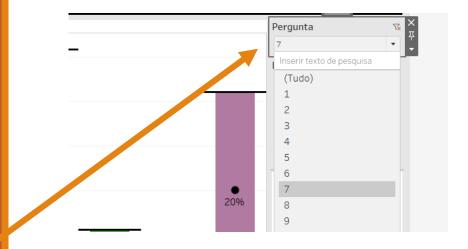
O dashboard apresentado anteriormente é de fácil visualização. Qualquer pessoa mesmo que não conheça estatística pode olhar para as barras e ver se uma resposta é significativa ou não com relação a população, bastando apenas observar se os intervalos de confiança se intersectam.

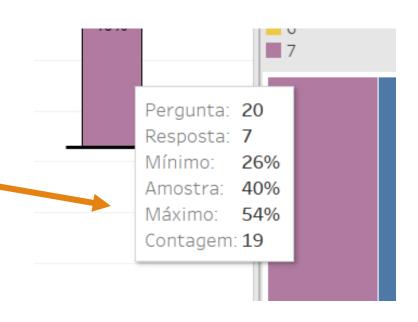
Qualquer pergunta pode ser analisada bastando selecionar.

O gráfico de árvore na lateral direita associa as áreas com as respostas de acordo com a porcentagem obtida nas respostas.

Passando o mouse sobre o gráfico é possível obter mais informações como contagem de resposta e os intervalos de confiança propriamente.

É possível fazer seleção por resposta clicando em uma resposta específica.





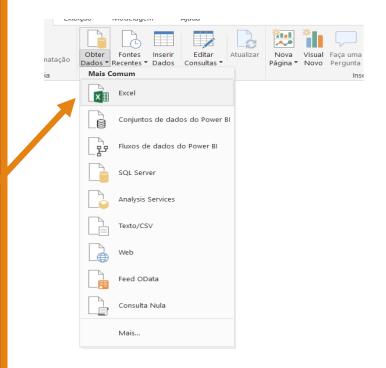
Gráficos com Power Bl

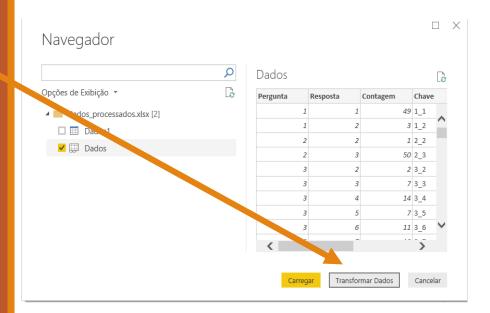
Importando os dados.

Com o Power BI aberto, clique em Obter Dados -> Excel.

Selecione o arquivo dados_processados.xlsx já utilizado anteriormente.

Selecione a planilha Dados e clique em Transformar Dados.



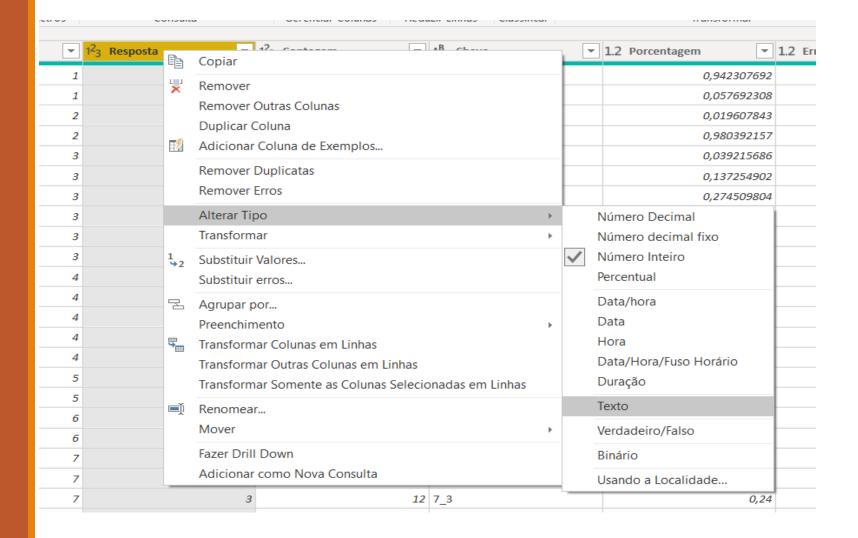


Alterando a coluna Resposta

Alterando o tipo de dado.

Selecione a coluna Resposta e clique com o botão direito do mouse.

Clique em alterar tipo e mude para texto.

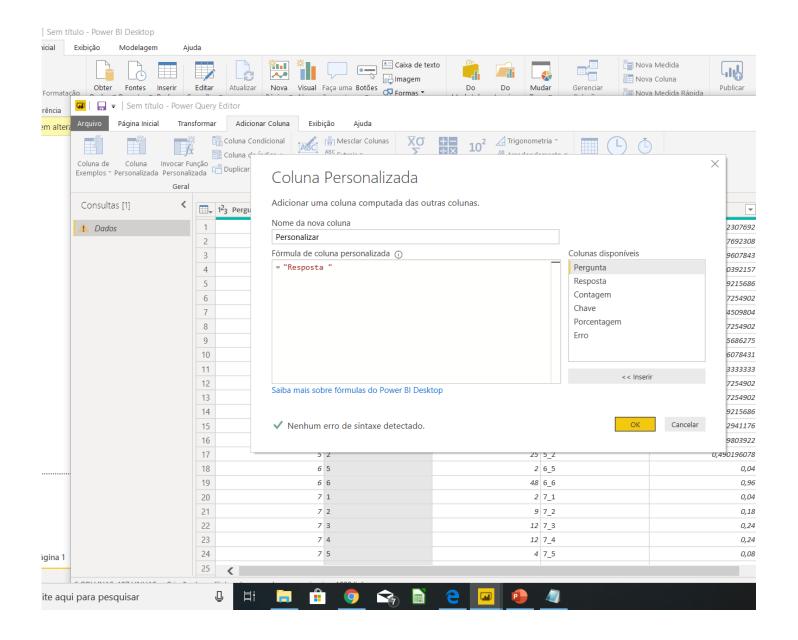


Alterando a coluna Resposta

Criando uma nova coluna Personalizar.

Clique em Adicionar Coluna -> Coliuna Personalizada.

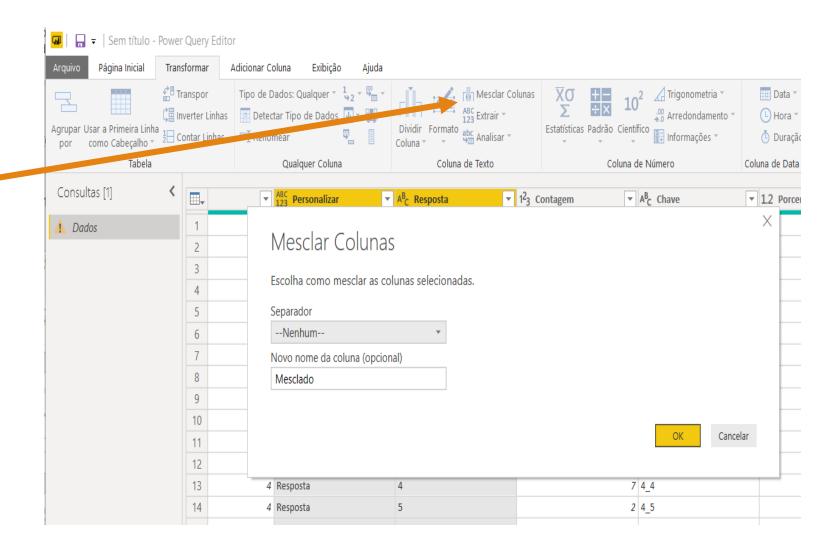
E configure igual a figura ao lado.



Mesclando as colunas

Selecione as colunas Personalizar e Mesclar.

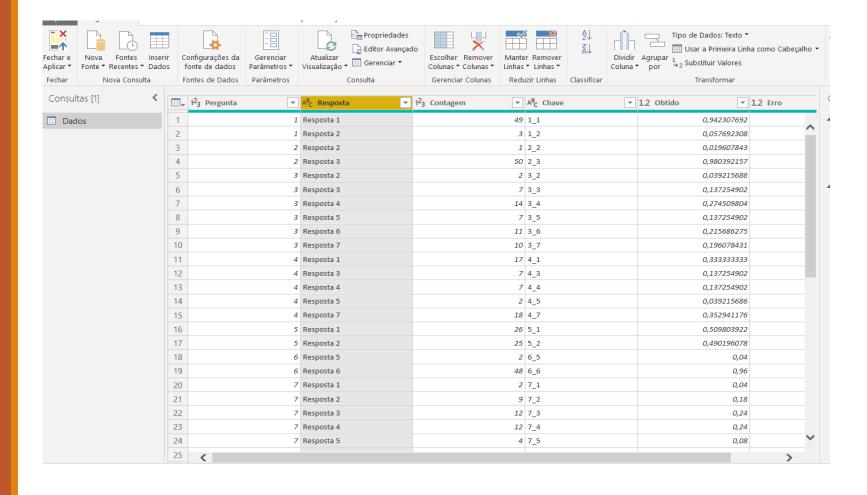
Clique em Mesclar Colunas.



Alterando o nome de Coluna

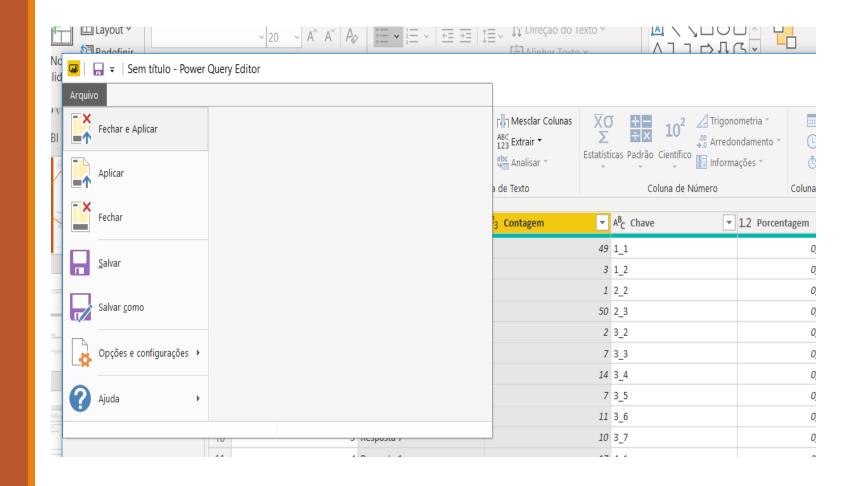
Altere o nome da coluna Porcentagem para Obtido.

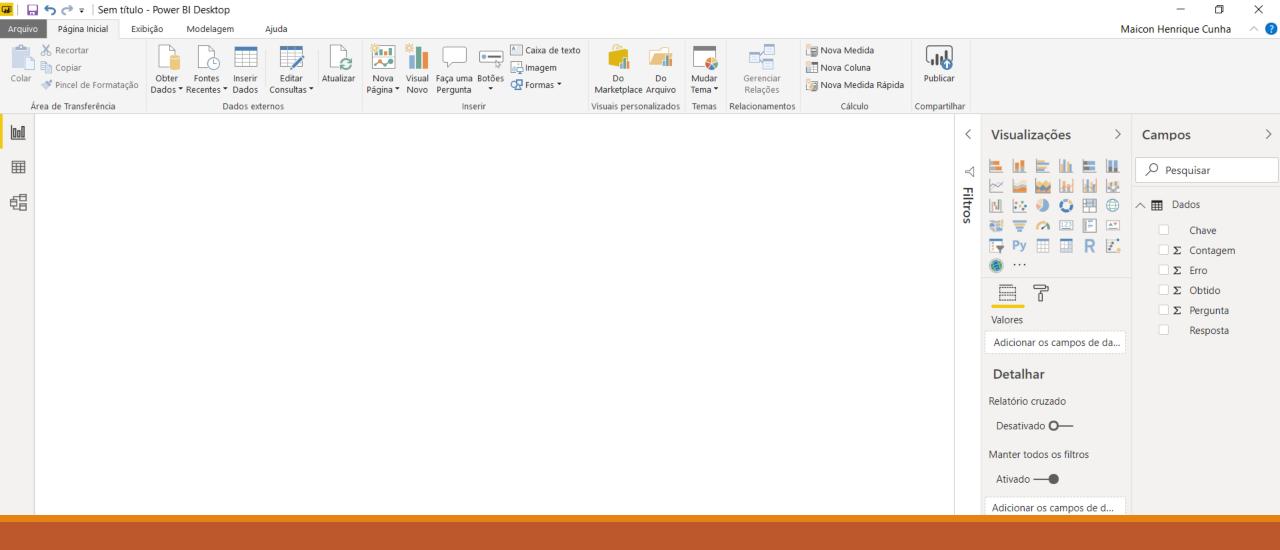
Basta clicar com o botão direito e renomear.



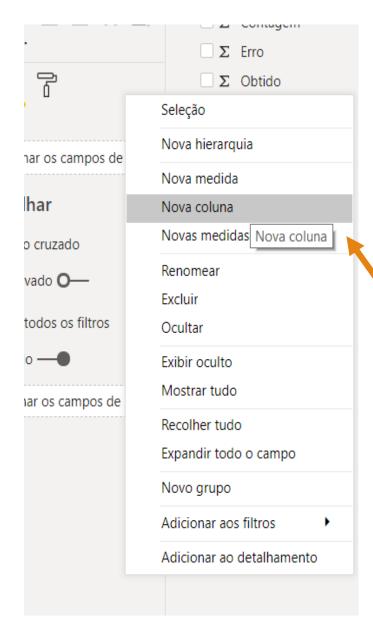
Fechando o Power Query

Clique em fechar e aplicar para aplicar as alterações feitas com os dados.

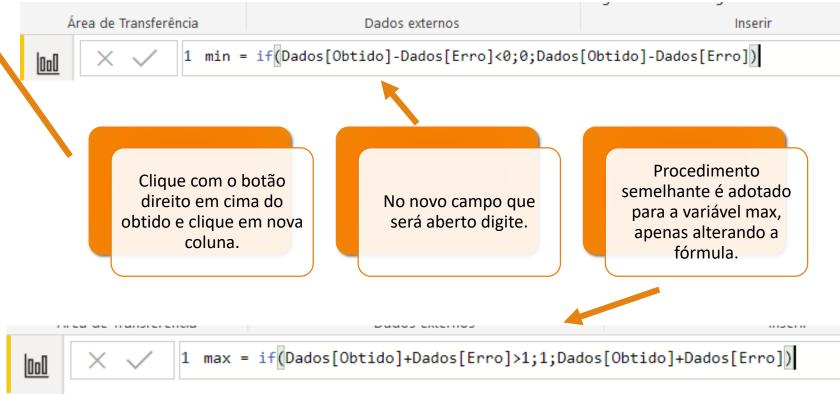




Visual do Power BI com os dados



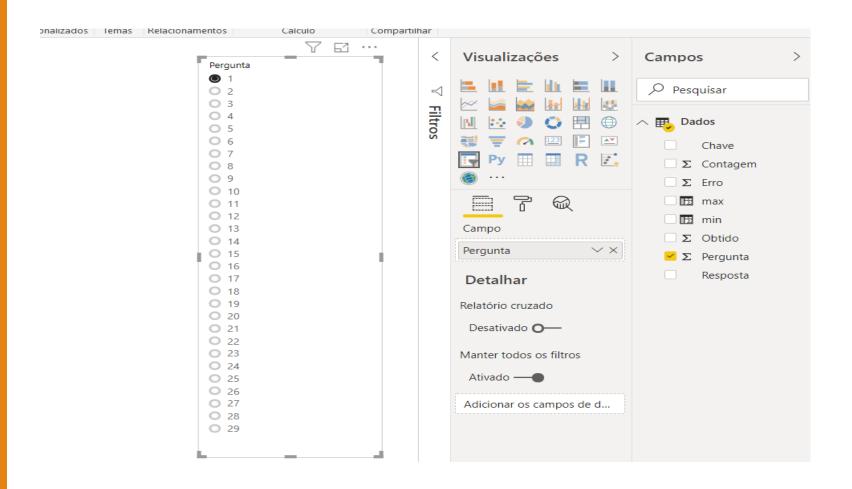
Criando as variáveis min e max



Adicionando um filtro para as perguntas

Adicione um filtro para as perguntas de modo a facilitar a visualização dos gráficos.

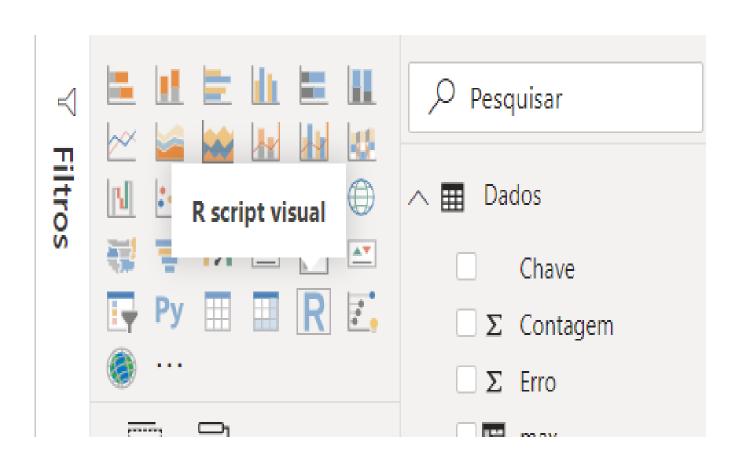
Basta clicar em segmentação de dados e arrastar a Pergunta para o campo de segmentação.



Criando um visual personalizado no R dentro do Power BI

O Power BI não possui um modelo para mostrar intervalos de confiança, por isso utilizaremos o R dentro do power BI.

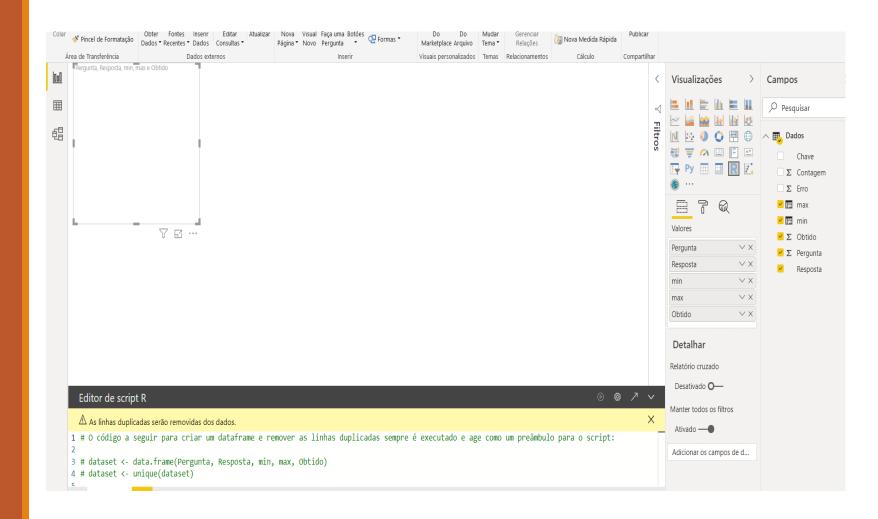
Desde que o R esteja instalado com os pacotes necessários, basta clicar em R script visual para editar um gráfico em R que será exibido no Power BI.



Adicionando variáveis no Script R

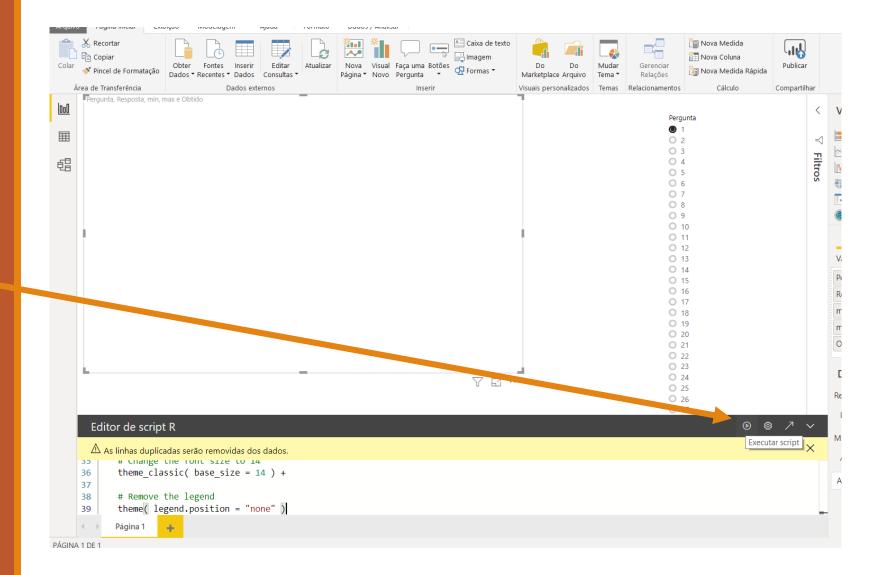
Arraste as seguintes variáveis para o espaço com um R indicando o sript R que será gerado:

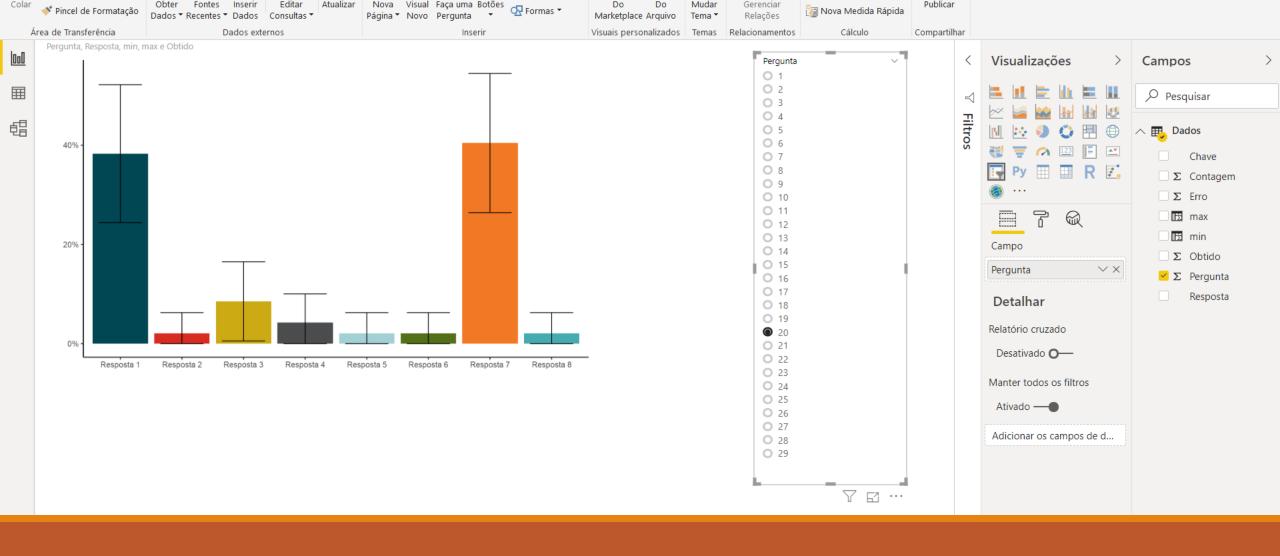
Pergunta, Resposta, min, max e Obtido.



Adicionando o script R

Copie o texto em <u>grafico erro.txt</u> dentro do campo Editor de Script R e clique em executar script.



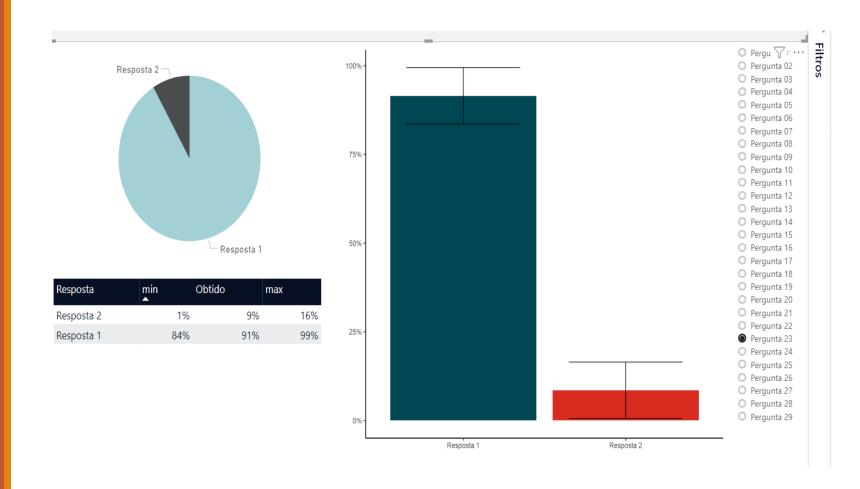


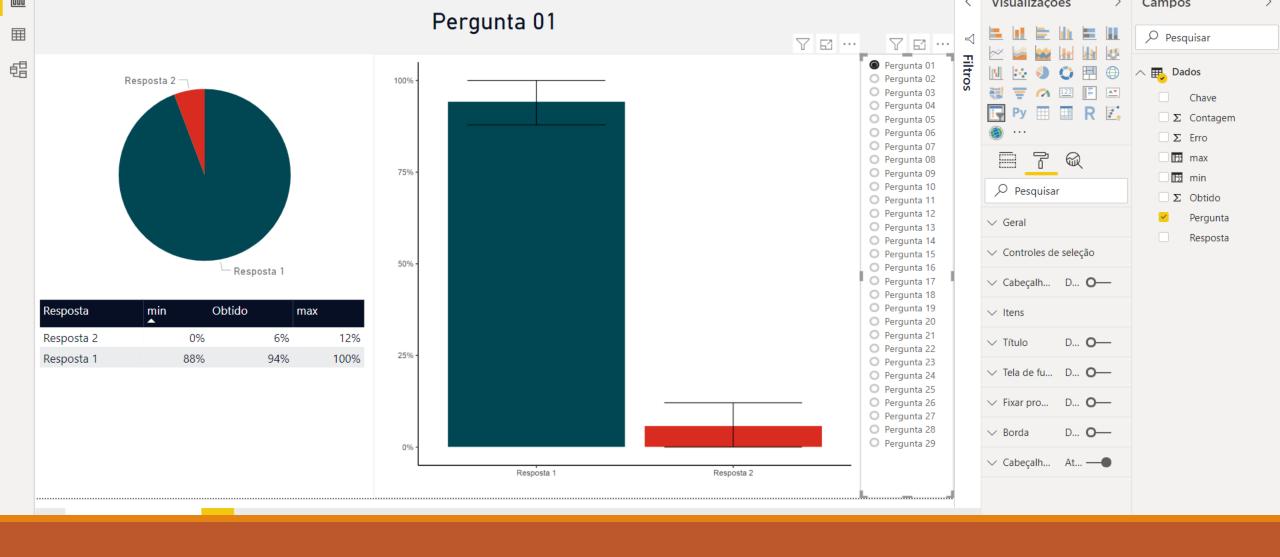
Mostrando o gráfico com as barras de erro

Observações sobre o gráfico feito em R

Existe uma linha no código R em que defino manualmente a cores para cada resposta. Esse código é necessário caso o gráfico no R venha a ser combinado com outros gráficos no Power BI para que as cores se mantenham as mesmas.

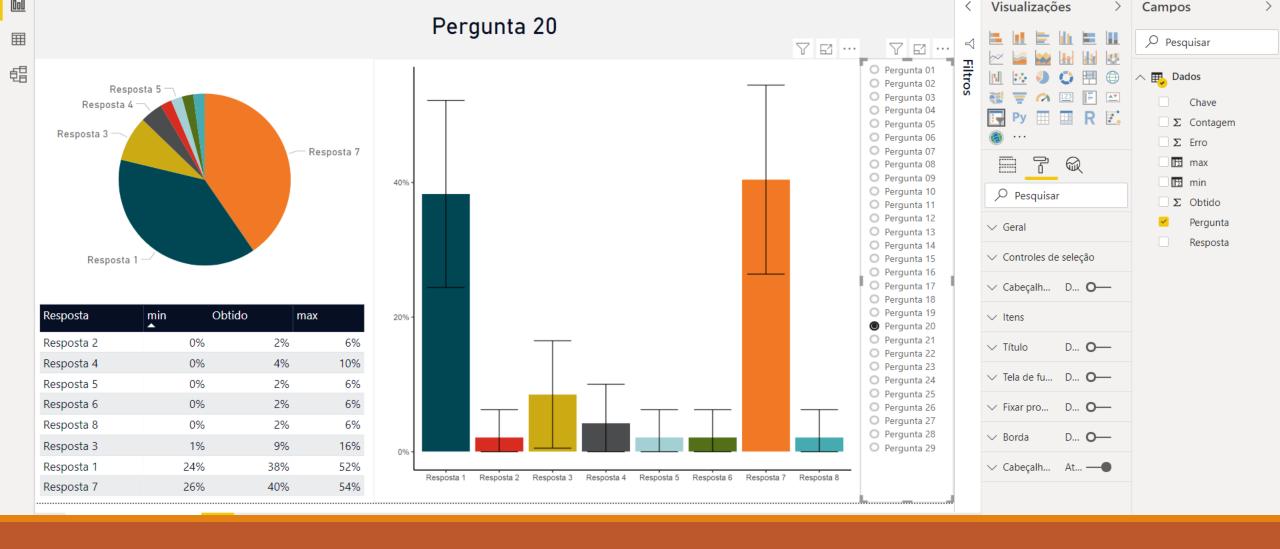
Segue um exemplo do que poderia acontecer caso as cores não fossem definidas manualmente.





Visualização final no Power BI

COM ALGUMAS MELHORIAS E INCREMENTOS PODE SER OBTIDO O SEGUINTE DASHBOARD NO POWER BI:



Visualização final no Power BI

COM ALGUMAS MELHORIAS E INCREMENTOS PODE SER OBTIDO O SEGUINTE DASHBOARD NO POWER BI:

Considerações finais



O cálculo do erro foi feito no Excel pela familiaridade que as pessoas tem com o software (poderia ser feito diretamente no Tableau ou no Power BI) e pela facilidade de alteração e manutenção, possibilitando interagir com os dados processados nas mais diversas plataformas.



Tanto o Power BI como o Tableau se mostraram eficientes para a construção do dashboards interativos a partir dos dados processados no Excel.



Qualquer pessoa pode facilmente escolher uma pergunta do questionário, olhar rapidamente e ver se há intersecção entre as barras de erro para identificar se uma resposta é significativa com relação à população ou não em ambos (Power BI e Tableau), facilitando a rotina do usuário final tomador de decisão.