

Visualização de dados com Tableau

II Seminário Abraji (Assoc. Bras. de Jornalismo Investigativo) – Março 2014

Guilherme R. Storck gstorck@gmail.com @gstorck

www.gazetadopovo.com.br/dados tableaupublic.com/profile/infografiagp dadosfinos.blogspot.com.br

visualização da informação infografia mapas jornalismo de dados arquitetura da informação design de interfaces dados abertos mobilidade





>> CRIME SEM CASTIGO

Consulte as estatísticas sobre assassinatos

A Gazeta do Povo montou um banco de dados interativo com números, gráficos e mapas sobre a investigação de mil homicídios dolosos ocorridos em Curitiba entre 2010 e 2013

04/08/2013 00:01	K			
Curtir 0	🈏 Twe	eetar 6	8+1 6	Comentários (0)
Мара	Gráficos	Linha do	tempo	

Mapa das 1000 mortes

http://goo.gl/98Z343

http://public.tableausoftware.com/views/HomicidiosCuritiba/Mapa

Reportagem

http://www.gazetadopovo.com.br/vidaecidadania/crime-sem-castigo/

A reportagem

Banco de dados montado a partir de inquéritos da Polícia Civil e de informações do Poder Judiciário

- 1 ano e meio
- mais 145 mil páginas (físicas) de inquéritos policiais

Período apurado:

- 2010 (inteiro)
- 2011, 2012 e 2013 (1° trim)

Equipe

Bruna M. Walter, José M. Lopes, Rogerio W. Galindo e Rosana Félix Ilustração e Diagramação: Osvalter Urbinati e Robson Vilalba

Endereço da morte DO (RIME Cidade Tipo de local de crime O que foi avaliado: Idade da vítima Inquéritos em 3 situações: Endereço da vítima • ainda em tramitação; Familiar • já arquivados sem solução; **Primeiro depoimento** • suspeito identificado e levado a Júri. Etnia 40 critérios de cada caso. aros • em que bairro morava a vítima, • qual foi a arma usada, exame do local de morte se havia alguém preso pelo crime e unhas (do crime) • quantas testemunhas foram ouvidas, enicas ANDAMENTO Resolvido • DA INVESTIGAÇão Indiciados

Passo a passo





Limpando os dados Selecionar informações úteis

comparáveis e quantificáveis

Datas, locais, número de tiros, arma utilizada

comparáveis e quantificáveis

bin pe

AL MODAND

Delegado

Ultima informação

Segundo delegado

Terceiro delegado

Outros delesados

Total de delesados

Data do inquérito

Local da monte

ind a dist

endareso darmana

Data do crime

Segredo de Justica

arturna

Gênero

Pácinas

HUMPHUM

Camilar

DOIRI

the thirty the

Alfred

Mau

RO

N.A.

PRIVACIDADE

No Tableau Public tudo fica disponível 201 para ser baixado Segundo delegado

DO I PI

comparáveis e quantificáveis

DIA DO

Whima

Gênero

Pácinas

Segredo de Justica

Ultima informação

Terceiro delegado

Outros delegados

Total de delegados

Data do inquérito

Local da monte

Data do crime

Delegado



Dados gerais > organizada rápida



Mais tempo para a limpeza demorada:





Limpando endereços

- editor de texto (busca avançada)
- google refine (mesclar ocorrências parecidas)





Google Refine



Antes de começar a parte demorada

Versão preliminar dos gráficos

- enviar para a equipe
- ajudar nas conclusões antes de escrever matérias
- Tableau é versátil: gerar gráficos e compartilhar





Mapeamento (geocoding)



- Bom pra encontrar número grande de endereços
- Não teve problema com limite de linhas
- Foi possível conferir os mapeamentos e corrigir quando necessário
- Equipe: Fabiano de Miranda e Lucas dos Santos







Google Fusion Tables KML

Long

XLS

- download KML
- extrair coordenadas
- incluir coordenadas na base



Produção do interativo



- Escolha das informações
- Geração dos gráficos
- Fazer gráficos/filtros funcionarem ao mesmo tempo

limpar

Produção do



- Escolha das info
- Geração dos gr
- Fazer gráficos/



Mapa Gráficos Linha do tempo

1000 homicídios em Curitiba

Veja os crimes analisados pela Gazeta do Povo, ocorridos em 2010 e nos 1º trimestres de 2011, 2012 e 2013

Passe o mouse sobre os pontos para mais informações. Para ampliar: use os controles (+) e (-) ou dê dois cliques em uma área vazia. Para arrastar: clique no mapa e segure até aparecer a mãozinha.



Mapa Gráficos Linha do tempo

1000 homicídios em Curitiba

Veja os crimes analisados pela Gazeta do Povo, ocorridos em 2010 e nos 1º trimestres de 2011, 2012 e 2013

Passe o mouse sobre as barras para mais informações.



Filtros simultâneos

Onde morava	
(All)	
PERFIL DO CRIM	4E
Bairro	
(All)	
Ano	
(All)	-]
(All)	
2010	
2011	
2012	
10040	
2013	

Local do crime

١		

Ŧ

Gênero (IIA) 🕙 Masculino) Feminino Envolvido c/ drogas (IIA) 💿) Sim Não ON.D. Passag. pela polícia (IIA) 💽

Sim

Não

ON.D.

Tooltips (mouseover)

Causa da morte: Arma de fogo 810 homicídios

-

Isso representa 81% dos homicíclice considerando os filtros s Se nenhuma seleção foi Av. F

Novo Mundo Av. República Argentina

Perfil do crime Ano: 2010 Mês: janeiro Causa: Arma de fogo Local: Via pública Tempo entre o crime e a abertura do inquérito: 68 dias

Perfil da vítima Etnia: Pardo Gênero: Masculino Idade: Até 20 os crimes no mapa ao dos meses analisados. ione um determinado ou use as setas ao lado criar uma animação.

<	2010, fevereiro	-
Ē	7	

Linha do tempo

Produção do mapa impresso



Produção do mapa impresso

- Escolha rigorosa das informações
- Layers de informação

Menos informações Melhor compreensão Mais informações Maior complexidade





Tableau: Apresentação

O que dá pra fazer com Tableau?

Mapas

Tendências

Pontos fortes

- Rápido pra visualizar
- Lida com muitos dados
- Publicar e compartilhar
- Versátil (roda em qualquer coisa)
- Grátis
- Ajuda (bons fóruns)

Pontos fracos

- Demora aprender
- Arquivos ficam públicos (na versão Public)
- Só windows (versão mac por vir)
- Lento

Download

+ableau++public

http://www.tableausoftware.com/public/download

Abrindo dados

Abrindo dados Abrir dados >

Create a viz now

Open your data (.xls or .txt)
 Drag & Drop to visualize

Abrindo dados Abrir dados > Escolher tipo de arquivo

😰 Tableau Public - Book1

File Data Help

Connect to Data

In a file

- Tableau Data Extract
- <u>Microsoft Access</u>
- Microsoft Excel
- <u>]</u> Text File

On a conver

Abrindo dados

Abrir dados > Escolher tipo de arquivo > Escolher planilha

Excel Workbook Connection	×
Step 1: Select an Excel workbook:	
eta\Curso Abraji POA\Exemplos\Dados\populacao_RS2010.xls	Browse
Step 2: Select the worksheet (table) to analyze:	
Single Table C Multiple Tables C Custom SQL	
metadados e outros	
População_RS2010	
I Step 2: Deep the data include field pamer in the first row?	
Step 3: Does the data include field names in the first row?	
Yes, the first row has field names in it.	
\bigcirc No, automatically generate names for the fields.	
Read if this is your first time connecting OK	Cancel

A interface do tableau

File Data Worksheet Dashboard A	Analysis Map Format Window	/ Help		
← → 🔳 🔰 🖬 🖬 🖬	K • 13 +≞ +∓ @ • At		Iormal 💌 🖮 🔟 🗸 - 🗍 💢 📗	Show Me
Data +	Pages	Columns		- 123 123 - 124 123 123 - 124 124 124 - 124 124 124 124 - 124 124 124 124 124 124 124 124 124 124
Dimensions 🔠 🔎 👻		Rows		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Abc Estado Abc Município Abc País	Filters		Drop field here	
Abc Measure Names				808 8988
	Marks Abc Automatic	Drop field Drop field here	Drop field here	
	Color O Abc Size Text		in the second	
	Detail Tooltip			
				Select or drag data Use the Shift or Ctrl key to
				select multiple fields
Measures				
# Area total (km ²)				

Worksheet Dashboard Analysis Map Format Window Help File Data

- Área total (km²) #
- Densidade demográfica (hab/km... #
- Geocodigo #
- Pop Relativa Urbana #
- Pop Relativa Urbana (sede muni... #
- Pop Urbana (sede municipal) #
- Pop Urbana (Total) #
- População Total #
- Number of Records =#
- Measure Values #

0]

-

173 1000

(11) 1000

1082 (022)
Conferir dados importados Verificar se o Tableau entendeu corretamente os dados

Data	3	
E.	População_RS2010 (p	opulacao
Dime	nsions	
Abc	Estado	
Abc	Município	
Abc	País	
Abc	Measure Names	

Measures

- # Área total (km²)
- # Densidade demográfica (hab/km...
- # Geocodigo
- # Pop Relativa Urbana
- # Pop Relativa Urbana (sede muni...
- # Pop Urbana (sede municipal)
- # Pop Urbana (Total)
- # População Total

Data	÷ p;
População_RS20	10 (populacao
Dimensions	- Q
Abc Estado	View data
Abc Município	

Conferir dados importados Verificar se o Tableau entendeu corretamente os dados

1	View Data: Popul	ação_RS2010 (popula	icao_F	(52010.xls)			
ſ	496 rows 🔿	Show Aliases					
	Estado	Município	País	Densidade d	Geocodigo	Number of Records	Pop Re
	Rio Grande do Sul	Pedras Altas	Brazil	1,61	4.314.175,00	1	2
	Rio Grande do Sul	Santa Margarida do Sul	Brazil	2,46	4.316.972,00	1	
	Rio Grande do Sul	Muitos Capões	Brazil	2,49	4.312.617,00	1	
	Rio Grande do Sul	São José dos Ausentes	Brazil	2,80	4.318.622,00	1	
	Rio Grande do Sul	Maçambará	Brazil	2,82	4.311.718,00	1	
	Rio Grande do Sul	Aceguá	Brazil	2,84	4.300.034,00	1	
	Rio Grande do Sul	Lavras do Sul	Brazil	2,95	4.311.502,00	1	
	Rio Grande do Sul	Itacurubi	Brazil	3,07	4.310.553,00	1	
	Rio Grande do Sul	Capão do Cipó	Brazil	3,08	4.304.655,00	1	
	Rio Grande do Sul	Capão Bonito do Sul	Brazil	3,33	4.304.622,00	1	

Conferir dados importados Verif dado

Basicamente, confira se está parecido com os dados do Excel:

of Records

Pop Re

• linhas preenchidas

🗄 Vie

496 r

Ric

Ric

Ric

Ric

Ric

Ric

- sem células vazias
- cabeçalhos reconhecidos
- números reconhecidos como números
- número total de linhas

Rio Grande do Sul	Lavras do Sul	Brazil	2,95	4.311.502,00
Rio Grande do Sul	Itacurubi	Brazil	3,07	4.310.553,00
Rio Grande do Sul	Capão do Cipó	Brazil	3,08	4.304.655,00
Rio Grande do Sul	Capão Bonito do Sul	Brazil	3,33	4.304.622,00

Criando uma visualização

Arraste uma *dimension* para o painel de linhas (*Rows*)

Pages			Columns			
			Rows	Município		
Filters						
			Município			
			Aceguá		Abc	
Marks			Água Santa		Abc	
Abc Aut	omatic		Agudo		Abc	
ADC AU	omauc		Ajuricaba		Abc	
	0	Abc 123	Alecrim		Abc	
Color	Size	Text	Alegrete		Abc	
	-		Alegria		Abc	

Arraste uma *measure* para o painel de colunas (*Columns*)



O painel Show me

- Visualização automática
- Sugestões de gráfico para os dados disponíveis
- Requisitos de dados para cada tipo de gráfico



For horizontal bars try 0 or more dimensions 1 or more measures

Escolha uma forma: Tabela



Pages	III Columns	
	IE Rows Município	
Filters		
	Município	
	Aceguá	4.394
Marks	Água Santa	3.722
Abc Automatic	Agudo	16.722
	Ajuricaba	7.255
() () () () () () () () () ()	Alecrim	7.045
Color Size Text	Alegrete	77.653
	Alegria	4.301
Detail Tooltip	Almirante Tamandaré do Sul	2.067
	Alpestre	8.027
Abc SUM(População Total)	Alto Alegre	1.848
	Alto Feliz	2.917
	Alvorada	195.673

Escolha uma forma: Tabela pintada



Pages	Columns	
	Rows Município	
Filters		
	Município	
-	Aceguá	4.394 🔺
Marks	Água Santa	3.722
Square	Agudo	16.722
	Ajuricaba	7.255
😔 🔿 Abç	Alecrim	7.045
Color Size Label	Alegrete	77.653
	Alegria	4.301
Detail Tooltip	Almirante Tamandaré do Sul	2.067
	Alpestre	8.027
SUM(População Total)	Alto Alegre	1.848
Abc 123 SUM(População Total)	Alto Feliz	2.917
	Alvorada	195.673

Escolha uma forma: Treemap



Pages	Column	s						
	Rows							
Filters								
		Porto Alegre	Santa Maria	Passo Fundo		Esteio	ljuí	
Marks								
Automatic								
Color Size Label			Gravataí	Sapucaia do Sul	Cruz Alta		São Borja	
				Uruguaiana		Rio		
Detail Tooltip			Viamão		-	-		
SUM(População Total)				Santa Cruz do Sul		_		
Abc 123 Município			Novo Hamburgo	Cachoeirinha				

Escolha uma forma: Bolhas (bubble chart)





Escolha uma forma: Bolhas (bubble chart)



Forma escolhida: Barras horizontais







O painel Marks

- Formatar os gráficos
- Cores, tamanhos, legendas, e mouse-overs
- Funciona com arrastar-e-soltar

	\square	(
Color	Size	Label
Detail	Tooltip	

• Ordenando lista (Sort)



Columns	SUM(População Total)	
Rows	Município F	
	Município	
	Porto Alegre	
	Caxias do Sul	
	Pelotas	
	Canoas	
	Santa Maria	
	Gravataí	
	Viamão	
	Novo Hamburgo	
	São Leopoldo	
	Rio Grande	

- Ordenando lista (Sort)
- Redimensionar tamanho da área



- Ordenando lista (Sort)
- Redimensionar tamanho da área
- Adicionar labels





- Ordenando lista (Sort)
- Redimensionar tamanho da área
- Adicionar labels
- Adicionar cor





Formatando: Antes & depois







No Tableau Public Salvar = Publicar na web (ou seja, tudo que é salvo fica disponível no ar)

🐺 T	ableau	ı Public - F			
File	Data	Workshee			
N	ew 🛛				
Open from Web					
C	ose				
Sa	ave to V	Web			
Sa	ave to	Web As			



Fazer log-in A conta é gratuita

Tableau Public Sign In	×
Logon	+ableau+public
Enter your Tableau Public email address and password to log on.	
Email: dadosfinos@gmail.com	
Password:	
Login Forgot your password?	
No account? Create one for FREE!	

Nomear arquivo

um bom nome é importante pra organização

um exemplo de nomenclatura: tipo da informação + periodo (+ fonte)

Save W	orkbook to the Public V	Veb	×
Name:	Rio Grande do Sul - Popul	ação	•
	Show Sheets as Tabs		
		Save	Cancel

Salvando pela segunda vez

Sempre pede para substituir o arquivo. Não tem problema.





Preview do arquivo salvo



Fim da parte básica.



Cruzando dados

Conectar dados novamente para abrir outra tabela:

Data > Connect data

Data	Worksheet	Dashboard	Analysis	Map Forma
Cor	nnect to Data			Ctrl+D
Pas	ste Data			Ctrl+V
Edi	t Relationship	s.,,		
Rep	place Data So	urce,		
V Pop	ulação_RS20	10 (populaca	o_RS2010.	.xls) +
Maniap	010		1	



Selecione a tabela que tem as coordenadas:



Data	а	\$	Agora
40	População_RS2010 (pop	pulacao	de da
ų.	RS LatLong (municipios	RS_LatL	(2 tab
Dime	nsions	III 🔎 👻	
Abc	Estado	c/ə	
Abc	EstadoCod		
Abc	EstadoSigla		
Abc	Mesorregiao		
Abc	MesorregiaoCod		æ
Abc	Microrregiao		
Abc	MicrorregiaoCod		
Abc	Município	60	
Abc	Regiao		
Meas	sures		
#	GEOCODIGO		
۲	LATITUDE		
۲	LONGITUDE		
=#	Number of Records		
#	Maasura Valuas		

Agora tem 2 fontes de dados conectadas (2 tabelas)

Crie uma nova worksheet para trabalhar com os novos dados

Worksheet	Dashboard	Analysis	۱
New Wor	ksheet	Ctrl+M	5
Clear		+	F
Actions			



New Worksheet

Dois cliques na Latitude e depois na Longitude para adicioná-las:

Data	¢	Columns
População_RS20	10 (populacao	Rows
😨 RS LatLong (mun	iicipios_RS_LatL	5
Dimensions	÷ ۹ 🎫	~ ~
Abc Estado		
Abc EstadoCod		5 moon
Abc EstadoSigla		
Abc Mesorregiao		-
Abc MesorregiaoCo	d	
Abc Microrregiao		
Abc MicrorregiaoCo	d	Corrie
Abc Município		
Abc Regiao		
Abc Measure Name:	5	
Measures		
# GEOCODIGO		
CATITUDE		
		4



Arraste o campo Município para o painel *Marks*



Volte na outra tabela para pegar um campo de lá:





Arraste a população em cima do botão Size no painel Marks



Ajustes Clique em *Size* para ajustar o tamanho das bolhas


Ajustes Clique em *Color* para adicionar transparência





Preparando os arquivos

Formato de dados que softwares entendem

- informações começam na célula A1
- 1^a linha: cabeçalhos
- linhas seguintes: cada linha deve corresponder a um item

	Г	Г		
				1
		E	Л	1
	E	Ш		

Cada linha, UM CASO, um registro, uma coisa

25454	 rear graning	
	-11	
Press	and so its	

Cada coluna, uma variável uma característica,



Cada planilha, um tema

Limpando uma tabela

- Manter apenas:
 - > dados principais
 - > 1 linha de cabeçalho
- Preencher linhas brancas (fill down)
- Adicionar todos os cabeçalhos
- Em alguns casos: Reformatar a estrutura (plugin do excel)

http://goo.gl/1igv6C

Remover

- introduções
- títulos
- fonte, data de publicação,
 observações
- totais

Limpando uma tabela

1	<u></u>	2	3	49	2					
1	Sinopse do Censo Demográfico 2010									
2	Tabela 2.1.23 - População residente, total, urbana total e urbana na sede municipal, em números absolutos e relativos, com indicação da área tot									
3	densi	dade demográfi	ca, segundo	os municípios - F	Rio Grande	do Sul - 2	2010			
4		População residente								
5		Absoluta				Relativa (%)				
6	Municípios		Urbana		ļ	Urb	ana			
7	Manicipios	Total	Total	Na sede municipal	Total	Total	Na sede			
8		TOTAL					municipal			
9										
10	Total	10 693 929	9 100 291	8 744 154	100,0	85,1	81,8			
11	Aceguá	4 394	1 059	1 059	100,0	24,1	24,1			
12	Água Santa	3 722	1 447	1 356	100,0	38,9	36,4			
13	Agudo	16 722	6 889	<mark>6 88</mark> 9	100,0	41,2	41,2			
14	Ajuricaba	7 255	4 108	3 872	100,0	56,6	53,4			
15	Alecrim	7 045	2 1 <mark>6</mark> 5	2 165	100,0	30,7	30,7			
16	Alegrete	77 653	69 594	69 009	100,0	89,6	88,9			
17	Alegria	4 301	1 585	1 353	100,0	36,9	31,5			
18	Almirante Tamandaré do Sul	2 067	839	642	100,0	40,6	31,1			
19	Alpestre	8 027	2 211	2 002	100,0	27,5	24,9			
20	Alto Alegre	1 848	743	662	100,0	40,2	35,8			

Limpando uma tabela

1	1	2	3	4	5	6	7				
1	Sinopse do Censo Demográfico 2010 Remover textos introdutórios urbana na sede mu observaçães to fontes com indicação da área total e										
2											
3	dens	idade demográ	ifica, segundo d	os municípios -	Rio Grande	do Sul - 2	2010				
4		População residente									
5			Absoluta	Relativa (%)							
6			Urb	ana	_	Urb	ana				
7	Municípios		Remov	er célu	las m	escla	das				
0		Total	Total	Na sede municipal	Total	Total	Na sede				
0			10001			1000	monioqui				
9								_			
10	Total	Rem	over to	tais 4 154	100,0	85,1	81,8				
11	Aceguá	4 394	1 059	1 059	100,0	24,1	24,1				
12	Água Santa	3 722	1 447	1 356	100,0	38,9	36,4				
13	Agudo	16 722	6 889	<mark>6 88</mark> 9	100,0	41,2	41,2				
14	Ajuricaba	7 255	4 108	3 872	100,0	56,6	<mark>53,4</mark>				
15	Alecrim	7 045	<mark>2 16</mark> 5	2 165	100,0	30,7	30,7				
16	Alegrete	77 653	69 594	69 009	100,0	89,6	88,9				
17	Alegria	4 301	1 585	1 353	100,0	36,9	31,5				
18	Almirante Tamandaré do Sul	2 067	839	642	100,0	40,6	31,1				
19	Alpestre	8 027	2 211	2 002	100,0	27,5	24,9				
20	Alto Alegre	1 848	743	662	100,0	40,2	35,8				

Importante: não jogue fora as informações!

• coloque esses dados em outra planilha

• mantenha uma cópia do arquivo original



1			2	3	4	2	0				
1	Sinopse do Censo Demográfico 2010										
2	Tabela 2.1.23 - População residente, total, urbana total e urbana na sede municipal, em números absolutos e relativos, com indicação da área total e densidade demográfica, segundo os municípios - Rio Grande do Sul - 2010										
3	População residente										
4			Absolute								
2			Absoluta			NGIa	LIVA (70,	Lichana			
0	Municípios			Un			urbana				
/			Total		Na sede municipal	Total		Na sede			
8				Total		T	otal	municipal			
9											
10	Total	1	0 693 929	9 100 291	8 744 154	100,0	85,1	81,8			
11	Aceguá		4 394	1 059	1 059	100,0	24,1	24,1			
12	Água Santa		3 722	1 447	1 356	100,0	38,9	36,4			
J.	1	2	3		4	5		6			
1	Município	População Total	Pop Urban (Total)	a Pop (sede	Urbana e municipal)	Pop Relativ Urbana	a Ur mi	p Relativa bana (sede unicipal)			
2	Aceguá	4394	10	59	1059	24,1010	5	24,10105			
3	Água Santa	3722	144	47	1356	38,8769	5	36,43203			
4	Agudo	16722	688	39	6889	41,1972	3	41,19723			
5	Ajuricaba	7255	410	08	3872	56,6230	2	53,37009			
6	Alecrim	7045	210	65	2165	30,7310	1	30,73101			
7	Alegrete	77653	695	94	69009	89,6217	8	88,86843			
8	Alegria	4301	15	35	1353	36,8518	9	31,4578			

Obrigado!