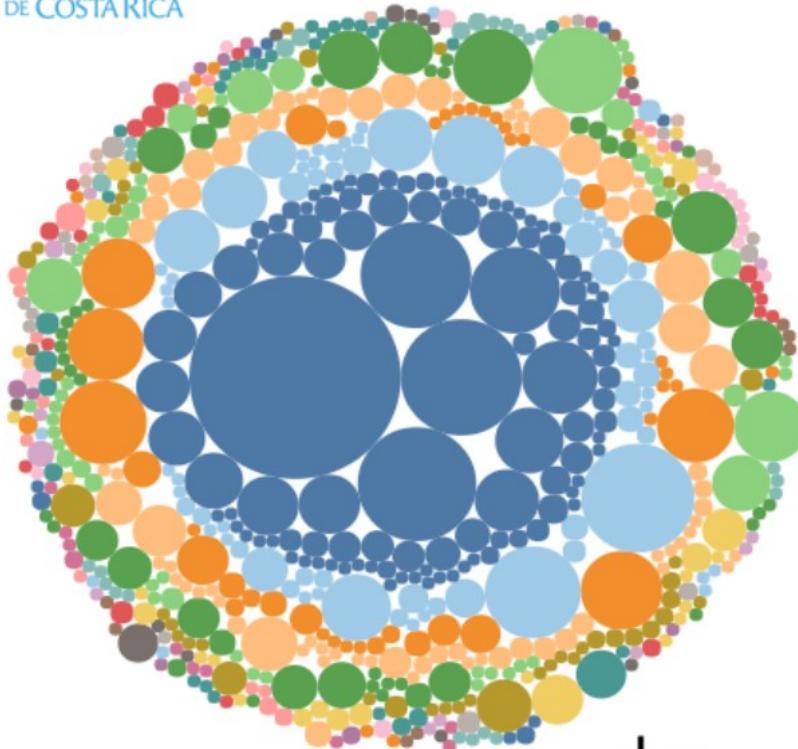




**Data  
Storytelling**

**WORKSHOP**



# Big Data y Transformación Digital !

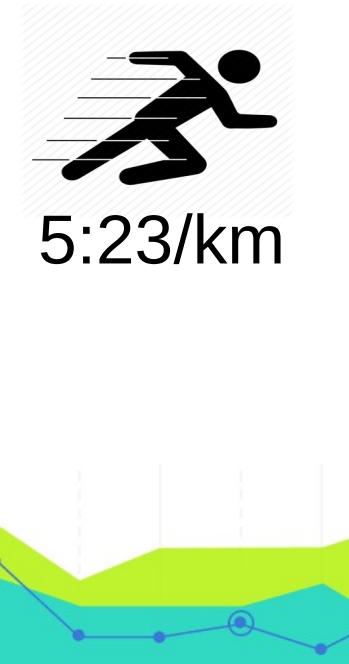
Jueves 15 de febrero

5 p.m.

Auditorio de la Escuela de Arquitectura.

# datelligence :::









@soniaardila1

**EL TIEMPO**

COLOMBIA BOGOTÁ MEDELLÍN CALI BARRANQUILLA BOCAJÁ LLANO MÁS CIUDADES

## 69 niños muertos por hambre en La Guajira

Las prácticas culturales y falta de servicios básicos aún pesan en la desnutrición de los menores.



EFE: POR: AGENCIA EFE

El viernes, la Fuerza Aérea Colombiana (FAC) transportó a dos niños wayuu, de 1 y 3 años, del hospital de Nazareth a La Guajira. Uno de ellos está grave.

Foto: Fuerzas Aéreas

RELACIONADOS: COLOMBIA LA GUAJIRA DESNUTRICIÓN INFANTIL COMUNIDADES WAYUU

Por: Redacción EL TIEMPO | 06 de noviembre 2016, 01:30 a.m.

A pesar de la intervención del Gobierno en La Guajira, los niños indígenas de este departamento se siguen muriendo por desnutrición. En el 2014, el Instituto Nacional de Salud reportó 48 menores de edad muertos por esta causa; el año pasado fueron 37, y en el 2016 ya van 69.

# EL HERALDO

## Gobierno reconoce que mortalidad infantil asociada a desnutrición subió 50% en La Guajira

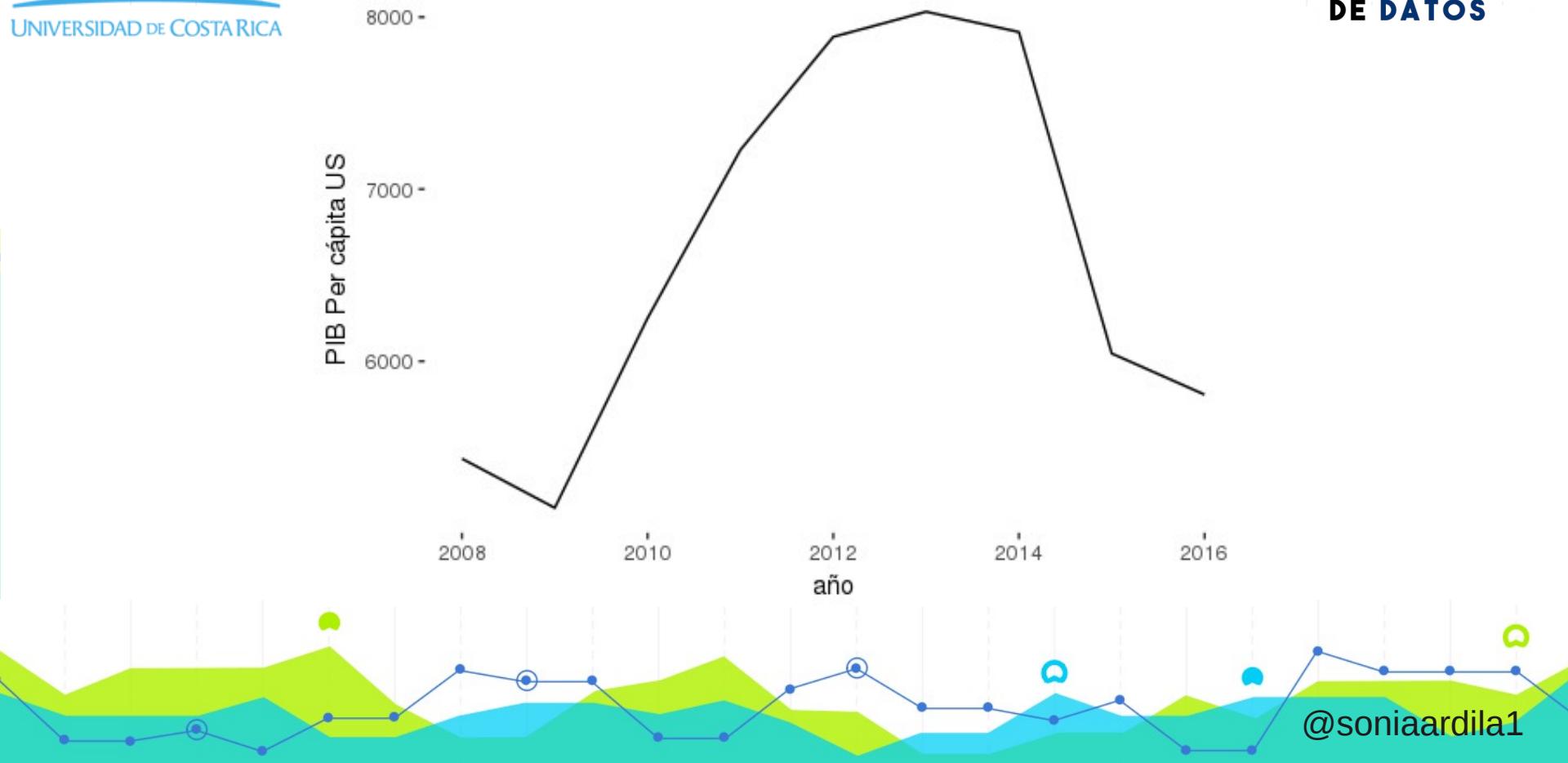
LA GUAJIRA / 15 de Febrero de 2016 - 21:50

En 2011 se registraron 23 casos de muertes infantiles por desnutrición, en 2013 un total de 40 y la cifra preliminar de 2015 es de 38 muertes, según Minsalud. Comunidades wayuu dicen que han muerto más menores.

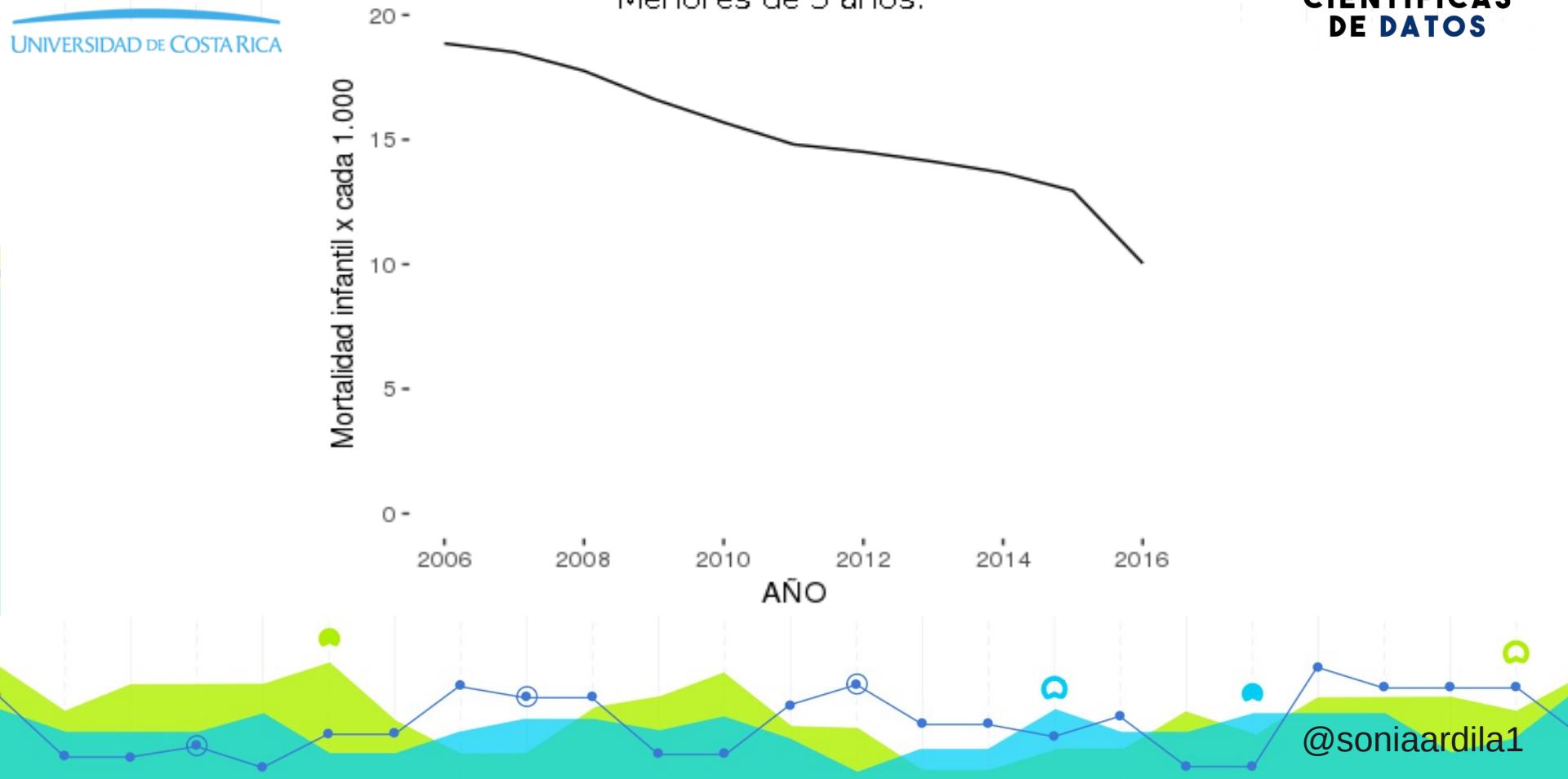




## PIB Per cápita Colombia

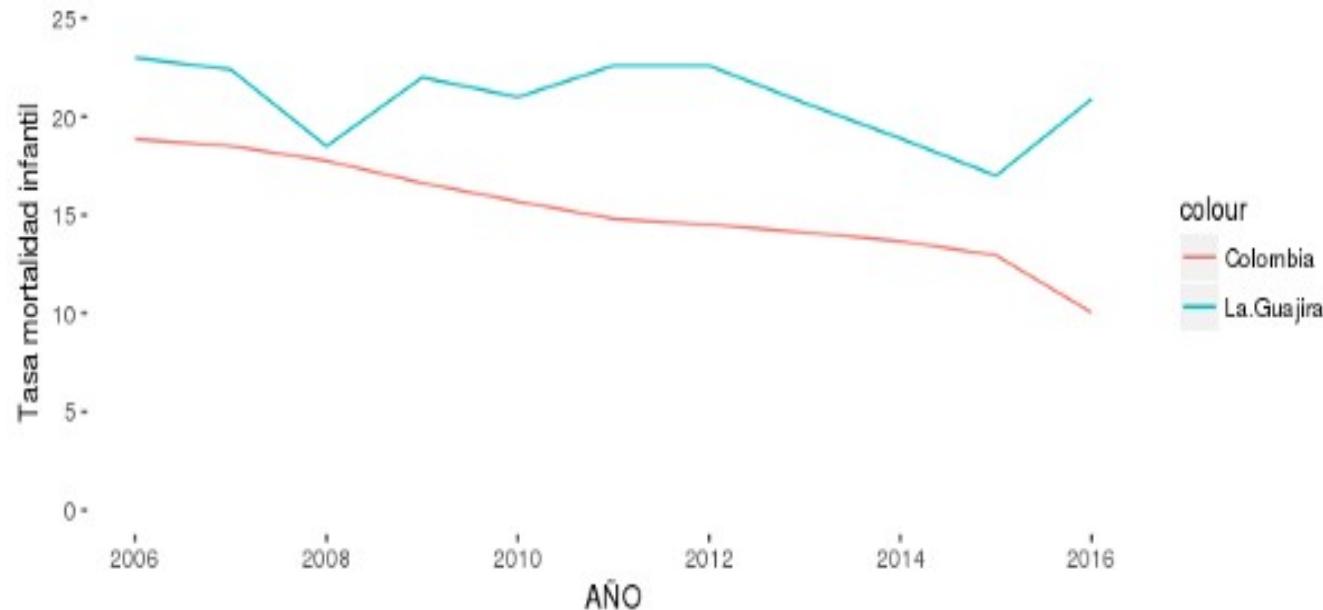


## Tasa Mortalidad Infantil por cada 1.000 nacidos. Menores de 5 años.

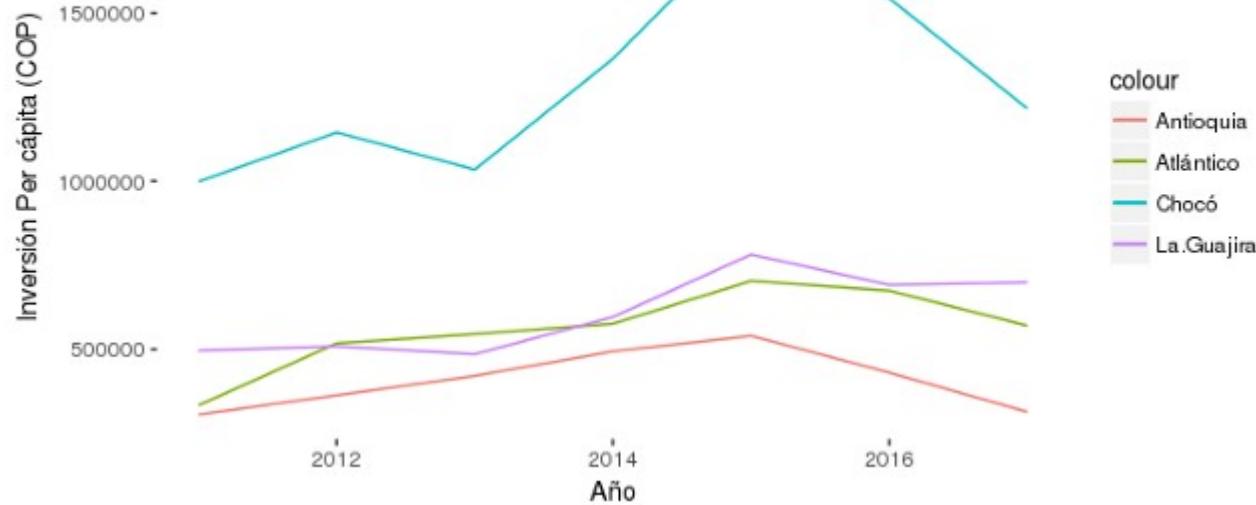


@soniaardila1

## Tasa mortalidad infantil x cada 1000 nacimientos. Menores 5 años

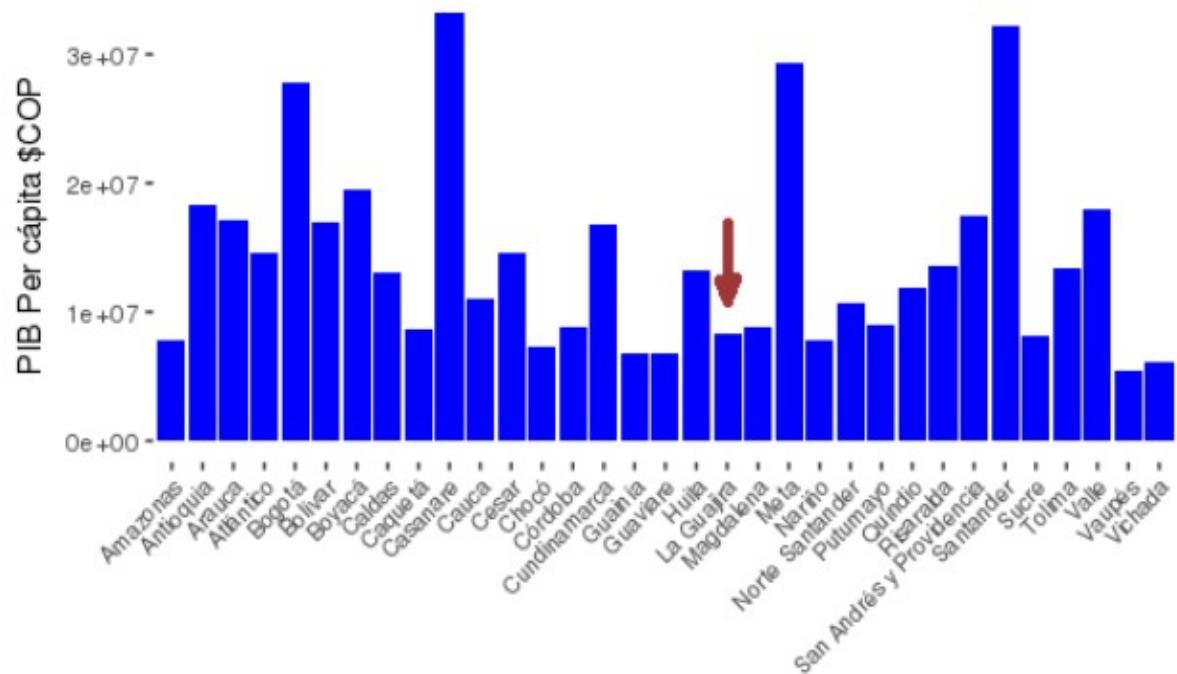


## Inversión Per cápita Cuatro Dptos Colombia 2011-2017



@soniaardila1

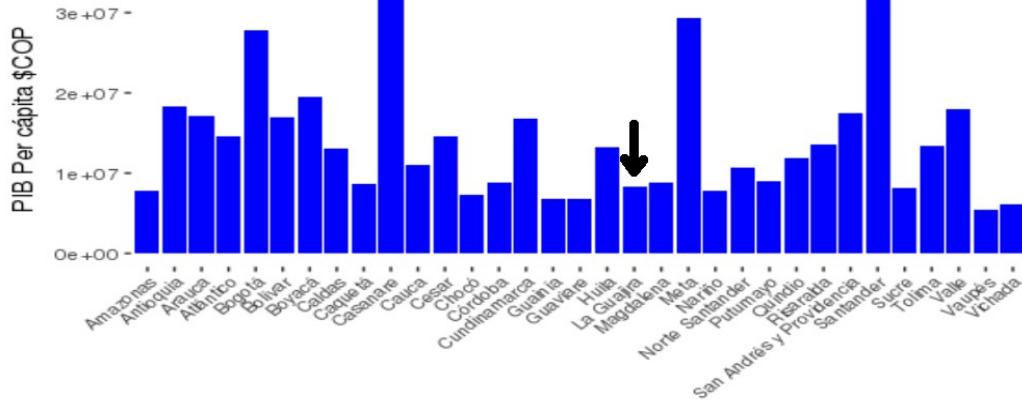
## PIB Per cápita Dptos Colombia 2016



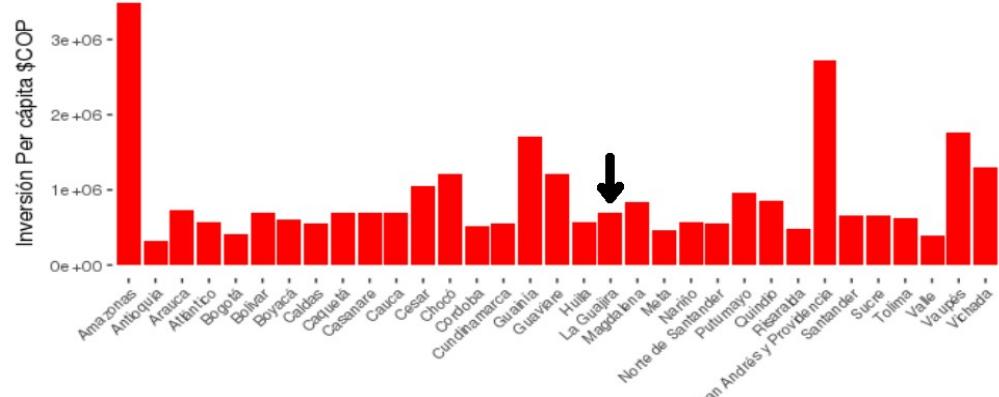
Dpto

@soniaardila1

## PIB Per cápita Dptos Colombia 2016

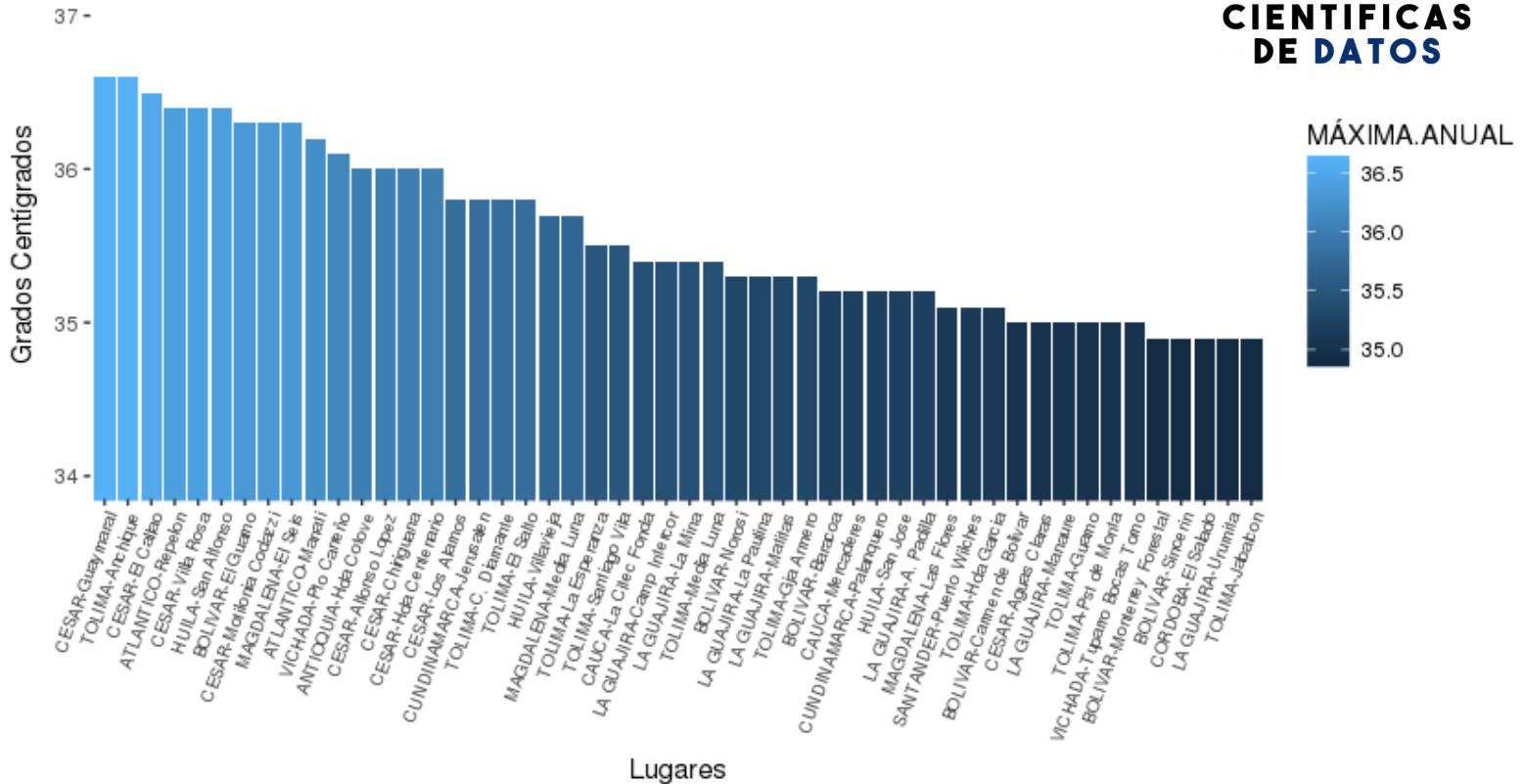


## Presupuesto Nacional-Inversión Per cápita Dptos Colombia 2017



@soniaardila1

## Los 50 lugares mas calientes de Colombia según temperatura máxima promedio anual 1981-2010




@soniaardila1

MEDIO AMBIENTE

IMPACTO SOCIAL

NEGOCIOS

TENDENCIAS

ES TENDENCIA:

Bolsas Plásticas

Groenlandia

Consultas Populares

Cambio Climático

[INICIO](#) / [IMPACTO](#) / [SEQUÍA Y HAMBRE EN LA GUAJIRA](#)

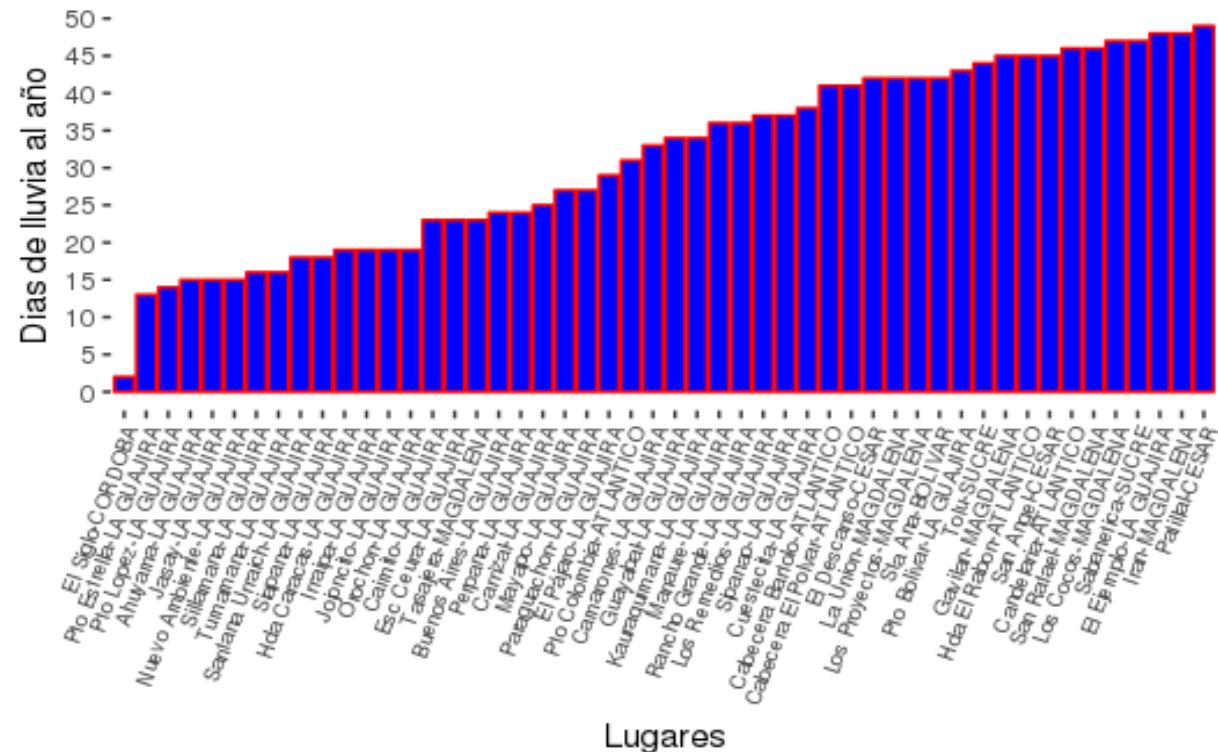
IMPACTO | 2016/08/05

## Sequía y hambre en La Guajira

Entre enero y marzo murieron 46 niños por desnutrición. La población indígena es quien más sufre la sequía y la hambruna que afectan a la región.



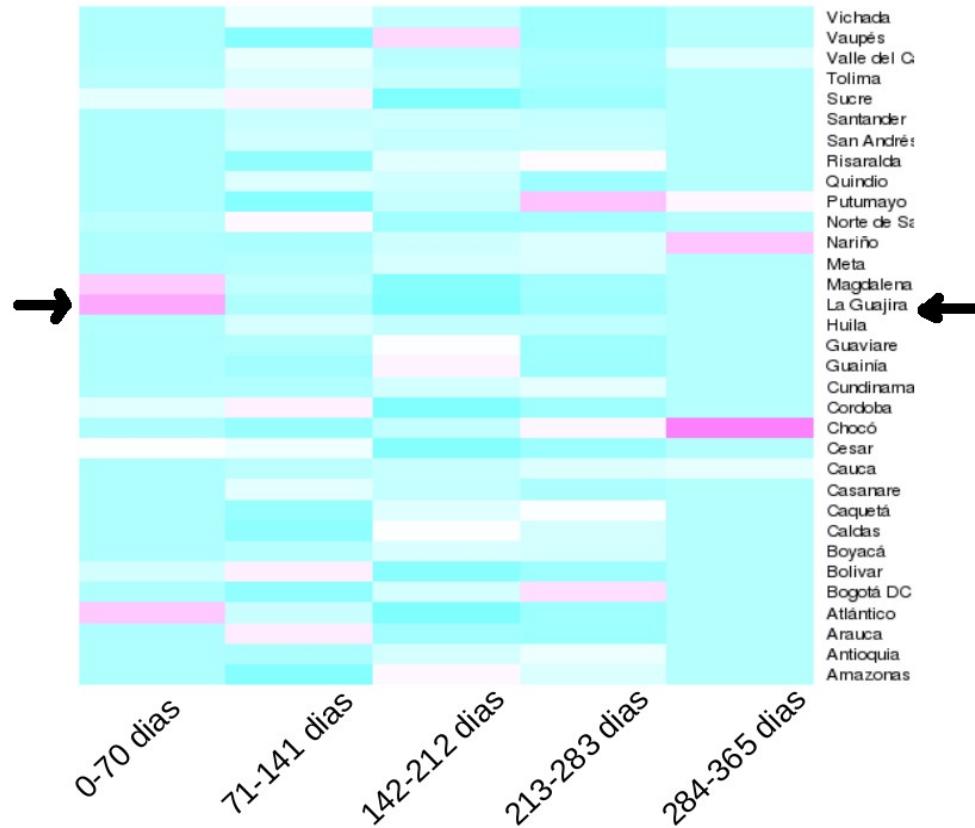
## Los 50 lugares de Colombia con menos días de lluvia en promedio anual 1981-2010



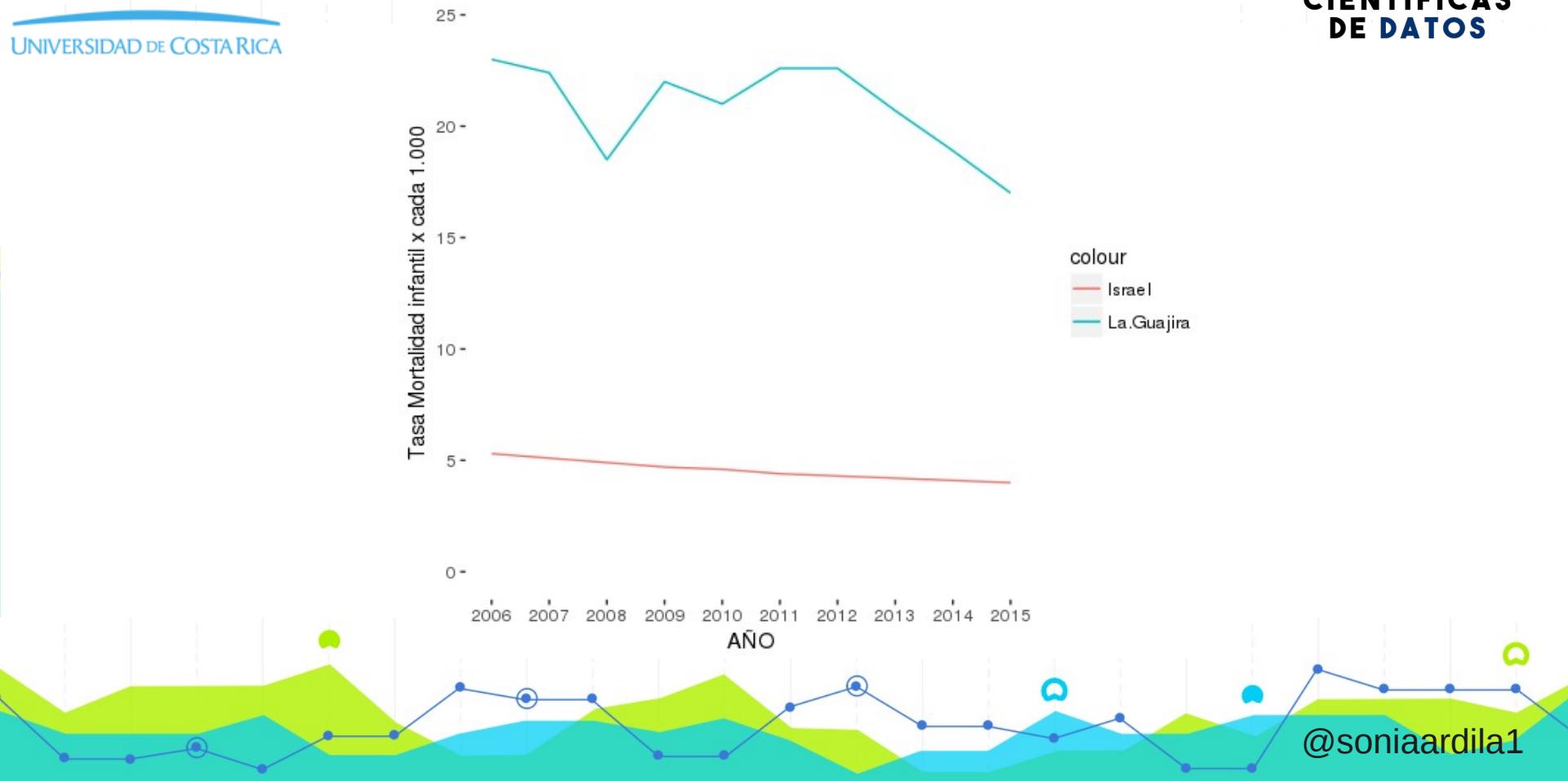
Lugares

@soniaardila1

## Días de lluvia al año por departamento. Promedios 1981 – 2010



## Mortalidad infantil x cada 1.000 nacidos menores de 5 años. Israel y La Guajira





sonia ardila

Data analyst in @datelligence . Data storyteller. @soniaardila1

Oct 2, 2017 · 5 min read

## La enfermedad de La Guajira.



Hans Rosling, el médico y estadístico sueco que estamos extrañando desde febrero de este año, solía repetir una frase cuando presentaba las cifras que reflejan el crecimiento ó decrecimiento de diferentes aspectos de un país. Esa

@soniaardila1



sonia ardila  
Feb 4, 2017 · 4 min read

Mi red de amigos en facebook. La historia detrás de un grafo...



Read more...



11



sonia ardila  
May 25, 2017 · 7 min read



Cómo escogemos nuestra nave espacial? Una cósmica historia de mercadeo!

Read more...

# Medium

@soniaardila1

App house from you, Carlos Pinto, and 3 others



sonia ardila

Data analyst in @datelligence . Data storyteller. @soniaardila1  
May 25, 2017 · 7 min read

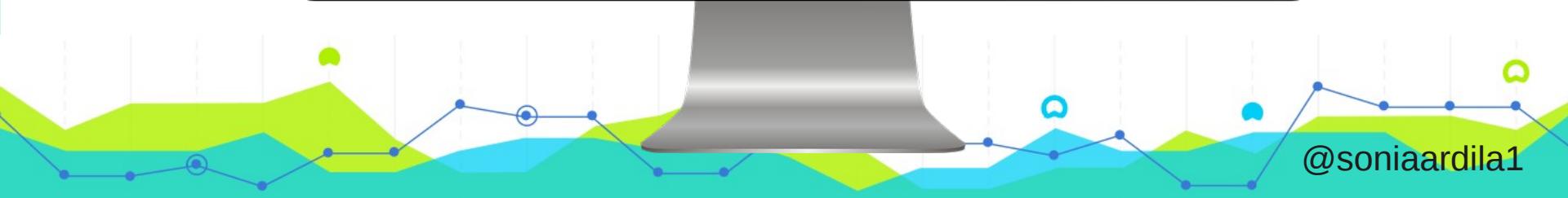
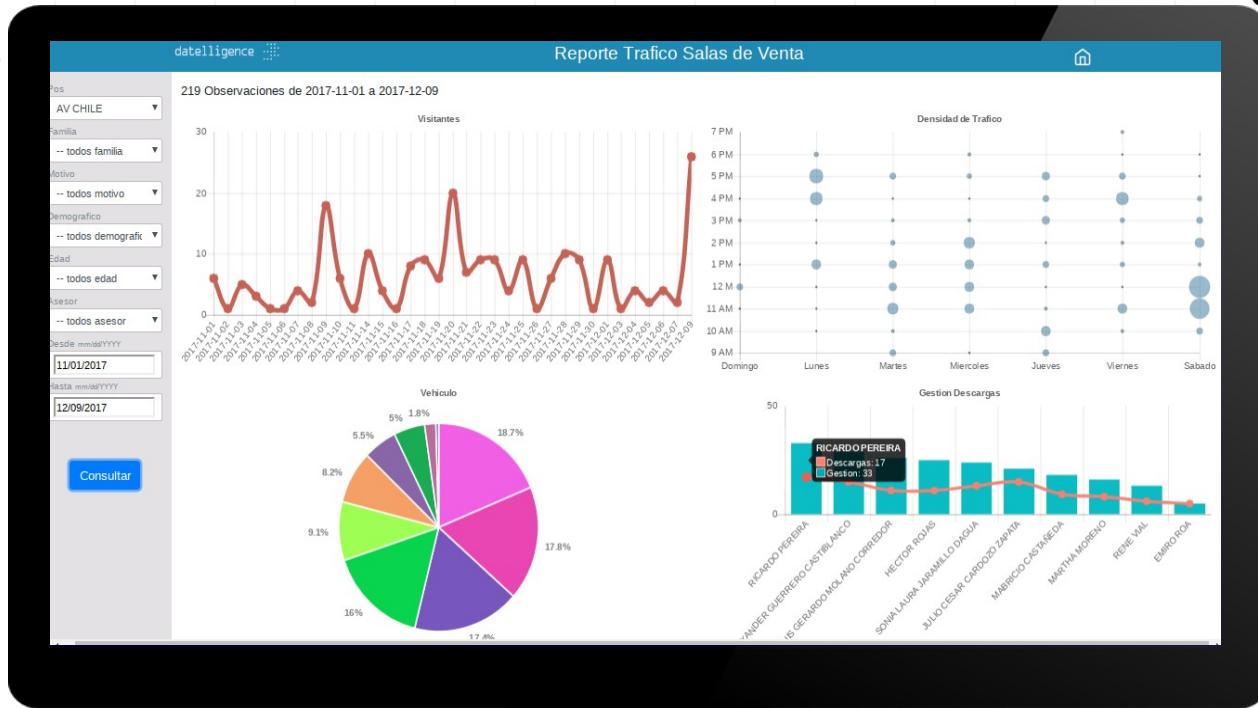
## Cómo escogemos nuestra nave espacial? Una cósmica historia de mercadeo!

Ver datos en tiempo real en un dashboard es clave para tomar decisiones. Si tomamos los datos de varios meses y los analizamos... aparecerá una historia!

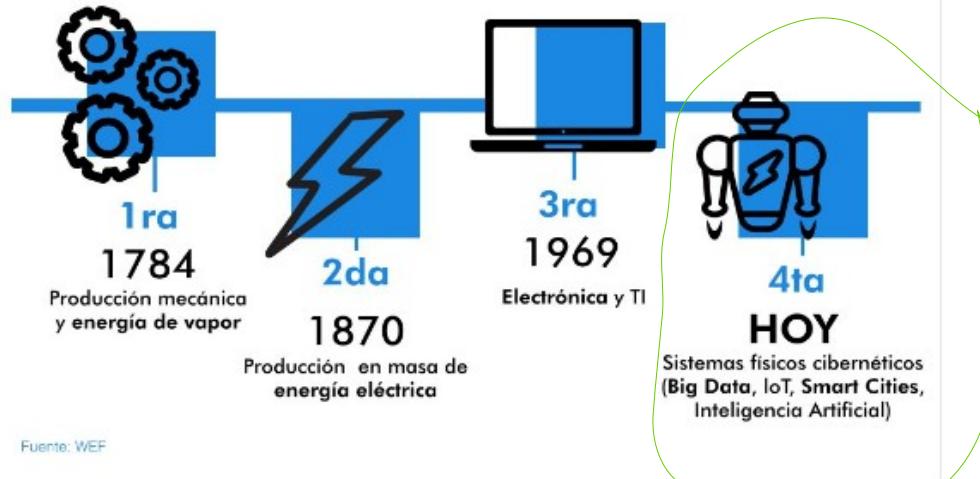


Esta historia leocurrea uno de nuestros clientes:

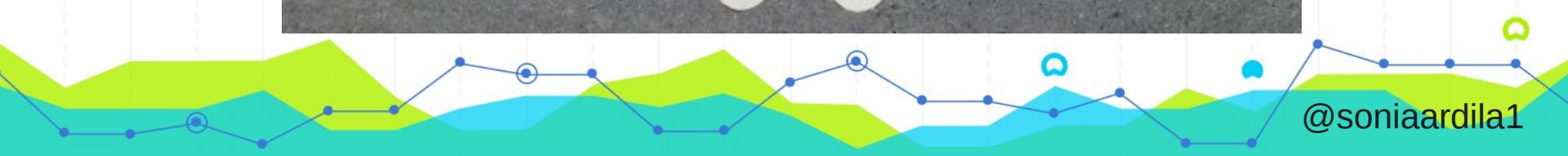
@soniaardila1

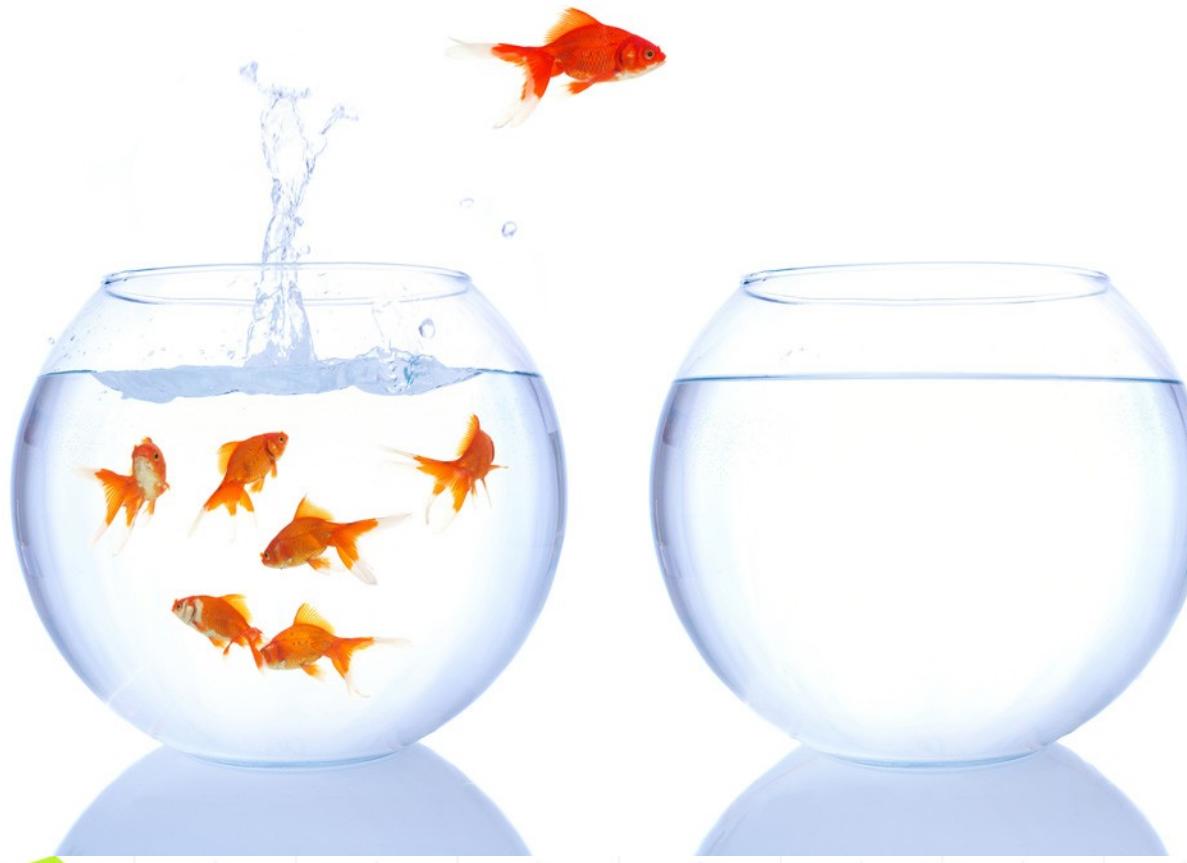


## CUARTA REVOLUCIÓN INDUSTRIAL



@soniaardila1







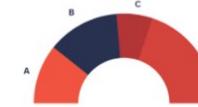
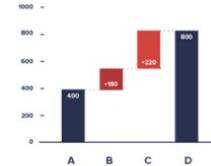
# People hear statistics, but they feel stories.

—Brent Dykes | Director of Data Strategy, Domo

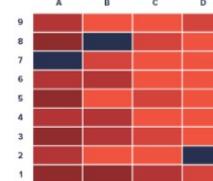
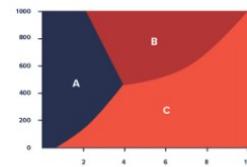




## Datos



## Visualización

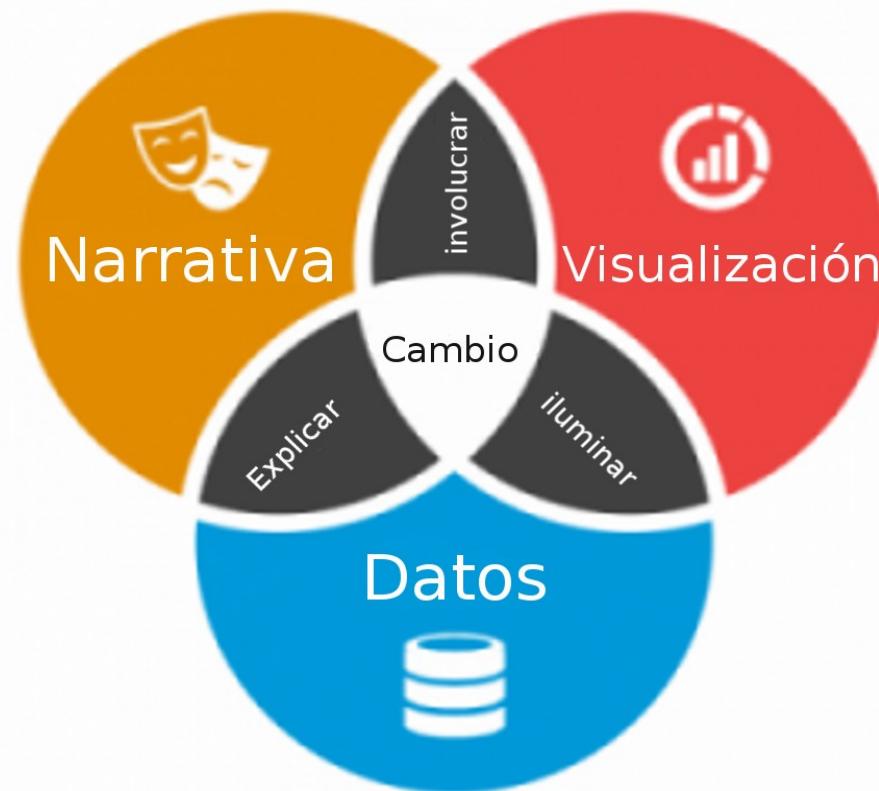


## Narrativa



@soniaardila1







# Data Storytelling

Quiénes son los protagonistas?

Cuál es el conflicto?

Cómo se puede resolver el conflicto?





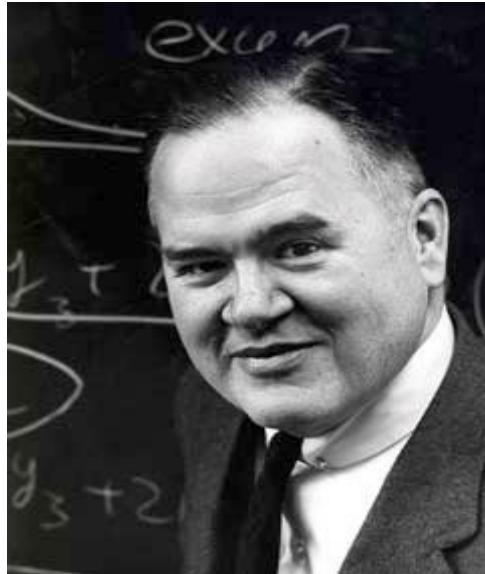
# Data Storytelling

## Acto 1

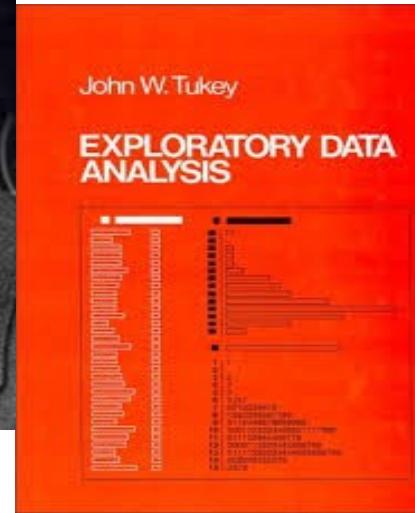
La audiencia conoce al protagonista (cómo funciona su mundo).  
EDA (Exploratory Data Analysis)



# John Tukey. 1.915 – 2.000



EDA

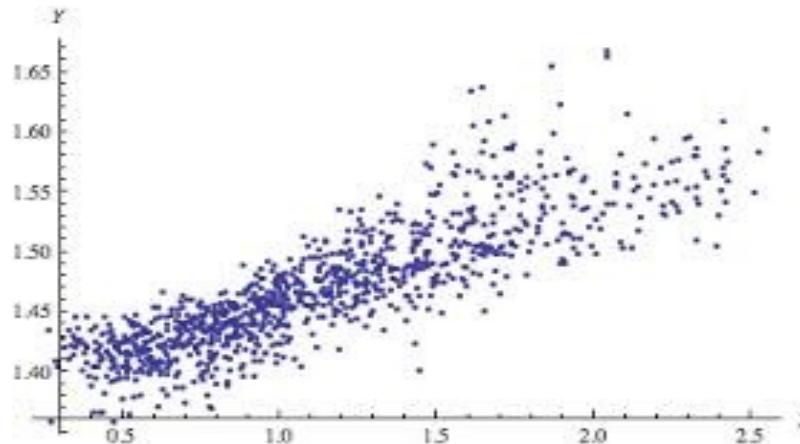


1.977



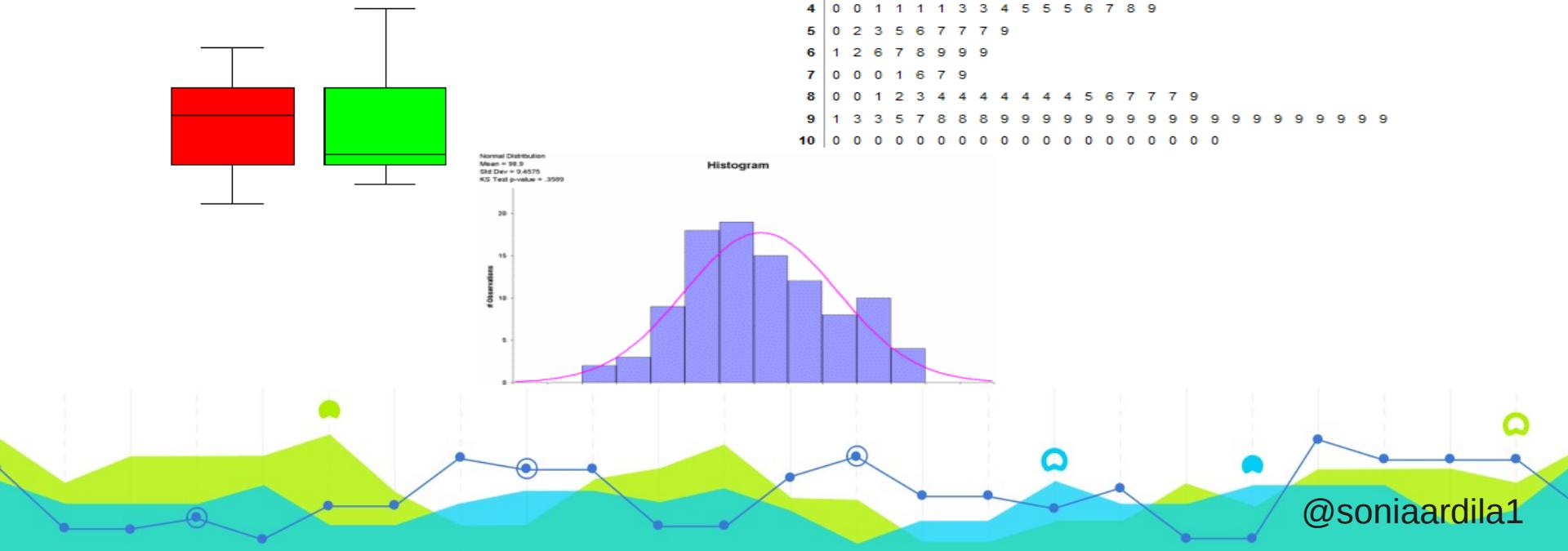
# 1. Presentación gráfica de los datos.

Symbol	Closing NAV	DIV	NAV Change	NAV Change %	Offer Price	10:00	11:00	12:00
FSAIX	27.13	+0.05	▲	+0.18	27.97	27.05	27.06	27.01
FSAVX	25.90	+0.61	▲	+2.41	26.70	25.62	25.65	25.66
FSRBX	34.63	+0.27	▲	+0.79	35.70	34.51	34.44	34.44
FBIOX	48.08	+1.11	▲	+2.36	49.57	47.72	47.81	47.94
FSLBX	43.65	+0.31	▲	+0.72	45.00	43.53	43.42	43.38
FBSOX	12.50	+0.13	▲	+1.05	12.89	12.43	12.42	12.42
FSCHX	40.81	+0.48	▲	+1.19	42.07	40.55	40.61	40.61
FDCPX	29.42	+0.93	▲	+3.26	30.33	29.02	29.07	29.10
FSHOX	27.93	+0.35	▲	+1.27	28.79	27.69	27.64	27.78
FSCPX	21.52	+0.22	▲	+1.03	22.19	21.40	21.38	21.38
FCYIX	13.00	+0.12	▲	+0.93	13.40	12.92	12.92	12.94
FSDAX	43.68	+0.49	▲	+1.13	45.03	43.30	43.34	43.37
FSDCX	12.95	+0.33	▲	+2.61	13.35	12.76	12.78	12.79
FSELX	32.25	+0.92	▲	+2.94	33.25	31.92	31.93	31.94

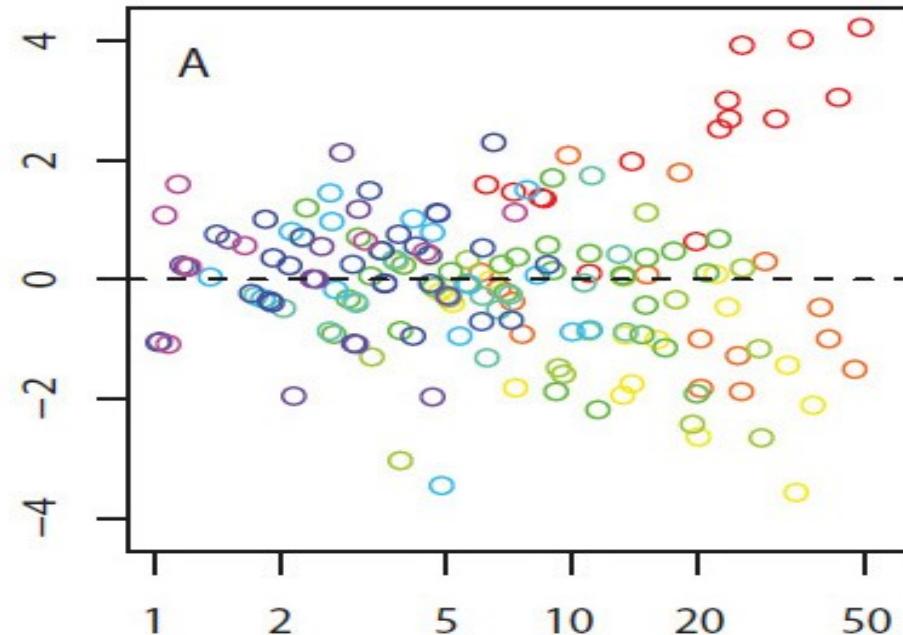


@soniaardila1

## 2. Permitir diferentes puntos de vista de la data.



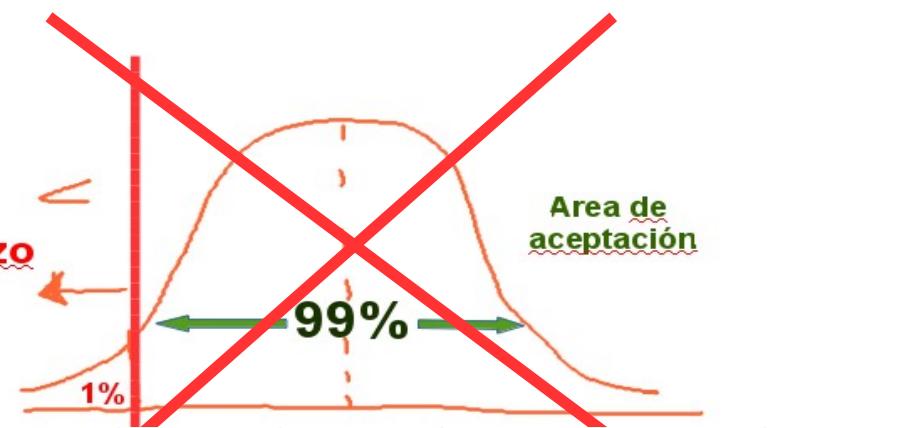
### 3. Búsqueda de simplicidad



~~$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$~~

~~| Nivel de confianza | Z <sub>alfa</sub> |
|--------------------|-------------------|
| 99,7%              | 3                 |
| 99%                | 2,58              |
| 98%                | 2,33              |
| 96%                | 2,05              |
| 95%                | 1,96              |
| 90%                | 1,645             |
| 80%                | 1,28              |
| 50%                | 0,674             |~~

**Area  
de  
rechazo**



@soniaardila1





	A	B
52		
53		
54	Punto de venta X Hora: 11 a.m. - 12	
55	# cajas	tiempo espera min
56	1	10
57	2	9
58	3	8,5
59	4	8
60	5	6
61	6	3
62	7	1,8
63	8	1
64		



@soniaardila1



1.00	0,00	6,00	0,00	2,00	0,00	0,00	0,00	0,00
0,00	0,00	13,20	0,00	3,25	0,00	2,00	0,00	0,00
0,00	0,00	19,50	0,00	4,00	0,00	4,50	0,00	0,00
24,00	0,00	8,25	0,00	0,00	0,00	0,00	0,00	0,00
24,00	0,00	9,50	0,00	0,00	1,00	0,00	0,00	0,00
0,00	0,00	6,90	0,00	2,50	0,00	0,00	0,00	0,00
0,00	0,00	24,00	0,00	2,50	0,00	0,00	0,00	0,00
22,00	0,00	3,50	0,00	2,00	0,00	0,50	0,00	0,00
39,00	0,00	21,00	0,00	0,00	0,00	1,00	0,00	0,00
0,00	0,00	2,00	0,00	0,75	0,00	0,00	0,00	0,00
0,00	8,40	2,00	0,00	0,00	0,00	0,00	0,00	0,00
0,00	7,00	0,00	17,75	3,00	5,00	3,00	0,00	0,00
0,00	2,40	2,75	0,00	0,75	0,50	3,00	0,00	0,00
0,00	7,00	0,00	19,00	6,00	3,00	3,00	0,00	0,00
0,00	5,75	9,50	19,00	6,00	3,00	3,00	0,00	0,00





Propiedad	Valor
Media de cada una de las variables x	9.0
Varianza de cada una de las variables x	11.0
Media de cada una de las variables y	7.5
Varianza de cada una de las variables y	4.12
Correlación entre cada una de las variables x e y	0.816
Recta de regresión	$y = 3 + 0.5x$



Propiedad	Valor
Media de cada una de las variables x	9.0
Varianza de cada una de las variables x	11.0
Media de cada una de las variables y	7.5
Varianza de cada una de las variables y	4.12
Correlación entre cada una de las variables x e y	0.816
Recta de regresión	$y = 3 + 0.5x$



Propiedad	Valor
Media de cada una de las variables x	9.0
Varianza de cada una de las variables x	11.0
Media de cada una de las variables y	7.5
Varianza de cada una de las variables y	4.12
Correlación entre cada una de las variables x e y	0.816
Recta de regresión	$y = 3 + 0.5x$



Propiedad	Valor
Media de cada una de las variables x	9.0
Varianza de cada una de las variables x	11.0
Media de cada una de las variables y	7.5
Varianza de cada una de las variables y	4.12
Correlación entre cada una de las variables x e y	0.816
Recta de regresión	$y = 3 + 0.5x$



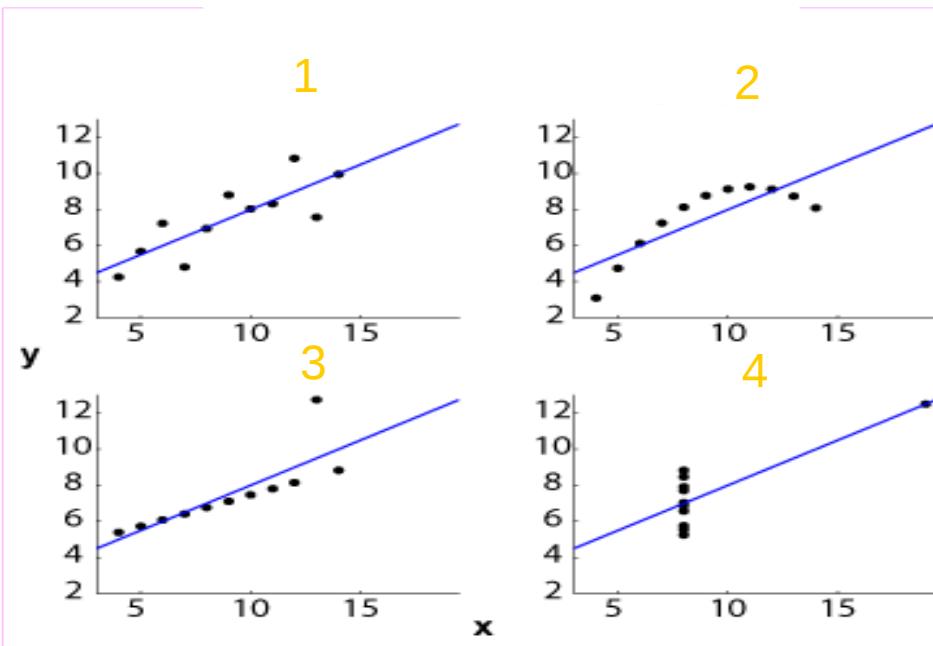
Propiedad	Valor
Media de cada una de las variables x	9.0
Varianza de cada una de las variables x	11.0
Media de cada una de las variables y	7.5
Varianza de cada una de las variables y	4.12
Correlación entre cada una de las variables x e y	0.816
Recta de regresión	$y = 3 + 0.5x$





1		2		3		4	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89



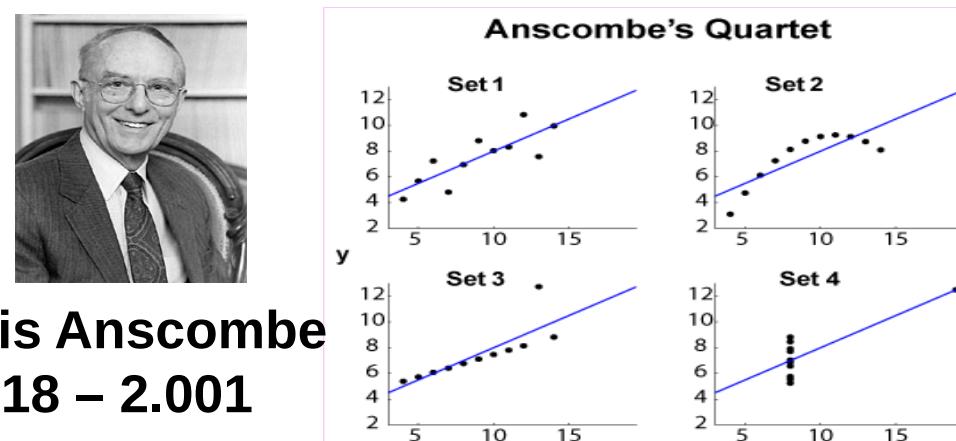


**Anscombe's quartet**

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89



**Francis Anscombe**  
 1.918 – 2.001



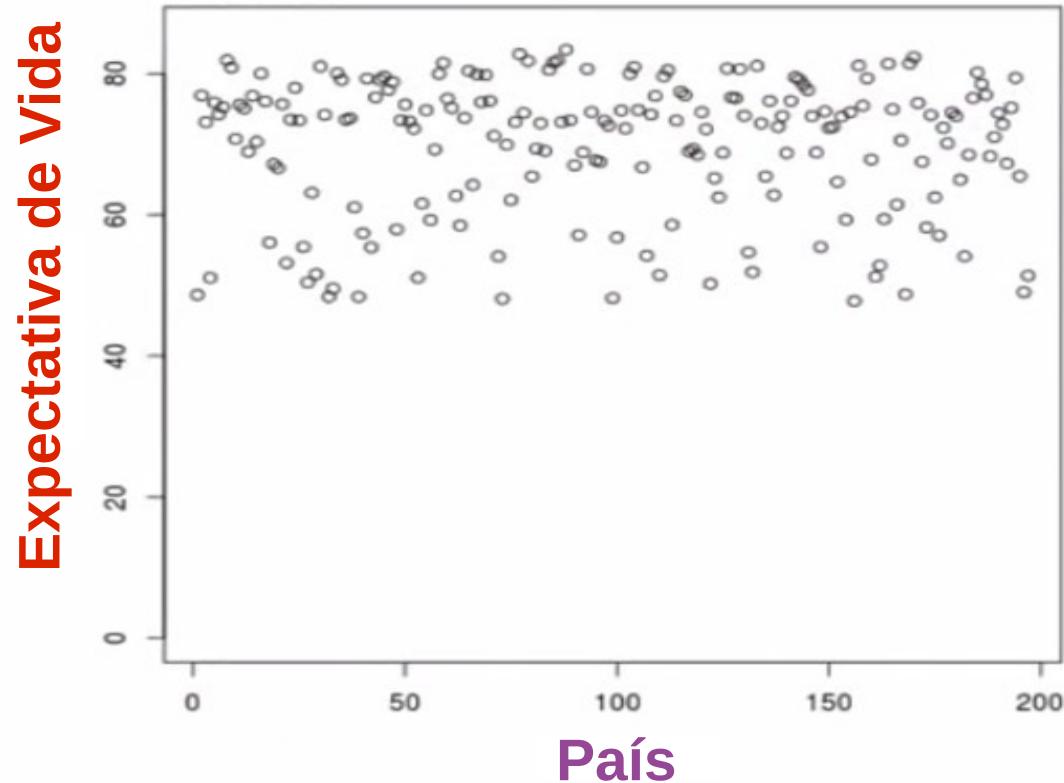
# Expectativa de vida 197 países 2.012



Afganistan: 48.6  
Alemania: 80.4  
Antillas Holandesas: 76.7  
Argentina: 75.9  
Australia: 81.9  
Belgica: 80  
Bolivia 66  
Brasil: 73.4  
Canada: 81  
Chile: 79  
China: 73.4  
Colombia: 73.7  
Corea: 80.6  
Corea del norte: 68.8

Costa Rica: 79.1  
Ecuador: 75.6  
Etiopia: 59.2  
Egipto: 73.2  
Francia: 81.5  
Haiti: 62  
Hong Kong, China: 82.7  
Irán: 72.9  
Irak: 69  
Israel: 81.6  
Italia: 81.8  
Japón: 83.3  
Lesoto: 48.1  
México: 76.9  
Holanda: 80.7  
Panamá: 76.1  
  
Perú: 73.9  
Rusia: 68.8  
Arabia Saudita: 73.9  
Singapur: 81.1  
Somalia: 51.2  
España: 81.4  
Santo Tomé y Príncipe: 64.7  
Sierra Leona: 47.8  
Suecia: 81.4  
Turkia: 73.9  
Reino Unido: 80.1  
Estados Unidos: 78.5  
Venezuela: 74.4  
Zimbabwe: 51.3



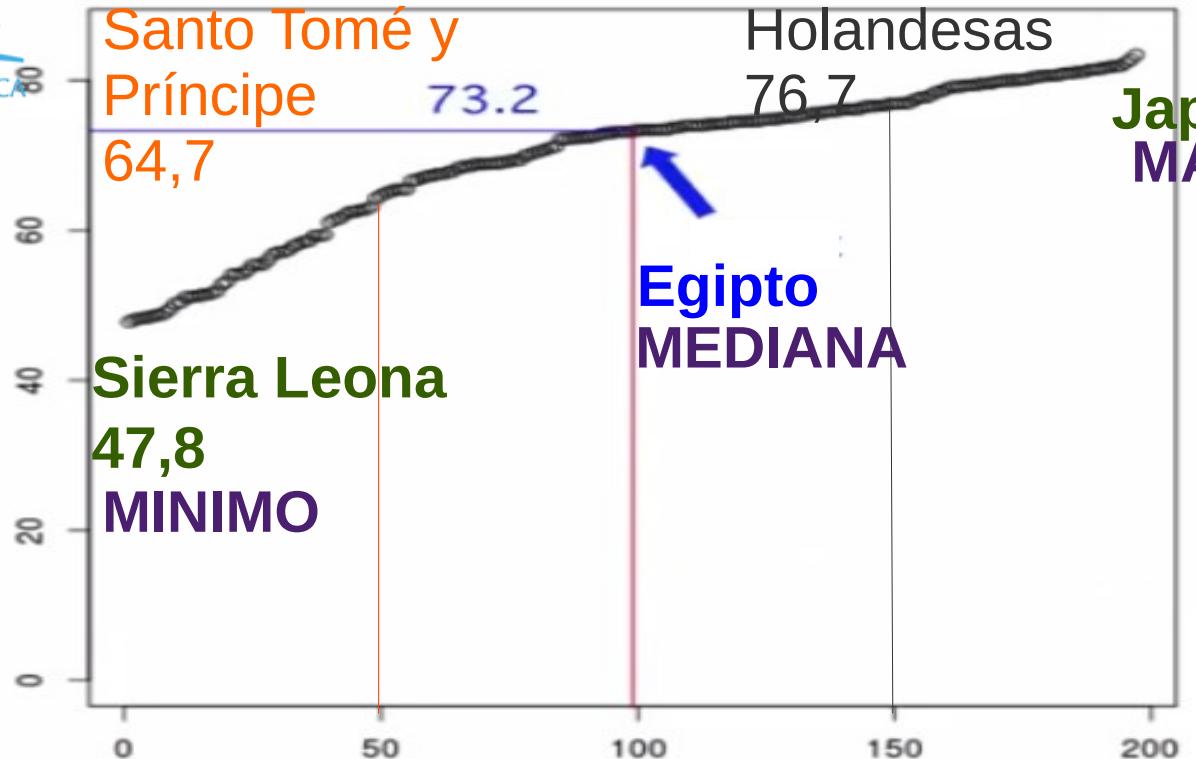


País



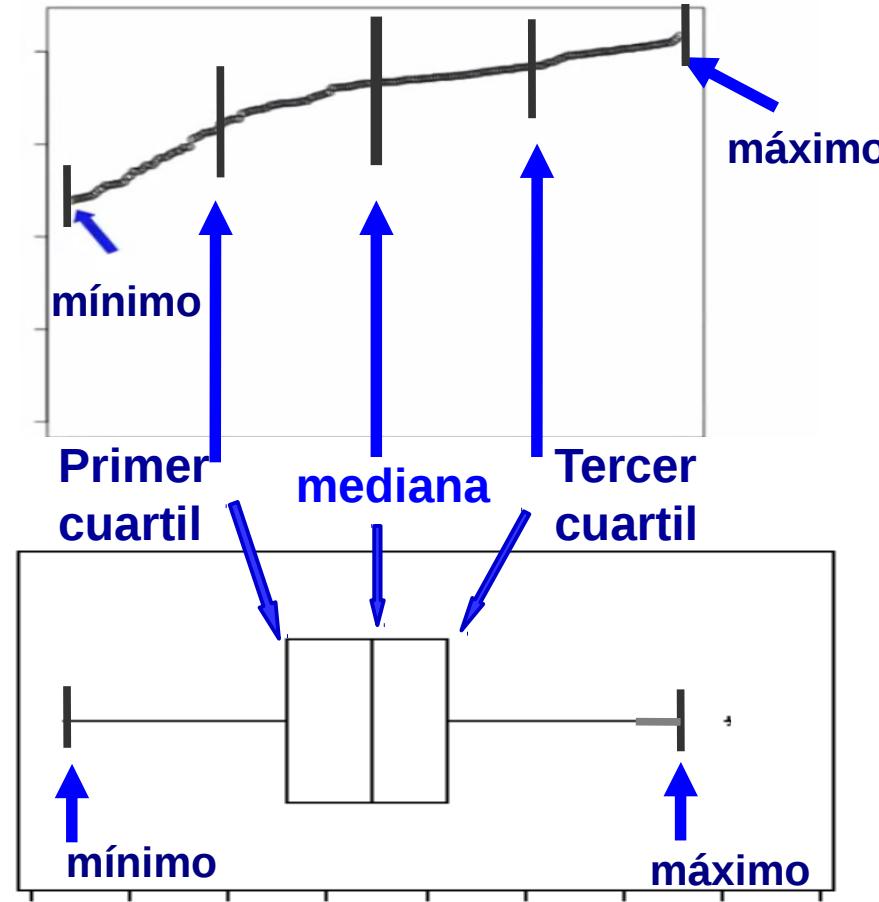
@soniaardila1

## Expectativa de Vida

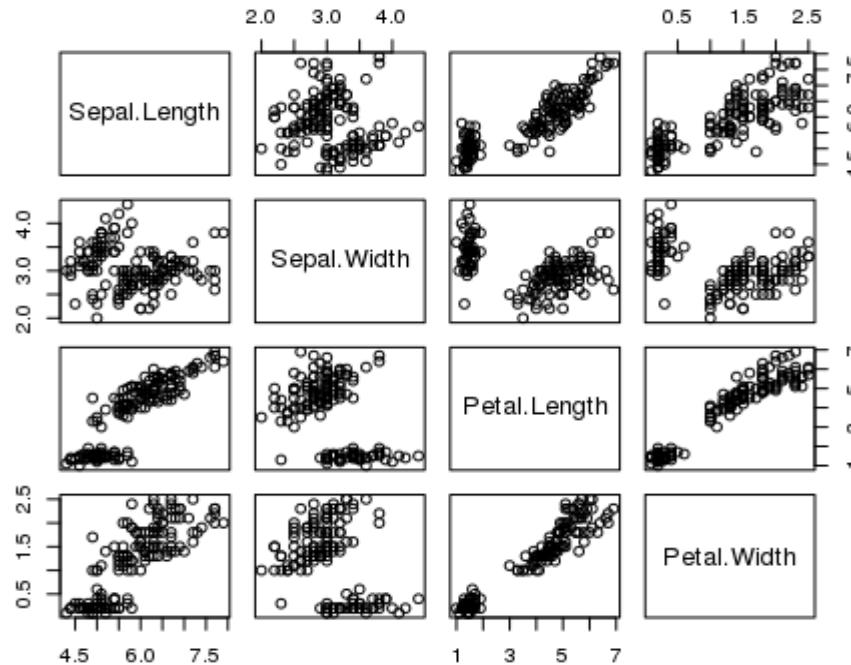


País

@soniaardila1



## SCATTER PLOT MATRIX



@soniaardila1



# Data Storytelling

**Acto 2**

Qué cambios queremos generar en la audiencia?  
Cómo queremos que se sienta?





sonia ardila

Data analyst in @datelligence . Data storyteller. @soniaardila1

Oct 2, 2017 · 5 min read

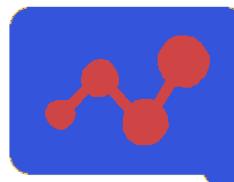
## La enfermedad de La Guajira.



Hans Rosling, el médico y estadístico sueco que estamos extrañando desde febrero de este año, solía repetir una frase cuando presentaba las cifras que reflejan el crecimiento ó decrecimiento de diferentes aspectos de un país. Esa

@soniaardila1

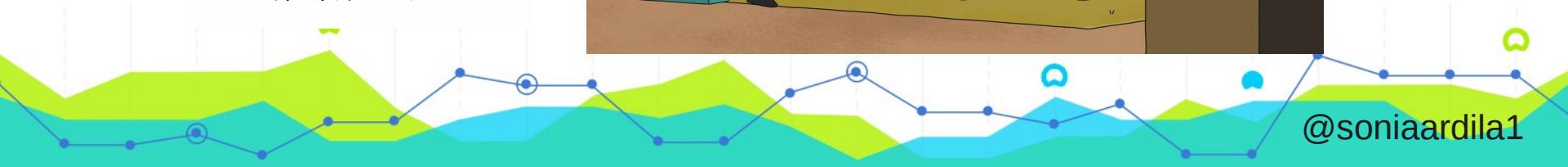
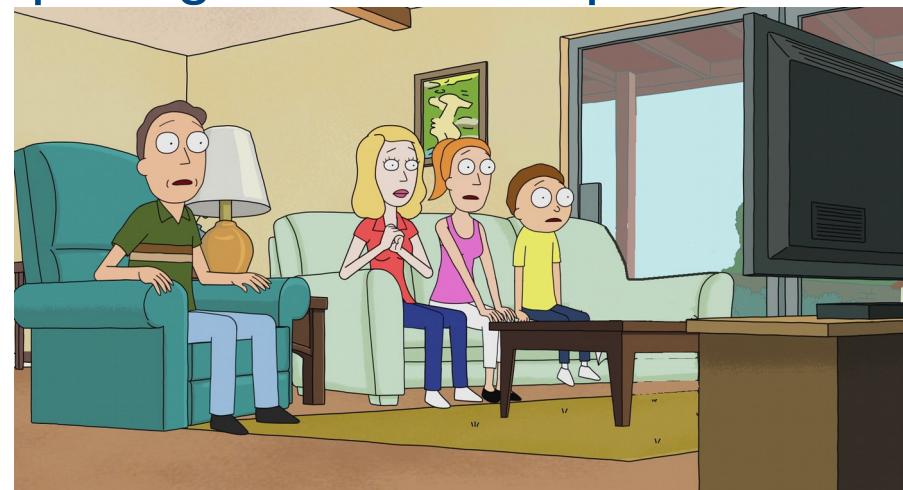


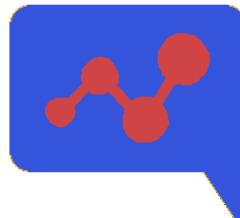


# Data Storytelling

## Acto 3

Presentamos un obstáculo para manejar durante la historia.  
Le tiramos piedras a los protagonistas: Complicaciones.  
Outliers.





# Data Storytelling

## Acto 4

La audiencia se transforma y empieza a comprender una realidad.



@soniaardila1

# <https://www.cran.r-project.org>



[CRAN](#)  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

*About R*  
[R Homepage](#)  
[The R Journal](#)

*Software*  
[R Sources](#)  
[R Binaries](#)  
[Packages](#)  
[Other](#)

*Documentation*  
[Manuals](#)  
[FAQs](#)  
[Contributed](#)

## The Comprehensive R Archive Network

### Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

### Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (Thursday 2017-09-28, Short Summer) [R-3.4.2.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)

### Questions About R

- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

What are R and CRAN?

@soniaardila1



## Choose Your Version of RStudio

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace. [Learn More about RStudio features.](#)

RStudio Desktop  
Open Source License

FREE

[DOWNLOAD](#)

[Learn More](#)

RStudio Desktop  
Commercial License

\$995 per  
year

[BUY](#)

[Learn More](#)

RStudio Server  
Open Source License

FREE

[DOWNLOAD](#)

[Learn More](#)

RStudio Server Pro  
Commercial License

\$9,995 per year

[DOWNLOAD](#)

[Learn More](#)

RStudio Server Pro +  
RStudio Connect  
Commercial License

\$29,995 per  
year

[TALK](#)

[Learn More](#)

Integrated  
Tools for R

[DOWNLOAD](#)

[Learn More](#)



[DOWNLOAD](#)

[Learn More](#)



[DOWNLOAD](#)

[Learn More](#)



[TALK](#)

[Learn More](#)



Priority  
Support



[DOWNLOAD](#)

[Learn More](#)



[TALK](#)

[Learn More](#)



Access via  
Web  
Browser



[DOWNLOAD](#)

[Learn More](#)



Pricing

FREE

\$995/yr

FREE

\$9,995/yr

\$29,995/yr

RStudio Desktop  
Open Source

RStudio  
Desktop  
Commercial

RStudio Server Open  
Source

RStudio Server Pro

RStudio Server Pro +  
RStudio Connect

[DOWNLOAD NOW](#)

[Learn More](#)

[BUY NOW](#)

[Learn More](#)

[DOWNLOAD NOW](#)

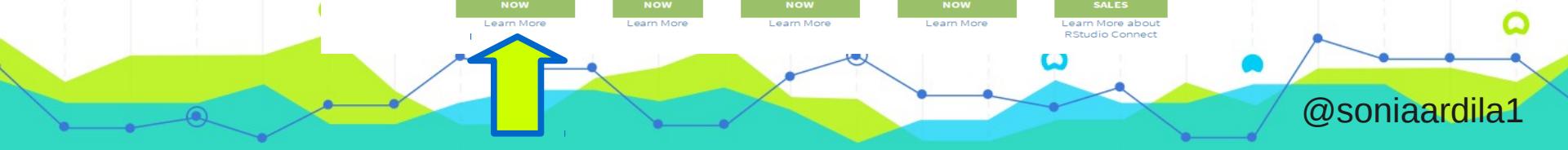
[Learn More](#)

[DOWNLOAD NOW](#)

[Learn More](#)

[CONTACT SALES](#)

[Learn More about  
RStudio Connect](#)





# Data Storytelling



Sandra Borda @sandraborda · Oct 16

Espero lluvia de unfollows y los atiendo de a uno, pero a mí esto me sigue pareciendo una exageración...

Translate from Spanish



@soniaardila1



País	Región	Días festivos/año
Colombia	Sur América	18
Argentina	Sur América	16
Chile	Sur América	15
Venezuela	Sur América	14
Brasil	Sur América	12
Finlandia	Europa	15
Turquía	Europa	14
Rusia	Europa	14
Malta	Europa	14
España	Europa	14
India	Asia	18
Japón	Asia	16
Corea del Sur	Asia	16
Libano	Asia	16
Tailandia	Asia	16
Panamá	Centro América	14
Belice	Centro América	12
Honduras	Centro América	12
Nicaragua	Centro América	12
Costa Rica	Centro América	11
El Salvador	Centro América	11
Guatemala	Centro América	11

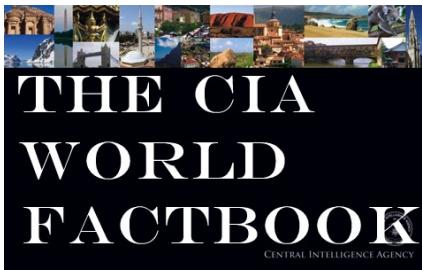
@soniaardila1

1	A	B	C	D	E	F	G	H	I
	País	Región	Días Festivos/año	Pib per cápita	% Desempleo	Muertes Vx100.000h	% Informalidad	Ranking felicidad	Puntaje felicidad
2	Colombia	Sur America	18	14100	9.2	37.6	64	36	6537
3	Brasil	Sur America	12	15200	11.5	30.53	37	22	6635
4	Argentina	Sur America	16	20000	6.6	7.62	47	24	6599
5	Chile	Sur America	15	24100	6.6	3.88	33	20	6652
6	Venezuela	Sur America	14	13800	6.9	47.04	48	82	5250
7	Finlandia	Europa	15	42200	9	2		5	7469
8	Turquía	Europa	14	24900	10.3	3.23	17	69	5500
9	Rusia	Europa	14	26500	5.7	9.21	21	49	5963
10	Malta	Europa	14	39900	5.3	0.7		27	6527
11	España	Europa	14	36400	19.4	0.64	23	34	6403
12	Tailandia	Asia	16	16900	0.6	6.98	43	32	6424
13	Libano	Asia	16	18500	6.8	2.03	20	88	5225
14	Corea del Sur	Asia	16	37700	3.7	1.1		55	5838
15	Japón	Asia	16	41300	3.1	0.28	13	51	5920
16	India	Asia	18	6600	3.5	5.75	85	122	4315
17	Belice	Centro America	12	8400	11.1	43.5	49	50	5956
18	Costa Rica	Centro America	11	16700	9.3	11.2	41	12	7079
19	El Salvador	Centro America	11	8800	7	93.1	56	45	6003
20	Guatemala	Centro America	11	8100	2.4	70.7	70.2	29	6454
21	Honduras	Centro America	12	5400	6.3	36	70.7	91	5181
22	Nicaragua	Centro America	12	5600	6.2	13.5	80	43	6071
23	Panamá	Centro America	14	23400	5.5	34.4	40	30	6452
24									

@soniaardila1

1	País	Región	Ranking felicidad	Puntaje felicidad	Ranking competitividad	puntaje competitividad	Ranking doing business	Puntaje doing business	Expectativa vida
2	Colombia	Sur America	36	6537	61	4.3	54	70.4	74
3	Brasil	Sur America	22	6635	81	4.06	116	57.7	75
4	Argentina	Sur America	24	6599	104	3.81	121	56.8	76
5	Chile	Sur America	20	6652	33	4.64	48	71.5	79
6	Venezuela	Sur America	82	5250	130	3.27	186	35.5	74
7	Finlandia	Europa	5	7469	10	5.44	10	81.1	81
8	Turquía	Europa	69	5500	55	4.39	55	69.2	75
9	Rusia	Europa	49	5963	43	4.51	51	71	71
10	Malta	Europa	27	6527	40	4.52	80	63.7	82
11	España	Europa	34	6403	32	4.68	33	74.9	83
12	Tailandia	Asia	32	6424	34	4.64	49	71.4	75
13	Líbano	Asia	88	5225	101	3.84	123	56.4	79
14	Corea del Sur	Asia	55	5838	26	5.03	4	83.9	82
15	Japón	Asia	51	5920	8	5.48	34	74.7	84
16	India	Asia	122	4315	39	4.52	130	54.7	68
17	Bélgica	Centro America	50	5956	123	3.52	120	56.8	70
18	Costa Rica	Centro America	12	7079	54	4.41	58	68.6	80
19	El Salvador	Centro America	45	6003	105	3.81	86	62.8	73
20	Guatemala	Centro America	29	6454	78	4.08	81	63.5	73
21	Honduras	Centro America	91	5181	88	3.98	110	58.1	73
22	Nicaragua	Centro America	43	6071	103	3.81	125	55.8	75
23	Panamá	Centro America	30	6452	42	4.51	69	65.7	78
24									





Pib Per cápita  
Tasa desempleo



Muertes violentas  
por cada 100.000 hab.



International  
Labour  
Organization

Tasa trabajo  
informal



Puntaje felicidad



Puntaje Competitividad



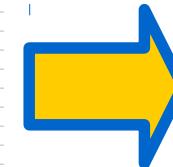
Puntaje Doing Business



Expectativa de vida



	A	B	C	D	E	F	G	H	I
1	País	Región	Días Festivos/año	PIB per cápita	% Desempleo	Muertes Vx100.000h	% Informalidad	Ranking felicidad	Puntaje felicidad
2	Colombia	Sur America	18	14100	9.2	37.6	64	36	6537
3	Brasil	Sur America	12	15200	11.5	30.53	37	22	6635
4	Argentina	Sur America	16	20000	6.6	7.62	47	24	6599
5	Chile	Sur America	15	24100	6.6	3.88	33	20	6652
6	Venezuela	Sur America	14	13800	6.9	47.04	48	82	5250
7	Finlandia	Europa	15	42200	9	2	5	7469	
8	Turquía	Europa	14	24900	10.3	3.23	17	69	5500
9	Rusia	Europa	14	26500	5.7	9.21	21	49	5963
10	Malta	Europa	14	39900	5.3	0.7		27	6527
11	España	Europa	14	36400	19.4	0.64	23	34	6403
12	Tailandia	Asia	16	16900	0.6	6.98	43	32	6424
13	Líbano	Asia	16	18500	6.8	2.03	20	88	5225
14	Corea del Sur	Asia	16	37700	3.7	1.1		55	5838
15	Japón	Asia	16	41300	3.1	0.28	13	51	5920
16	India	Asia	18	6600	3.5	5.75	85	122	4315
17	Bélgica	Centro America	12	8400	11.1	43.5	49	50	5956
18	Costa Rica	Centro America	11	16700	9.3	11.2	41	12	7079
19	El Salvador	Centro America	11	8800	7	93.1	56	45	6003
20	Guatemala	Centro America	11	8100	2.4	70.7	70.2	29	6454
21	Honduras	Centro America	12	5400	6.3	36	70.7	91	5181
22	Nicaragua	Centro America	12	5600	6.2	13.5	80	43	6071
23	Panamá	Centro America	14	23400	5.5	34.4	40	30	6452
24									



Festivos.csv





# GitHub

<https://github.com/sap0408>





Your pinned repositories have been updated. Drag and drop to reorder them.



## sonia ardila

sap0408

Data Scientist en formación  
permanente! Data Storyteller!

@datelligence

Colombia

sap0408@gmail.com

<http://www.datelligence.com>

### Overview

Repositories 8

Stars 1

Followers 1

Following 0

### Pinned repositories

Customize your pinned repositories

#### [opendata-CancerColombia-1999-2011](#)

Data del Instituto nacional de cancerología.

★ 1

#### [TALLER-DATA-JOURNALISM](#)

Datasets para explorar y encontrar noticias !

★ 1

#### [concesionarios](#)

#### [DIAS-FESTIVOS-PAISES](#)

Países con más días festivos en cuatro regiones del planeta.

#### [OpenData](#)

Forked from eledero/OpenData

Datos Abiertos BOG

R

#### [Taller-Data-Storytelling](#)

Dataset, Slides Taller Sonia Ardila Data storytelling.



# dataset : Festivos

Import Dataset

Name	Festivos	Input File
Encoding	Automatic	Pais,Continente,Dias Festivos/año,Pib per cápita,Dias vía Colombia,America,18,14100,,9.2,37.6,64,36,6537,61,4.3,53 Brasil,America,12,15200,30,11.5,30.53,37,22,6635,81,4.06 Argentina,America,16,20000,,6.6,7.62,47,24,6599,104,3.81 Chile,America,15,24100,,6.6,3.88,33,20,6652,33,4.64,57,7 Venezuela,America,14,13800,,6.9,47.04,48,82,5250,130,3.2 Finlandia,Europa,15,42200,30,9.2,,5,7469,10,5.44,13,81 Turquia,Europa,14,24900,,10.3,3.23,17,69,5500,55,4.39,69 Rusia,Europa,14,26500,,5.7,9.21,21,49,5963,43,4.51,40,71 Malta,Europa,14,39900,,5.3,0.7,,27,6527,40,4.52,76,82 España,Europa,14,36400,30,19.4,0.64,23,34,6403,32,4.68,3 Tailandia,Asia,16,16900,15,0.6,6.98,43,32,6424,34,4.64,4 Libano,Asia,16,18500,,6.8,2.03,20,88,5225,101,3.84,126,7 Corea del Sur,Asia,16,37700,15,3.7,1.1,,55,5838,26,5.03, Janón,Asia,16,41300,20,3.1,0.28,13.51,5920,8.5,48,34.84
Heading	<input checked="" type="radio"/> Yes <input type="radio"/> No	
Row names	Automatic	
Separator	Comma	
Decimal	Period	
Quote	Double quote (")	
Comment	None	
na.strings	NA	
<input checked="" type="checkbox"/> Strings as factors		

Data Frame

Pais	Continente	Dias.Festivos.año	Pib.per.c
Colombia	America	18	14100
Brasil	America	12	15200
Argentina	America	16	20000
Chile	America	15	24100
Venezuela	America	14	13800
Finlandia	Europa	15	42200
Turquia	Europa	14	24900
Rusia	Europa	14	26500
Malta	Europa	14	39900
España	Europa	14	36400
Tailandia	Asia	16	16900
Libano	Asia	16	18500
Corea del Sur	Asia	16	37700
Janón	Asia	16	41300

Import Cancel


 @soniaardila1

RStudio

File Edit Code View Plots Session Build Debug Tools Help

Project: (None)

PCA.R x DiasFestivos.R\* x Untitled1 x BOXPLOT.R x CONGESTION >

Import Dataset From Local File... From Web URL...

Data

- filter 1419 obs. of 13 variables
- filter2 236 obs. of 3 variables
- graphData 236 obs. of 3 variables
- SATURNO 1419 obs. of 13 variables

Values

grafica List of 9

Environment History

Import Dataset From Text (base)... From Text (readr)... From Excel... From SPSS... From SAS... From Stata...

Environment History Connections

Files Plots Packages Help Viewer

Console ~/



@soniaardila1

## Import Dataset



Name

Festivos

Input File

```
Pais,Región,Días Festivos/año,Pib per cápita.% Desempleo.Muert
Colombia,Sur America,18,14100,9.2,37.6,64,36,6537,61,4,3,54,70
Brasil,Sur America,12,15200,11.5,30.53,37,22,6635,81,4,06,116,
Argentina,Sur America,16,20000,6.6,7,62,47,24,6599,104,3,81,12
Chile,Sur America,15,24100,6.6,3,88,33,20,6652,33,4,64,48,71,5
Venezuela,Sur America,14,13800,6.9,47,04,48,82,5250,130,3,27,1
Finlandia,Europa,15,42200,9,2,,5,7469,10,5,44,10,81,1,81
Turquia,Europa,14,24900,10,3,3,23,17,69,5500,55,4,39,55,69,2,7
Rusia,Europa,14,26500,5,7,9,21,21,49,5963,43,4,51,51,71,71
Malta,Europa,14,39900,5,3,0,7,,27,6527,40,4,52,80,63,7,82
España,Europa,14,36400,19,4,0,64,23,34,6403,32,4,68,33,74,9,83
Tailandia,Asia,16,16900,0,6,6,98,43,32,6424,34,4,64,49,71,4,75
Libano,Asia,16,18500,6,8,2,03,20,88,5225,101,3,84,123,56,4,79
Corea del Sur,Asia,16,37700,3,7,1,1,,55,5838,26,5,03,4,83,9,82
Japón,Asia,16,41300,3,1,0,28,13,51,5920,8,5,48,34,74,7,84
```

Encoding Automatic

Heading  Yes  No

Row names Automatic

Separator Comma

Decimal Period

Quote Double quote ("")

Comment None

na.strings NA

 Strings as factors

Data Frame

Pais	Región	Días.Festivos.año	Pib.per.cá
Colombia	Sur America	18	14100
Brasil	Sur America	12	15200
Argentina	Sur America	16	20000
Chile	Sur America	15	24100
Venezuela	Sur America	14	13800
Finlandia	Europa	15	42200
Turquia	Europa	14	24900
Rusia	Europa	14	26500
Malta	Europa	14	39900
España	Europa	14	36400
Tailandia	Asia	16	16900
Libano	Asia	16	18500
Corea del Sur	Asia	16	37700
Japón	Asia	16	41300

Import

Cancel

1:1 (Top Level)

Console ~/

&gt;

@soniaardila1

RStudio

File Edit Code View Project Workspace Plots Tools Help

New >

- R Script Ctrl+Shift+N
- Text File
- R Sweave
- R Markdown
- R HTML
- R Documentation

Open File... Ctrl+O

Recent Files >

Save Ctrl+S

Save As... Save All

Print...

Close Ctrl+W

Close All Ctrl+Shift+W

Quit R... Redistribute it under the terms of the GNU General Public License() for support but running in an English locale.

R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

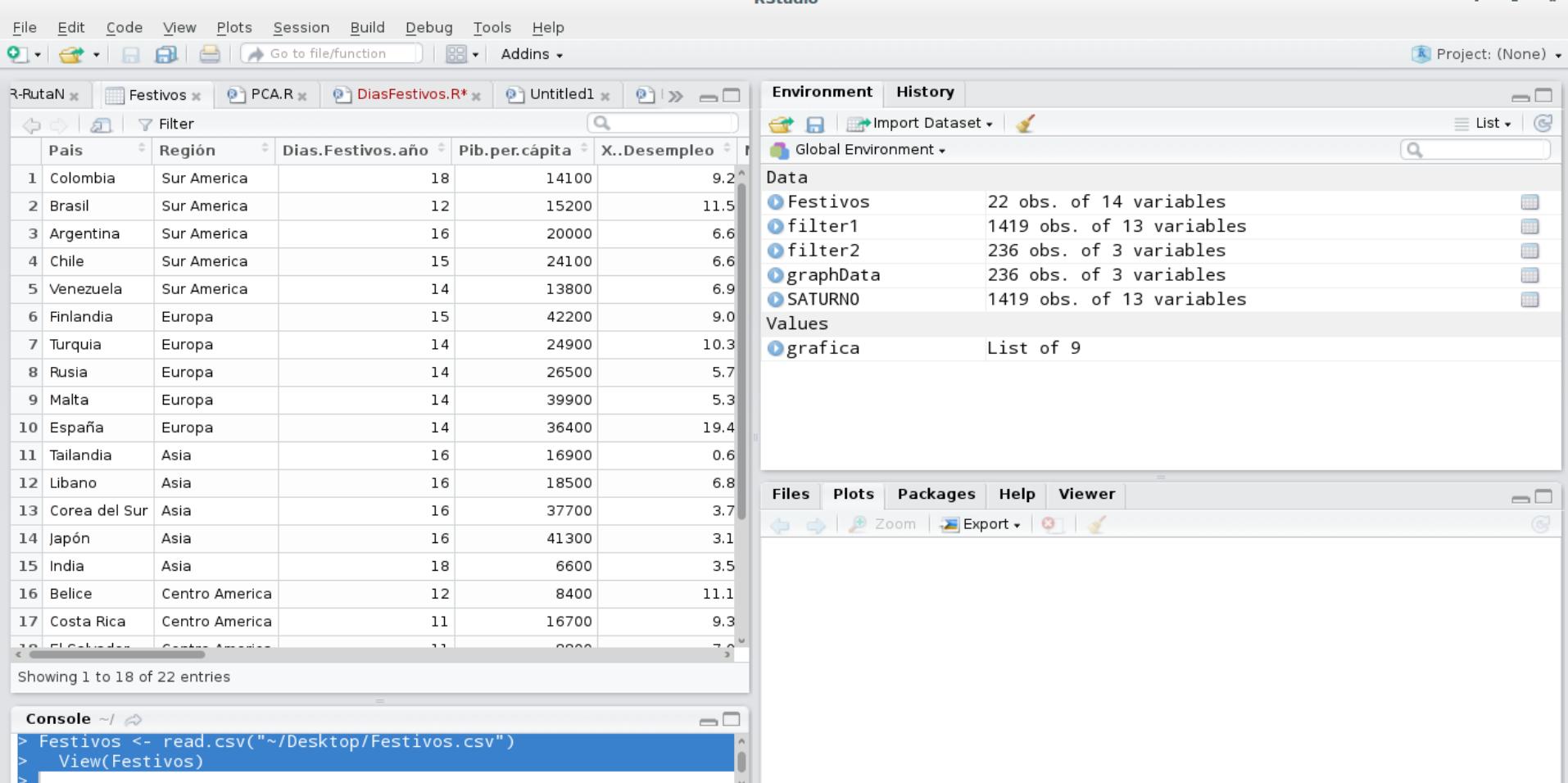
> |

Project: (None)

Workspace History Load Save Import Dataset Clear All

Files Plots Packages Help Zoom Export Clear All





@soniaardila1

PCAR.R x DiasFestivos.R\* x Festivos x Untitled1\* x BOXPLOT. x

Source on Save Run Source

```
1 #sin cargar ningún paquete
2 Festivos
```

2:1 (Top Level) R Script

## Console ~/

```
> Festivos
```

	Pais	Región	Dias.Festivos.año	Pib.per.cápita
1	Colombia	Sur America	18	14100
2	Brasil	Sur America	12	15200
3	Argentina	Sur America	16	20000
4	Chile	Sur America	15	24100
5	Venezuela	Sur America	14	13800
6	Finlandia	Europa	15	42200
7	Turquia	Europa	14	24900
8	Rusia	Europa	14	26500
9	Malta	Europa	14	39900
10	España	Europa	14	36400
11	Tailandia	Asia	16	16900
12	Libano	Asia	16	18500
13	Corea del Sur	Asia	16	37700

## Environment History

Import Dataset

Global Environment

## Data

Festivos	22 obs. of 14 variables
filter1	1419 obs. of 13 variables
filter2	236 obs. of 3 variables
graphData	236 obs. of 3 variables
SATURNO	1419 obs. of 13 variables

## Values

grafica List of 9

## Files Plots Packages Help Viewer

Zoom Export



```

PCA.R x DiasFestivos.R* x Festivos x Untitled1* x BOXPLOT.R x CONGESTIONCIUDADESMUNDO.R x >> Run Source & 
Source on Save & 
1 #sin cargar ningún paquete
2 Festivos
3
4 View(Festivos)
5 str(Festivos)

```

5:1 (Top Level) R Script

## Console ~/

```

> View(Festivos)
> str(Festivos)
'data.frame': 22 obs. of 14 variables:
$ País          : Factor w/ 22 levels "Argentina","Belice",...
$ Región        : Factor w/ 4 levels "Asia","Centro America",...
$ Dias.Festivos.año: int 18 12 16 15 14 15 14 14 14 ...
$ Pib.per.cápita: int 14100 15200 20000 24100 13800 42200 24900 26500 39900 ...
$ X..Desempleo  : num 9.2 11.5 6.6 6.6 6.9 9 10.3 5.7 5.3 19.4 ...
$ Muertes.Vx100.000h: num 37.6 30.53 7.62 3.88 47.04 ...
$ X..Informalidad: num 64 37 47 33 48 NA 17 21 NA 23 ...
$ Ranking.felicidad: int 36 22 24 20 82 5 69 49 27 34 ...
$ Puntaje.felicidad: int 6537 6635 6599 6652 5250 7469 5500 5963 6527 6403 ...
$ Ranking.competitividad: int 61 81 104 33 130 10 55 43 40 32 ...
$ puntaje.competitividad: num 4.3 4.06 3.81 4.64 3.27 5.44 4.39 4.51 4.52 4.68 ...
$ Ranking.doing.business: int 54 116 121 48 186 10 55 51 80 33 ...
$ Puntaje.doing.business: num 70.4 57.7 56.8 71.5 35.5 81.1 69.2 71 63.7 74.9 ...
$ Expectativa.vida    : int 74 75 76 79 74 81 75 71 82 83 ...

```

## Environment History

Import Dataset

Global Environment

## Data

Festivos	22 obs. of 14 variables	
filter1	1419 obs. of 13 variables	
filter2	236 obs. of 3 variables	
graphData	236 obs. of 3 variables	
SATURNO	1419 obs. of 13 variables	

## Values

grafica List of 9

## Files Plots Packages Help Viewer

Zoom

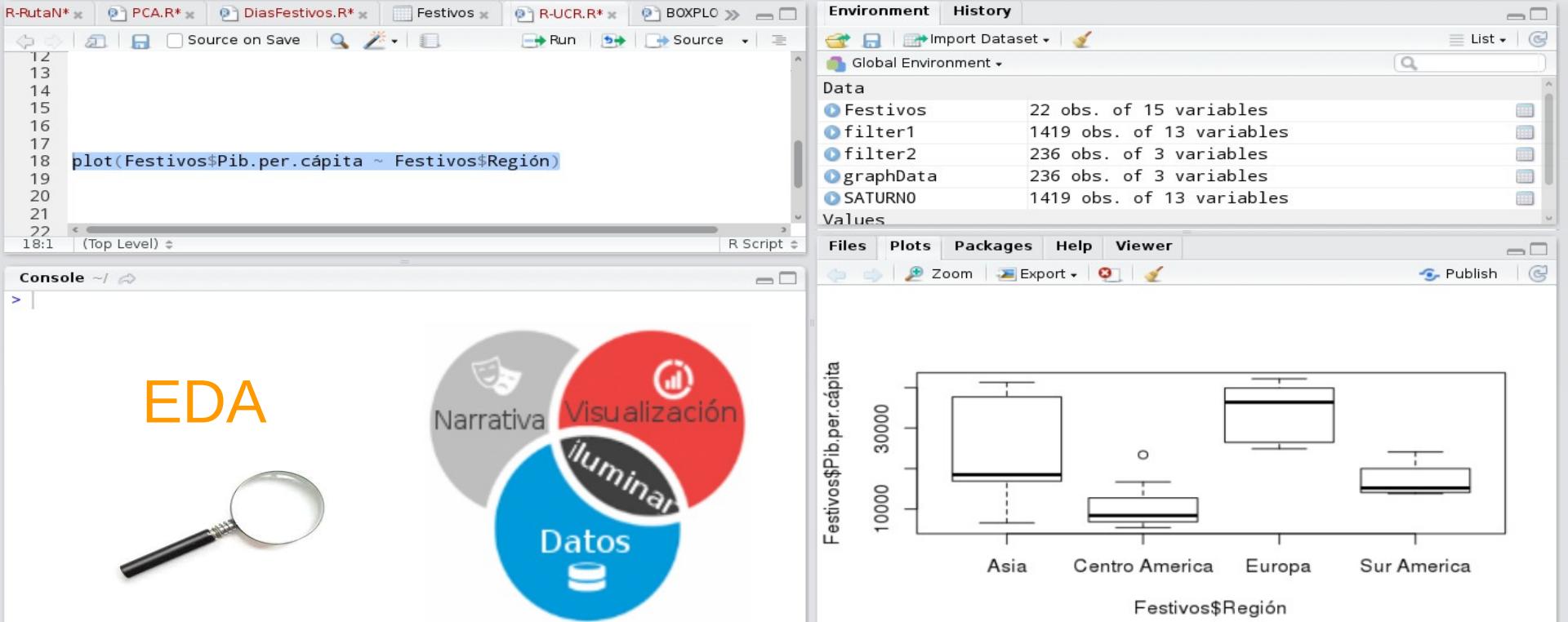
Export

@soniaardila1

File Edit Code View Plots Session Build Debug Tools Help

Go to file/function

Project: (None)



Project: (None)

R-RutaN\* PCA.R\* DiasFestivos.R\* Festivos R-UCR.R\* BOXPLOT.R

Source on Save Run Source

```

17
18 plot(Festivos$Pib.per.cápita ~ Festivos$Región)
19
20 plot(Festivos$Pib.per.cápita ~ Festivos$Región,
21       xlab= "Región", ylab = "Pib Per cápita", main="Pib Per cápita en
22       países con más días festivos en cuatro regiones")
23
24
25
26
27
28

```

20:1 (Top Level) R Script

Environment History

Import Dataset Global Environment

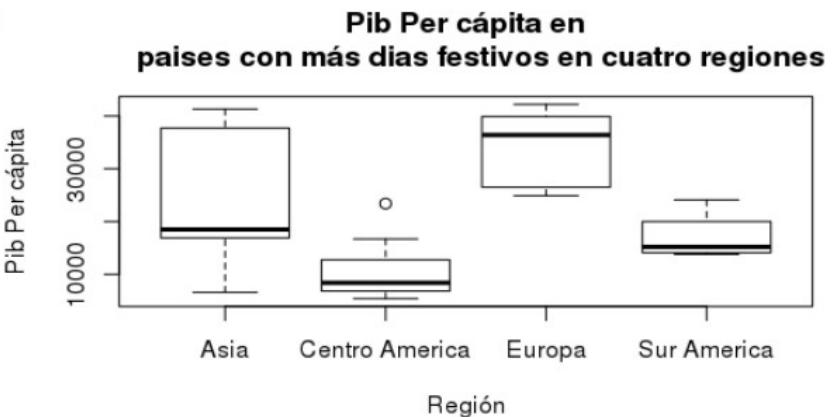
Data

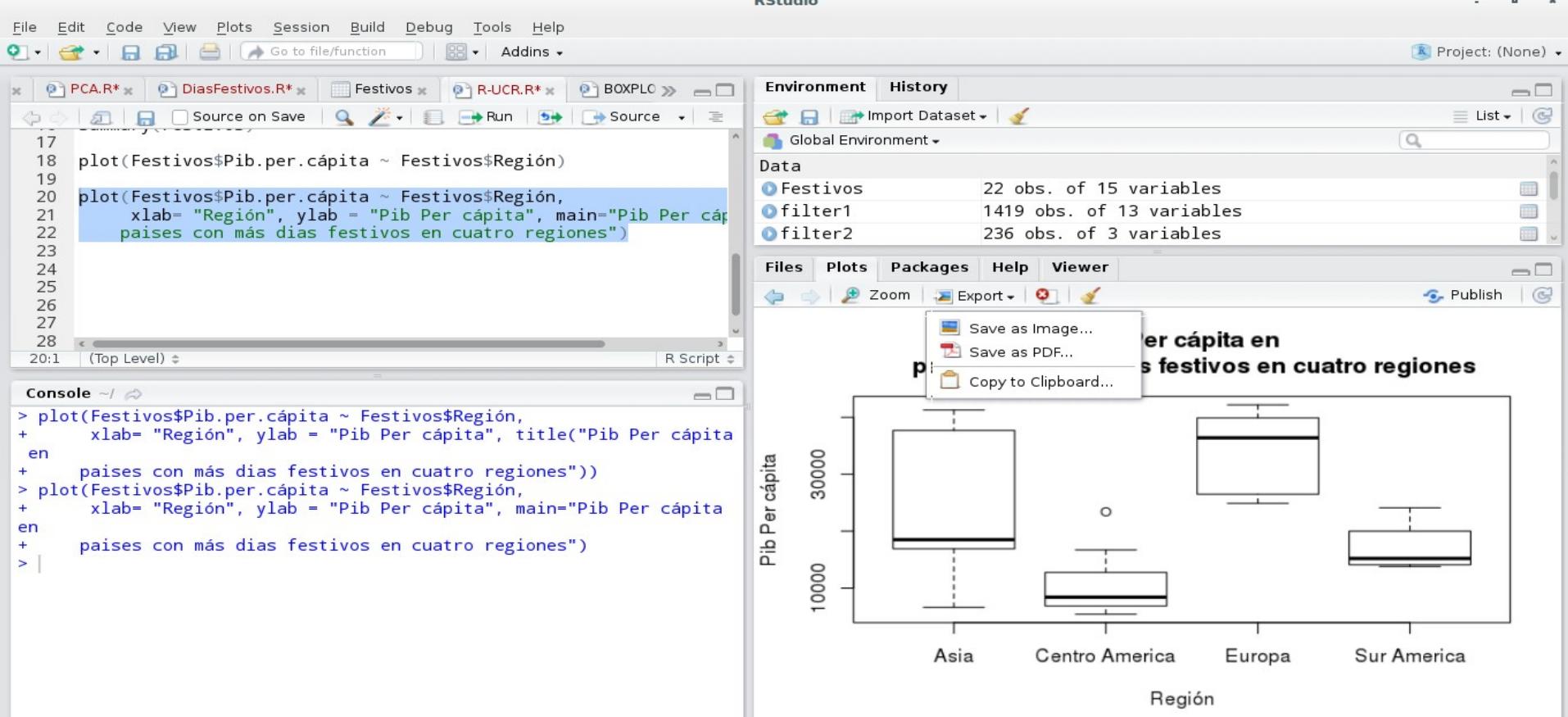
- Festivos 22 obs. of 15 variables
- filter1 1419 obs. of 13 variables
- filter2 236 obs. of 3 variables
- graphData 236 obs. of 3 variables
- SATURNO 1419 obs. of 13 variables

Values

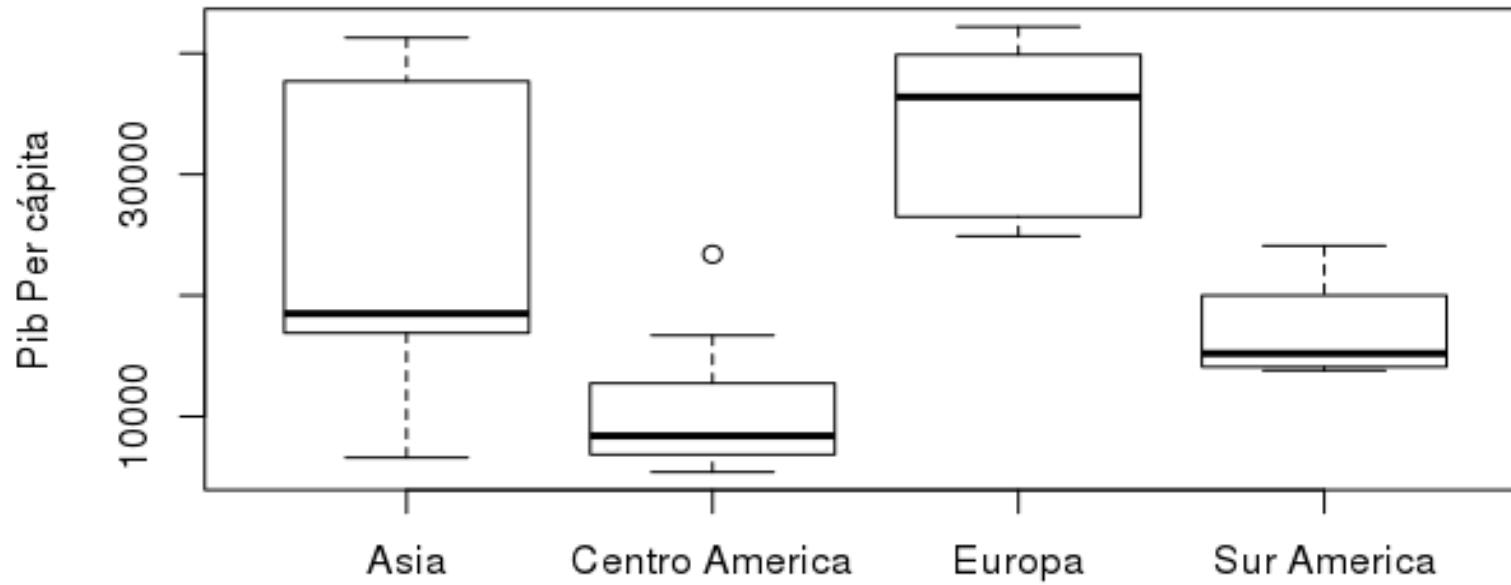
Files Plots Packages Help Viewer

Zoom Export Publish





## Pib Per cápita en países con más días festivos en cuatro regiones



@soniaardila1

File Edit Code View Plots Session Build Debug Tools Help

Go to file/function

Project: (None)

JINDO.R\* R-UCR.R\* Festivos taller.R\* conferenciaunitec.R\* >

```
37
38
39
40
41
42
43
44
45
46
47
48
```

41:1 (Top Level) R Script

Console ~/

```
> library("ggplot2", lib.loc = "~/R/x86_64-pc-linux-gnu-library/3.3")
> |
```

Environment History

Import Dataset |

Global Environment

Data

filter1	1419 obs. of 13 variables
filter2	236 obs. of 3 variables
graphData	236 obs. of 3 variables
SATURNO	1419 obs. of 13 variables

Values

grafica	List of 9
---------	-----------

Files Plots Packages Help Viewer

Install | Update |

Name	Description	Version
------	-------------	---------

User Library

<input type="checkbox"/>	assertthat	Easy pre and post assertions.	0.1
<input type="checkbox"/>	BH	Boost C++ Header Files	1.60.0-2
<input type="checkbox"/>	colorspace	Color Space Manipulation	1.2-6
<input type="checkbox"/>	DBI	R Database Interface	0.4-1
<input type="checkbox"/>	dichromat	Color Schemes for Dichromats	2.0-0
<input type="checkbox"/>	digest	Create Compact Hash Digests of R Objects	0.6.9
<input type="checkbox"/>	dplyr	A Grammar of Data Manipulation	0.5.0
<input checked="" type="checkbox"/>	ggplot2	An Implementation of the Grammar of Graphics	2.1.0

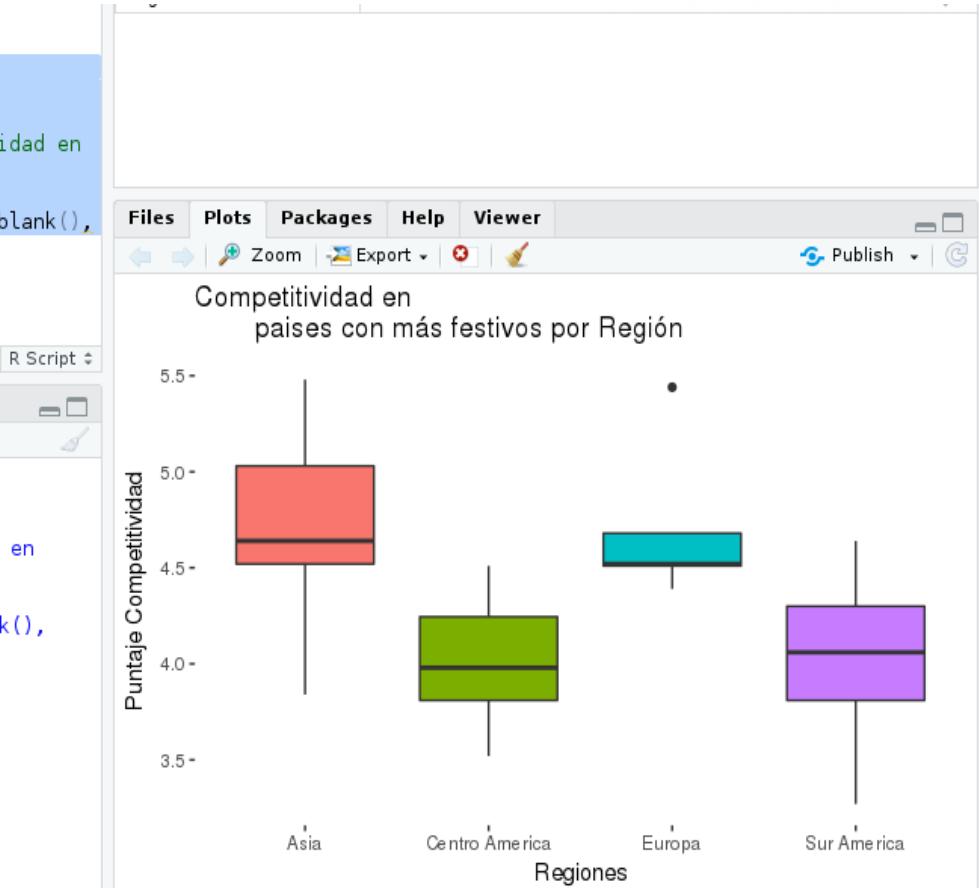


```

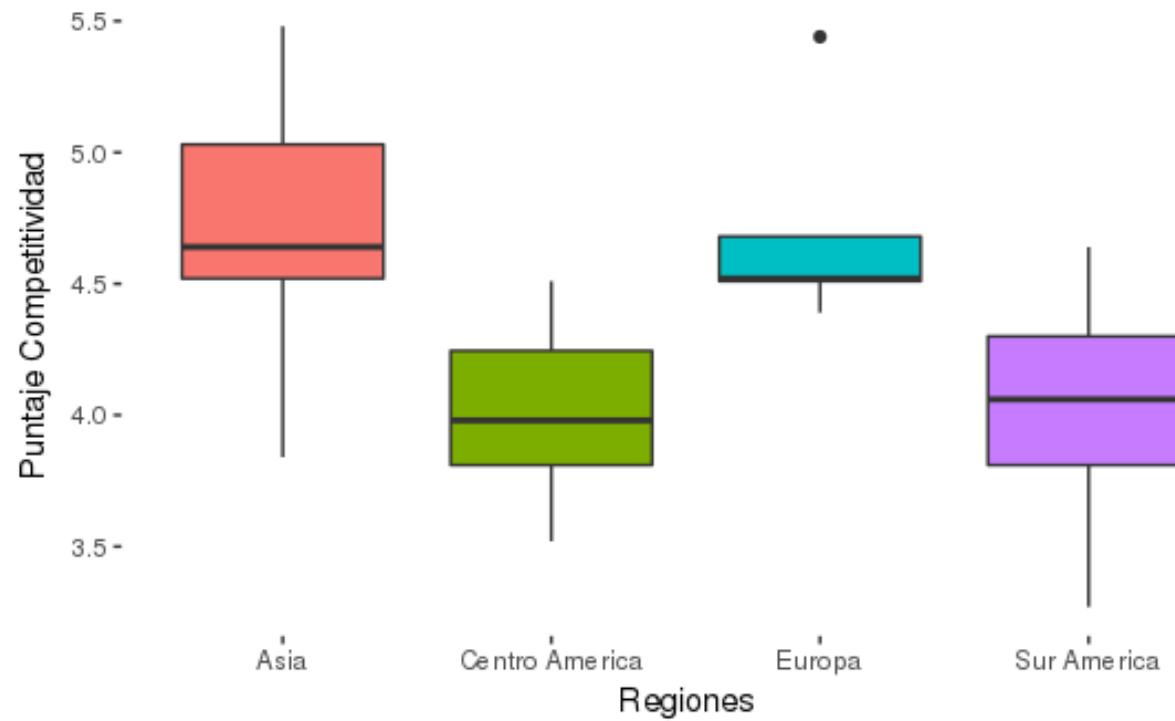
6
7
8 grafica <- ggplot(Festivos, aes(x= Región, y = puntaje.competitividad,
9   fill= Región)) +
10 geom_boxplot()+
11 labs (x= "Regiones", y="Puntaje Competitividad", title = "Competitividad en
12     países con más festivos por Región")+
13 theme(legend.position = "none");grafica +
14 theme(panel.background = element_blank(),panel.grid.major = element_blank(),
15       panel.grid.minor = element_blank())

```

8:1 (Top Level) ▾



## Competitividad en países con más festivos por Región



@soniaardila1

```

23
24
25 plot(Festivos$Pib.per.cápita ~ Festivos$Dias.Festivos.año,
26       xlab= "Días Festivos", ylab = "Pib Per cápita",
27       main="Relación Días Festivos - Pib Per cápita en países
28       con más días festivos en cuatro regiones")
29
30
31
32
33
34
35:1 (Top Level) R Script

```

## Console

```

> plot(Festivos$Pib.per.cápita ~ Festivos$Dias.Festivos.año,xlab= "Días Festivo
s",
+ ylab = "Pib Per cápita", main="Relación Días Festivos - Pib Per cápita en pa
íses
+ con más días festivos en cuatro regiones")
>

```

## Environment History

Import Dataset | Global Environment

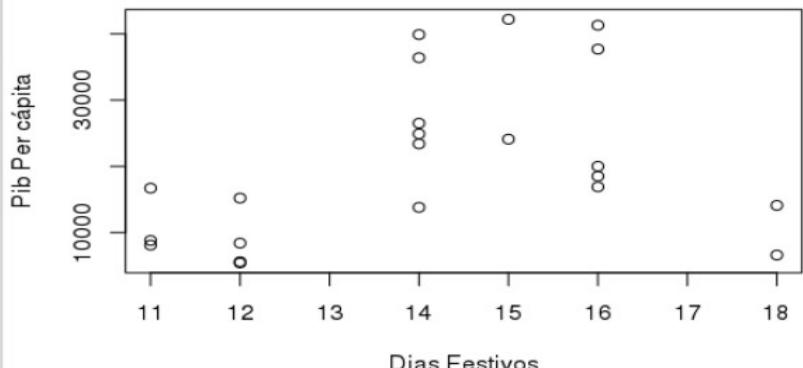
## Data

Festivos	22 obs. of 15 variables
filter1	1419 obs. of 13 variables
filter2	236 obs. of 3 variables

## Files Plots Packages Help Viewer

Zoom Export | Publish

### Relación Días Festivos - Pib Per cápita en países con más días festivos en cuatro regiones



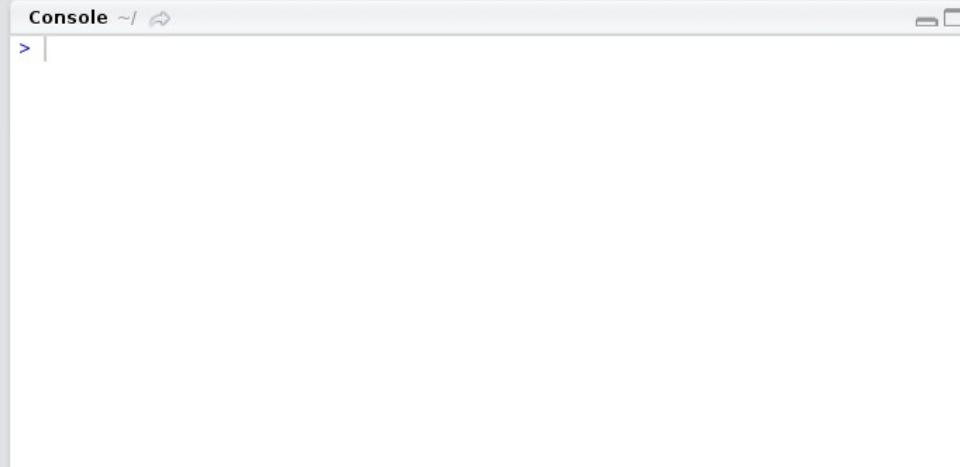
@soniaardila1

OT.R\* CONGESTIONCIUDADES MUNDO.R\* Festivos R-UCR.R\* taller.R\* Addins

Source on Save Run Source

```
46  
47  
48 ggplot(Festivos, aes(Dias.Festivos.año,Pib.per.cápita))+  
  geom_point(aes(color=factor(Región)))+  
  xlab("Días festivos") + ylab("Pib per cápita") +  
  ggtitle ("Relación días festivos -Pib per cápita en países con  
    más días festivos en cuatro regiones")  
53  
54  
55  
56  
57
```

(Top Level) R Script



Environment History

Import Dataset

Global Environment

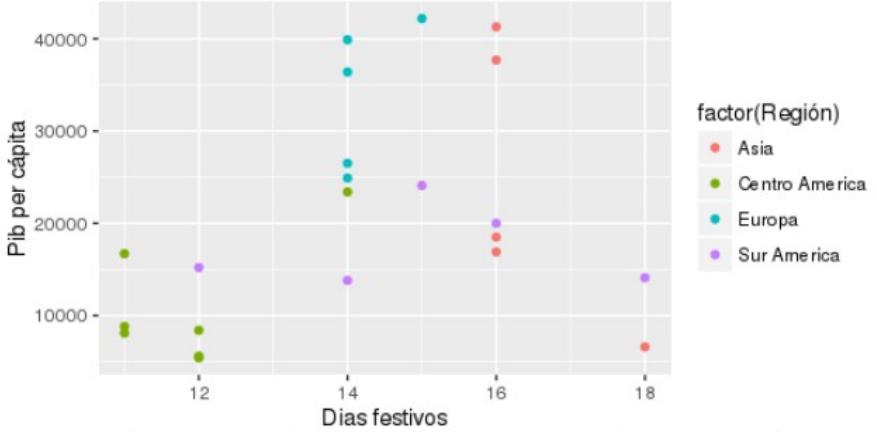
Data

1 Festivos	22 obs. of 14 variables
1 filter1	1419 obs. of 13 variables
1 filter2	236 obs. of 3 variables
1 graphData	236 obs. of 3 variables
1 SATURNO	1419 obs. of 13 variables

Files Plots Packages Help Viewer

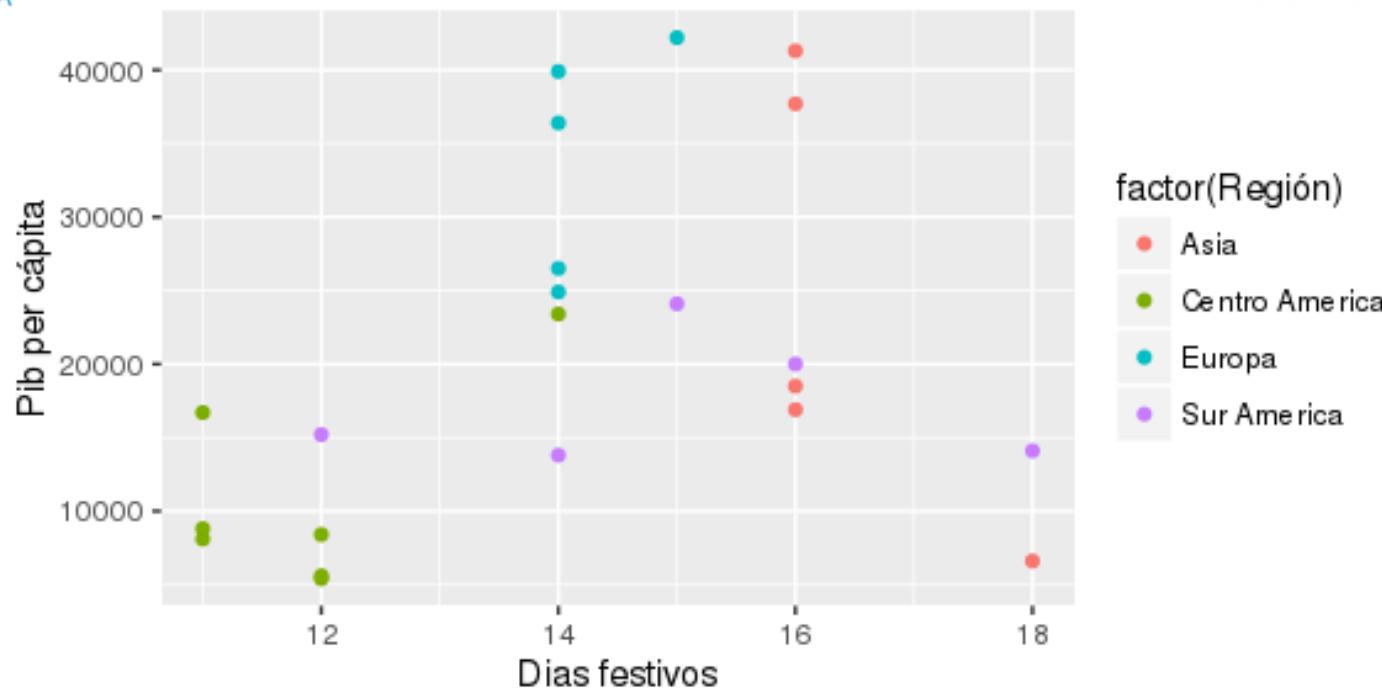
Zoom Export

Relación días festivos -Pib per cápita en países con  
más días festivos en cuatro regiones



@soniaardila1

## Relación días festivos -Pib per cápita en países con más días festivos en cuatro regiones



@soniaardila1

OT.R\* CONGESTIONCIUDADES MUNDO.R\* Festivos R-UCR.R\* taller.R\* confere... Run Source

```

38
39 ggplot() + geom_histogram(data = Festivos, aes(x= Dias.Festivos.áño),
40   fill="blue", color="red", binwidth = 1) +
41   labs(y="Conteo países", x="Días festivos al año",
42     title="Países con mas días festivos al año en cuatro regiones
43     del mundo") +
44   theme(panel.background = element_blank(), panel.grid.major = element_blank(),
45     panel.grid.minor = element_blank())
46
47
48
49
39:1 (Top Level) R Script

```

Console ~

```

> ggplot() + geom_histogram(data = Festivos, aes(x= Dias.Festivos.áño),
+   fill="blue", color="red", binwidth = 1) +
+   theme(panel.background = element_blank(), panel.grid.major = element_blank(),
+     panel.grid.minor = element_blank())
> ggplot() + geom_histogram(data = Festivos, aes(x= Dias.Festivos.áño),
+   fill="blue", color="red", binwidth = 1) +
+   labs(y="Conteo países", x="Días festivos al año",
+     title="Países con mas días festivos al año en cuatro regiones
+     del mundo") +
+   theme(panel.background = element_blank(), panel.grid.major = element_blank(),
+     panel.grid.minor = element_blank())
>

```



## Environment History

Import Dataset | Global Environment

### Data

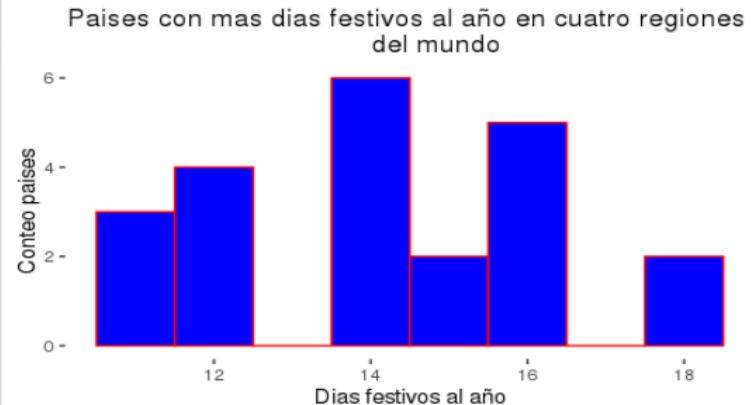
Festivos	22 obs. of 14 variables
filter1	1419 obs. of 13 variables
filter2	236 obs. of 3 variables
graphData	236 obs. of 3 variables
SATURNO	1419 obs. of 13 variables

### Values

graphics List of 0

Files Plots Packages Help Viewer

Zoom Export



@soniaardila1

```

31
32 cor(Festivos[,3:7])
33
34
35
36
37
38
39
40
41
32:1 (Top Level) ⇡

```

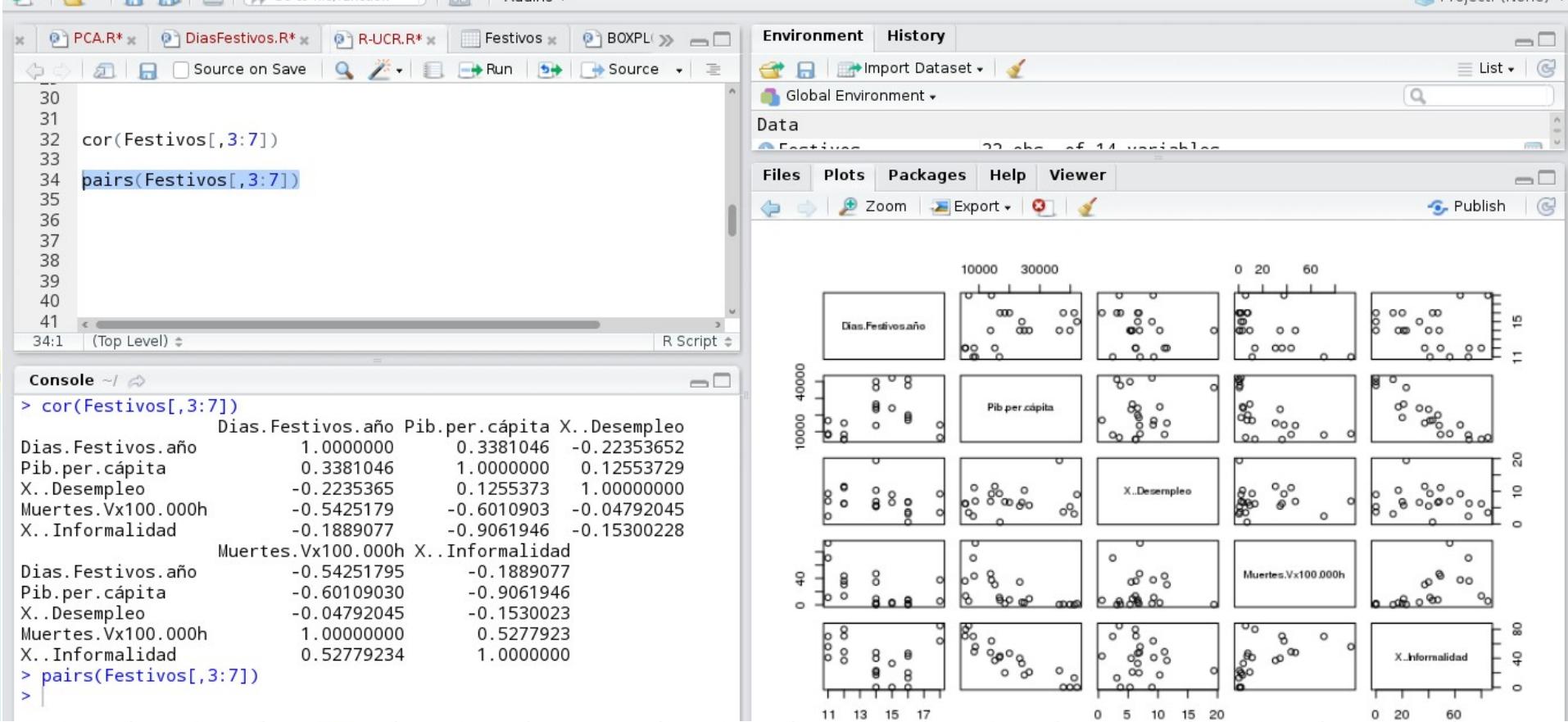
**Console** ↵

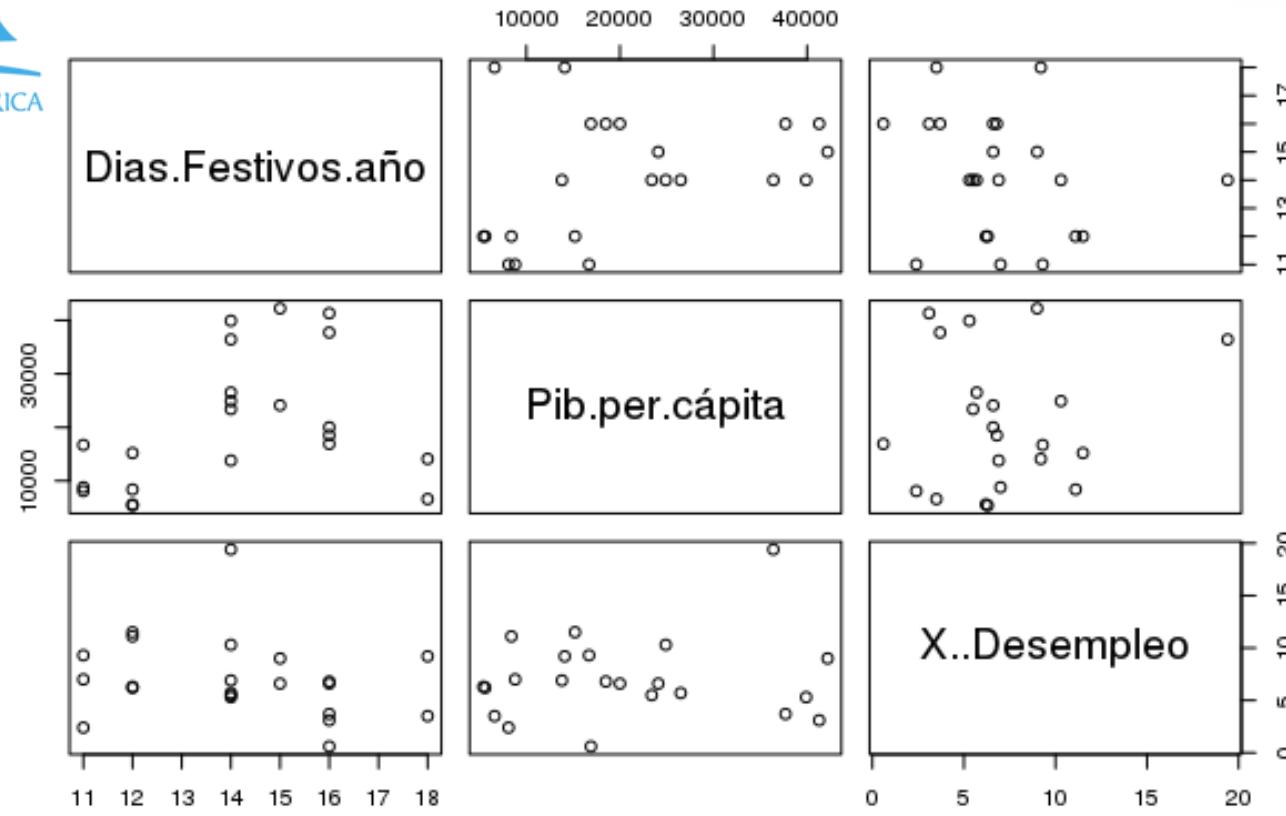
```
> cor(Festivos[,3:7])
```

	Dias.Festivos.año	Pib.per.cápita	X..Desempleo
Dias.Festivos.año	1.0000000	0.3381046	-0.22353652
Pib.per.cápita	0.3381046	1.0000000	0.12553729
X..Desempleo	-0.2235365	0.1255373	1.00000000
Muertes.Vx100.000h	-0.5425179	-0.6010903	-0.04792045
X..Informalidad	-0.1889077	-0.9061946	-0.15300228
	Muertes.Vx100.000h	X..Informalidad	
Dias.Festivos.año	-0.54251795	-0.1889077	
Pib.per.cápita	-0.60109030	-0.9061946	
X..Desempleo	-0.04792045	-0.1530023	
Muertes.Vx100.000h	1.00000000	0.5277923	
X..Informalidad	0.52779234	1.0000000	

>







Console Terminal × element\_blank()

```
> install.packages("car")
Installing package into '/home/sap/R/x86_64-pc-linux-gnu-library/3.4'
(as 'lib' is unspecified)
also installing the dependencies 'minqa', 'nloptr', 'RcppEigen', 'lme4', 'SparseM'
, 'MatrixModels', 'pbkrtest', 'quantreg'

trying URL 'https://cran.rstudio.com/src/contrib/minqa_1.2.4.tar.gz'
Content type 'application/x-gzip' length 53548 bytes (52 KB)
=====
downloaded 52 KB

trying URL 'https://cran.rstudio.com/src/contrib/nloptr_1.0.4.tar.gz'
Content type 'application/x-gzip' length 353957 bytes (345 KB)
=====
downloaded 345 KB

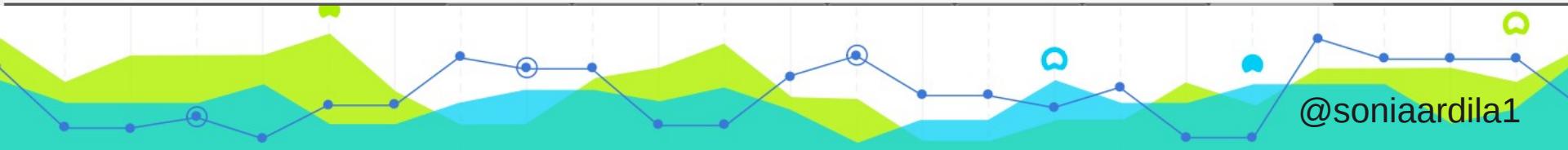
trying URL 'https://cran.rstudio.com/src/contrib/RcppEigen_0.3.3.3.1.tar.gz'
Content type 'application/x-gzip' length 1627022 bytes (1.6 MB)
=====
downloaded 1.6 MB

trying URL 'https://cran.rstudio.com/src/contrib/lme4_1.1-15.tar.gz'
```

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<b>User Library</b>		
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> BH	Boost C++ Header Files	1.65.0-1
<input type="checkbox"/> bindr	Parametrized Active Bindings	0.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
<input checked="" type="checkbox"/> car	Companion to Applied Regression	2.1-6
<input type="checkbox"/> cli	Helpers for Developing Command Line Interfaces	1.0.0
<input type="checkbox"/> colorspace	Color Space Manipulation	1.3-2
<input type="checkbox"/> crayon	Colored Terminal Output	1.3.4
<input type="checkbox"/> dichromat	Color Schemes for Dichromats	2.0-0
<input type="checkbox"/> digest	Create Compact Hash Digests of R Objects	0.6.14
<input type="checkbox"/> dplyr	A Grammar of Data Manipulation	0.7.4
<input checked="" type="checkbox"/> ggplot2	Create Elegant Data Visualisations Using the Grammar of Graphics	2.2.1
<input type="checkbox"/> ggvis	Interactive Grammar of Graphics	0.4.3
<input type="checkbox"/> glue	Interpreted String Literals	1.2.0
<input type="checkbox"/> gtable	Arrange 'Grobs' in Tables	0.2.0

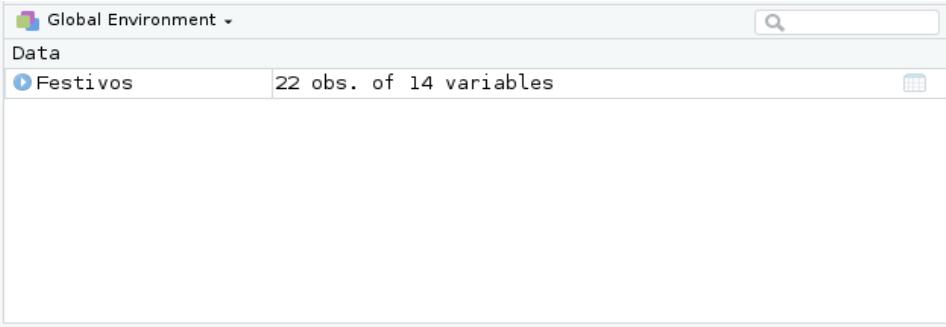


```
1 install.packages("car")
2
3 scatterplotMatrix(~ Dias.Festivos.año +
4 Pib.per.cápita + X..Desempleo, data = Festivos)
```

3:1 (Top Level) R Script

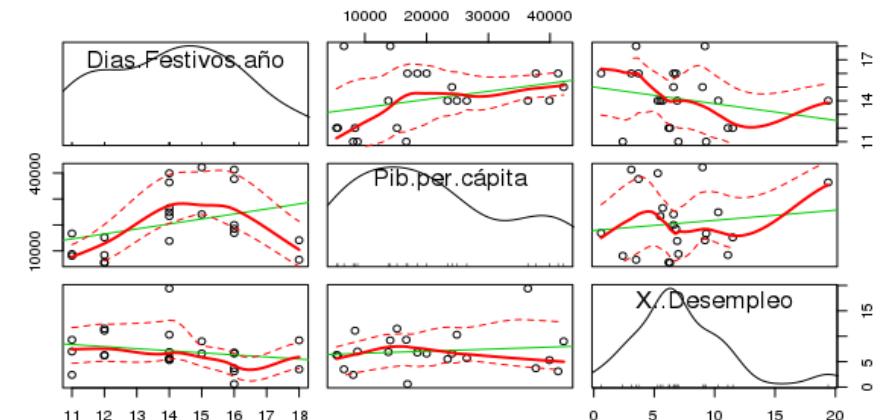
Console Terminal

```
> scatterplotMatrix(~ Dias.Festivos.año +
+ Pib.per.cápita + X..Desempleo, data = Festivos)
>
```

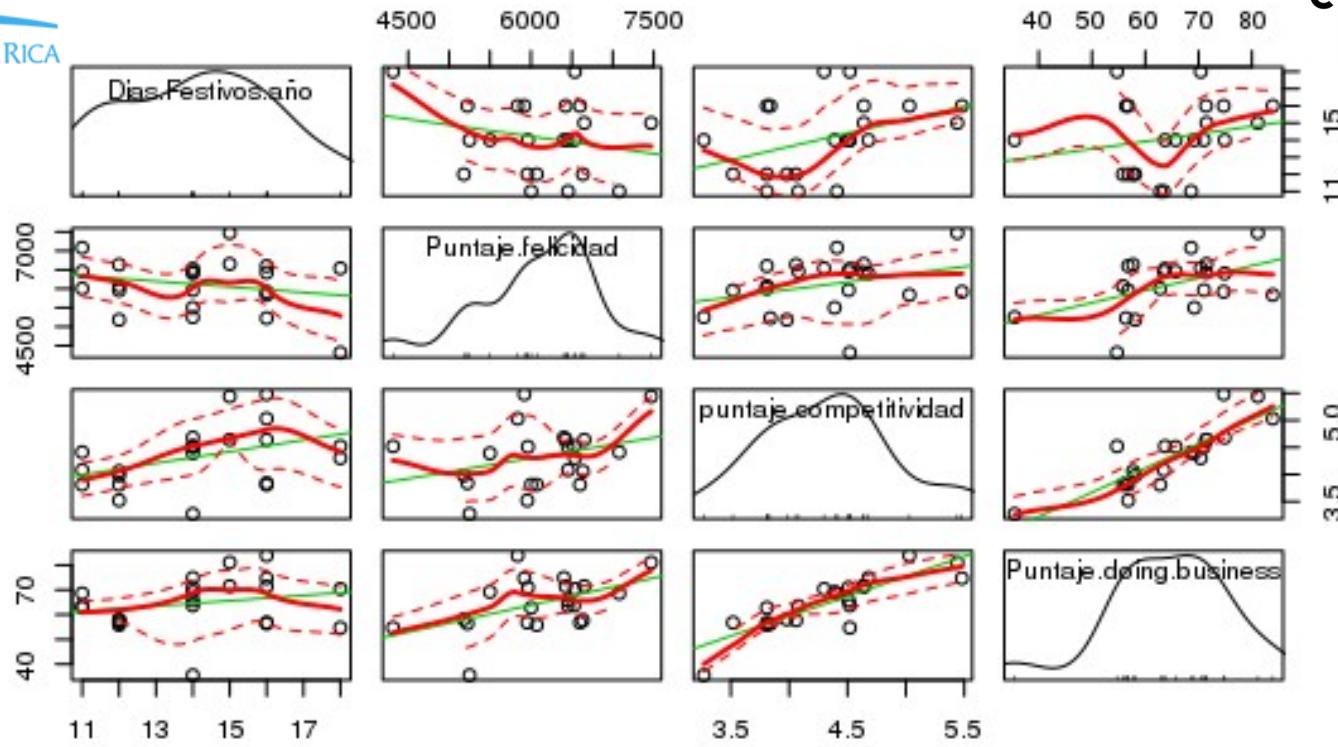


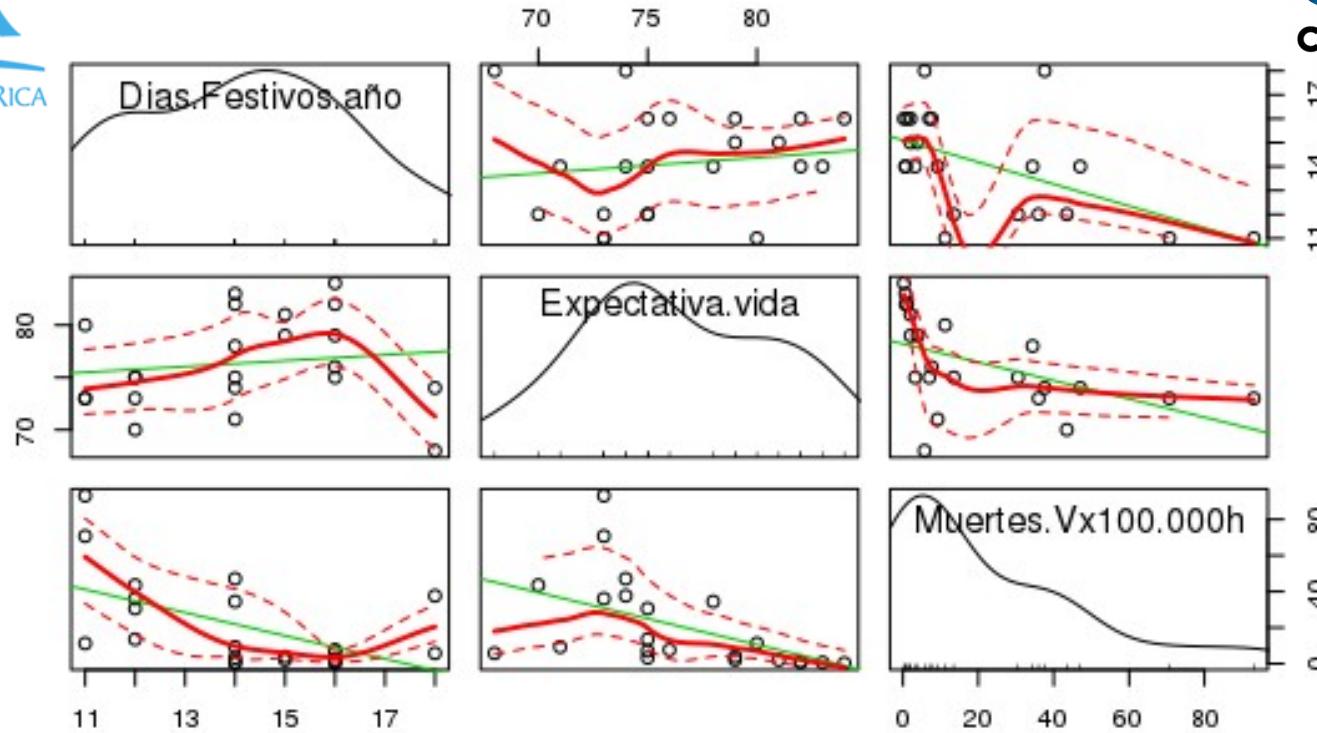
Files Plots Packages Help Viewer

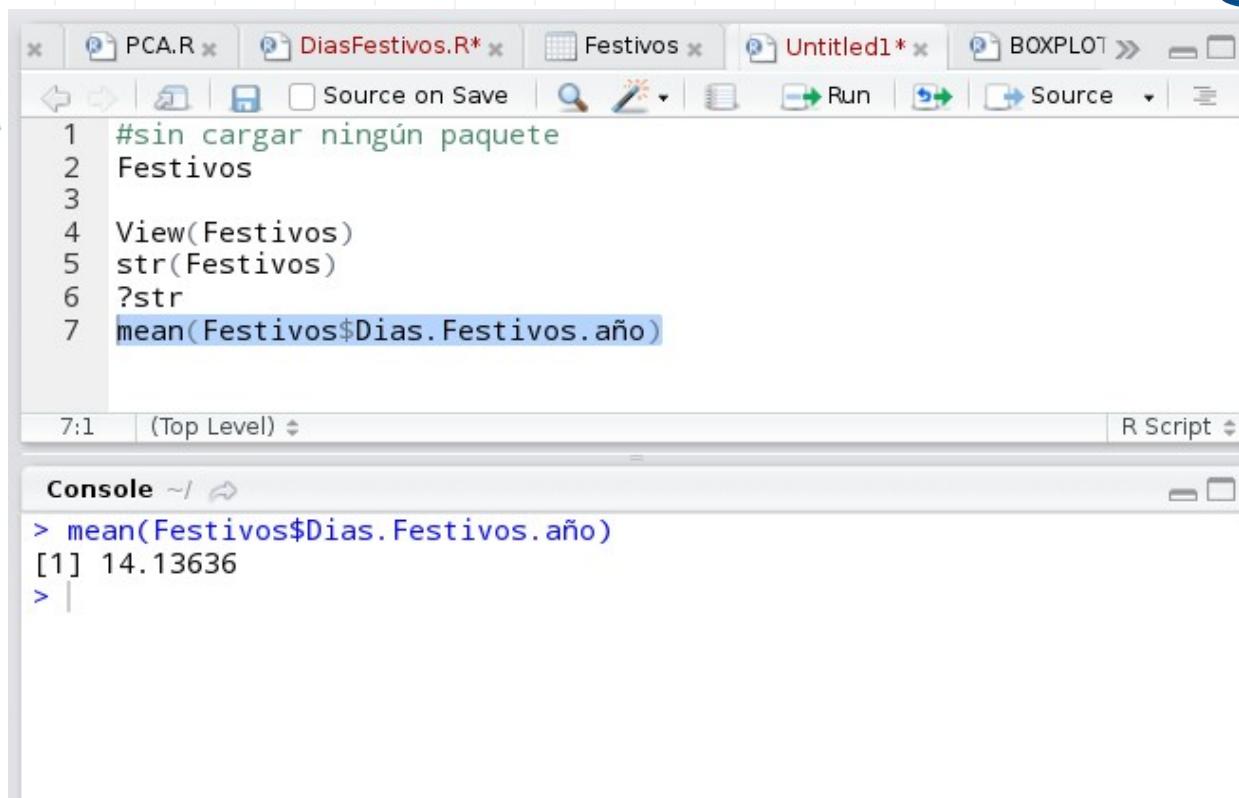
Zoom Export



@soniaardila1







The screenshot shows the RStudio interface. The top panel is a script editor with the following R code:

```
1 #sin cargar ningún paquete
2 Festivos
3
4 View(Festivos)
5 str(Festivos)
6 ?str
7 mean(Festivos$Dias.Festivos.año)
```

The bottom panel is a console window showing the execution of the last line of code:

```
> mean(Festivos$Dias.Festivos.año)
[1] 14.13636
>
```



```
4 View(Festivos)
5 str(Festivos)
6 ?str
7 mean(Festivos$Dias.Festivos.año)
8 mean(Festivos[,3])
9 mean(Festivos$Dias.Festivos.año[Festivos$Región == "Centro America"])
10
11
12
```

9:1 (Top Level) R Script

## Console ~/ ↻

```
> mean(Festivos$Dias.Festivos.año[Festivos$Región == "Centro America"])
[1] 11.85714
> |
```



```
R-RutaN x PCA.R* x DiasFestivos.R* x Festivos x R-UCR.R* x BOXPLOT.R x >> Source on Save Run Source
```

```
5 str(Festivos)
6 ?str
7 mean(Festivos$Dias.Festivos.año)
8 mean(Festivos[,3])
9 mean(Festivos$Dias.Festivos.año[Festivos$Región == "Centro America"])
10
11 #diferencia de cada país frente al promedio de dias festivos
12
13 Festivos$Diff <- Festivos$Pib.per.cápita - mean(Festivos$Pib.per.cápita)
14
15
16
```

13:1 (Top Level) R Script

### Console ~/

```
> mean(Festivos$Dias.Festivos.año[Festivos$Región == "Centro America"])
[1] 11.85714
> Festivos$Diff <- Festivos$Pib.per.cápita - mean(Festivos$Pib.per.cápita)
>
```



```
12  
13 Festivos$Diff <- Festivos$Pib.per.cápita - mean(Festivos$Pib.per.cápita)  
14 print(Festivos)
```

14:1 (Top Level) ▾

Console ~/ ↗

	Puntaje.doing.business	Expectativa.vida	Diff
1	70.4	74	-6559.0909
2	57.7	75	-5459.0909
3	56.8	76	-659.0909
4	71.5	79	3440.9091
5	35.5	74	-6859.0909
6	81.1	81	21540.9091
7	69.2	75	4240.9091
8	71.0	71	5840.9091
9	63.7	82	19240.9091
10	74.9	83	15740.9091
11	71.4	75	-3759.0909
12	56.4	79	-2159.0909
13	83.9	82	17040.9091
14	74.7	84	20640.9091
15	54.7	68	-14059.0909
16	56.8	70	-12259.0909
17	68.6	80	-3959.0909
18	62.8	73	-11859.0909
19	63.5	73	-12559.0909
20	58.1	73	-15259.0909
21	55.8	75	-15059.0909

```

12 Festivos$Diff <- Festivos$Pib.per.cápita - mean(Festivos$Pib.per.cápita)
13 print(Festivos)
14
15 summary(Festivos)
16
17
18

```

16:1 (Top Level) ▾

R Script ▾

Console ~/ ↵

> summary(Festivos)

	Pais	Región	Dias.Festivos.año	Pib.per.cápita	X..Desempleo
Argentina	: 1	Asia :5	Min. :11.00	Min. : 5400	Min. : 0.600
Belice	: 1	Centro America:7	1st Qu.:12.00	1st Qu.:10050	1st Qu.: 5.350
Brasil	: 1	Europa :5	Median :14.00	Median :17700	Median : 6.600
Chile	: 1	Sur America :5	Mean :14.14	Mean :20659	Mean : 7.091
Colombia	: 1		3rd Qu.:16.00	3rd Qu.:26100	3rd Qu.: 9.150
Corea del Sur:	1		Max. :18.00	Max. :42200	Max. :19.400
(Other)	:16				

	Muertes.Vx100.000h	X..Informalidad	Ranking.felicidad	Puntaje.felicidad
Min.	: 0.280	Min. :13.00	Min. : 5.00	Min. :4315
1st Qu.	: 2.330	1st Qu.:28.00	1st Qu.: 27.50	1st Qu.:5858
Median	: 8.415	Median :43.00	Median : 39.50	Median :6237
Mean	:20.954	Mean :45.15	Mean : 46.18	Mean :6112
3rd Qu.	:35.600	3rd Qu.:60.00	3rd Qu.: 54.00	3rd Qu.:6534
Max.	:93.100	Max. :85.00	Max. :122.00	Max. :7469
NA's	:3			

	Ranking.competitividad	puntaje.competitividad	Ranking.doing.business
Min.	: 8.00	Min. :3.270	Min. : 4.00
1st Qu.	: 35.25	1st Qu.:3.875	1st Qu.: 49.50
Median	: 54.50	Median :4.400	Median : 74.50
Mean	: 63.18	Mean :4.330	Mean : 79.23



sonia ardila

Dec 24, 2017 · 7 min read

## Colombia con más días del Sagrado Corazón y menos días del trabajo.



[Read more...](#)



51



Todo empezó al ver que Colombia e India tienen la mayor cantidad de días festivos al año en el mundo: 18 días. Un tweet de la politóloga @sandraborda sobre este tema tratado por la consultora mercer, me puso a pensar:

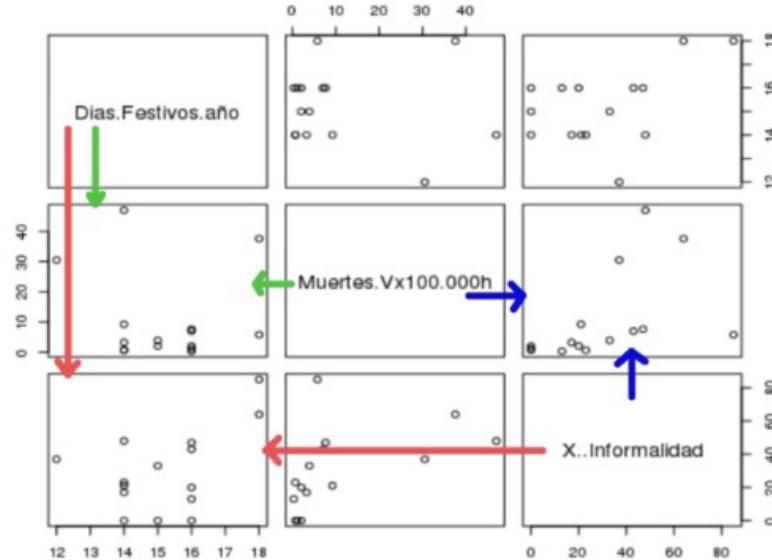


Lo primero que pensé fue que con todo lo que falta por hacer en este país, 18 son demasiados días y tanto tiempo libre no aporta al crecimiento del país. En la tabla observamos los cinco países con más festivos al año en tres continentes.

@soniaardila1

País	Continente	Días Festivos/año
Colombia	America	18
Brasil	America	12
Argentina	America	16
Chile	America	15
Venezuela	America	14
Finlandia	Europa	15
Turquía	Europa	14
Rusia	Europa	14
Malta	Europa	14
España	Europa	14
Tailandia	Asia	16
Líbano	Asia	16
Corea del Sur	Asia	16
Japón	Asia	16
India	Asia	18

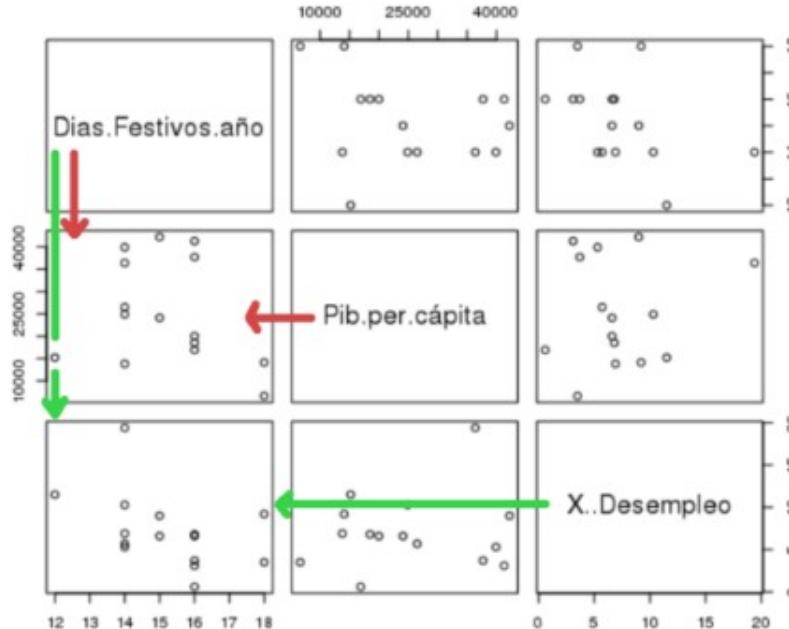
15 países con más días festivos al año en el mundo en tres continentes. 2.017



En la matriz vemos que a más días festivos, mayor trabajo informal pero con una correlación no muy fuerte (cor 0,46). Las flechas señalan el eje en donde se ubica la variable. También se observa que a más trabajo informal, más muertes violentas por cada 100.000 habitantes (cor 0,5) y aquí hay una gran variación ya que según worldlifeexpentancy Colombia tiene 38 muertes violentas al año por cada 100.000 habitantes, Venezuela 47 y países como Finlandia tiene 2, Corea 1.1, Malta 0.7 y Japón 0.28. Gráficamente aparece una correlación positiva mucho más clara (cuadro con flechas azules), lo que lleva a identificar la variable trabajo informal como una variable más nociva que la variable días festivos. Por otro lado al ver la correlación entre los días festivos y la cantidad de muertes violentas, vemos que no hay relación, ya que aunque se genera una correlación negativa (cor—0.10) este valor más cercano al cero no nos lleva a decir que a más días festivos menos muertes violentas.

Sorprendentemente dos variables que pensé que crucificarían completamente a los días festivos para poder decir tranquilamente que 18 días festivos al año no son buenos para el crecimiento del país, no lo hicieron.





Resulta que analizando los datos de estos 15 países, al correlacionar los días festivos con el pib per cápita aparece una tímidamente correlación negativa (-0.25), que no permite decir lo que me encantaría decir y es que a más días festivos menor pib per cápita. Y si vemos la gráfica (cuadro con flechas rojas) observamos que los puntos no parecen descender muy claramente.



Al abordar la variable pib per cápita aparece de forma invisible la variable población y llama la atención que según Ciafactbook, el pib per cápita de Colombia es más del doble que el pib per cápita de India y la diferencia en el tamaño de la población es enorme (48 millones frente a 1.200 millones). Llama la atención porque de estos 15 países, hay unos con mayor población que Colombia y un mayor pib per cápita. Es el caso de Rusia y Brasil. Esto lleva a ver a India con otros ojos mas allá de su tamaño. Además, esta India con un pib per cápita menor al de Colombia es la misma India que Bloomberg destaca como la próxima potencia superando a China.

Bloomberg ▾

Superpower India to Replace China as Growth Engine

# Superpower India to Replace China as Growth Engine

India's economy has youth on its side.

By Michael Heath

September 17, 2017, 5:29 PM GMT-5

From Benchmark

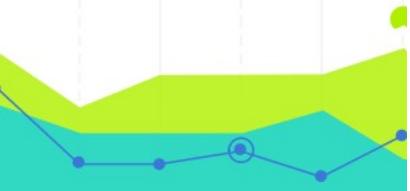


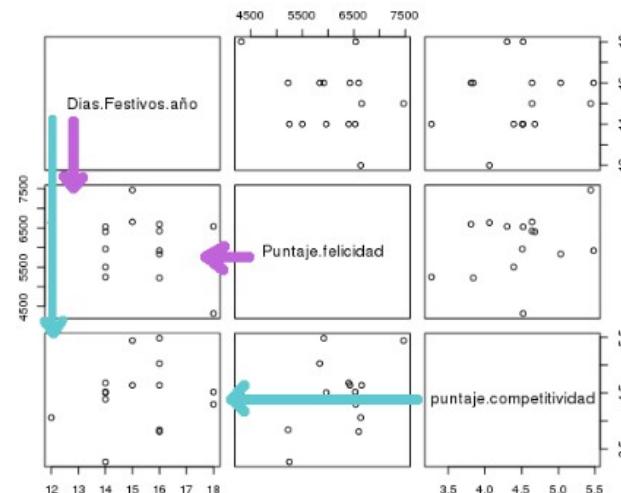
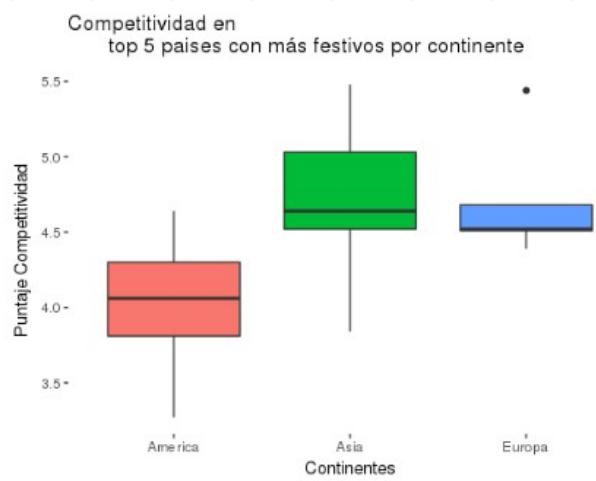
El desempleo no ayudó mucho porque su correlación con los días festivos fue de solo -0.44. Y es que con países como Japón con un desempleo de 3.1% ó Corea del Sur con 3.7% y los dos con 16 días festivos al año, no hay argumentos para atacar a los días festivos. Es más, Japón tenía 15 días festivos al año y desde el 2.016 empezó a celebrar un nuevo festivo: el “día de la montaña” para completar los 16 días.

La cantidad de días festivos al año si puede tener efectos en la economía de un país y en la calidad de vida de sus habitantes pero claramente esta relación no siempre es lineal y no necesariamente porque se trabaje más días al año, la economía de un país será más fuerte. Esto me lleva a pensar en un tema cultural, en la forma de ver la vida. Tal vez por esas diferencias culturales a pesar de tener Japón un gran número de festivos al año, muchos japoneses trabajan muchas horas extras y no se toman los días libres correspondientes, según indica la consultora Mercer y por ese motivo y también para incentivar el consumo, estrenaron un nuevo festivo.

Será que los días festivos aportan a la felicidad de los ciudadanos de estos 15 países? No. La correlación es del - 0.3 (World Happiness Report).

Y antes de ver la relación de competitividad con felicidad y días festivos, es muy evidente que los puntajes en competitividad de los países con más días festivos, son menores en el continente americano que en Asia ó en Europa.





@soniaardila1

Vemos que de estos 15 países, solo dos tienen un pib per cápita inferior al de Colombia (Venezuela y la misteriosa India) y hay cinco países con más del doble del pib per cápita de Colombia, lo que indica que se trata de economías mucho más fuertes y después de ver el caso de Japón y sus habitantes que trabajan horas extra y no se toman sus días libres, se podría pensar que en este y otros países del grupo, hay otras variables que se correlacionarían positivamente con la felicidad pero no se trata de los días festivos.

Colombia se encuentra en el quinto lugar en este grupo de países en cuanto a puntaje de felicidad con 6537 puntos. En primer lugar está Finlandia con 7469 puntos y en último lugar en estos 15 países se encuentra India con 4315 puntos. Japón y Corea también tienen calificaciones inferiores a las de Colombia: 5920 y 5838 respectivamente.

En cuanto a competitividad, aunque la gráfica nos muestra una correlación positiva entre la cantidad de días festivos y el puntaje de competitividad del país según la clasificación del Foro Económico mundial, la correlación es solo del 0.17 y como en este top 15 por días festivos hay países más desarrollados y de otras culturas distintas a la latinoamericana, me inclino a pensar que esa otra forma de ver el mundo es lo que los hace más competitivos y por eso no es necesario que tengan menos días libres para que sus países progresen.

Bloomberg menciona a un país con 18 días festivos al año como próxima potencia mundial y no se trata de Colombia.



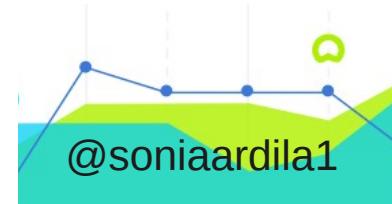
@soniaardila1

Y aunque Colombia supera a India en pib per cápita, en tener menos trabajo informal, en el ranking doing business, en expectativa de vida y en felicidad, India tiene menos desempleo que Colombia (3.5% vs 9.2% según cifras del banco mundial en 2016), tiene solo 6 muertes violentas por cada 100.000 habitantes cuando Colombia tiene 38 y tiene un puntaje de competitividad superior a Colombia (4.5 vs 4.3 y por ranking, India es puesto 39 vs puesto 61 de Colombia). Un país con este alto nivel de competitividad es un país enfocado y no de forma gratuita ha logrado tener su gran Silicon Valley en la ciudad de Bangalore ni de forma gratuita hace parte de los BRIC countries ó BRIC economies, término acuñado por Jim O'Neill en 2.001 y que se refiere a países con altos niveles de desarrollo. Originalmente eran BRICS (Brasil, Rusia, India, China y Sur Africa). Ya ha salido Sur Africa y el tiempo nos dirá si la “B” sobrevive.

Japón con su incremento en días festivos me parece un buen ejemplo para descifrar si 18 días festivos son malos ó buenos para Colombia.

Pareciera que cuando el mundo se ve de diferente manera, la cantidad de días libres al año no afecta negativamente en el crecimiento de un país.

Por eso no sorprende noticias como la muerte de un médico japonés de 105 años que tenía actividades en su agenda para los siguientes dos años.

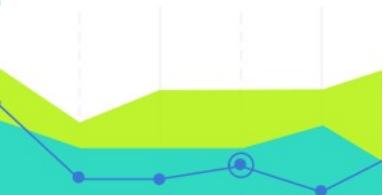


By SAM ROBERTS JULY 25, 2017



Trabajar, producir y crecer pareciera ser parte de la cultura de los países desarrollados de esta lista.

Japón se encuentra entre los países con mayor edad de retiro: 65 años tanto para hombres como para mujeres. Y un japonés jubilado vive en promedio 19 años después de su jubilación según cifras de Bloomberg. Más que la comida ó el ejercicio, el estar activo parece ser el secreto. El secreto para el bienestar de los ciudadanos y de todo el país.



Después de ver los datos y ver que la cantidad de días festivos no inciden en los resultados negativos de las economías y tampoco en los positivos como en la felicidad, creo que la forma de ver el mundo de estos 15 países puede justificar sus resultados. Y esa forma de ver el mundo tiene que ver con las diferentes culturas y es probable que también con la religión.

Así que como Colombia no ha hecho la tarea y aún no somos lo suficientemente competitivos, ayudaría tener menos tiempo libre. Pero además de reducir los días festivos (Méjico tiene solo siete), es urgente un cambio de mentalidad para que los colombianos piensen en grande y se esfuerzen más, aprendan más y produzcan más todos los días del año.

Sonia Ardila

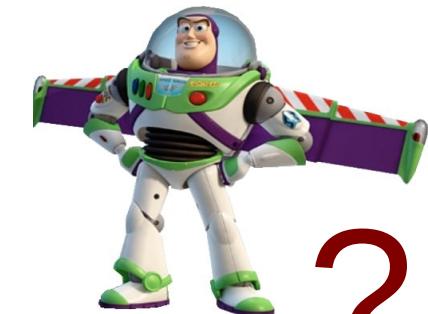
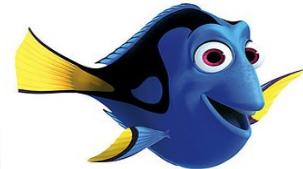
@soniaardila1



@soniaardila1



# Data Storytelling



?





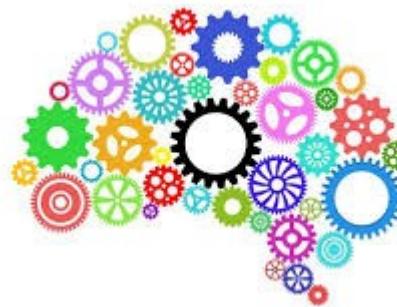
# Data Storytelling



@soniaardila1



# Data Storytelling



SOFT  
SKILLS

## Publicación de la mejor historia!



INICIO    EVENTOS    NOSOTROS    SERVICIOS    PORTAFOLIO    DATA STORYTELLING LAB

### DATA STORYTELLING LAB

#### BIENVENIDOS AL DATA STORYTELLING LAB !



"La gente lee estadísticas pero sienten historias". Brent Dykes. Director de Data Strategy en Domo.



Sonia Ardila

18/10/2017



#### LA ESTRUCTURA DE LA HISTORIA !

Explorando los datos para planear nuestra historia!



Sonia Ardila

28/10/2017



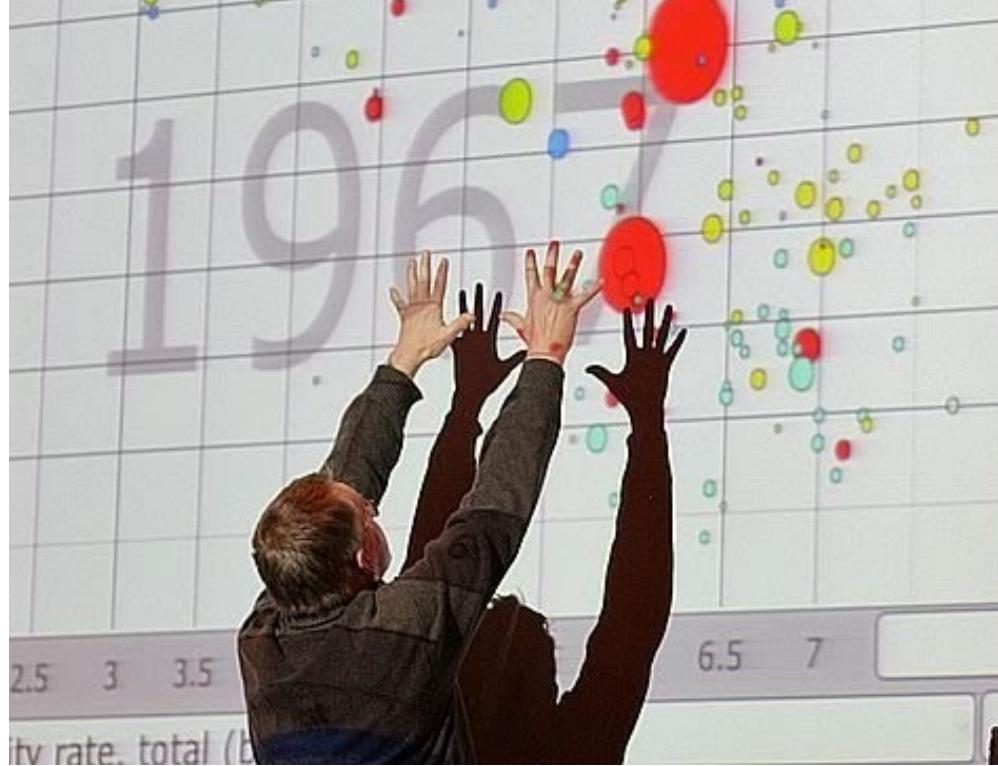
#### BUSCANDO AL PROTAGONISTA CON EDA !

Análisis exploratorio de datos para...

Protagonistas ?

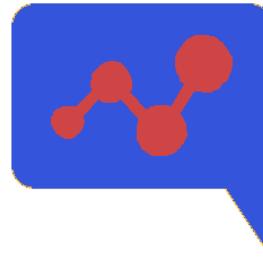
Conflictos ?

Solución  
al conflicto ?



[https://youtu.be/OT9poH\\_D2lw](https://youtu.be/OT9poH_D2lw)





# Data Storytelling

# Gracias !

Sonia Ardila  
[sap@datelligence.com](mailto:sap@datelligence.com)  
[@soniaardila1](https://twitter.com/soniaardila1)

