

TALLER DE DATOS ABIERTOS TRABAJANDO DATOS JUDICIALES

Lic. Germán Tessmer

Sub-Director

Observatorio Económico Social | UNR

Lic. Patricio Almeida Gentile

Analista

Observatorio Económico Social | UNR



El experimento de Linda

Linda tiene 31 años, es soltera, no tiene pelos en la lengua y es muy inteligente. Se licenció en Filosofía. Cuando estudiaba, le preocupaban mucho los problemas de discriminación y justicia social, y también participó en manifestaciones antinucleares

¿A que categoría pertenece?

1. Linda trabaja de cajera en un banco
2. Linda es profesora de una escuela primaria
3. Linda es vendedora de seguros
4. Linda trabaja de cajera en un banco y es un miembro activo del movimiento feminista



¿Por qué preocuparse por la limpieza de datos?

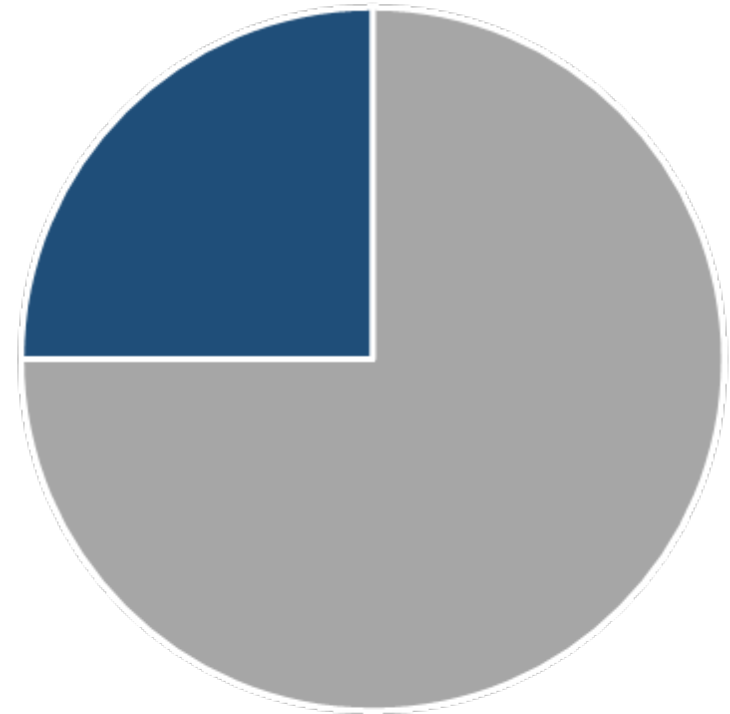


Recolección

Limpieza

Análisis

Reporte



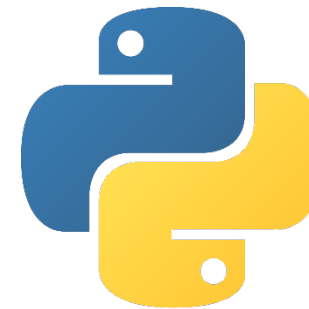
■ Limpieza de Datos

■ Todo lo demás



- Entender la estructura
- Mirar los datos
- Visualizarlos

¿Cómo?





R es un entorno y lenguaje de programación con un enfoque al análisis estadístico.

- Es conocido por especialistas.
- Su uso requiere aprender a programar.
- Brinda una solución práctica a grandes conjuntos de datos.
- Es gratuito, tanto el núcleo como las librerías de paquetes.
- Brinda mayor velocidad de procesamiento



Excel es una aplicación de hojas de cálculo utilizada en tareas financieras y contables.

- Es más conocido y usado.
- Es de uso intuitivo al inicio.
- Brinda una solución práctica a pocos conjuntos de datos.
- Es pago, tanto el núcleo del programa, como la mayoría de los plugins.
- Se ralentiza con el mayor uso de datos y gráficos.



Principios del formato "tidy"

Nombre	Edad	Color de Ojos	Altura
Juan	34	Otros	1,78
Alicia	55	Azul	1,64
Tomás	76	Marrón	1,56
Cleto	19	Otros	1,83

Variable

Observación

Hay 3 reglas que se deben cumplir para considerar que un *data frame* este ordenado:

- Cada variable debe tener su propia columna
- Cada observación debe tener su propia fila
- Cada valor deber tener su propia celda



Síntomas (I): Los encabezados de columnas son valores, no variables

Nombre	Edad	Marrón	Azul	Otros	Altura
Juan	34	0	0	1	1,78
Alicia	55	0	1	0	1,64
Tomás	76	1	0	0	1,56
Cleto	19	0	0	1	1,83

Nombre	Edad	Color de Ojos	Altura
Juan	34	Otros	1,78
Alicia	55	Azul	1,64
Tomás	76	Marrón	1,56
Cleto	19	Otros	1,83



Síntomas (II): Las variables están guardadas tanto en filas como en columnas

Nombre	Medidas	Cantidad
Juan	nro_perros	2
Juan	nro_gatos	5
Juan	nro_aves	0
Cleto	nro_perros	3
Cleto	nro_gatos	0
Cleto	nro_aves	1

Nombre	nro_perros	nro_gatos	nro_aves
Juan	2	5	0
Cleto	3	0	1



Síntomas (III): Múltiples variables guardadas en la misma columna

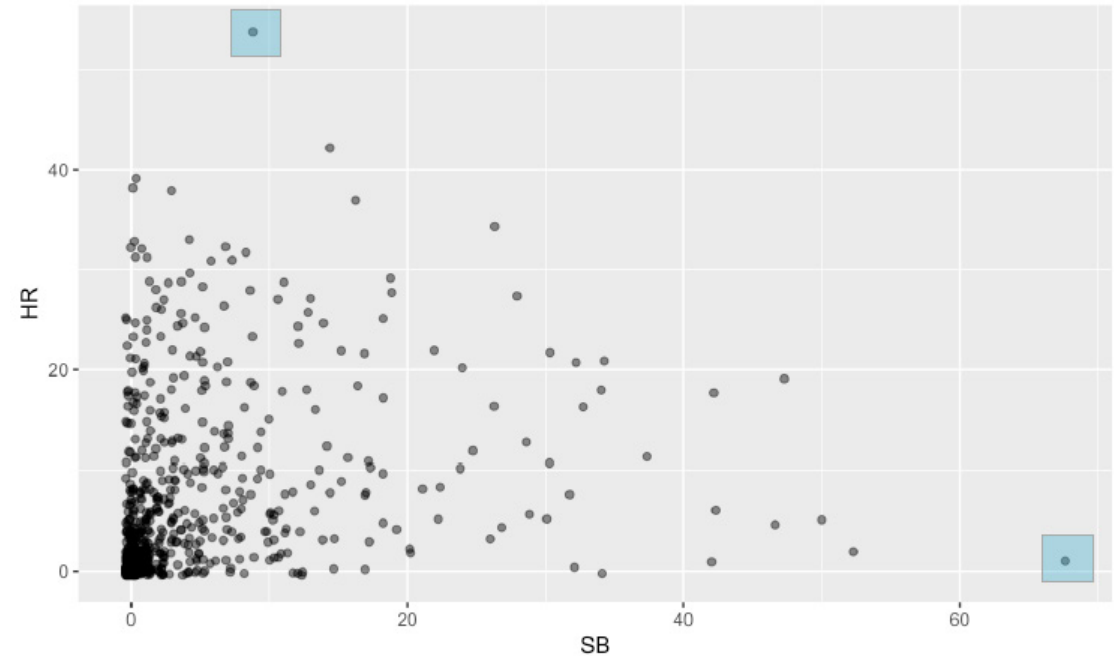
Nombre	Sexo_Edad	Color de Ojos	Altura
Juan	M_34	Otros	1,78
Alicia	F_55	Azul	1,64
Tomás	M_76	Marrón	1,56
Cleto	M_19	Otros	1,83

Nombre	Sexo	Edad	Color de Ojos	Altura
Juan	Masculino	34	Otros	1,78
Alicia	Femenino	55	Azul	1,64
Tomás	Masculino	76	Marrón	1,56
Cleto	Masculino	19	Otros	1,83



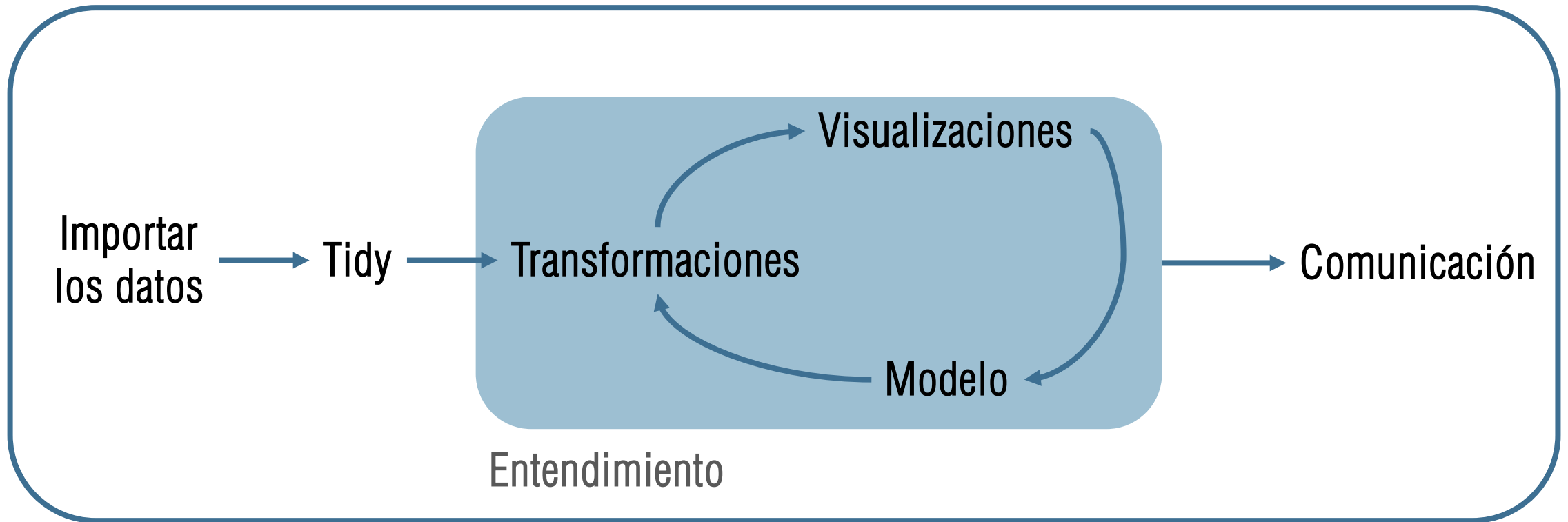
Outliers (fuera de serie)

- Valores extremos que se encuentran distantes de otros valores
- Varias causas:
 - Validez de las mediciones
 - Variabilidad en la medición
 - Errores en experimentos
 - Errores en el ingreso de la información
- Deben ser descartados o retenidos, dependiendo de la causa



- Pueden aparecer en múltiples formas
 - Valores tan extremos que no pueden ser plausibles (persona con edad de 348 años)
 - Valores que no tienen sentido (persona con edad negativa)
- Múltiples causas:
 - Errores de medición
 - Errores de ingreso de información
 - Codificación especial (por ejemplo, que 99 signifique no sabe / no contesta)
- Generalmente deben ser reemplazados o removidos.





Programa



The image shows a browser window displaying the website <https://www.santafe.gov.ar/datosabiertos/>. The browser's address bar and menu are visible. The website header includes the Santa Fe logo and the text "GOBIERNO ABIERTO". Below the header, there are navigation tabs: "TRANSPARENCIA", "PARTICIPACIÓN", "COLABORACIÓN", and "DATOS". The main content area shows a data table on a tablet screen. The table has columns for "Puesto", "Fecha", "Especie", "Cantidad", "Kg", and "Precio". The text "DATOS ABIERTOS" is overlaid in large white letters on the tablet screen. Social media icons for Twitter, Facebook, and YouTube are visible in the top right corner of the website.

SANTA FE
GOBIERNO ABIERTO

TRANSPARENCIA PARTICIPACIÓN COLABORACIÓN DATOS

DATOS ABIERTOS

Ponemos a disposición de los ciudadanos datos vinculados a la administración pública, libres de barreras tecnológicas, legales o económicas para su reutilización. En esta sección podés acceder a información significativa sobre la provincia y la gestión pública, mediante una visualización online o descargando los datos.





[Inicio](#) / [meta-ogp](#) / Dataset

Temas ^

- Administración pública (2)
- Salud (1)

Tipo de contenido ^

- Dataset** x

Etiquetas (Tags) ^

- meta-ogp** x
- justicia (14)
- penal (5)

17 resultados

Buscar

Ordenar por

Fecha de modificación v

En orden

Descendente v

Filtrar

Reiniciar



Actividad judicial civil

Justicia y Seguridad

Estadísticas sobre sentencias de primera y segunda instancia del poder judicial provincial. Años 2016 y 2017.

4x csv



Archivo Editar Ver Historial Marcadores Herramientas Ayuda

Datos Abiertos | Gobierno Abie X Área de cobertura | Datos Abier X Datos Abiertos | Gobierno Abie X Datos Abiertos | Gobierno Abie X Datos Abiertos | Gobierno Abie X +

https://www.santafe.gov.ar/datosabiertos/dataset/actividad Buscar

SANTAFE TRANSPARENCIA PARTICIPACIÓN COLABORACIÓN DATOS

Estadísticas sobre sentencias de primera y segunda instancia del poder judicial provincial.

Años 2016 y 2017.

Otros accesos

La información de esta página (metadata del dataset) también está disponibles en estos formatos.

[JSON](#) [RDF](#)

a través de la [API de DKAN](#).

Compartir

[Twitter](#)

[LinkedIn](#)

[Reddit](#)

Datos y Recursos



Causas ingresadas al fuero civil

[Previsualizar](#)

[Descargar](#)



Resoluciones del fuero civil

[Previsualizar](#)

[Descargar](#)



Sentencias de 1ra instancia - civil

[Previsualizar](#)

[Descargar](#)



Sentencias de 2da instancia - civil

[Previsualizar](#)

[Descargar](#)



Causa: es una demanda que se hace en busca de lograr reconocimiento de derechos

Resolución: es la solución que se encuentra al problema

Sentencia: es la resolución judicial que pone fin a una causa en la que se da respuesta a todo lo solicitado por las partes.

Sentencia de primera instancia: es la dictada por un juez o tribunal que pone fin a un caso específico.

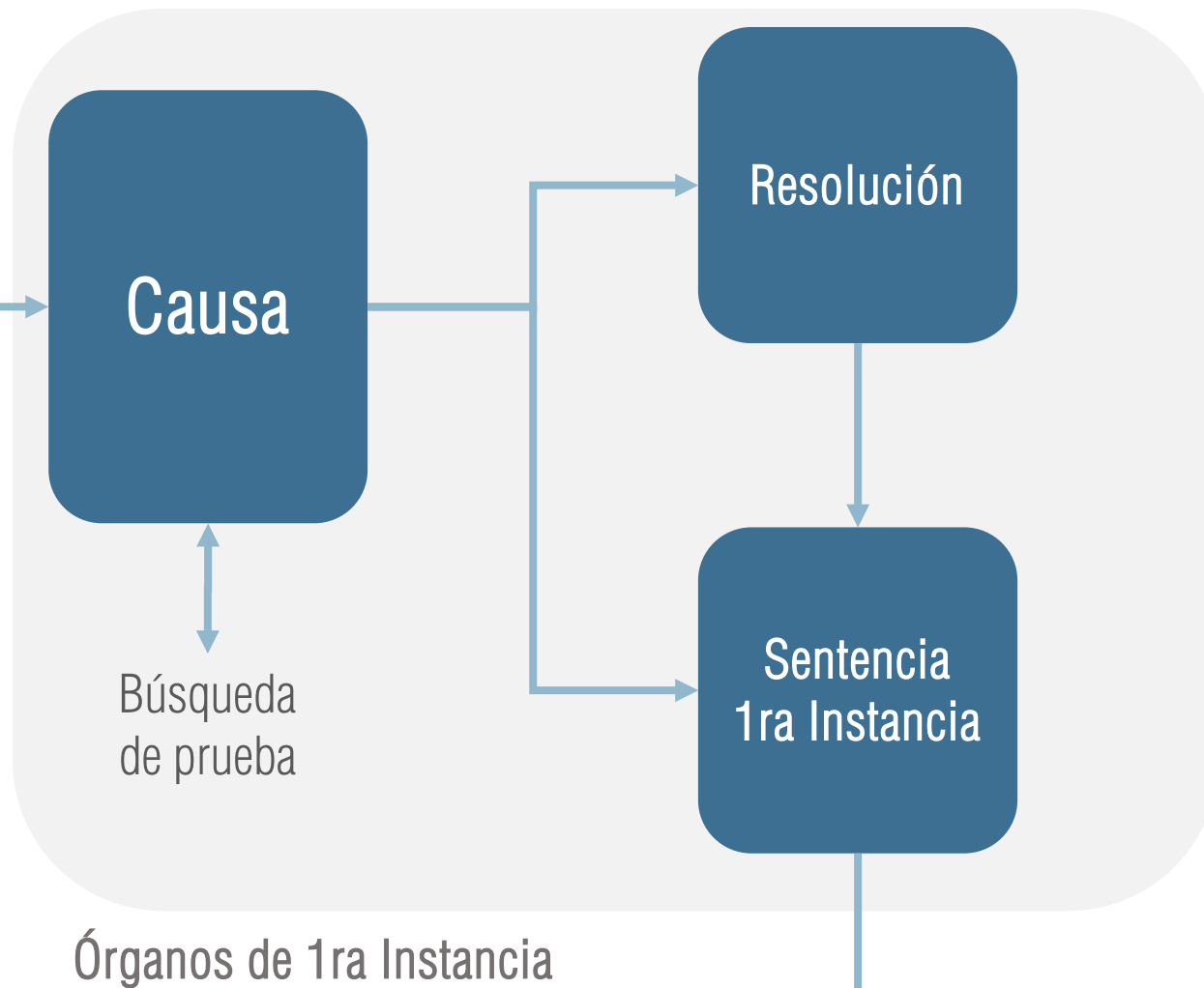
Sentencia de segunda instancia: la dictan los órganos jurisdiccionales superiores al de primer instancia.



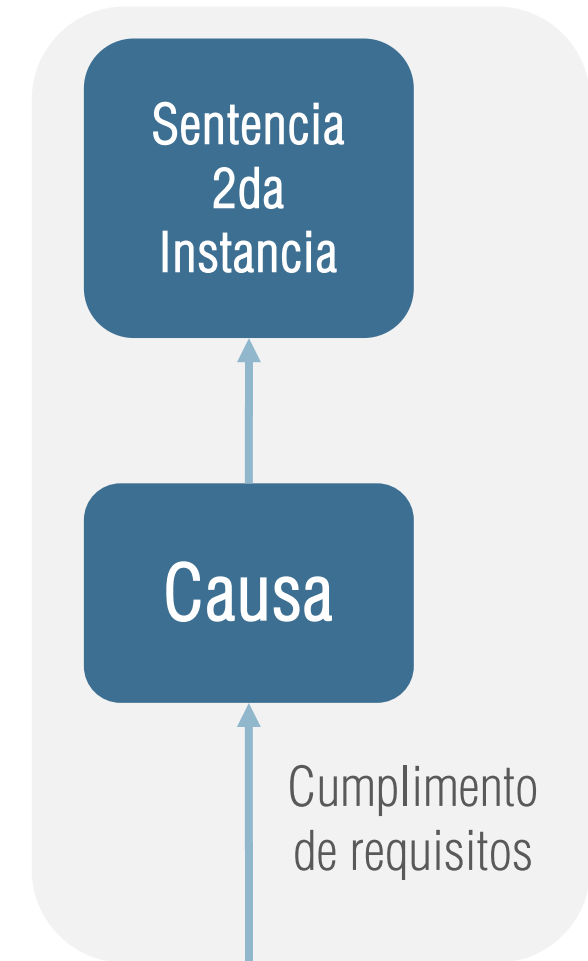
Procedimiento judicial en lo civil: procedimiento

Demanda

Si no es desestimada



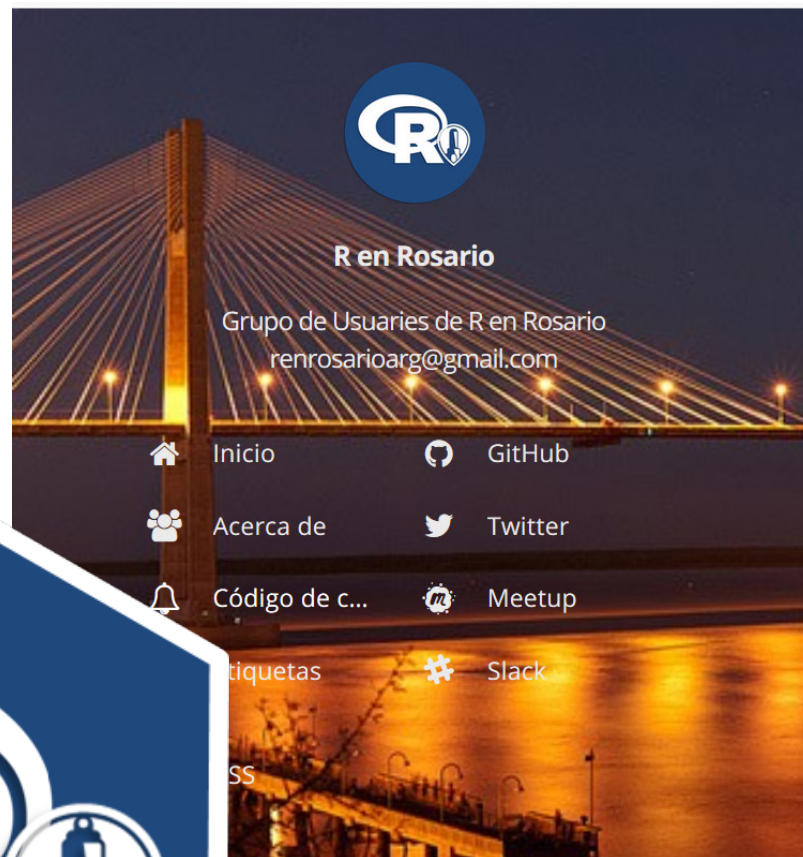
Órganos de 1ra Instancia



Órganos de 2da Instancia

Apelación





Taller sobre Visualización con ggplot2

8 Marzo 2019 en meetup

Celebrando Día Internacional de la Mujer y su rol en la ciencia de datos, tuvimos el primer encuentro del año y nos dedicamos a aprender cómo realizar gráficos con ggplot2.

[Seguir leyendo](#)



Segundo meetup: Trabajo reproducible en R

6 Diciembre 2018 en meetup

Nico Ballarini nos visitó y enseñó a usar R para que nuestro trabajo sea reproducible.

[Seguir leyendo](#)

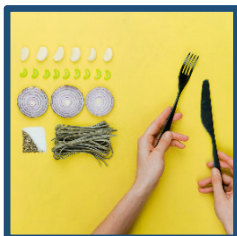
<https://renrosario.rbind.io/>



[@renrosarioarg](https://twitter.com/renrosarioarg)



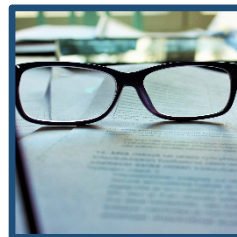
PARA SABER MÁS...



6 ideas para entender cómo piensa un economista



Políticas públicas, ¿cómo saber que funciona (y que no)?



No todo es un número



La educación en Santa Fe y Rosario: resultados Aprender 2016



Gnomos, calzoncillos y... correlación





¡GRACIAS!



german.tessmer@unr.edu.ar



@Gerkandar



@OesUnr



Rosario=

