

# Visual Analytics

Communicating data-driven insights  
through data visualization techniques  
and useful dashboards

Mikel Madina &

Miren Berasategi

[miren.berasategi@deusto.es](mailto:miren.berasategi@deusto.es)



# 0. Introduction



# 0.1 Key points

- **Data driven:** as seen in Onieva's and Lorenzo's lectures
- **Insights:** que usen las características gráficas
- **Data visualization techniques:** para obtener las los insights
- **Dashboards:** as *situattion awareness* tools

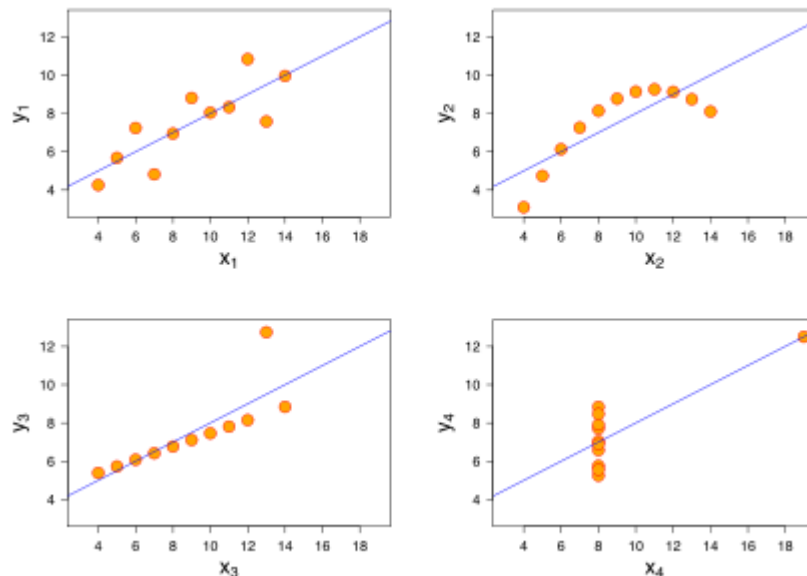
Tableau Desktop para practicar



# 0.2 Why use visualizations

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



Property	Value	Accuracy
Mean of x	9	exact
Sample variance of x	11	exact
Mean of y	7.50	to 2 decimal places
Sample variance of y	4.125	plus/minus 0.003
Correlation between x and y	0.816	to 3 decimal places
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively

## Anscombe's Quartet

### Speaker notes

Baina honek ez du esan nahi teknika estatistikoak alde batera utzi behar direnik; batak besteari lagundu behar diote (estatistika tradizionalak bisualizazioari eta alderantziz)





# 0.2 Why use visualizations



## Speaker notes

Honek erantzungo lioke “insights” parteari:

- *Targets* lehenbizi (benetako insighten oinarriak)
- Ondoren *actions*
  - Search (ze target eta non)
  - Query (identifikatu target, baldin eta badago)
  - Azkenik analyze, komunikatu eta datu/modelo berriak sortu



# Tableau

(oinarrizko ezaugarriak) data.xls

Fake data for online marketing tools and Goals



# Tableau 0.1

Tableau - Libro1

Archivo Datos Servidor Ventana Ayuda

Conexiones [Añadir](#)

datos  
Microsoft Excel

Hojas

Hoja1

Nueva unión

Hoja1 (datos)

Conexión  
☒ En tiempo real ☐ Extraer

Filtros  
0 | [Añadir](#)

Ordenar campos Orden de fuente de datos ☐ Mostrar alias ☐ Mostrar campos ocultos 24 filas

Fuente	Trimestre	Gastado	Visitas	Ganado	Objetivo
Adwords	20160101	1.000	50.000	900	1.500
Twitter	20160101	200	8.500	1.300	1.000
Facebook	20160101	500	20.000	800	1.500
Adwords	20160401	1.000	48.000	1.200	1.500
Twitter	20160401	300	9.000	1.400	1.000
Facebook	20160401	750	21.500	1.400	1.500
Adwords	20160701	1.000	50.000	1.500	1.500
Twitter	20160701	400	10.000	1.000	1.000
Facebook	20160701	750	23.000	200	1.500

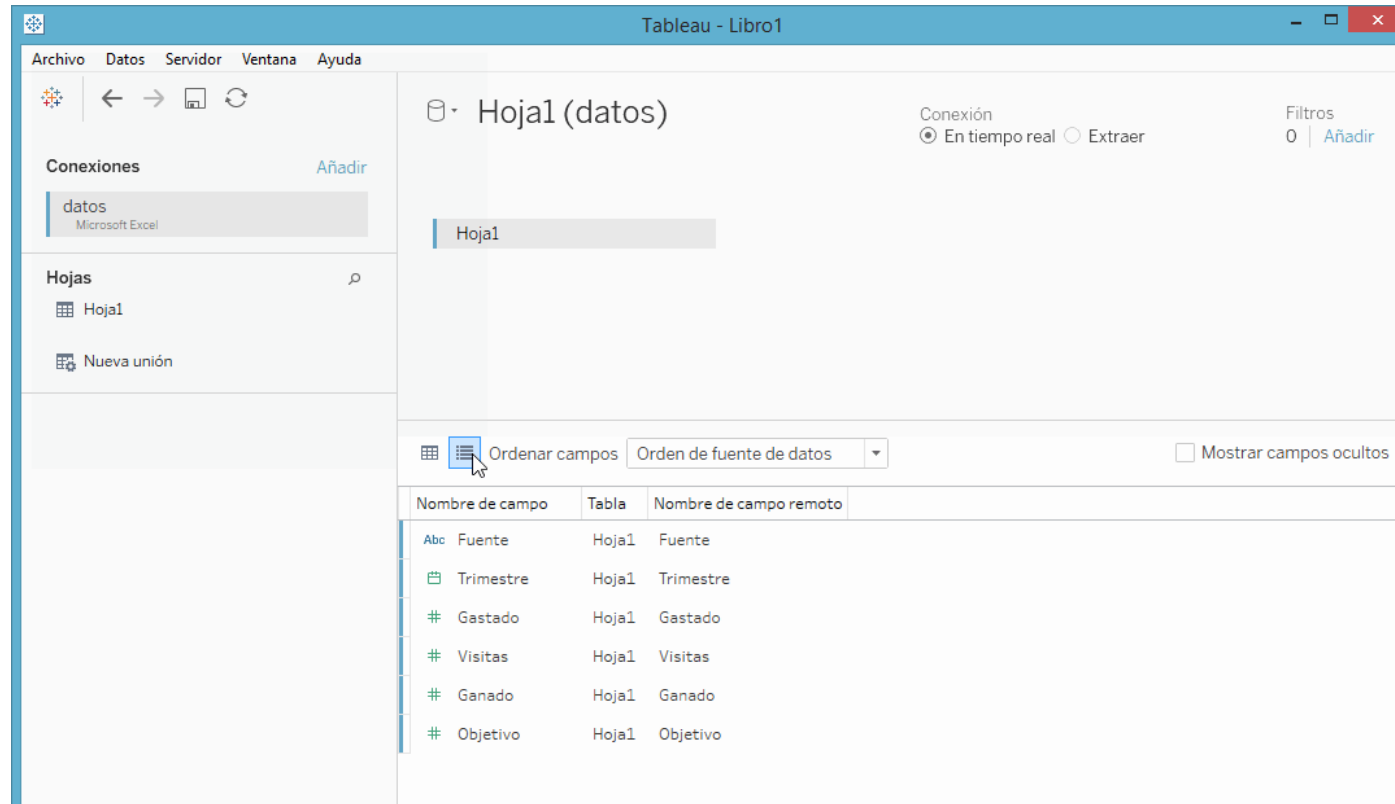
Fuente de datos Hoja 1

Ir a la hoja de trabajo

Load data



# Tableau 0.2



## Speaker notes

- datuak tableaura kargatu
- “data” bistan variableak pixkat aztertu
  - sinbolotxoa eta kolorea, tableauk datua nola interpretatu duen jakiteko
  - ze variable dauden, gero hortik datu kalkulatuak ateratzeko, adibidez irabazitakoa / gastatutakoa oinarritzko KPI modura





# 1. Graphics



# 1.1 Reminder: variable types

- Quantitative
  - Discrete
  - Continuous
- Qualitative
  - Ordinal
  - Nominal



# 1.1 Reminder: variable types

## A question of time

Spatial and time/hour variables are special variable types.

**Time variables** are specially complex:

- are there 365 days in every year? 30 days in every month? 24 hours in every day?
- *timezones* make it even more complex to use hours or time of day



# 1.1 Reminder: variable types

## A question of time

Time may be used as a continuous or as a qualitative variable.

- as a qualitative variable, it has a hierarchy: year > month > (week >) day > hour > minute
- but different hierarchies may be necessary: bimonthly publications, multiple work shifts in a day...





# 1.2 Mapping variables to graphics

Qualitative <b>Nominal</b>	Qualitative <b>Ordinal</b>	Quantitative <b>Interval, Ratio</b>
Position	Position	Position
Colour (Hue)	Pattern (Density)	Size (Length)
Pattern (Texture)	Colour (Lightness)	Angle/Slope
Connection/Edge	Colour (Hue)	Size (Area)
Containment	Pattern (Texture)	Size (Volume)
Pattern (Density)	Connection/Edge	Pattern (Density)
Colour (Lightness)	Containment	Colour (Lightness)
Symbol/Shape	Size (Length)	Colour (Hue)
Size (Length)	Angle/Slope	Pattern (Texture)
Angle/Slope	Size (Area)	Connection/Edge
Size (Area)	Size (Volume)	Containment
Size (Volume)	Symbol/Shape	Symbol/Shape

Speaker notes

Grafikoagoa egitea komeniko litzateke



# 1.3 Graphs typology

Variable types and insights (Munzner's “targets”) as *ingredients*

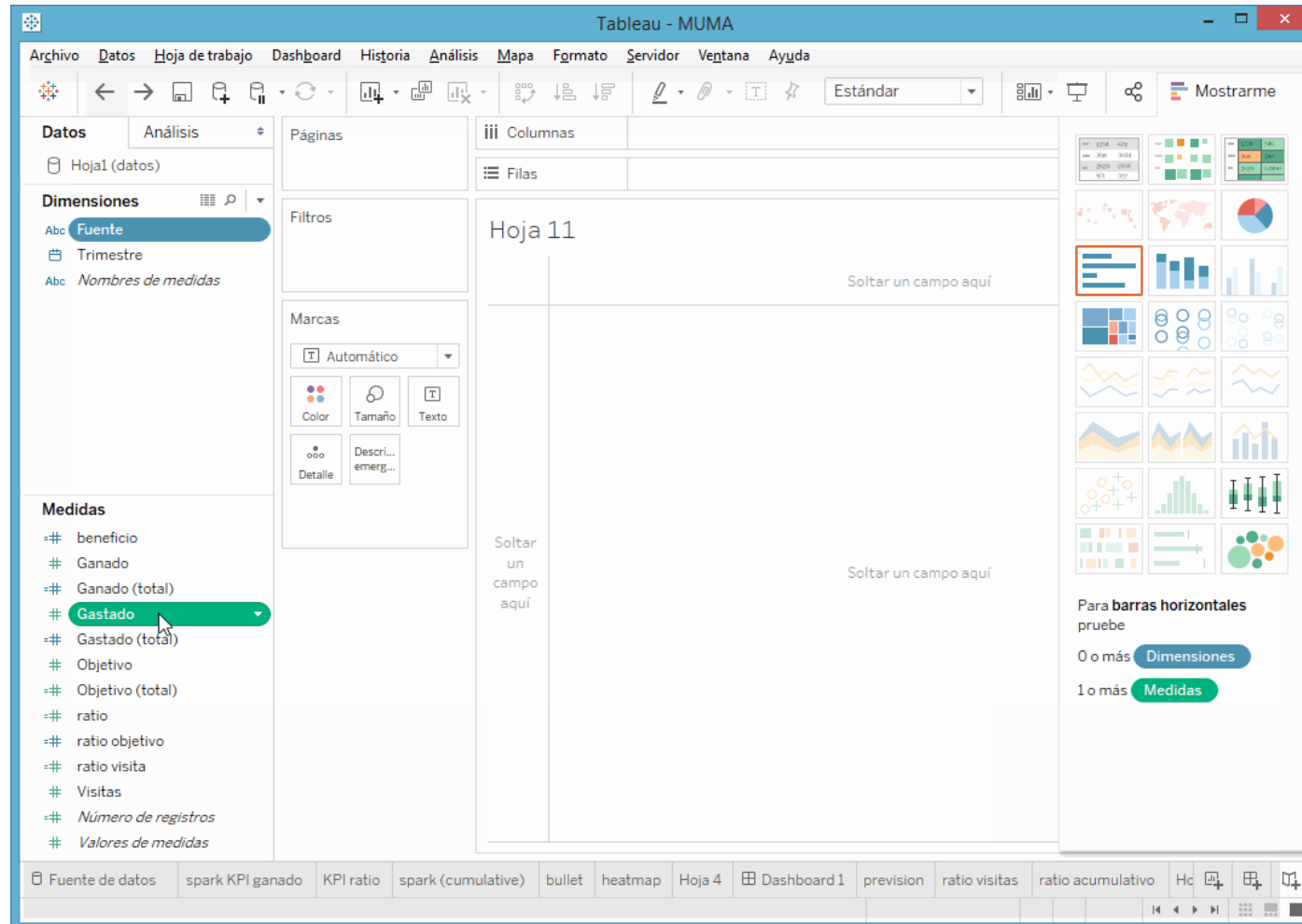
The Data Visualisation Catalogue

Speaker notes

Si eso Kirken taula sartu, baina taula sortzeko kodea nahiko lioso da



# Tableau: Show me



Show me

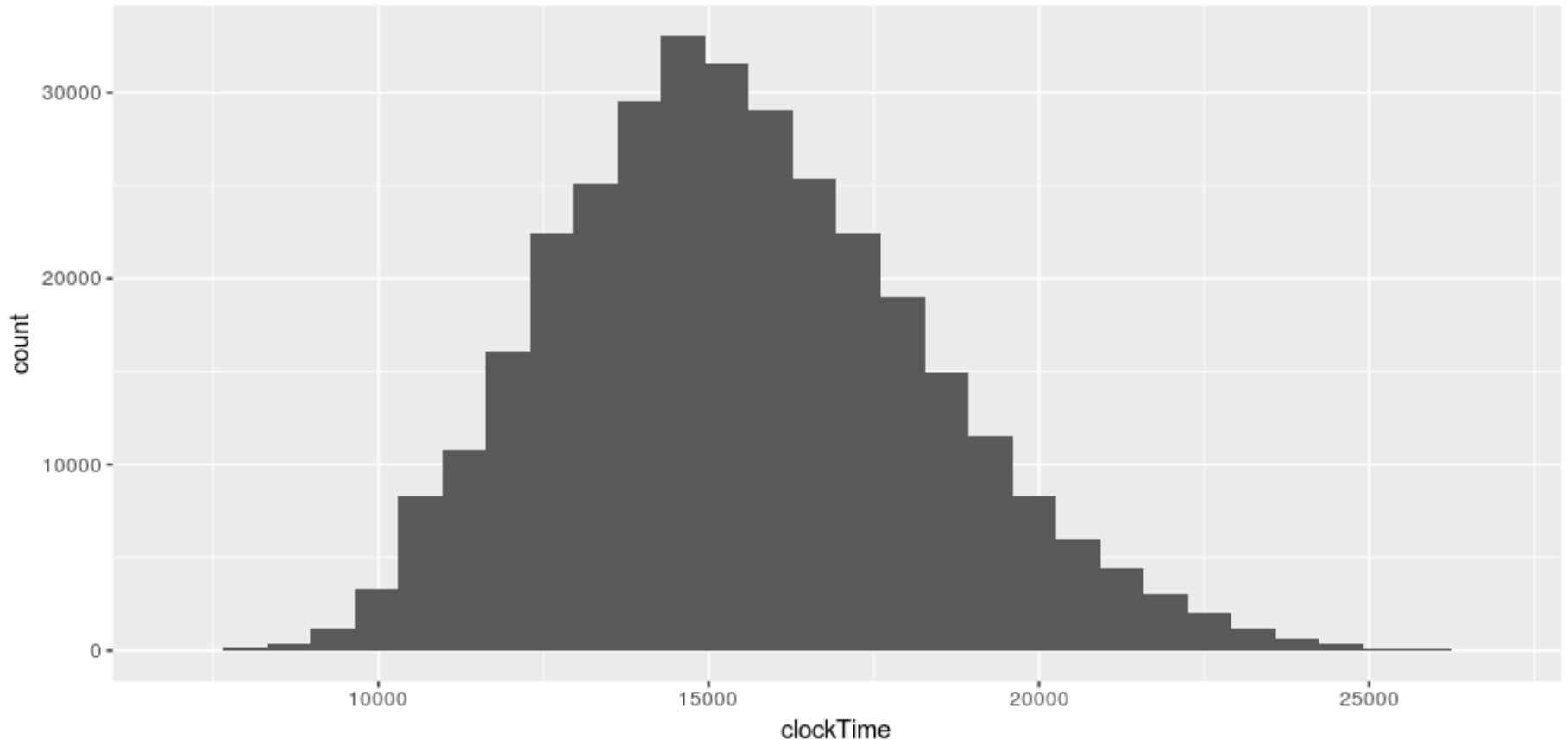


## 2. Provide easier analysis





## 2.1 Change default settings

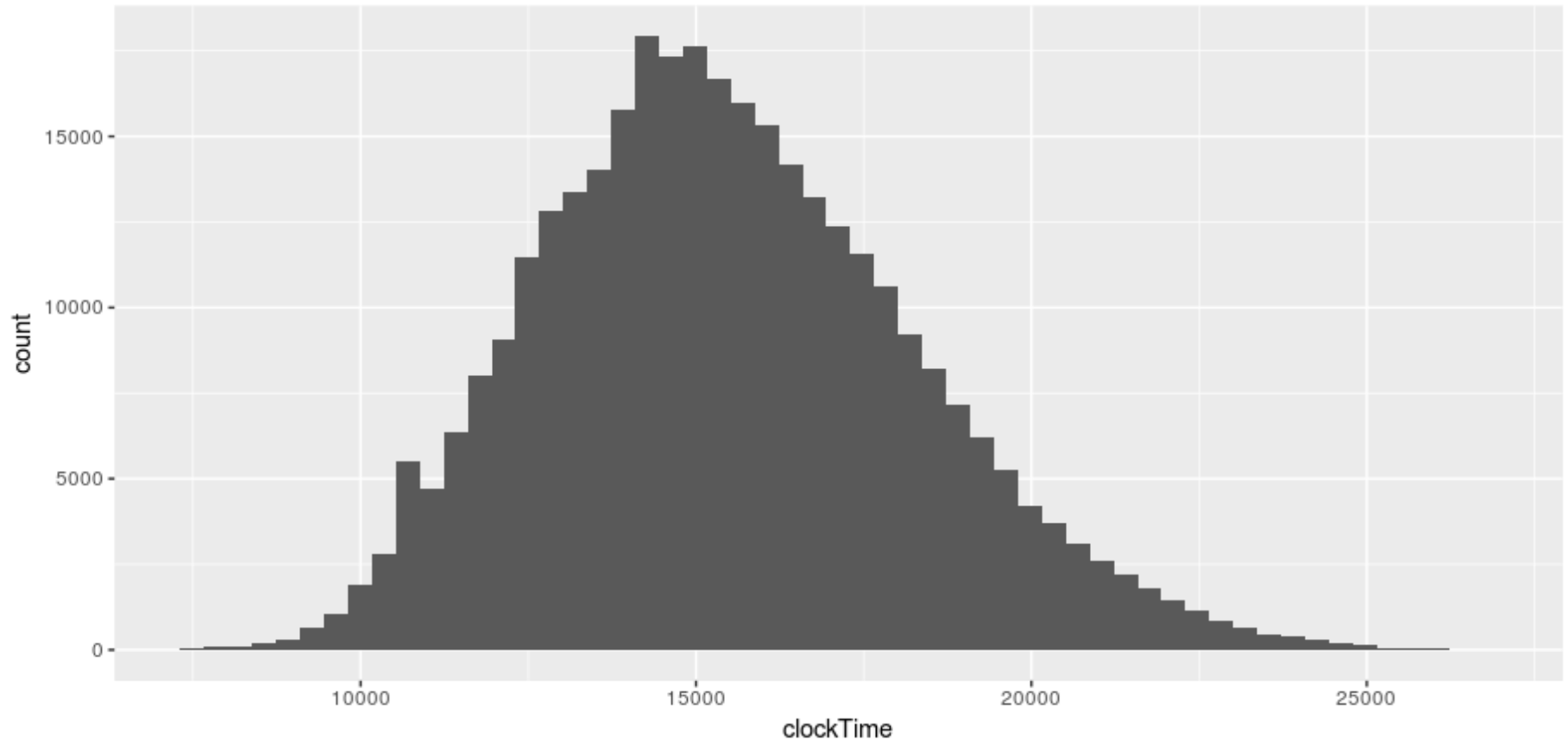


Speaker notes

Data: Berlin Marathon

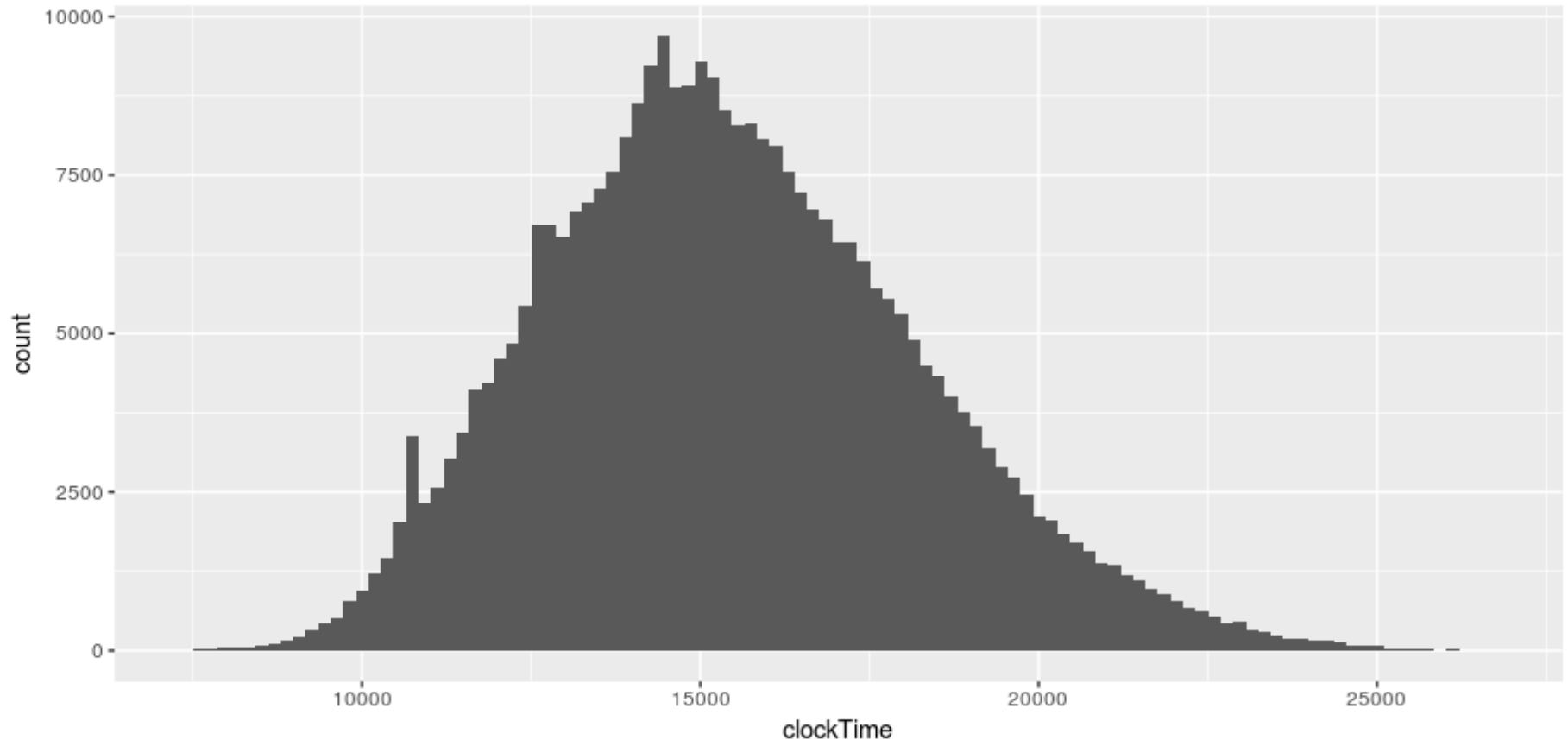


## 2.1 Change default settings



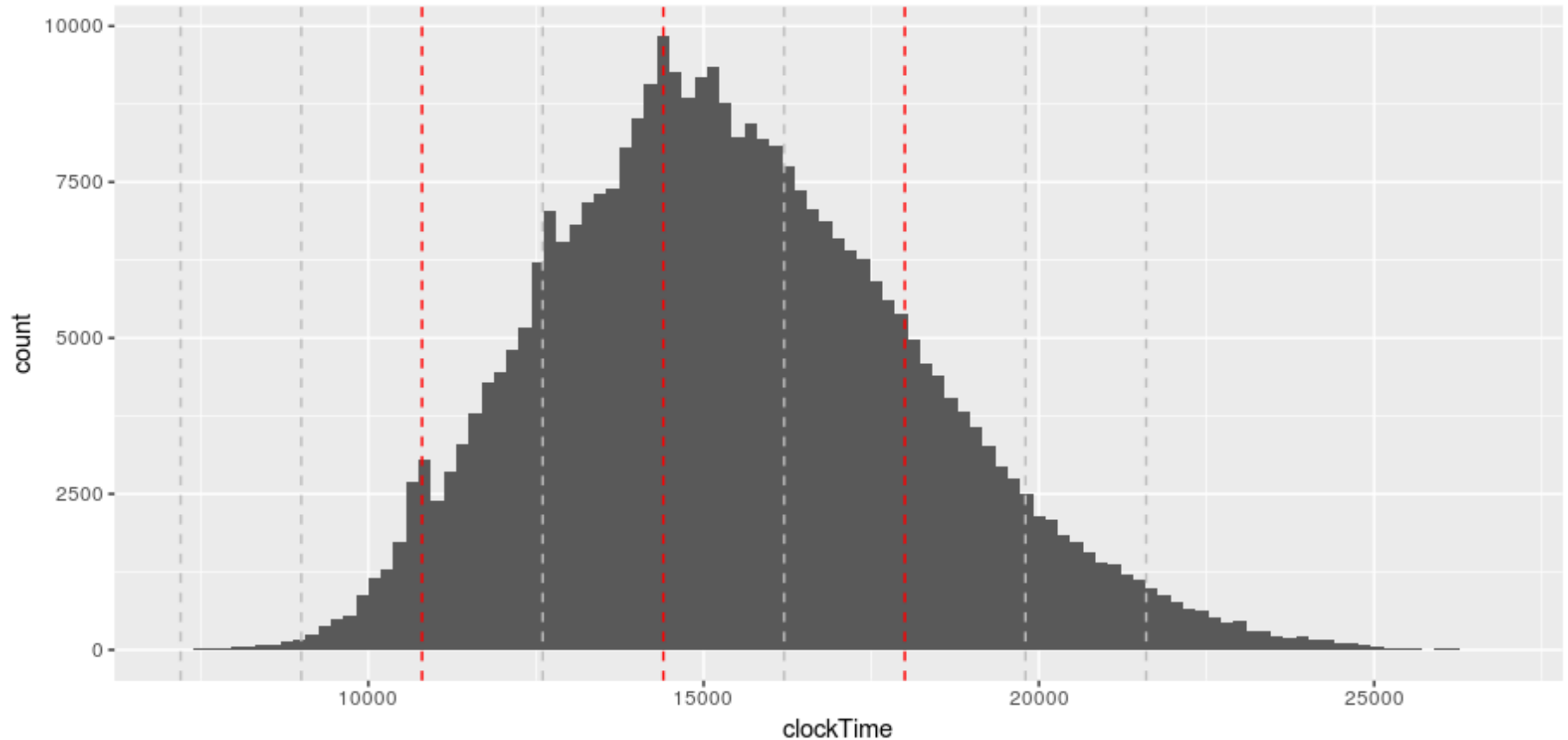


## 2.1 Change default settings





## 2.1 Change default settings







## 2.2 Make simpler graphs

*Data-ink is the non-erasable core of the graphic, the non-redundant ink arranged in response to variation in the numbers represented. (Tufte 1983)*



## 2.2 Make simpler graphs

A step-by-step example: [Data looks better naked](#)



[Speaker Deck](#)

Talk by [Joey Cherdarchuk](#)

[Full Screen](#)

**Remove**  
to improve  
(the **data-ink** ratio)



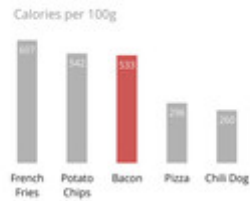
[Previous Slide](#) [Next Slide](#)

[Previous](#) [Next](#)

[share](#)



After



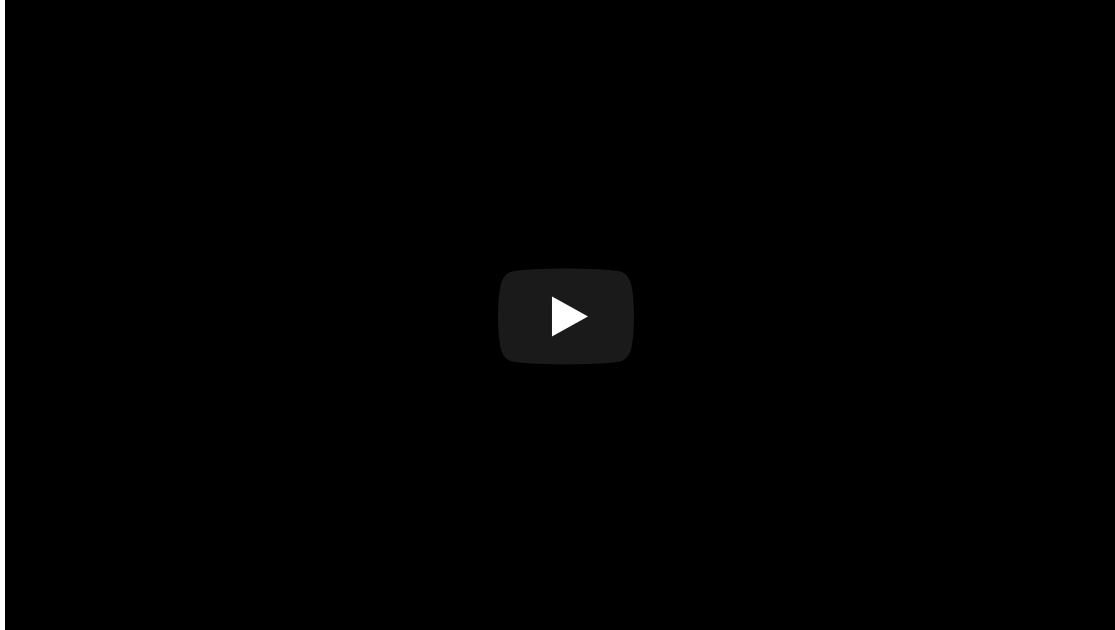
## Speaker notes

Embedeatzea zaila izan da eta diapoak pasatzerakoan izan daiteke efekto arraroren bat egitea



## 2.2 Make simpler graphs

(denborakin ikusteko)



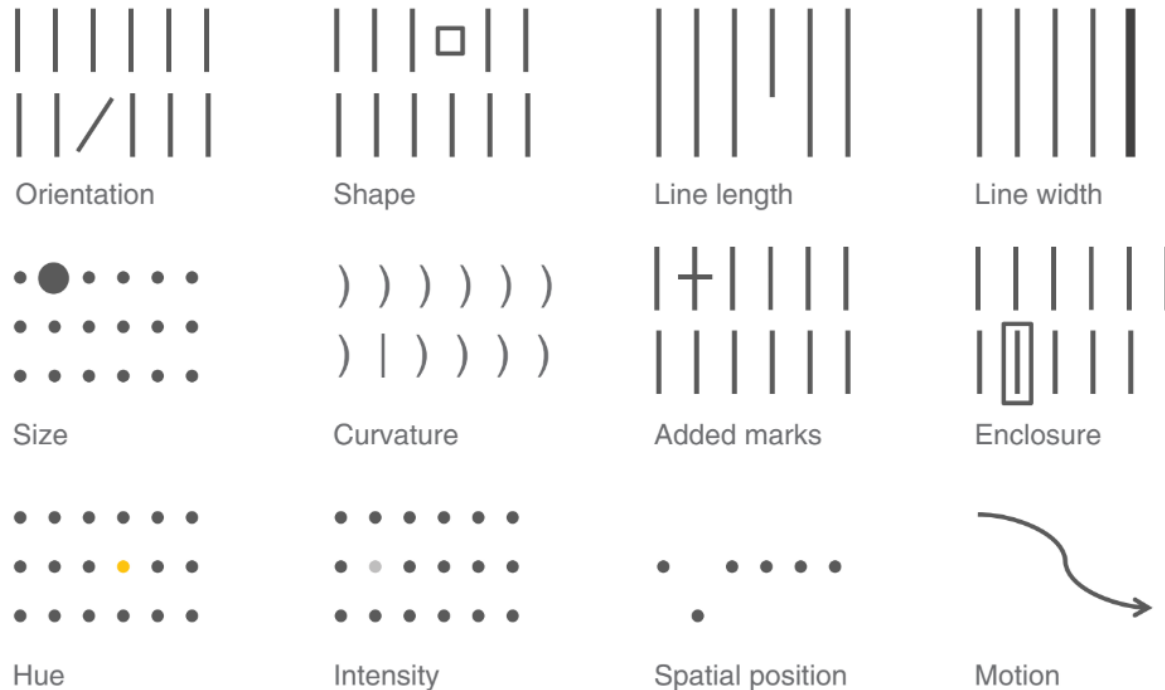
Nussbaumer, [Declutter Your Data Visualizations](#)





# 2.3 Highlight observations

## Preattentive attributes



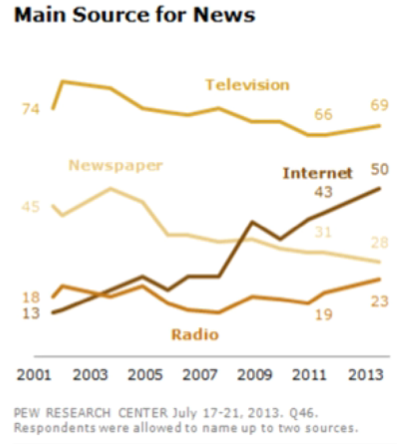
**FIGURE 4.4** Preattentive attributes

Source: Adapted from Stephen Few's *Show Me the Numbers*, 2004.



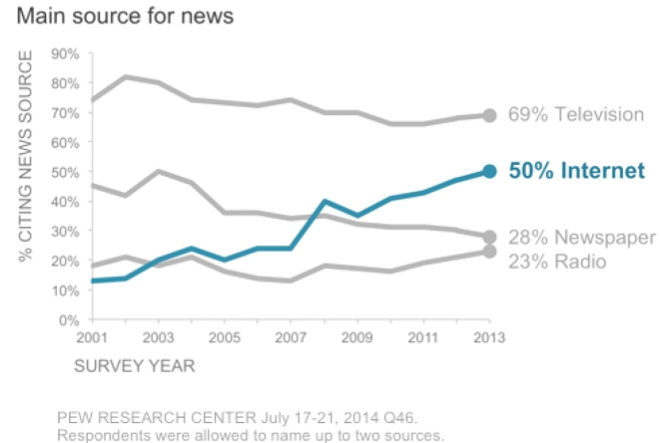
# 2.3 Highlight observations

**1. More Americans get news online...** 50% of the public now cite source for national and international news 📺, still below television, newspapers and radio. (Report)



## More Americans get news online

50% of the public cite the **internet** as a main source for national & international news. This remains below television, but is far above newspapers and radio.



Source: <http://www.pewresearch.org/fact-tank/2013/10/16/12-trends-shaping-digital-news/>  
© 2010 - 2016 Cole Nussbaumer Knaflic. All rights reserved.

storytelling  data

Nussbaumer, Do you see it? The importance of contrast when communicating with data [video]



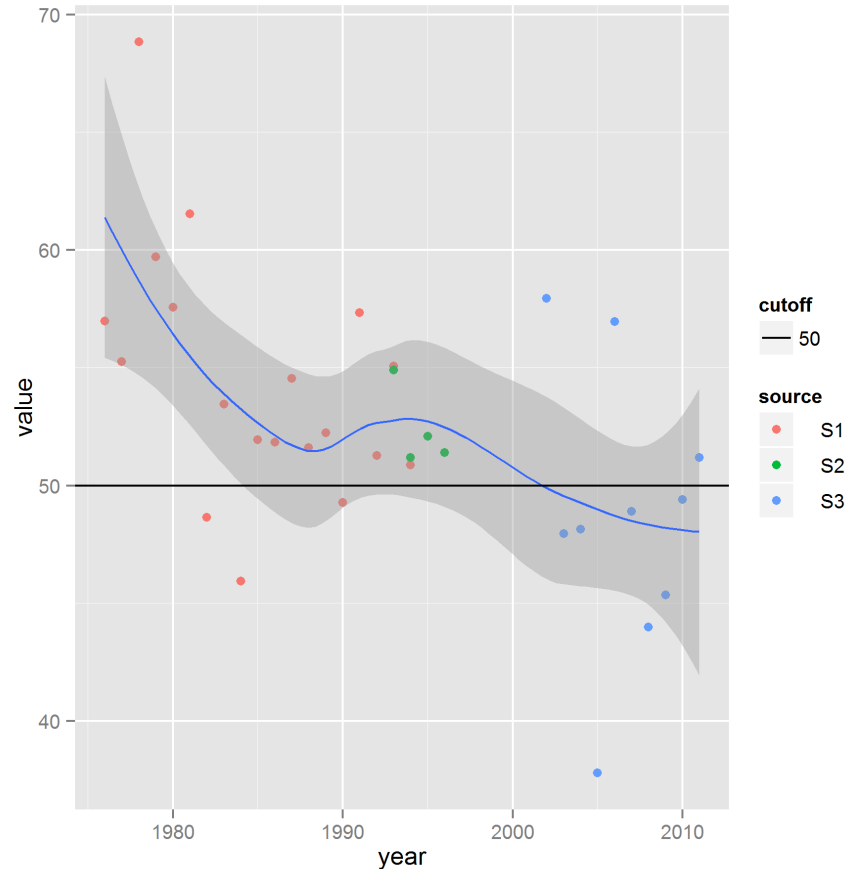
## 2.4 Add variables (as context)

- Adding preexisting variables (con **mesura**)
- Creating conditional variables from preexisting variables
  - binaries or with few levels are best
  - example of calculated field or variable: weekend date



## 2.5 Add statistical information

- statistical summaries
- models



source



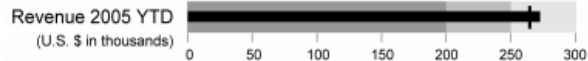


# Tableau: (not so) basic graphs

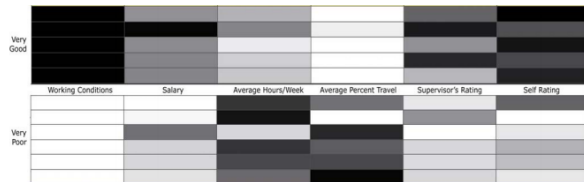
- Sparklines (Tufte 2006)



- Bulletgraphs (Few 2007)



- Heatmaps (Few 2006)

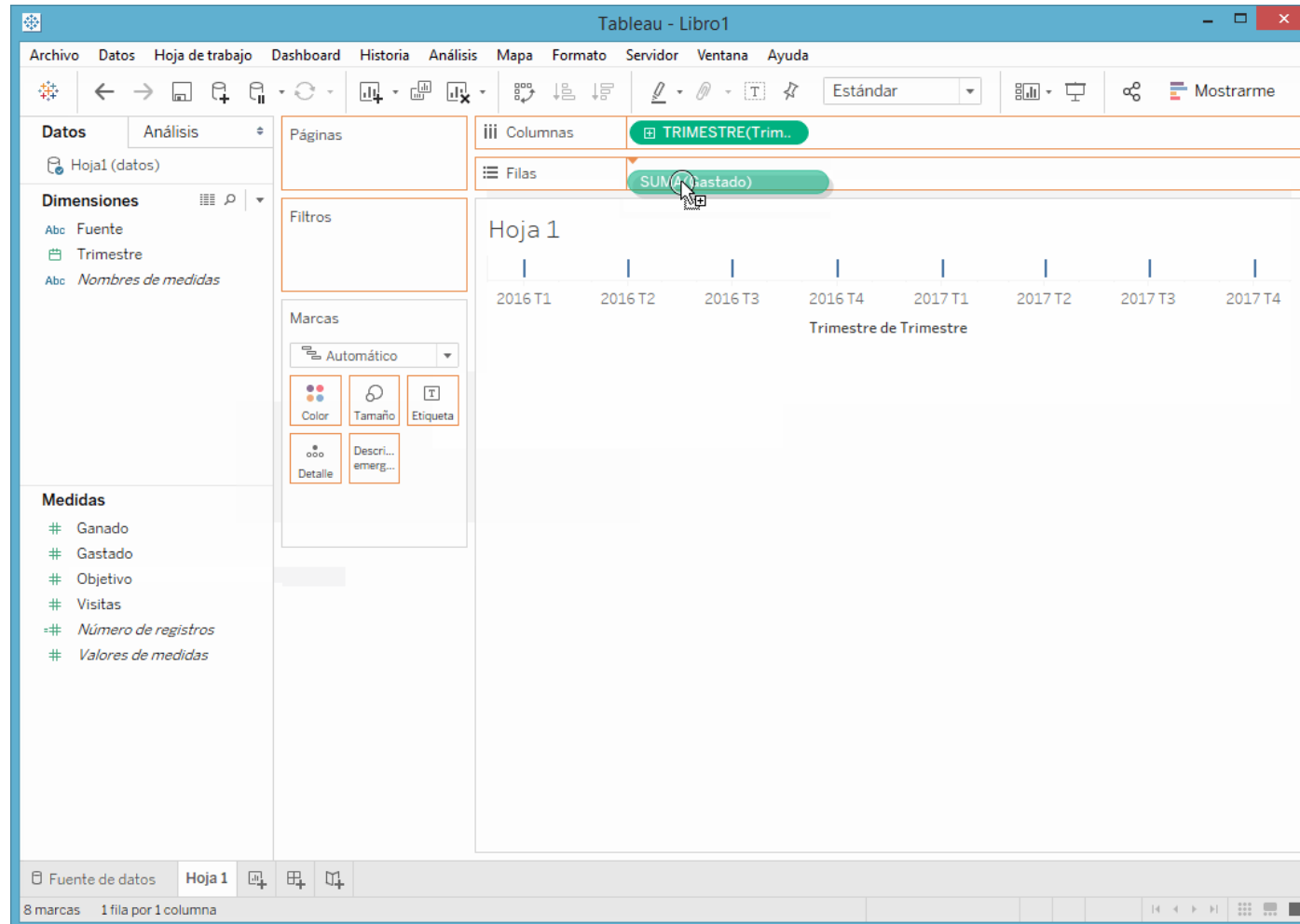


## Speaker notes

Nahi badezu bat azaldu eta ondoren tableaun egin, ala hiruak azaldu ta gero hiruak tableaun egin



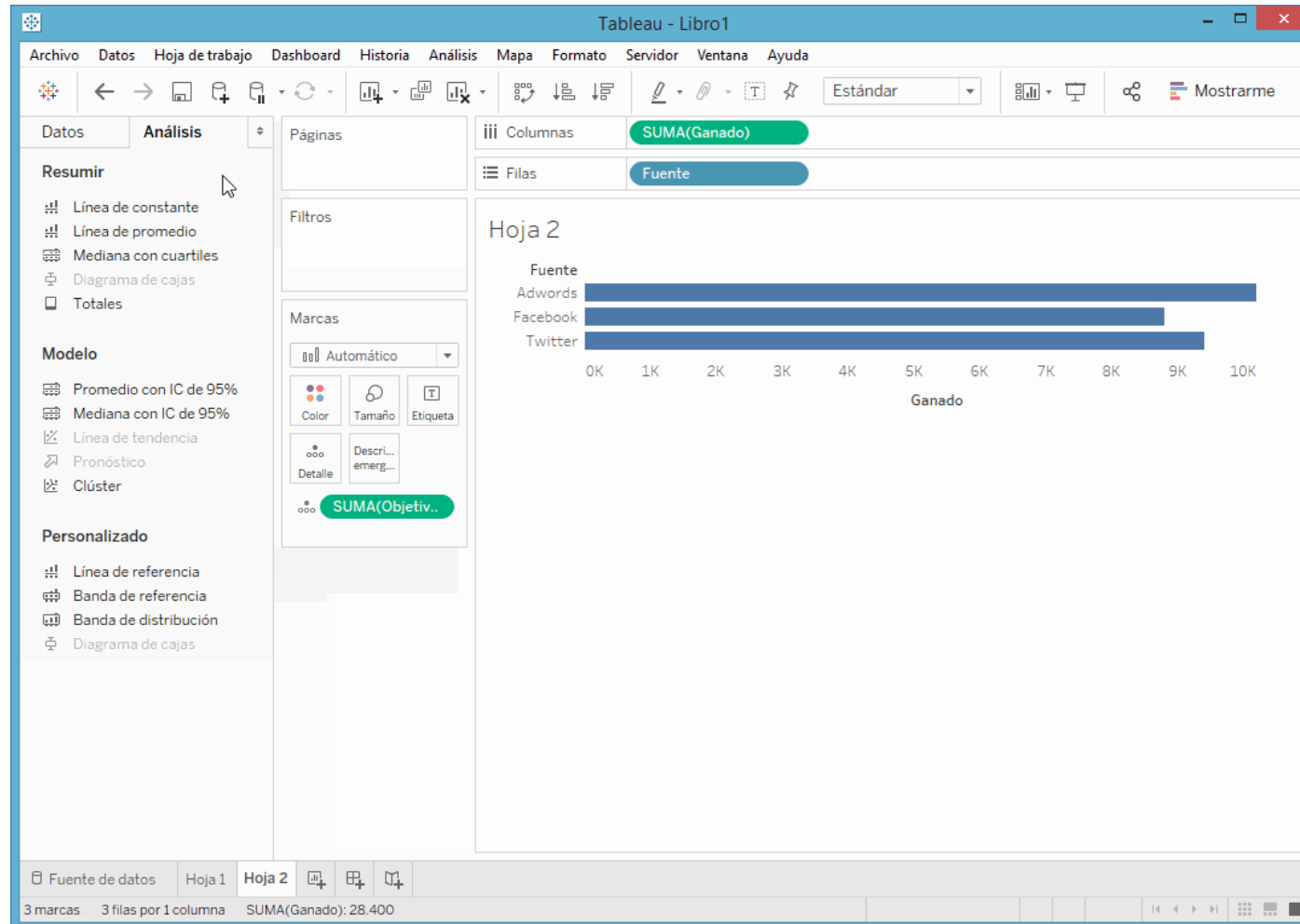
# Tableau 1.1: Timelines



Timelines



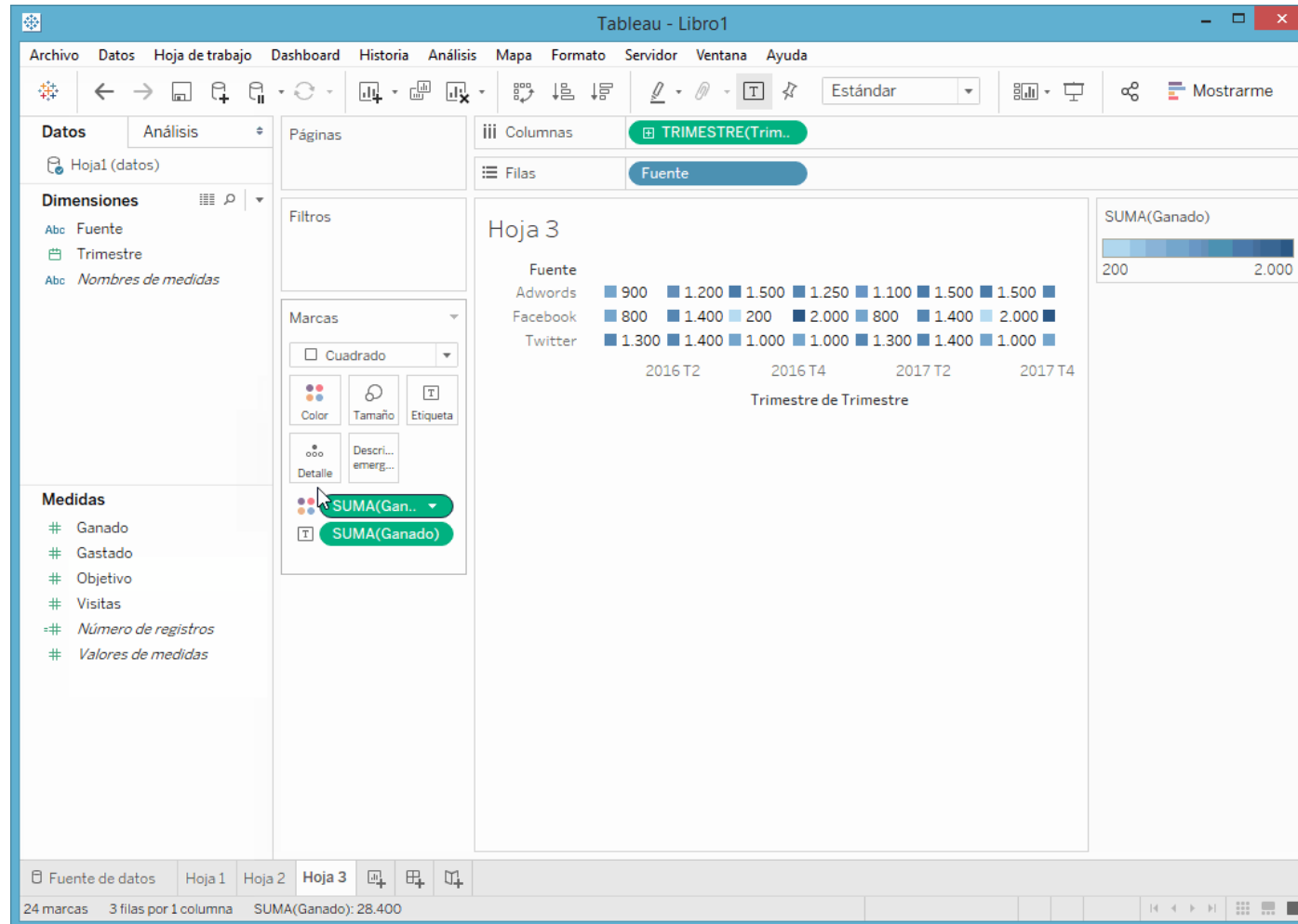
# Tableau 1.2: Bulletgraphs



Bulletgraphs



# Tableau 1.2: Heatmaps

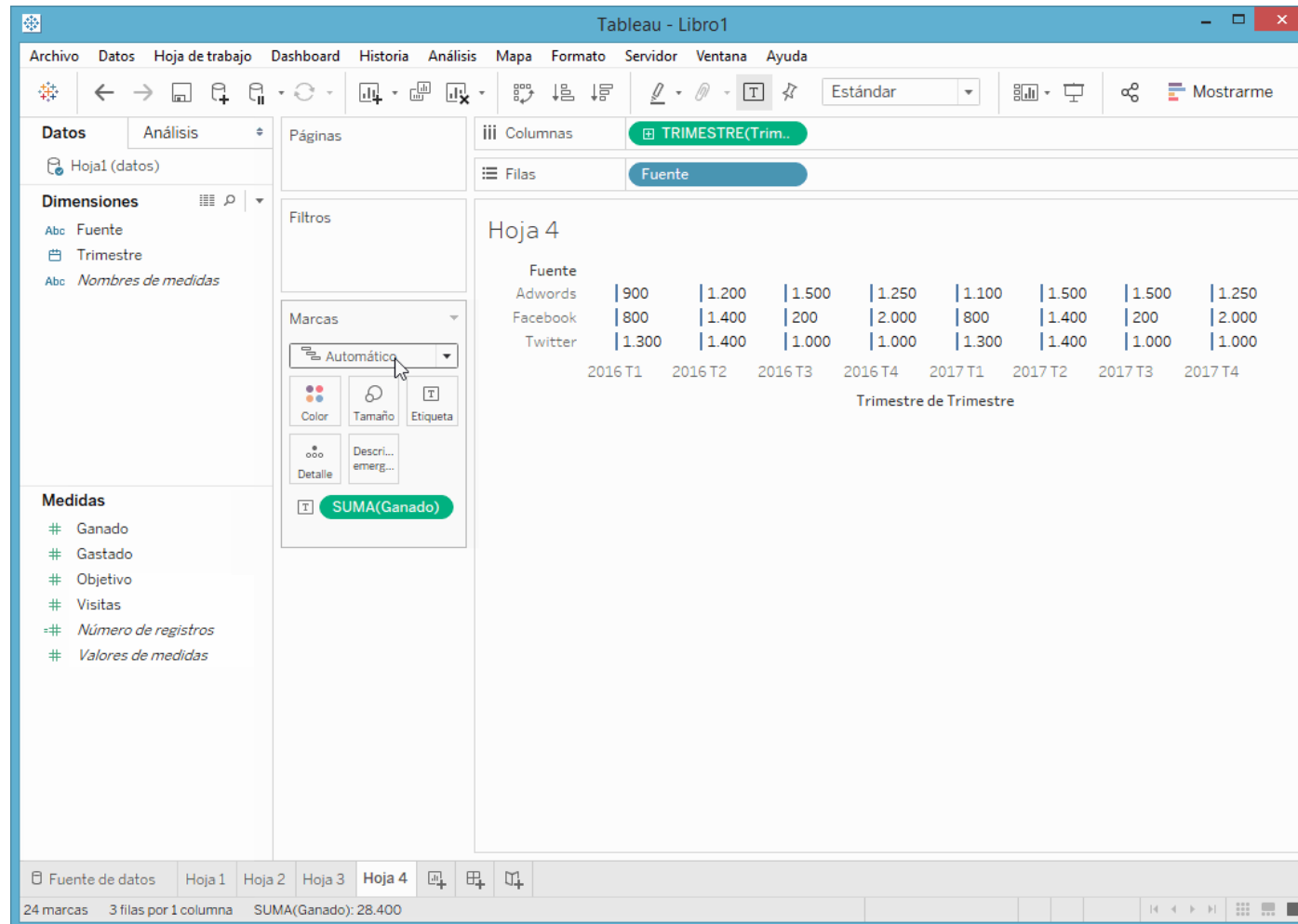


Heatmaps





# Tableau 2.1: automatic aggregation



Automatic aggregation



# Tableau 2.2: calculated fields

The screenshot shows the Tableau interface with a data source named 'Fuente' and a calculated field named 'Ratio Ganado/Gastado' being created. The calculated field formula is  $[Ganado] / [Gastado]$ . The background view shows a table with columns for 'Fuente' and 'Trimestre' (2016 T1, 2016 T2, 2016 T3, 2016 T4, 2017 T1, 2017 T2, 2017 T3, 2017 T4) and rows for 'Adwords', 'Facebook', and 'Twitter'. The status bar at the bottom indicates 24 marcas, 3 filas por 1 columna, and SUMA(Ganado): 28.400.

Tableau - Libro1

Archivo Datos Hoja de trabajo Dashboard Historia Análisis Mapa Formato Servidor Ventana Ayuda

Datos Análisis

Hoja1 (datos)

Dimensiones

- Abc Fuente
- Trimestre
- Abc Nombres de medidas

Medidas

- # Ganado
- # Gastado
- # Objetivo
- # Visitas
- # Número de registros
- # Valores de medidas

Páginas

Filtros

Marcas

Texto

Color Tamaño Texto

Detalle Descripción emergente

SUMA(Ganado)

Columnas

TRIMESTRE(Trimestre)

Filas

Fuente

Hoja 4

Fuente

	2016 T1	2016 T2	2016 T3	2016 T4	2017 T1	2017 T2	2017 T3	2017 T4
Adwords	900	1.200	1.500	1.250	1.100	1.500	1.500	1.250
Facebook	800	1.400	200	2.000	800	1.400	200	2.000
Twitter	1.300	1.400	1.000	1.000	1.300	1.400	1.000	1.000

Trimestre de Trimestre

Ratio Ganado/Gastado

$[Ganado] / [Gastado]$

Aplicar Aceptar

24 marcas 3 filas por 1 columna SUMA(Ganado): 28.400

Calculated fields



# Tableau 2.3: aggregation on calculated fields

Tableau - Libro1

Archivo Datos Hoja de trabajo Dashboard Historia Análisis Mapa Formato Servidor Ventana Ayuda

Datos

Hoja1 (datos)

Dimensiones

- Fuente
- Trimestre
- Nombres de medidas

Medidas

- Ganado
- Gastado
- Objetivo
- Ratio Ganado/Gastado
- Visitas
- Número de registros
- Valores de medidas

Columnas

AÑO(Trimestre)

Filas

Fuente

Trimestre

Fuente	2016	2017
Adwords	4,85	5,35
Facebook	5,73	4,40
Twitter	15,67	9,90

Hoja 4

6 marcas 3 filas por 2 columnas SUMA(Ratio Ganado/Gastado)



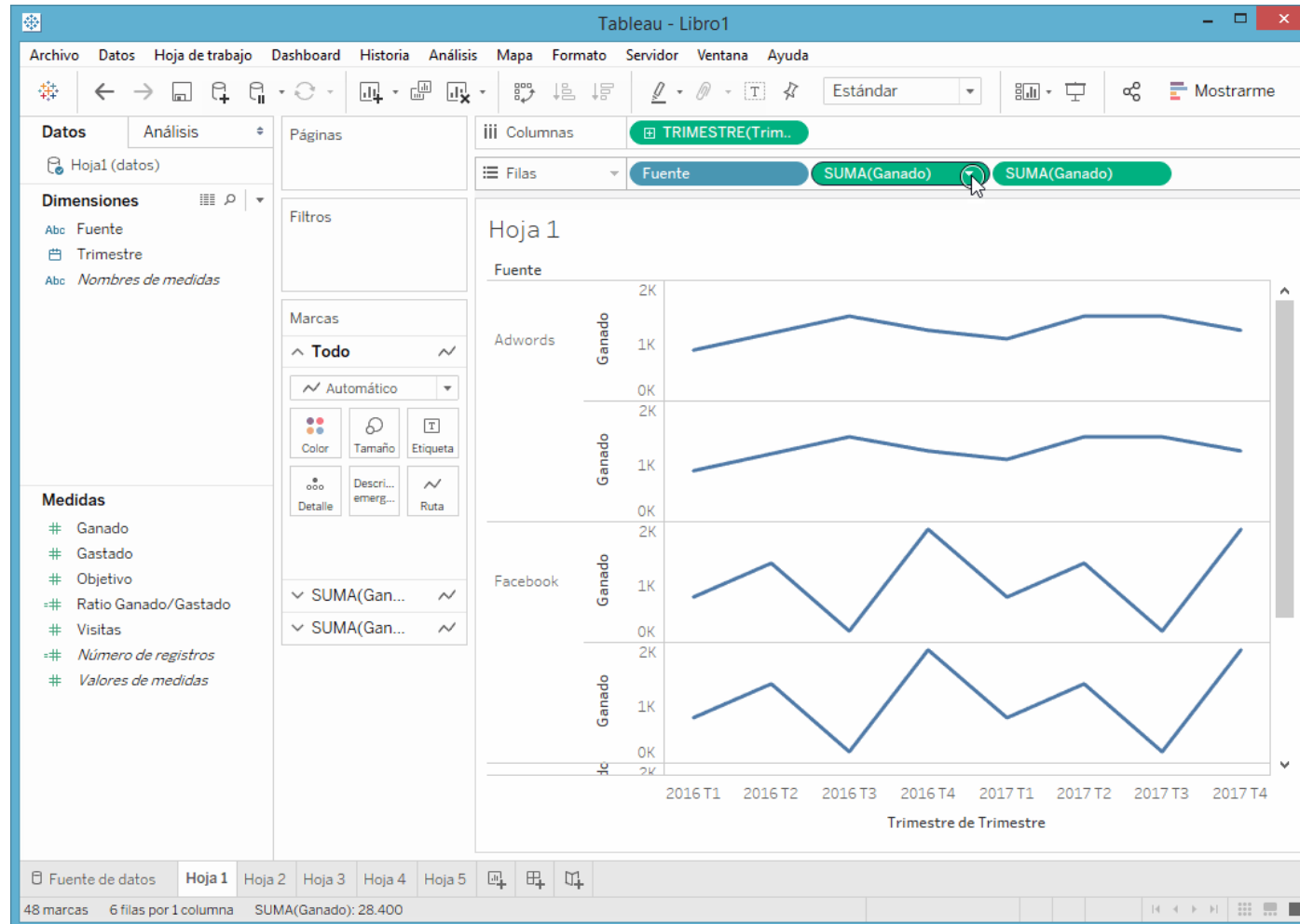
Aggregation on calculated fields





# Tableau 2.4: adding KPIs to timelines

## 1





# Tableau 2.5: adding KPIs to timelines

## 2

The screenshot shows the Tableau interface with a dashboard titled 'Hoja 1'. The main view displays a line chart for 'Ganado (KPI)' over time, with a data point for 'Adwords' showing 10.200. A KPI card for 'Gastado (KPI)' is overlaid on the chart, showing a formula error: 'El cálculo contiene errores'. The formula is 'LOOKUP(window\_sum)'. A list of functions is displayed, including 'WINDOW\_STDEV', 'WINDOW\_STDEV', and 'WINDOW\_SUM'. A tooltip for 'WINDOW\_SUM' indicates it calculates the sum of values in the window. The bottom status bar shows '24 marcas', '3 filas por 1 columna', and 'SUMA(Ganado): 28.400'.

Tableau - Libro1

Archivo Datos Hoja de trabajo Dashboard Historia Análisis Mapa Formato Servidor Ventana Ayuda

Datos: Hoja1 (datos)

Dimensiones: Fuente, Trimestre, Nombres de medidas

Columnas: TRIMESTRE(Trim..)

Filas: Fuente, Ganado (KPI), SUMA(Ganado)

Hoja 1

Fuente Ganado (KPI)

Adwords 10.200

Ganado

2K 1K

Gastado (KPI)

LOOKUP(window\_sum)

f WINDOW\_STDEV  
f WINDOW\_STDEV  
f WINDOW\_SUM

Indica la suma de valores que hay en la ventana

El cálculo contiene errores

Aplicar Aceptar

Todo

Escribir texto de búsqueda...

INT  
ISDATE  
ISFIELDNAME  
ISMEMBER  
ISNULL  
ISUSERNAME  
LAST  
LEFT  
LEN  
LN  
LOG  
LOOKUP  
LOWER

LOOKUP (expresión, [compensación])

Indica el valor de la expresión dada en una fila objetivo, especificada como compensación relativa desde la fila actual. Use FIRST() +n y LAST()-n para un objetivo relativo a las filas primera/última en la división. Si se omite la compensación, la fila Comparar con debe configurarse en el menú de campo. Indica NULL si la fila objetivo no se puede determinar.

Trimestre de Trimestre

Fuente de datos Hoja 1 Hoja 2 Hoja 3 Hoja 4 Hoja 5

24 marcas 3 filas por 1 columna SUMA(Ganado): 28.400



# 3. Dashboards



## 3.1 Dashboards for *situation awareness*

Few (liburua, [laburpena 2007](#))

*The term “dashboard” refers to a single screen information display that is used to monitor what’s going on in some aspect of the business.*

- Perception of own’s environment
- Comprehension of it’s meaning
- Projection of that understanding into the future





## 3.2 Do's : Principles you should follow

- Use flicker and sound to grab attention
- Encourage active thinking about the data, not just passive reaction to alarms
- Don't over-automate actions to the point where people become disengaged
- Provide smooth and simple means to respond
- Provide a common picture for the whole team
- Support projections for proactive responses
- Match the mental model

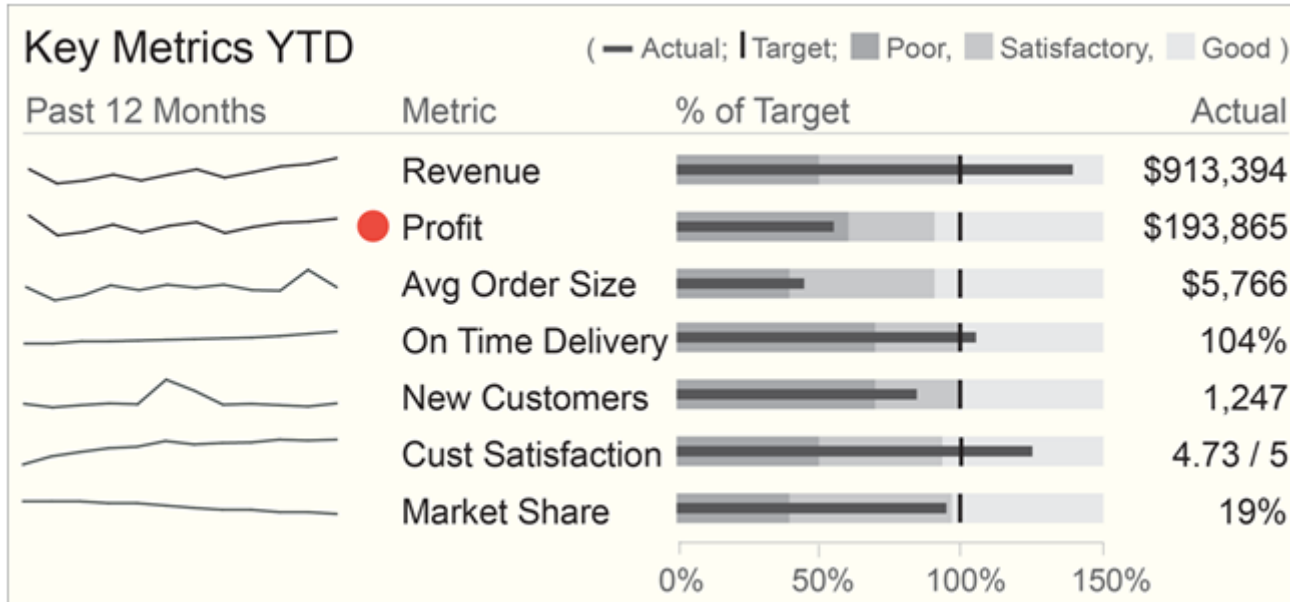


## 3.3 ... and don't's: Design problems you should avoid

- Too much complexity
- Too many alert conditions
- Alerts that cannot be differentiated
- Overwhelming visuals
- Distracting visuals
- Inappropriate visual salience
- Mismatch between information and its visual representation
- Indirect expression of measures
- Not enough context



## 3.4 Few's few examples



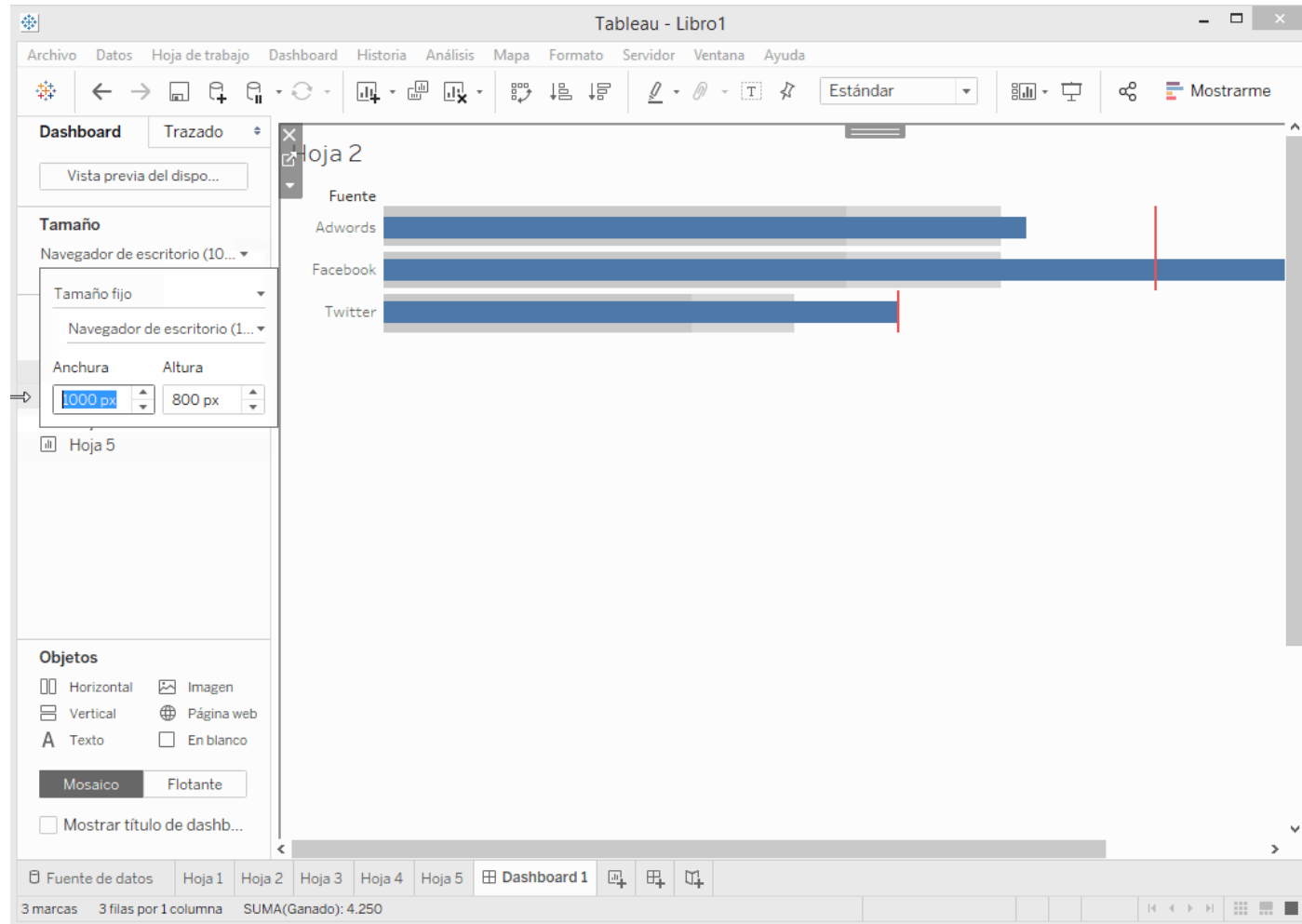


# Dashboards in Tableau





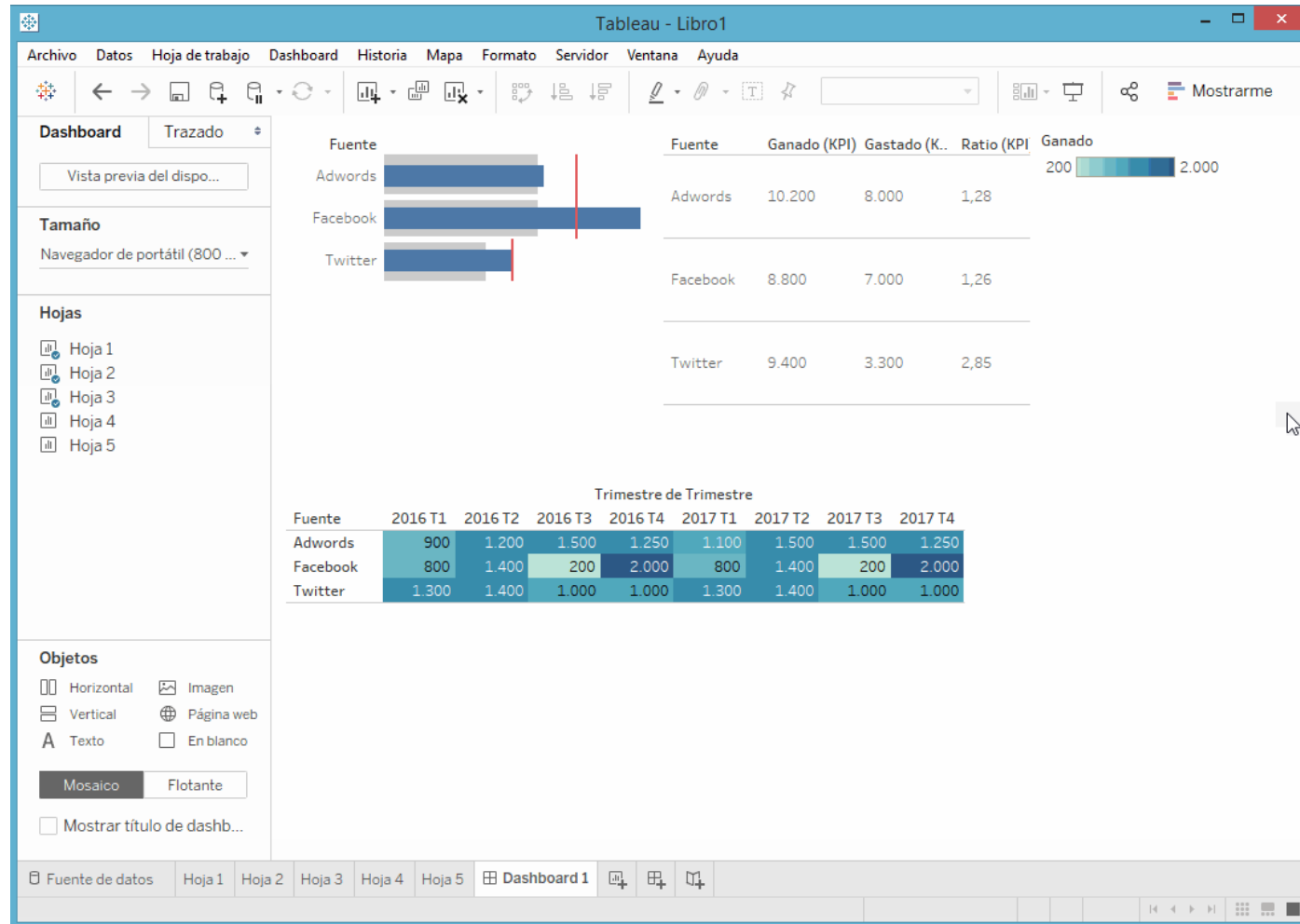
# Tableau 2.1: basic dashboard



Basic dashboard



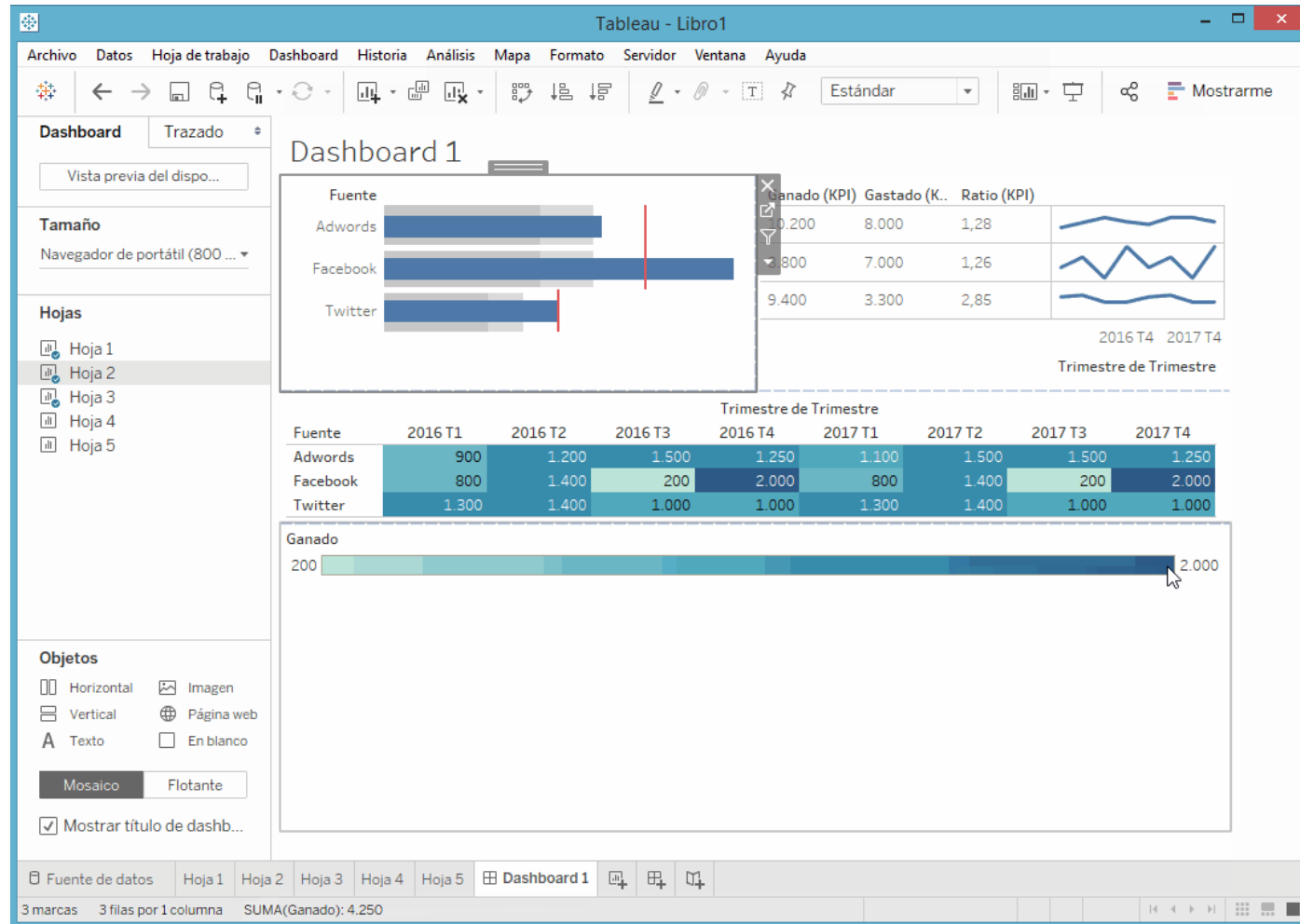
# Tableau 2.2: basic formatting



Basic formatting



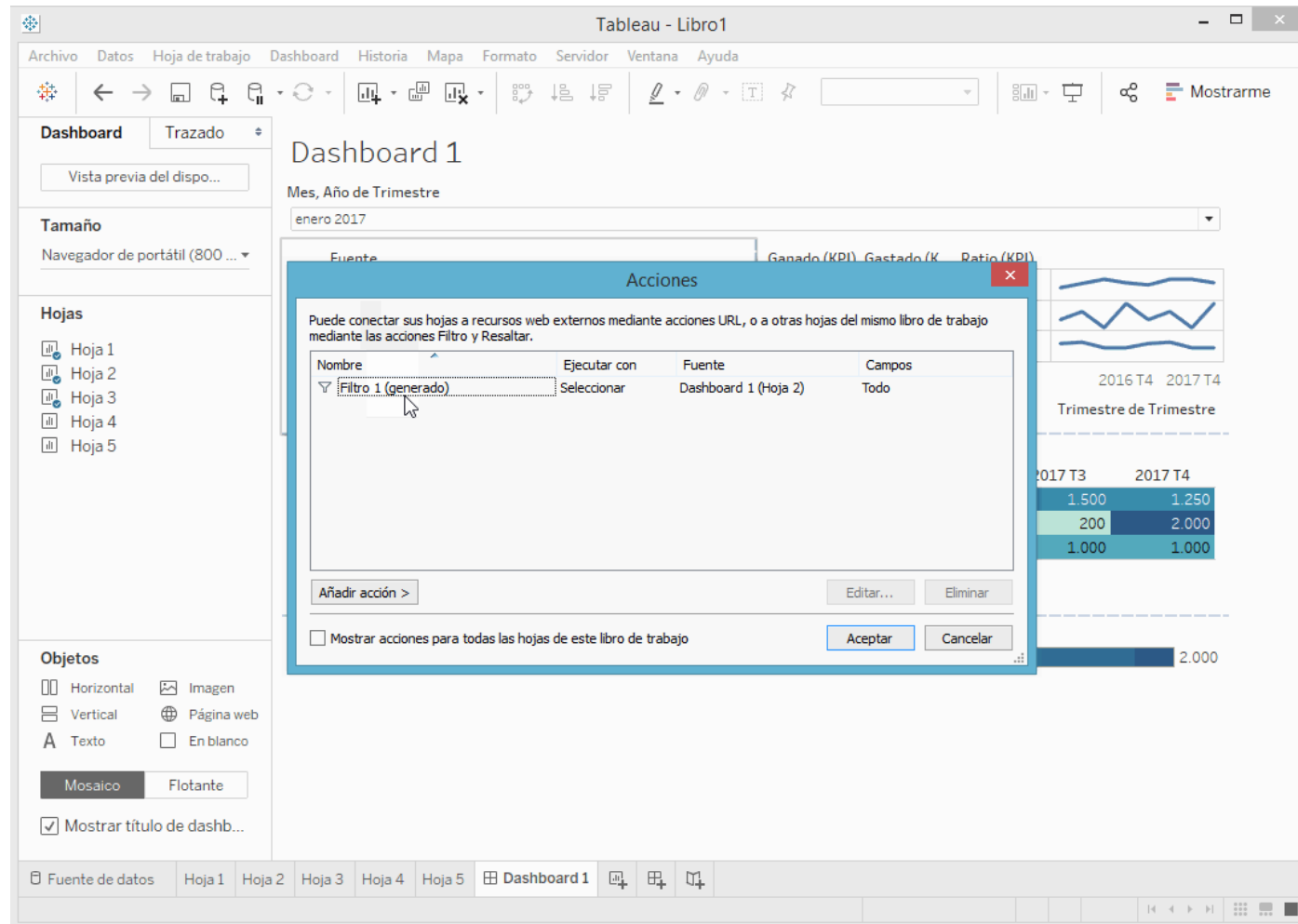
# Tableau 2.3: show filters



Show filters



# Tableau 2.4: highlight action

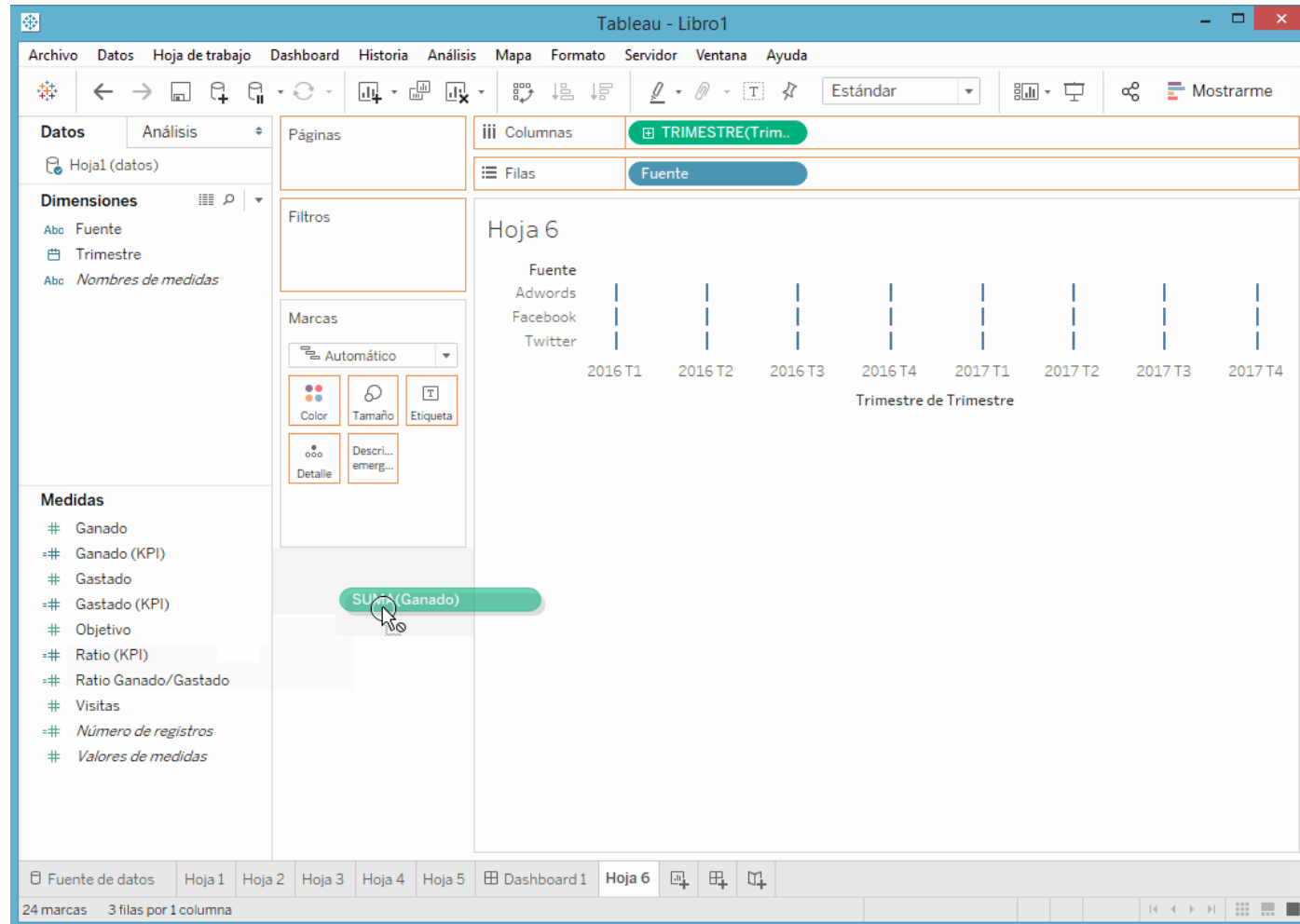


Highlight action





# Tableau 2.5: filter action



Filter action



# References



# Thank you!

Miren Berasategi  
[miren.berasategi@deusto.es](mailto:miren.berasategi@deusto.es)



# License



Copyright © 2018 University of Deusto

This work (except for the quoted images, whose rights are reserved to their owners) is licensed under the Creative Commons "Attribution-ShareAlike" License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/>

