

VISUALITZACIÓ DE DADES

Dra. Vanessa Serrano, Universitat Rovira i Virgili

Dr. Jordi Cuadros, IQS Universitat Ramon Llull

Maig, 2023

#ProDigital

TEMARI

1. Principis bàsics de visualització
2. Visualitzacions efectives amb eines habituals
3. Conceptualització dels gràfics
4. Visualització de dades simples

VISUALITZACIÓ DE DADES

3. Conceptualització dels gràfics

- Gramàtica de gràfics
- Implementació a R

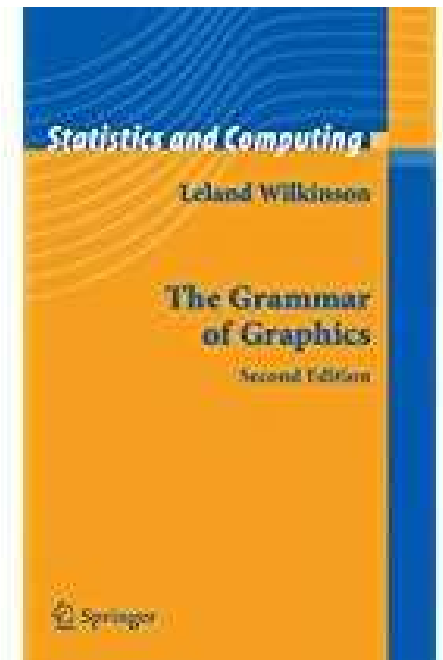
GRAMÀTICA DE GRÀFICS

GRAMÀTICA DE GRÀFICS

«A grammar of graphics is a tool that enables us to concisely describe the components of a graphic. Such a grammar allows us to move beyond named graphics (e.g., the 'scatterplot') and gain insight into the deep structure that underlies statistical graphics.»

Wickham, H. (2010)

<http://vita.had.co.nz/papers/layered-grammar.html>



GRAMÀTICA DE GRÀFICS

Una gramàtica de gràfics implica entendre un gràfic com un conjunt d'**elements** que es combinen d'acord a unes **regles**.

Exemples d'elements d'un gràfic:

- Dades
- Ajustaments de posició
- Escales
- Sistemes de coordenades
- ...

EL PAQUET DE R GGPLOT2

- Un paquet és una col·lecció de funcions de R, dades i codi compilat en un format ben definit.
- El directori de R on s'emmagatzemen els paquets s'anomena llibreria.
- R té un conjunt de paquets estàndards.
 - <https://www.rstudio.com/products/rpackages/>
 - <https://cran.r-project.org/web/packages/>
- El paquet ggplot2 està basat en la gramàtica de gràfics (Wilkinson, 2005).

IMPLEMENTACIÓ A R

INSTAL·LACIÓ DE R I RSTUDIO

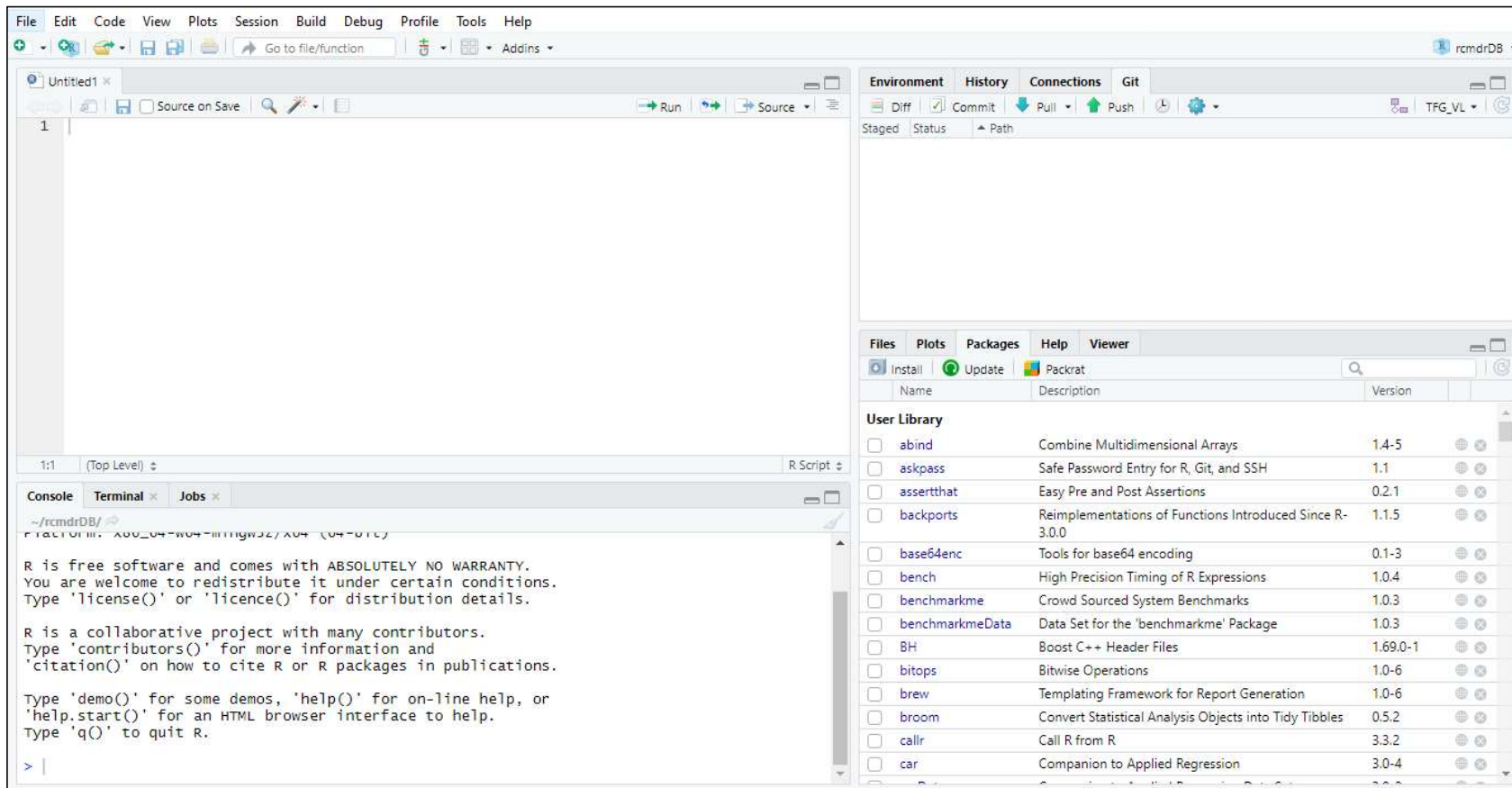
R

- Entorn de computació estadística
- <https://www.r-project.org/>

RStudio

- És un entorn de desenvolupament integrat per a R amb una interfície gràfica d'usuari
- <https://www.rstudio.com/>

INSTAL·LACIÓ DE R I RSTUDIO



INTRODUCCIÓ AL PAQUET GG PLOT2

Primer de tot cal instal·lar i carregar el paquet ggplot2

- `install.packages("ggplot2")`
- `library("ggplot2")`

<http://ggplot2.tidyverse.org/reference/>

<http://www.r-graph-gallery.com/portfolio/ggplot2-package/>

Hi ha alguns paquets addicionals que complementen ggplot2: ggthemes, ggExtra, GGally, ggfittext, treemapify, ggrepel, ggmap, gganimate, ggspectra, ggmosaic...

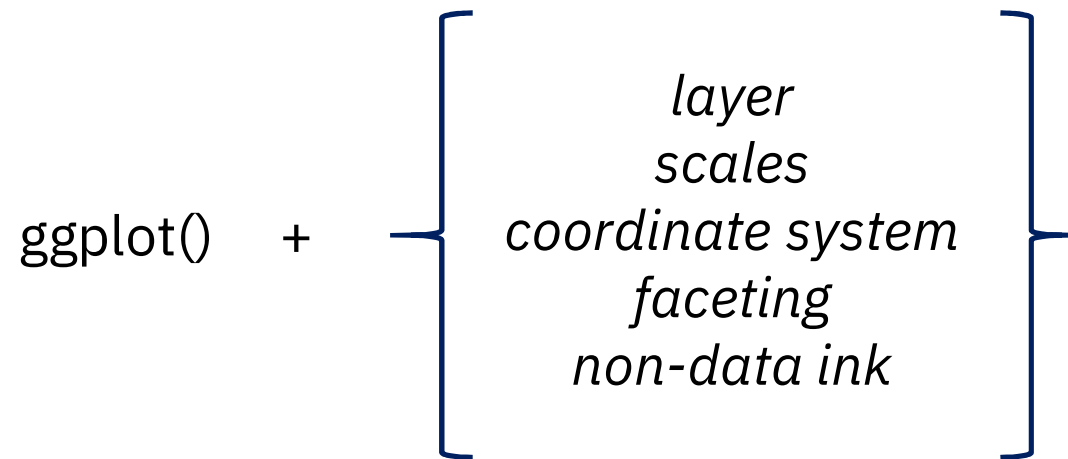
- <https://exts.ggplot2.tidyverse.org/>

ELEMENTS DE GGLOT2

Un gràfic amb ggplot2 pot tenir els següents elements:

- Elements principals:
 - *Data, Aesthetic mapping, Geometric objects, Scales, Coordinate system, Faceting, Non-data ink (themes, titles, legend, notes)*
- Elements addicionals
 - *Stats, Position adjustments*

LA GRAMÀTICA DE GGPLOT2



LA GRAMÀTICA DE GGPLOT2

Un gràfic de ggplot2 està format per **capes**.

Una capa de ggplot2 pot estar formada pels elements següents:

- Dades (*data*)
- Mapatge estètic (*aesthetic mapping*)
- Geometries (*geometric objects*)
- Estadístiques (*stats*, hi ha valors per defecte)
- Ajustos de posició (*position adjustments*, hi ha valors per defecte)

Les capes es poden superposar.

LA GRAMÀTICA DE GGPLOT2

Dades (*data*)

És la taula que conté les dades a representar.

Mapatge estètic (*aesthetic mapping*)

Estableix les relacions entre les variables de la taula de dades i les propietats estètiques del gràfic (mida, color, posició de la x i la y...).

Objectes geomètrics (*geometric objects*)

Estableix la forma com es representen les dades (tipus d'objectes a representar).

MAPATGE ESTÈTIC

Principals mapatges estètics per a **variables quantitatives**

- posició (x, y, xmin, xmax, ymin...)
- Mida
- color (col, fill)

Principals mapatges estètics per a **variables qualitatives**

- Posició
- color (col, fill...)
- Forma
- mida

MAPATGE ESTÈTIC

Aesthetic	Description
x	X axis position
y	Y axis position
colour	Colour of dots, outlines, or other shapes
fill	Fill colour
size	Diameter of points, thickness of lines
alpha	Opacity
linetype	Line dash pattern
label	Text on a plot or axes
shape	Shape

OBJECTES GEOMÈTRICS

Evolució d'una variable

- Line diagram: `geom_line()`

Distribució de dues variables

- Scatter plot: `geom_point()`
- Bar plot: `geom_bar()`
- Boxplot: `geom_boxplot()`
- Histogram: `geom_histogram()`

<https://www.rstudio.com/wp-content/uploads/2016/11/ggplot2-cheatsheet-2.1.pdf>

OBJECTES GEOMÈTRICS

Common plot types	Geometries
Scatter plots	point
	jitter
	abline
Bar plots	histogram
	bar
	errorbar
Line plots	line

OBJECTES GEOMÈTRICS

Relació entre dos variables

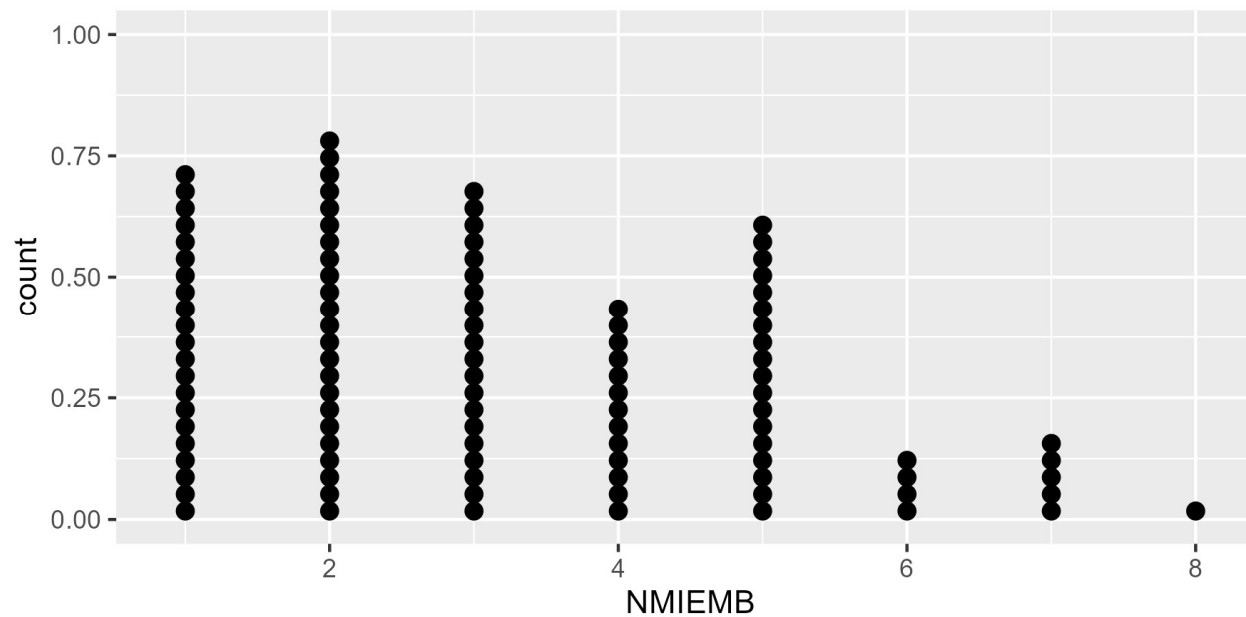
- Scatter plot: `geom_point()`
- Boxplot: `geom_boxplot()`

Distribució de una sola variable

- Dot plot: `geom_dotplot()`
- Bar plot: `geom_bar()`
- Boxplot: `geom_boxplot()`
- Histogram: `geom_histogram()`

EXAMPLE - ELEMENTS PRINCIPALS

```
ggplot(df1CEUTA, aes(x=NMIEMB)) +  
  geom_dotplot(dotsize=0.5)
```



ELEMENTS ADDICIONALS

Elements additional de ggplot2:

- Escales (*scales*)
- Sistema de coordenades (*coordinate system*)
- Facetes (*faceting*)
- Altres (*non-data ink: themes, titles, legend, notes*)

ELEMENTS ADDICIONALS

Escales (*scales*)

Serveix per a controlar les relacions entre les variables del conjunt de dades i les propietats estètiques del gràfic.

Exemples:

- Graduació dels eixos
- Ajustaments de les escales de color
- Escales no lineals

ELEMENTS ADDICIONALS

Sistema de coordenades (*coordinate system*)

Permet ajustar les opcions del sistema de coordenades, és a dir, com es calculen les posicions.

Exemples:

- Límits de zoom o de visualització
- Escales de la mateixa mida
- Rotació del sistema de coordenades

ELEMENTS ADDICIONALS

Facetes (*faceting*)

Correspon a la creació de diferents minigràfics per a cada valor d'una o vàries variables.

Opcions:

- `facet_wrap(...)`
- `facet_grid(...)`

ELEMENTS ADDICIONALS

Temes (*non-data ink: themes*)

Permeten ajustar les característiques de presentació de diferents elements del gràfic.

- Existeixen diferents temes creats que responen a formes estandarditzades de representació: `theme_bw()`, `theme_few()`, `theme_minimal()`...
- `element_black()` s'usa per a excloure un element.
- La resta d'elements s'ajusten com `element_text()`, `element_line()`, `element_rect()`, `unit()`, `margin()`.

ELEMENTS ADDICIONALS

Altres (*non-data ink*)

- Títols: labs(...)
- Llegendes: s'eliminen des del theme, es modifiquen amb guides
- Notes: per afegir text a un gràfic

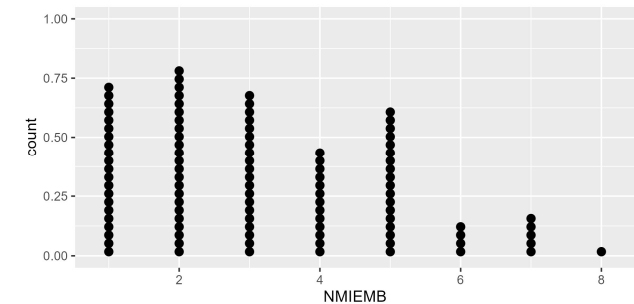
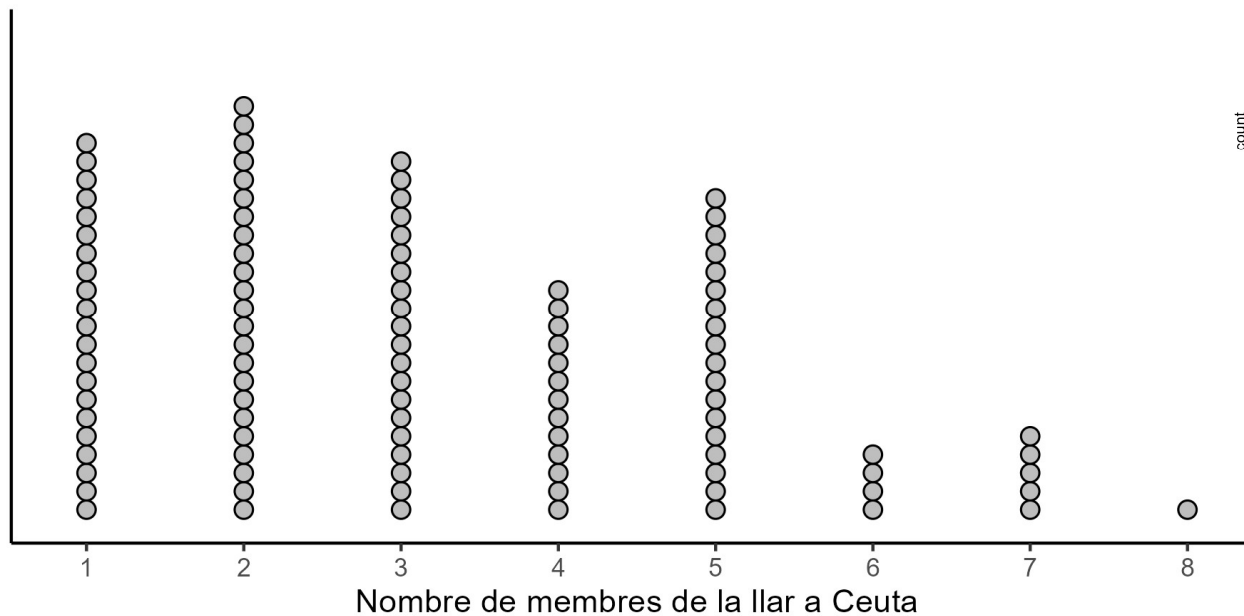
EXEMPLE - ELEMENTS ADDICIONALS

```
ggplot(df1CEUTA, aes(x=NMIEMB)) +  
  geom_dotplot(dotsize=0.5, col="black", fill="grey")+  
  scale_y_continuous(NULL, breaks=NULL)+  
  scale_x_continuous(breaks=1:max(df1CEUTA$NMIEMB))+  
  xlab("Nombre de membres de la llar a Ceuta")+  
  theme_classic()
```

Escales

Títols

Temes



Dra. Vanessa Serrano, Dr. Jordi Cuadros
vanessa.serrano@urv.cat, jordi.cuadros@iqs.url.edu