

# Az R adatelemzési nyelv alapjai III.

Egészségügyi informatika és biostatisztika

Gézi András  
gezi@mit.bme.hu

# Graphics with ggplot2

# Data visualization

- **Exploratory data visualization**

Exploring the data visually to find patterns among the data entities

- **Explanatory data visualization**

Showcasing the identified patterns using simple graphs

# Why ggplot2?

- de facto standard visualization package
- designed for print-quality graphics in seconds
- consistent underlying grammar of graphics
- plot specification at a high level of abstraction
- very flexible
- theme system for polishing plot appearance
- mature and complete graphics system
- many users, active mailing list



# Grammar of graphics

- The basic idea: independently specify plot building blocks and combine them to create just about any kind of graphical display you want.
- Building blocks:
  - data
  - aesthetic mapping
  - geometric object (~ layers)
  - statistical transformations
  - scales
  - coordinate system
  - position adjustments
  - faceting
  - themes

# Working with the grammar

- Three required components
  - **Data** – The raw material of the visualization.
  - **Aesthetics** – The mappings of the data to the visualization.
  - **Layers** – How to render the data and aesthetics to the screen. Typically „geom\_” functions.

# Data: Titanic



PassengerId	Passenger ID
Survived	Passenger Survival Indicator
Pclass	Passenger Class
Name	Name
Sex	Sex
Age	Age
SibSp	Number of Siblings/Spouses Aboard
Parch	Number of Parents/Children Aboard
Ticket	Ticket Number
Fare	Passenger Fare
Cabin	Cabin
Embarked	Port of Embarkation



# Credits

- <https://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html>
- <https://www.edureka.co/blog/ggplot2-tutorial/>
- <https://tutorials.datasciencedojo.com/introduction-to-data-visualization-with-r-and-ggplot2/>

# Further reading

- Hadley Wickham  
*ggplot2: Elegant Graphics  
for Data Analysis*  
Springer 2016

