

Visualisaties in R

Workshop DAIR seminar

Martine Jansen | Fontys

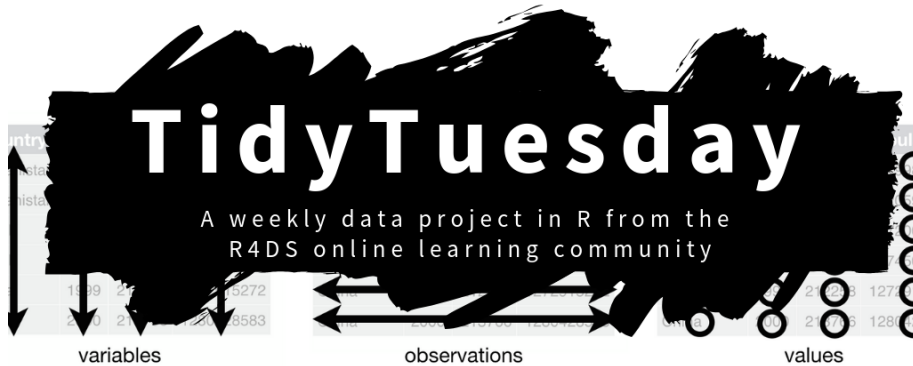
Veerle van Son | Avans

woensdag 10 november, 2021

Wat gaan we doen

- Korte presentatie: Tidy Tuesday
- Compacte uitleg: (gg)plots in R
- Zelf: aan de slag met een dataset

Tidy Tuesday



- Sinds 1 april 2018
- "safe and supportive forum for individuals to practice their wrangling and data visualization skills independent of drawing conclusions"
- Elke week een ruwe dataset en artikel
- Exploreer de data en visualiseer
- Deel plaatje en code
- 16273 tweets door 3510 personen
- Lees meer op: <https://github.com/rfordatascience/tidytuesday>
- en #tidytuesday op twitter

Voorbeeld Tidy Tuesday

- 17 augustus 2021
- Voice Interaction Computer Star Trek
- Data set met 2214 regels en 14 kolommen

TT Voice Interaction Computer Star Trek 1/2

variable	class	description
name	character	ID name
char	character	The name of the speaking character
line	character	The complete line of dialog (may contain more speech than the speech interaction). Parentheticals are directions and not spoken.
direction	character	Stage directions as written in the episode script
type	character	The type of interaction, see detailed definitions below
pri_type	character	The primary interaction type as defined by the below ranking
domain	character	The domain of interaction, see detailed definitions below
sub_domain	character	The sub-domain of interaction, generally a specific setting, see below
nv_resp	logical	On a person's line when the computer completes the query but without speaking a response.
interaction	character	The actual speech interaction. May be shorter or longer than the line of dialog
char_type	character	Either Person or Computer
is_fed	logical	Indicates whether an interaction is with the standard Enterprise system (true) or another (false)
error	logical	The interaction resulted in an error

TT Voice Interaction Computer Star Trek 2/2

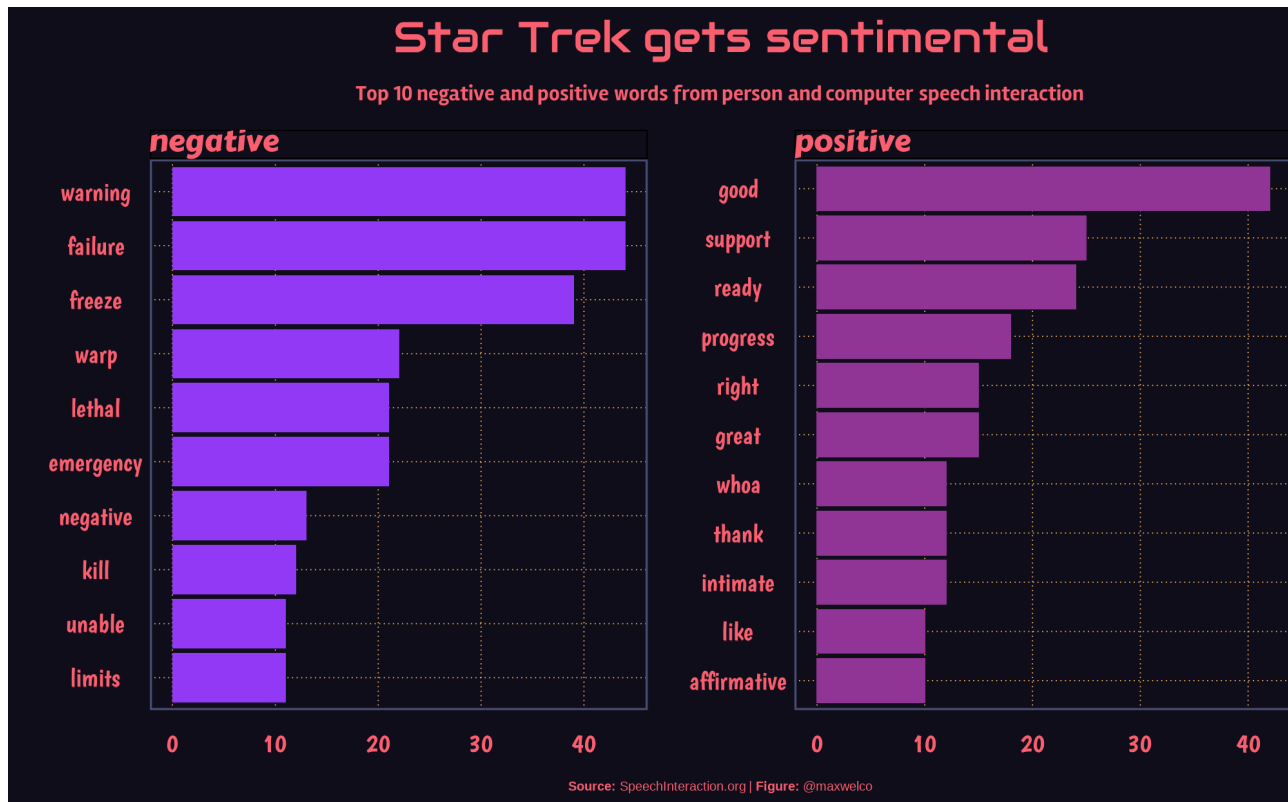
Oliveira

Eilerts

Rennie

Hudson

Gephard



ggplot's anatomie

Gebhard met annotatie

mogelijke code

Eilerts met annotatie

mogelijke code

Verschillende soorten visualisaties in R

<https://www.r-graph-gallery.com/>

De data die wij gaan visualiseren

- Afkomstig van DUO
- 107522 regels met 17 variabelen
- per vo-school, per jaar, en per profiel:
 - het aantal examenkandidaten
 - het aantal geslaagden
 - het aantal gezakten
 - gemiddeld cijfer op lijst
 - gemiddeld cijfer schoolexamen
 - gemiddeld cijfer centraal examen

Aan de slag

- Ga naar <https://github.com/MMJansen/DAIR2021>
- Klik op de groen knop "Code", en kies Download Zip
- Unzip het bestand

De volgende stappen veronderstellen de aanwezigheid van R en RStudio:

- Vanuit de verkenner, dubbelklik **DAIRpres.Rproj**
- Vanuit het files overzicht rechtsonder in RStudio, dubbelklik **Geslaagden.Rmd**