Becoming "TidyR" ov

Data Visualization De

in an Online Commur

Practice

PRESENTER:

er Time:

evelopment

nity of

Da

ski
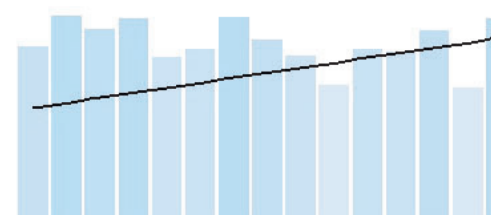
ta visuali

ls can be

knitr

broom

R

zation

plumber

develope

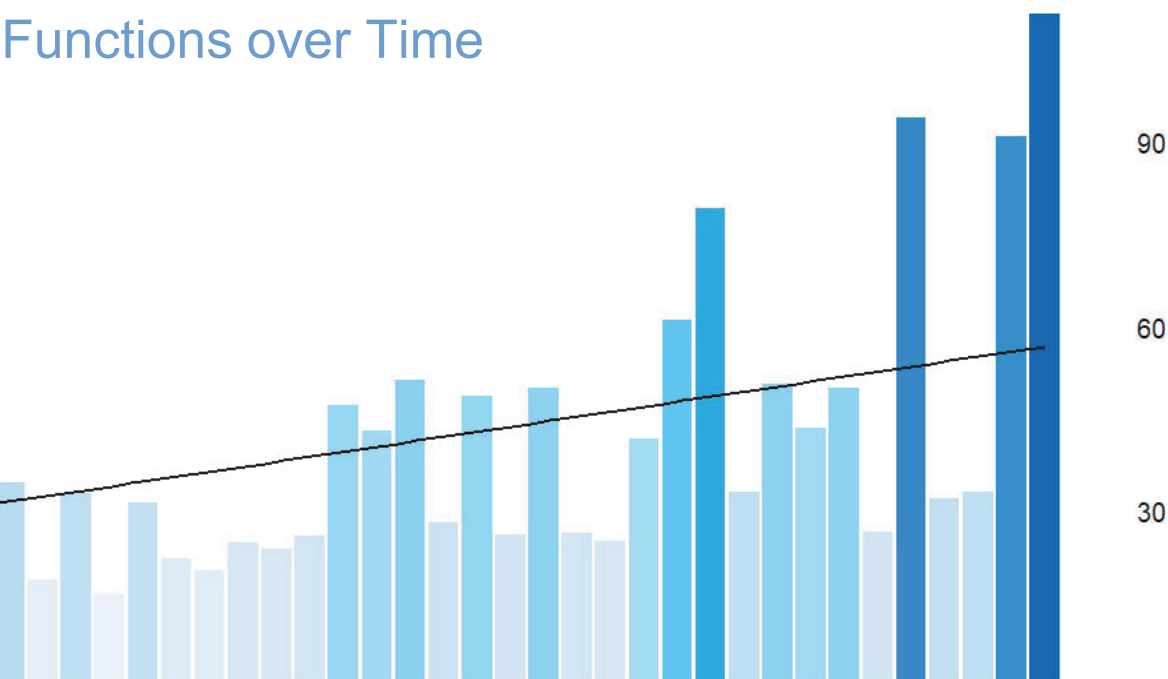# Cont

Over time, contributors

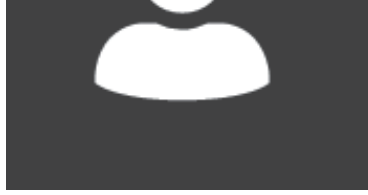This suggests their coc

Average Number of

r = .54, *p* < .001

tent Analysis

write longer code and use more functions.
de may become more sophisticated over time.

Functions over Time

Anthony Schm

Evaluation, Statistics,

aschmi11@utk.edu / anthonys

AUTHORS
Schmidt, Anthony (ESM)
Rosenberg, Joshua (TPTE

Communities of Pra

- Groups of people who share a commo
  in activities that promote domain-spe

#TidyTuesday

- The #TidyTuesday hashtag is a commu
  those interested in data visualization a
- Every Tuesday, a unique dataset is pos
  hashtag #TidyTuesday.
- Participants use the dataset to create
  R programing language
- Participants then share their visualiza
  to their code (on Github).

R

idt

and Measurement

schmidt.co

£)

actice

n interest and participate
cific learning.

unity of practice for
and data science.
sted on Twitter using the

visualizations using the

tions on Twitter <u>and</u> link

skr
byl
#Ti
Tw

con

his can be
participat
dyTuesda
itter-base
mmunity .

...developed...

...ting in the

...ay

...ed

0 10

Communicat

geom_label(), geom_curve(), str_wr

Data Cleaning/Wr

mutate(), filter(), case_

Data Visualiza

theme(), geom_path(), aes(),

Qual

...ion

...rap(), arrow()
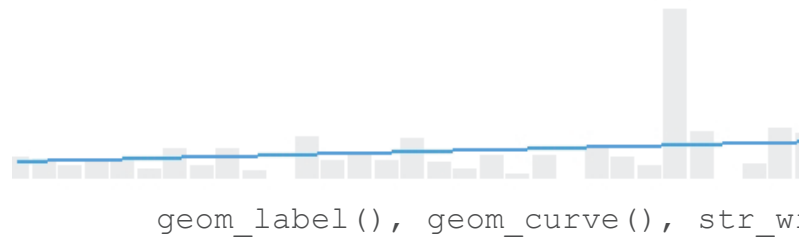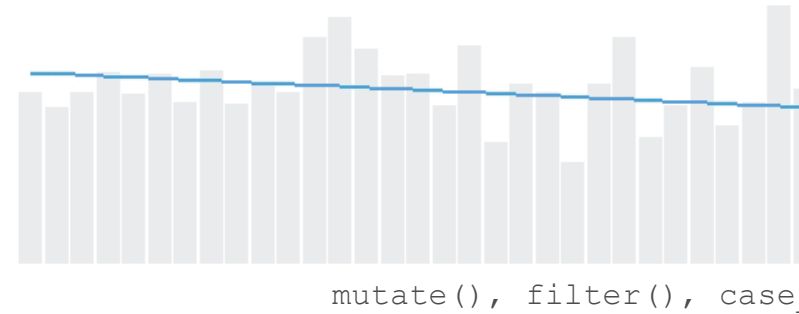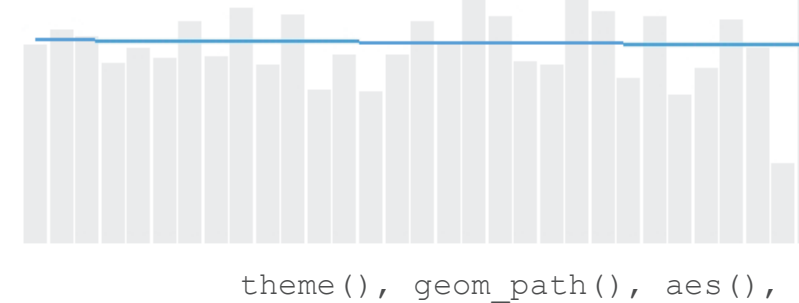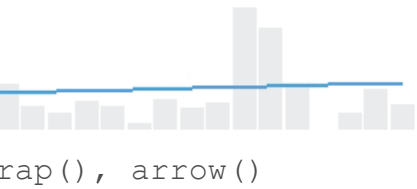
Functions related to communication are increasing.

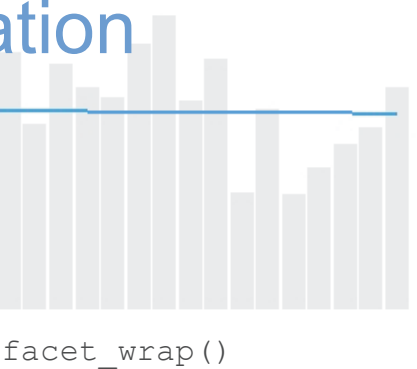Annotations and a focus on textual elements increasingly become important.

...rangling

..._when()
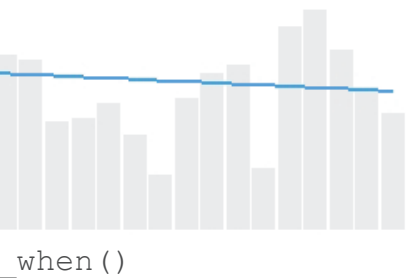
Data cleaning, wrangling, and munging functions decrease.

This suggest less processing of data or using more efficient code.

...ation

...facet_wrap()

Functions related to visualization remain mostly constant.

This suggests visualizations remain the primary focus of coding.

...itative Analyses

R

- R is a statistical programming language that is used for data analysis, modeling, machine learning, and visualization.

```
library(ggp
iris %>%
  ggplot()
  geom_poin
    (x=Sepa
     y=Sepa
```

Research Purpose

- Understand how #TidyTuesday functi community of practice.
- Determine if participation in #TidyTu visualization skill development.

Data / Methods

- Data: One year of #TidyTuesday twee tidytuesday.rocks app
- Descriptive analytics (frequencies, me
- Qualitative thematic coding
- Content analysis of code contribution

Descriptives

4,418   Total number of Tweets

- C

  m

- Vi

  re

2,428   Unique contributions

```
plot2)
         +
     t(aes
l.Length,
al.Width))
```
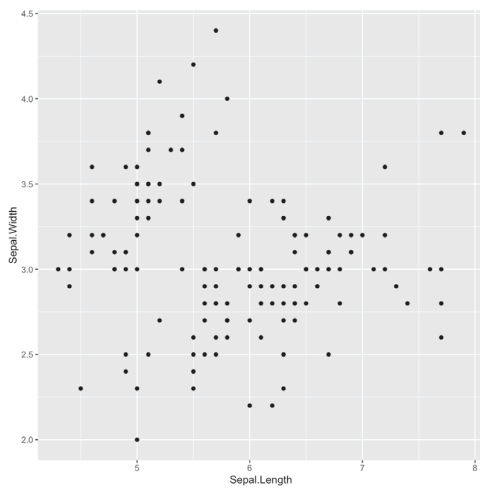

ions as an online

esday can foster data

ts from

eans, correlations)

s (using TidyCode)

Findings

ode becomes **longer**, using
ore functions, over time.
isualization functions
emain constant.

Density
0.03

0.02

0.01

0.00

#TidyTuesda

Age distribution of Amusement Park injuries in Texas

There is a peak in injuries among children aged around 10

From age 35 onwards, injuries sharply decrease (probably attendance to amusement parks too!)
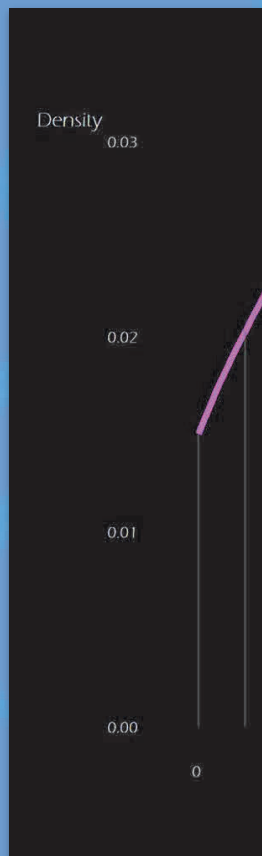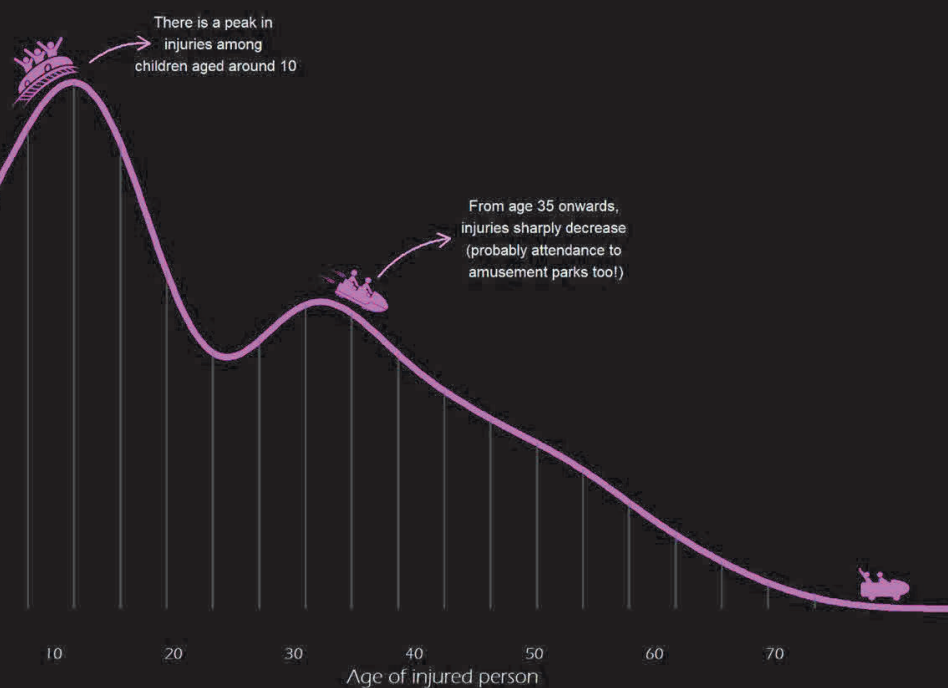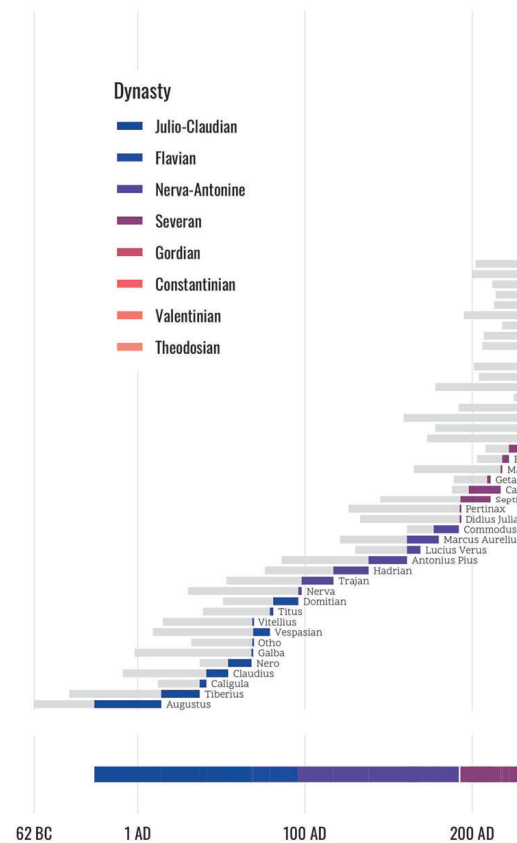
Age of injured person

#tidytuesday by @ariamsita, data: data.world

When In Rome: The Game Of Imperial Throne

Illustrated below is a timeline of the life and reigns of
The light grey bar depicts the liftime of the emperor,
the duration of their reign. An overall timeline by dyr
Unsurprisingly, the majority of emperors reign endin

Dynasty
- Julio-Claudian
- Flavian
- Nerva-Antonine
- Severan
- Gordian
- Constantinian
- Valentinian
- Theodosian

M
Geta
Ca
Septi
Pertinax
Didius Julia
Commodus
Marcus Aureliu
Lucius Verus
Antonius Pius
Hadrian
Trajan
Nerva
Domitian
Titus
Vitellius
Vespasian
Otho
Galba
Nero
Claudius
Caligula
Tiberius
Augustus

62 BC   1 AD   100 AD   200 AD

y Examples
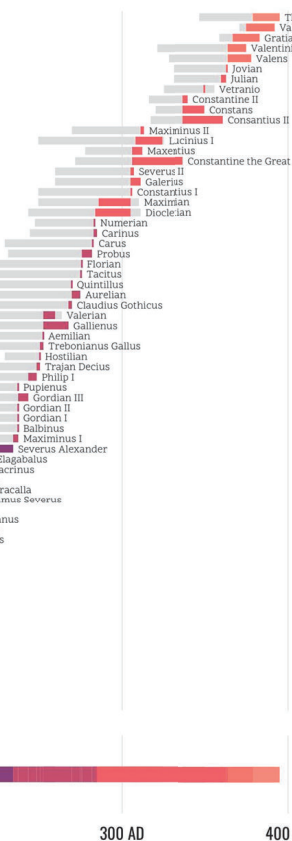
Nuclear explosions, 1945-1998

Total number of nuclear explosions by year, grouped by country and by atmospheric or und
type, as defined in the sipri report (2000)

Atmospheric

Partial Test Ban Treaty (1963)
Bans nuclear weapon tests in the atmosphere,
in outer space and under water

Comprehensive Nuclear-Test-Ban Treaty (1996)
Bans all nuclear explosions,
for both civilian and military purposes,
in all environments

China United Kingdom
France USA
India Soviet Union
Pakistan

Underground
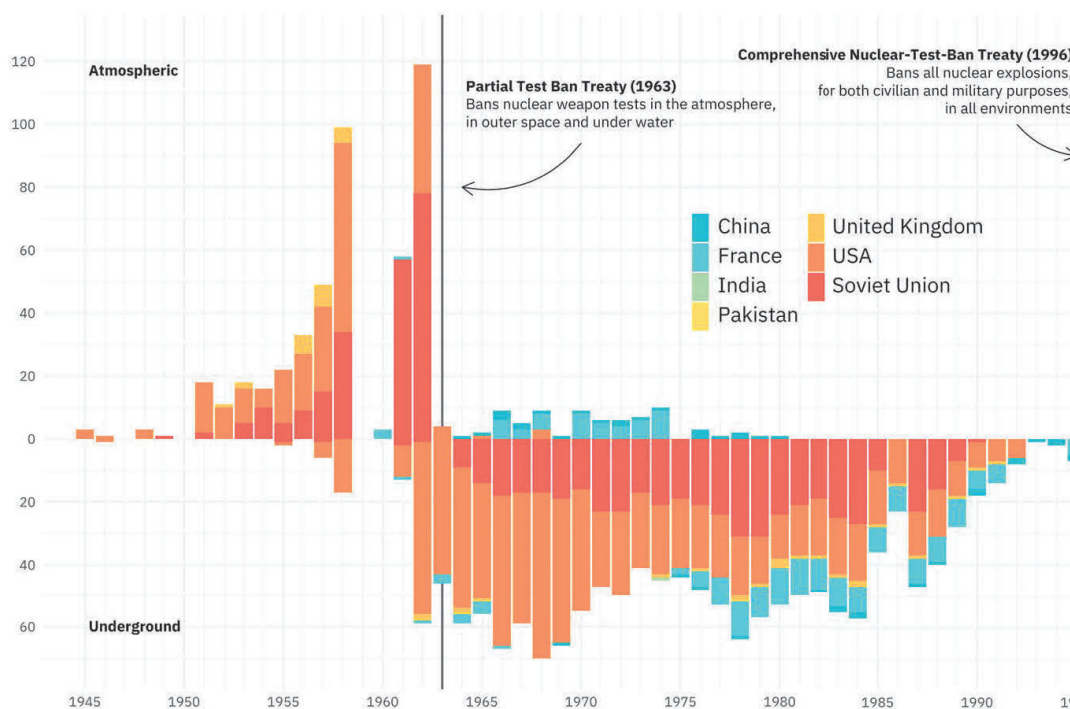
Source: Stockholm International Peace Research Institute | Graphic: Ge

Major Themes in Tweets

Emotion – 5%
Learning – 7%

Sharing Visualization – 10

Engaging Community – 10

Discussing Findings In Dat

Explaining Visualization –

N = 150

erground

995

eorgios Karamanis
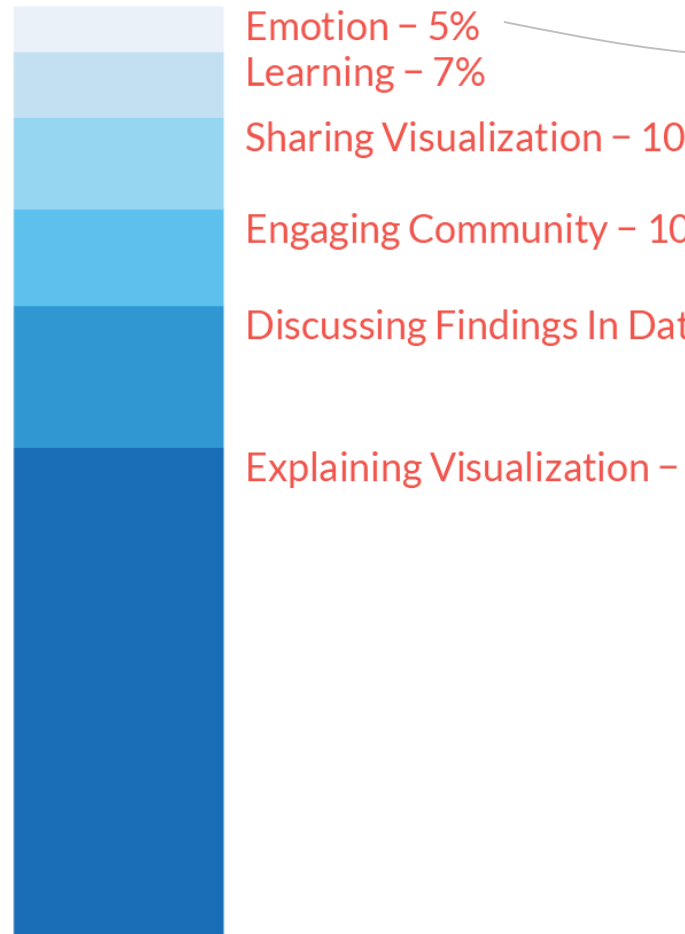
This thread highlights one of my favo
parts of #TidyTuesday, spontaneous
friendly collaboration!!

Sara continues to do awesome deep
on the #tidytuesday datasets! Lots o
content here and on her GitHub.
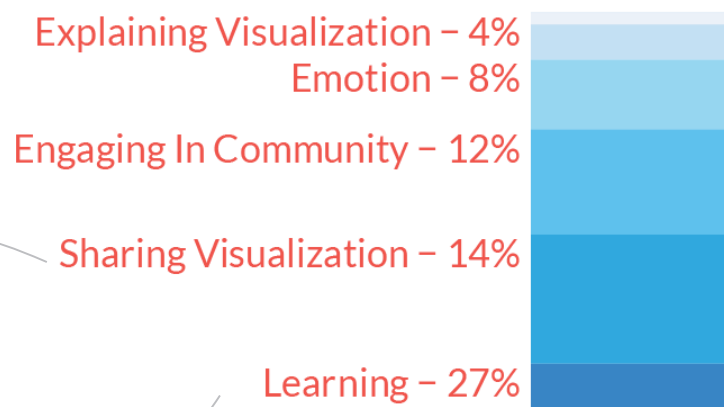
Thanks to #TidyTuesday for making

%

%

a – 15%

52%

First ever #TidyTuesday submission! I've always been reluctant to share, because I see so many great submissions. But today I've done some new cool things (first time I made a map in R!!) and I'm quite happy with the result.

My #tidytuesday viz this week is inspired by awesome people who make scarves from data. Can't knit, but I can geom_linerange

#tidytuesday I used to think all nobel peace price winners are grey old guys. Turns out, they are getting younger compared to the other folks.

#TidyTuesday turned into tidy Saturday this week! Looking at plastic pollution around the globe. I made a waffle chart of the five countries with most mismanaged plastic waste.

Major Themes in Retweets

orite
and

Explaining Visualization – 4%
Emotion – 8%

dives
f great

Engaging In Community – 12%

Sharing Visualization – 14%

Learning – 27%

800 Unique contributors

46.7 Mean contributions per week

5.93 Mean contributions per user

.95 Correlation between likes and retweets

ata wrangling functions

ecrease.

ommunication functions

ave **increased**.

itial *tweets* **explain**

sualizations, describe

ndings, or engage the

ommunity.

etweets **celebrate** others'
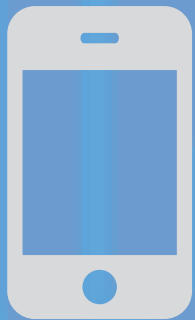
ork, discuss learning, or

are others' visualizations

nd resources.

onclusion: Participation can

elp develop and improve

sualization skills.

Take a picture to
download the **poster** and **short pap**

upp

er

Thanks to #TidyTuesday for making
easy to learn/use #R and to get to kn
others! #grateful

Jake's dataviz work as part of #TidyT
is always beautiful. This one is especi
great. Check out the legend he made
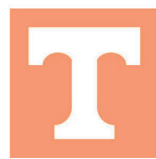ggplot.

Making
Data Sci

RESEARCH

t so
ow

Tuesday
ially
using

Celebrating Others – 35%

N = 150

ence
count()

GROUP

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE