Raincloudplots 2.0. A robust and transparent data visualisation tool

Rogier A. Kievit
Professor of Developmental Neuroscience
rogier.kievit@radboudumc.nl
How, and why, do cognitive abilities change across the lifespan?

(SEM) Model based developmental cognitive neuroscience
Today

• Background
  • Data visualisation and (open) science
• What are Raincloudplots?
• History and development
• Future plans (NWO Open Science Fund)
Done well, data visualization is the single most powerful way scientists can communicate messages broadly.
Nightigale (1858)

Whewell (1836)

Snow (1854)
However:
There are lot of bad graphs

Very, VERY bad graphs
Too much information
Too little information...

Figure 1 Empirical coverage of CIs for the relative-risk parameter $\beta$ of haplotype 01100. Results are based on 10,000 simulated data sets with the same haplotype frequencies as the FUSION data. Haplotype 01100 has a multiplicative effect on disease risk, with $\beta = 0.35$. 
Ugly
The ‘extra’ category

The Bicycle of Education
Our goal:
Improve ‘everyday’ data visualisation
Data visualization should **show the data, not hide it**

#barbarplots

The goal: A single plot that has all these benefits
Raincloudplots
What are raincloud plots?

[...]“raincloud plots” can visualize raw data, probability density, and key summary statistics such as median, mean, and relevant confidence intervals in an appealing and flexible format with minimal redundancy.[...]
History (2018-)

Micah Allen @micahgallen - Mar 15, 2018
Now I'm getting somewhere!

Tom Rhys Marshall @TomRhysMarshall
I wrote (something like) a matlab implementation of @neuroconscience's 'raincloud plots'.
git.fmrrib.ox.ac.uk/marshall/publi...

Introducing Raincloud Plots!

COGNITIVE NEUROSCIENCE / OPEN SCIENCE

MARCH 15, 2018  MICAH  12 COMMENTS

The neural determinants of age-related changes in fluid intelligence: a pre-registered, longitudinal analysis in UK Biobank [version 2; peer review: 3 approved]
Rogier A. Kievit, Delia Fuhrmann, Gesa Sophia Borgeest*, Ivan L. Simpson-Kent*, Richard N. A. Henson
From tweets to papers

Micah Allen: R
Kirstie Whitaker: Jupyter Notebooks / github
Davide Poggiali: R
Tom Rhys Marshall: Python
Matlab

@rogierK can I reuse this joke?
Davide Poggiali: do your local repos have a /figs/tutorial_matlab folder in them?
Tom Rhys Marshall: mine not
Davide Poggiali: argh
Tom Rhys Marshall: so I guess I need to 'add'?
(heya – I’m online if I can help w git)
Kirstie Whitaker: yes please Kirstie
in my local version of the repo there's a folder /figs/tutorial_matlab with all the figures in it for the matlab tutorial
The paper

• Raincloud plots rationale
• Walk-through tutorial in 3 languages:
  • R
  • Python
  • Matlab
# Rainclouds with boxplots

```r
p6 <- ggplot(simdat,aes(x=group,y=score, fill = group, colour = group))+
gem::geom_flat_violin(position = 'position_nudge(x = .25, y = 0)',adjunct = 2, trim = FALSE)+
gem::geom_point(position = 'position_jitter(width = .15), size = .25')+
gem::geom_boxplot(aes(x = as.numeric(group)+0.25, y = score),outlier.shape = NA, alpha = 0.3, width = .1, colour = "BLACK")+
ylab('Score')+xlab('Group')+coord_flip()+theme_cowplot()+guides(fill 1 = FALSE, colour = FALSE)+
scale_colour_brewer(palette = "Dark2")+
scale_fill_brewer(palette = "Dark2")+
ggtitle("Figure R6: Change in Colour Palette")
ggsave('6boxplots.png', width = w, height = h)
```

```r
p9 <- ggplot(simdat,aes(x=group,y=score, fill = group, colour = group))+
gem::geom_flat_violin(position = 'position_nudge(x = .25, y = 0)',adjunct = 2, trim = TRUE)+
gem::geom_point(position = 'position_jitter(width = .15), size = .25')+
gem::geom_boxplot(aes(x = as.numeric(group)+0.25, y = score),outlier.shape = NA, alpha = 0.3, width = .1, colour = "BLACK")+
ylab('Score')+xlab('Group')+coord_flip()+theme_cowplot()+guides(fill 1 = FALSE, colour = FALSE)+
scale_colour_brewer(palette = "Dark2")+
scale_fill_brewer(palette = "Dark2")+
ggtitle("Figure R9: Complex Raincloud Plots with Facet Wrap")
ggsave('9facetplot.png', width = w, height = h)
```

**Figure R9: Complex Raincloud Plots with Facet Wrap**

<table>
<thead>
<tr>
<th>Group</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>high</td>
</tr>
<tr>
<td>Group2</td>
<td>low</td>
</tr>
</tbody>
</table>

![Rainclouds with boxplots](6boxplots.png)

![Complex Raincloud Plots with Facet Wrap](9facetplot.png)
Python

# Changing orientation
dx="group"; dy="score"; ort="v"; pal = "Set2"; sigma = .2
f, ax = plt.subplots(figsize=(7, 5))
pt.RainCloud(x = dx, y = dy, data = df, palette = pal,
             bw = sigma, width_viol = .5, ax = ax, orient = ort)

# Load in the repeated data
df_rep = pd.read_csv("repeated_measures_data.csv", sep="",
                    header=None)
df_rep.columns = ["score", "timepoint", "group"]

# Plot the repeated measures data
dx = "group"; dy="score"; dhue="timepoint"
orht="h"; pal="Set2"; sigma = .2
f, ax = plt.subplots(figsize=(12, 5))
pt.RainCloud(x = dx, y = dy, hue = dhue, data = df_rep,
             palette = pal, bw = sigma, width_viol = .7, ax = ax,
             orient = ort, alpha = .65, dodge = True, pointplot = True,
             move = .2)
Matlab

```matlab
% example 1
f0 = figure('Position', fig_position);
subplot(1, 2,1)
h1 = raincloud_plot(d1, 'box_on', 1, 'color', cb(1,:), 'alpha', 0.5,...
   'box_dodge', 1, 'box_dodge_amount', .15, 'dot_dodge_amount', .15,...
   'box_col_match', 0);
h2 = raincloud_plot(d2, 'box_on', 1, 'color', cb(4,:), 'alpha', 0.5,...
   'box_dodge', 1, 'box_dodge_amount', .35, 'dot_dodge_amount', .35,
   'box_col_match', 0);
legend([h1(1), h2(1)], {'Group 1', 'Group 2'});
title('A) Dodge Options Example 1')
set(gca, 'XLim', [-.055 .15]); box off

% example 2
subplot(1, 2, 2)
h1 = raincloud_plot(d1, 'box_on', 1, 'color', cb(1,:), 'alpha', 0.5,...
   'box_dodge', 1, 'box_dodge_amount', .15, 'dot_dodge_amount', .35,...
   'box_col_match', 1);
h2 = raincloud_plot(d2, 'box_on', 1, 'color', cb(4,:), 'alpha', 0.5,...
   'box_dodge', 1, 'box_dodge_amount', .35, 'dot_dodge_amount', .75,...
   'box_col_match', 1);
legend([h1(1), h2(1)], {'Group 1', 'Group 2'});
title('B) Dodge Options Example 2')
set(gca, 'XLim', [-.055 .15]); box off

% save
print(f7, fullfile(figdir, '7Rain5.png'), '-dpng');

Figure M10
Repeate measures raincloud plot - some aesthetic options

Figure M7
A) Dodge Options Example 1
B) Dodge Options Example 2
```
The revision: New and improved

Inspired by the raincloud approach by @micahgallen et al., I wrote some #rstats code to flip rainclouds combined with intra-individual trends. Optimized scripts will be on my GitHub soon. Thanks all for the encouraging feedback!

1x1 & 2x2 RM designs
Raincloudplots: Impact so far

- >4 million reached on twitter
- Used in psychology, neuroscience, chemistry, meteorology, philosophy
- ~90,000 views
- ~700 github stars
Why did it succeed?

• They look good
• No overclaiming of novelty: A tool, not a discovery or invention
• They are almost always better than the default alternative
• ‘Maximal statistical information’ / ‘Inference at a glance’
• Accessible
  • 3 languages
  • Tutorials start simple

Our goal here is not to propose a totally novel invention, but rather to make a powerful visualization strategy freely, easily, and transparently available across commonly used platforms. To this end, similar but distinct plotting strategies include beanplots (Kampstra, 2008), estimation plots (Ho et al., 2018), pirateplots (Phillips, 2016), sinaplots (Sidiropoulos et al., 2018), stripcharts (Chambers, 2017), beeswarm plots (Eklund, 2016), and many others.

Ellison AM. Exploratory data analysis and graphic display. Design and Analysis of Ecological Experiments. 1993; 14–45.
Raincloudplots 2.0. A robust and transparent data visualisation tool

• Fantastic initiative
  • Bonus: Open Science track record (10%)
• 0.2 FTE: Jordy van Langen
• 0.6 FTE: Recruiting now

What is our plan?
Work Package 1

1.1
• Improve & extend R-package “raincloudplots”
  https://github.com/jorvlan/raincloudplots
• Community engagement (user-feedback)
• Submit to CRAN (R-package hosting server)

1.2
• Create ShinyApp “raincloudplots” to easily drag & drop data
Work Package 2 – Outreach & education

• Organize Workshops (hybrid format)
  • Step-by-step tutorials
• Break-out rooms: R/Python/Matlab/JASP
• Dates:
  • 30-sept-2022 (14:00 CET)
  • 11-nov-2022 (14:00 CET)
  • 17-feb-2023 (14:00 CET)
• Free Stickers!
Work Package 3 – JASP statistics

- Collaboration & integration
- Used by 226 universities
- 59 countries
- 30,000 monthly downloads
- Gateway to coding
# Timeline

<table>
<thead>
<tr>
<th>Task</th>
<th>1 - 3</th>
<th>3 - 5</th>
<th>5 - 7</th>
<th>7 - 9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-package improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WP1: research assistant</td>
</tr>
<tr>
<td>R-shiny development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WP1: research assistant</td>
</tr>
<tr>
<td>Organize virtual workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WP2: research assistant</td>
</tr>
<tr>
<td>Collaborate with JASP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WP3: research assistant</td>
</tr>
<tr>
<td>Collect user-feedback R-package</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Main applicant</td>
</tr>
<tr>
<td>Collect user-feedback Shiny app</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Write results report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- WP1: research assistant
- WP2: research assistant
- WP3: research assistant
- Main applicant
Raincloudplots 2.0. A robust and transparent data visualisation tool
Thank you!