DATA VISUALIZATION
making your figures pretty
by showing the patterns
and removing the clutter

Simone Des Roches
“The signal is the truth. The noise is what distracts us from the truth.”

Nate Silver
FiveThirtyEight

“Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space”

Edward R Tufte
The Visual Display of Quantitative Information
<table>
<thead>
<tr>
<th>time</th>
<th>trial</th>
<th>count</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>treatment</td>
<td>9.271</td>
<td>0.750</td>
</tr>
<tr>
<td>20</td>
<td>treatment</td>
<td>8.403</td>
<td>0.336</td>
</tr>
<tr>
<td>30</td>
<td>treatment</td>
<td>7.016</td>
<td>0.658</td>
</tr>
<tr>
<td>40</td>
<td>treatment</td>
<td>7.269</td>
<td>0.579</td>
</tr>
<tr>
<td>50</td>
<td>treatment</td>
<td>9.613</td>
<td>0.574</td>
</tr>
<tr>
<td>10</td>
<td>control</td>
<td>1.250</td>
<td>0.759</td>
</tr>
<tr>
<td>20</td>
<td>control</td>
<td>0.377</td>
<td>0.618</td>
</tr>
<tr>
<td>30</td>
<td>control</td>
<td>0.764</td>
<td>0.404</td>
</tr>
<tr>
<td>40</td>
<td>control</td>
<td>4.372</td>
<td>0.532</td>
</tr>
<tr>
<td>50</td>
<td>control</td>
<td>2.651</td>
<td>0.344</td>
</tr>
</tbody>
</table>

something added at time = 25
silencing the noise

- **remove box:** `plot(..., bty="n")`
- **axes cover all values:** `ylim=c(0, 10)`
- **text is upright:** `las=1`

**Base:**

- **remove grid:** `+ theme_classic()`
- **reduce digits:** `+ scale_y_continuous(breaks = seq(0, 10, 2))`

**Ggplot:**

- **remove box:** `plot(..., bty="n")`
- **axes cover all values:** `ylim=c(0, 10)`
- **text is upright:** `las=1`
silencing the noise

**solid points:** plot(\ldots, pch=16, \\
**shading:** rect(x1, y1, x2, y2, col="lightgrey", border=NA)

shading: + geom_rect(mapping=aes(x1, x2, y1, y2), fill="grey", colour="white")
silencing the noise

**solid points:** `plot(..., pch=16,`  

**colour:** `col = c("grey50", "black")[as.numeric(trial)]`  

**colour:** `+ scale_color_manual(values=c("darkblue", "firebrick"))`
silencing the noise

**text size:** plot(..., cex.axis=1, cex.lab=1.5)

**position text:** text(x,y,"text", srt=90, cex=0.8, col="grey50")

**position text:** geom_text(label="something added", size=4, x=24, y=5, col="grey50", angle=90)
library("plotrix")

remove error bar "caps": plotCI(..., sfrac=0)

remove error bar "caps": geom_errorbar(aes(ymin, ymax), width=0)

show data?

indicate significance?
remove repetitive legend*, add text:
```
text(x, y, "text")
```
*(or offload legend info to caption)

make legend order consistent:
```
+ guides(col=guide_legend(title=NULL, reverse=T))
```
FIGURE COHESION

• consistently assign the same colours, points, & line types to the same groups across different figures

• optimize plot arrangement, align axes & match scales for multi-panel figures
WHAT ARE YOUR FIGURES FOR?

i.e. how will they be viewed?

presented

BIG TEXT
thick lines
CLEAR LEGEND
colour

printed*

OPTIMIZE ARRANGEMENT
legend vs caption?
BLACK & WHITE FRIENDLY

*or read digitally
RESOURCES

Tufte, Edward. 1983. The Visual Display of Quantitative Information

Wilke, Claus O. Fundamentals of Data Visualization
http://serialmentor.com/dataviz/