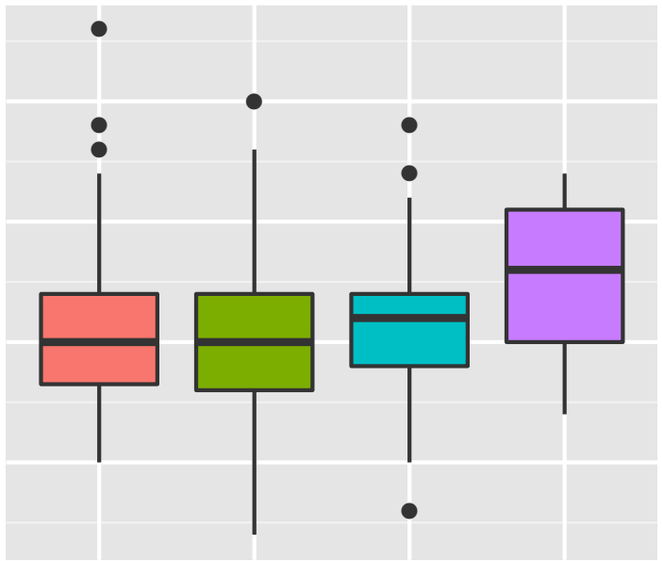
Statistics on Street Corners

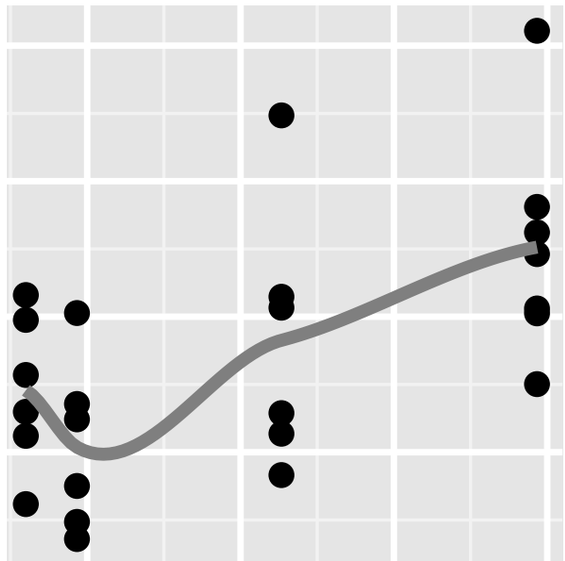
Using Crowd-Sourcing to Conduct Statistical Inference with Exploratory Data Analysis

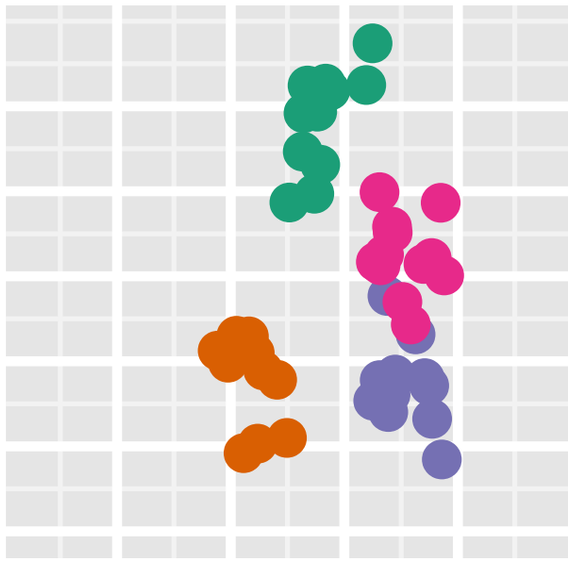
Di Cook
Econometrics and Business Statistics,
Monash University

Joint work with Heike Hofmann, Mahbub Majumder, Niladri Roy Chowdhury, Hadley Wickham, Andreas Buja, Debby Swayne, Eun-kyung Lee, Yifan Zhao, Tengfei Yin, Lendie Follett, Eric Hare, Adam Loy, Susan Vanderplas, Nathaniel Tomasetti

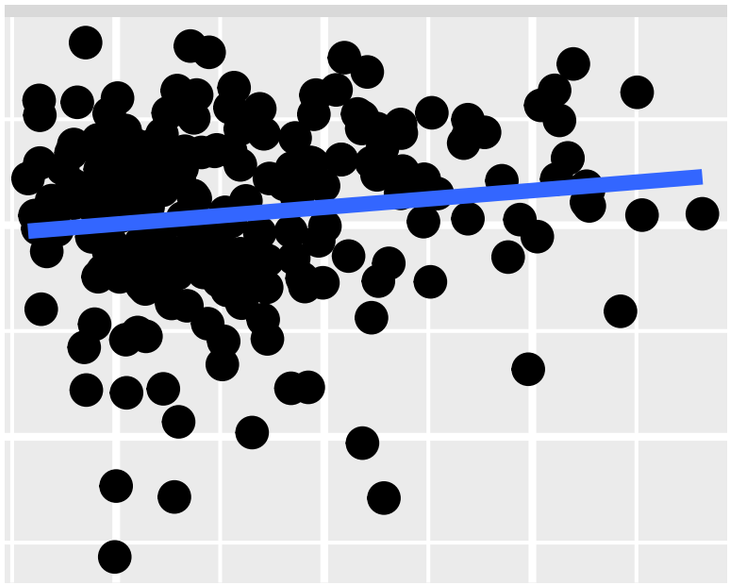
Biometrics by the Border NOV 2017

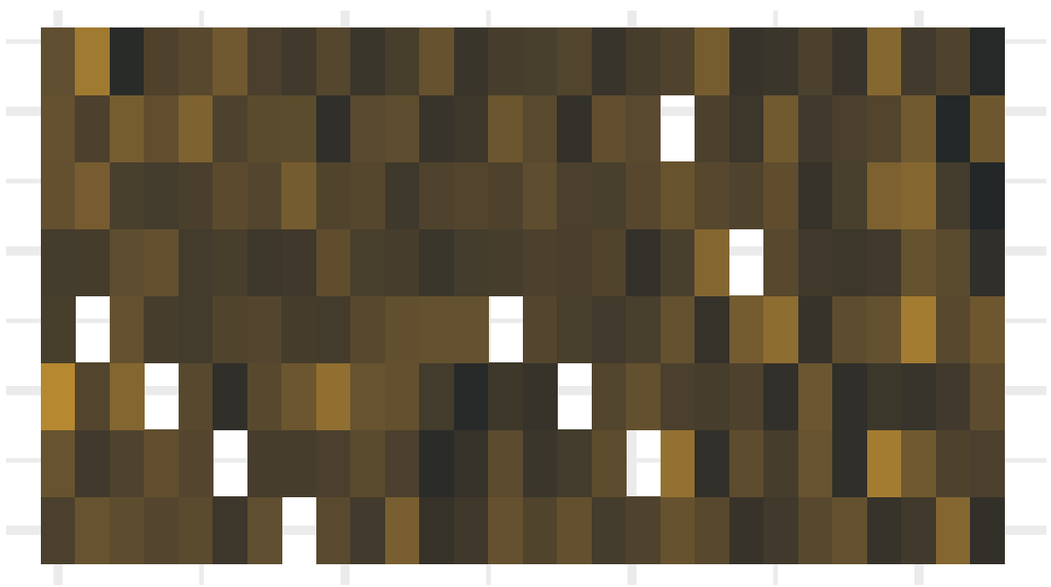


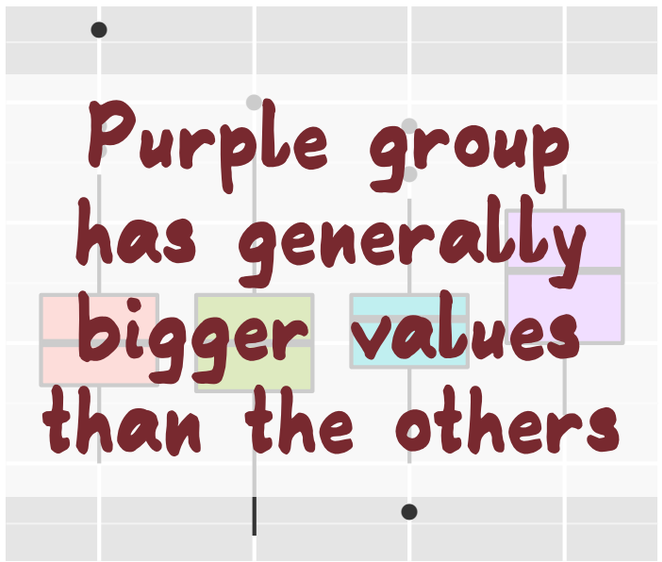


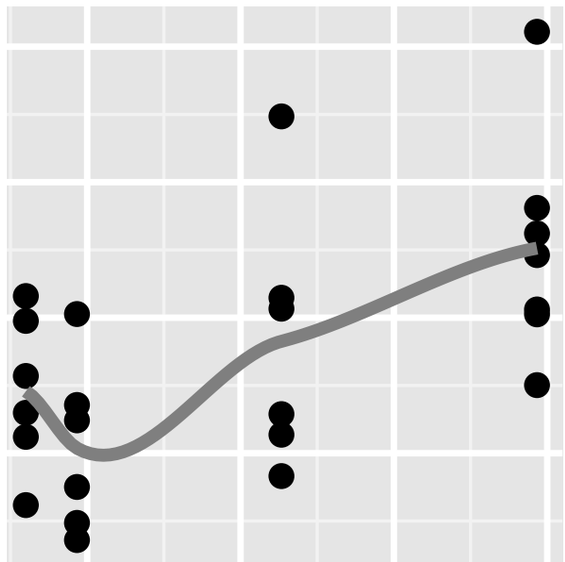


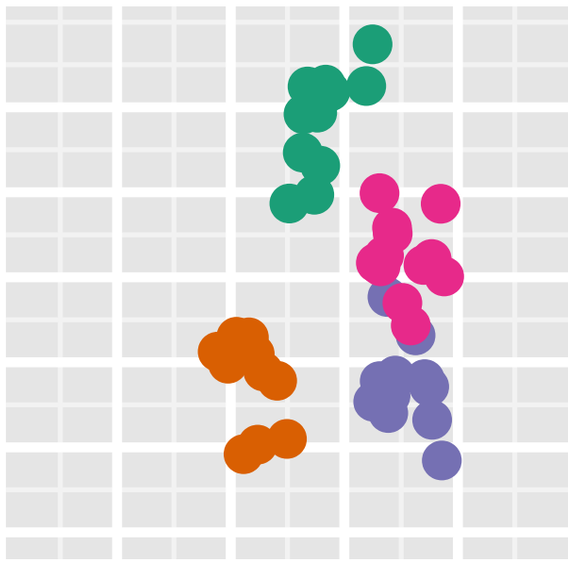
What do you see?



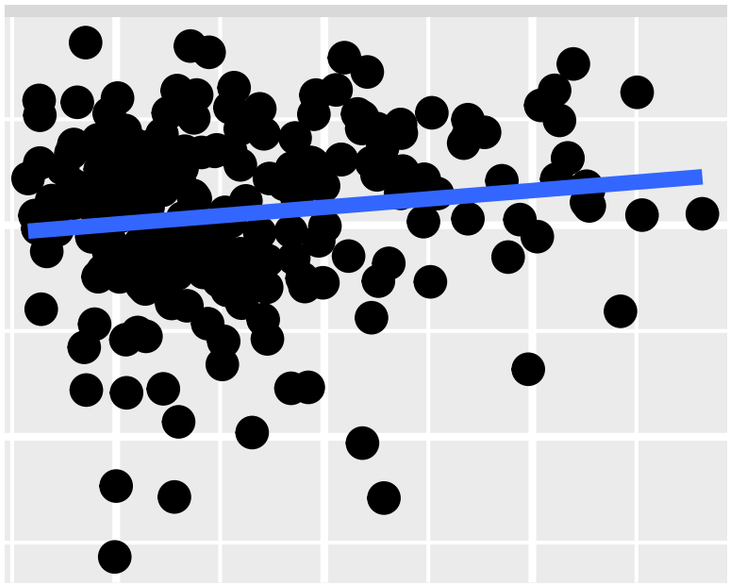


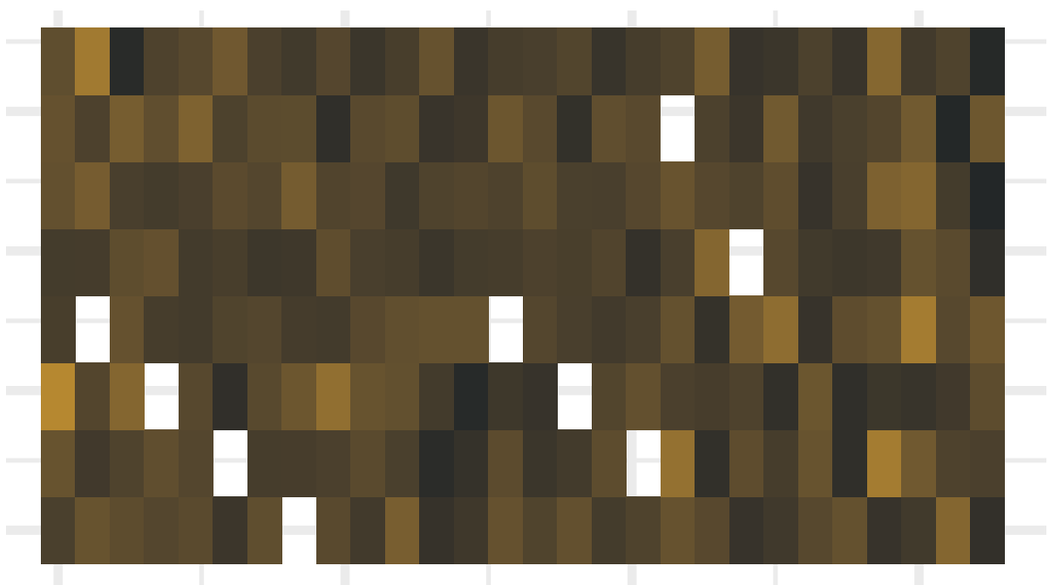


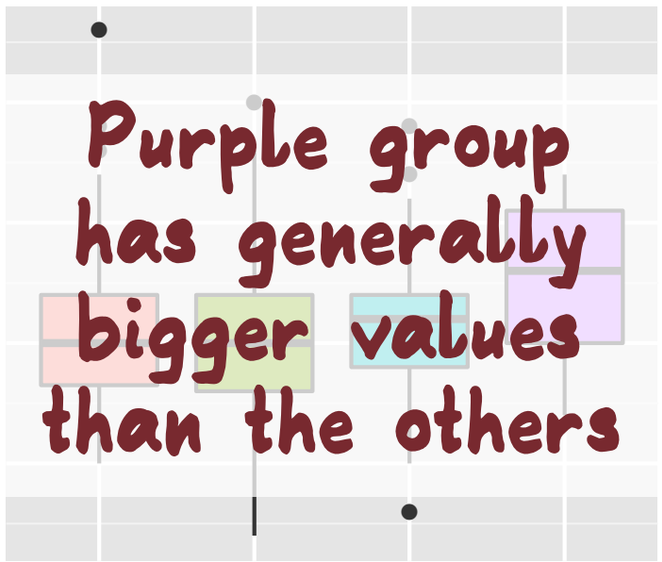


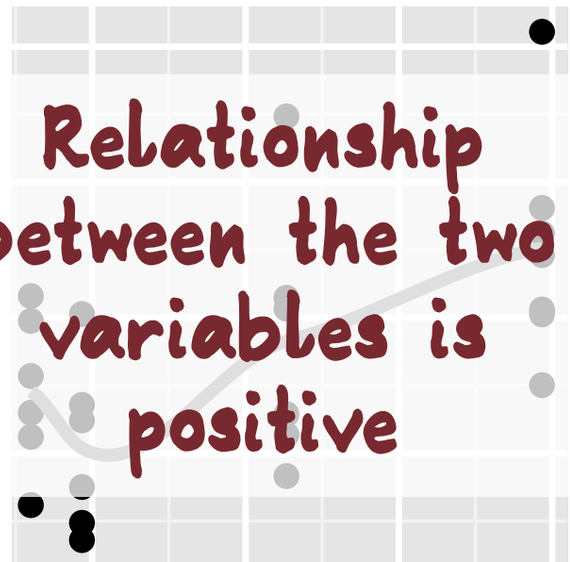


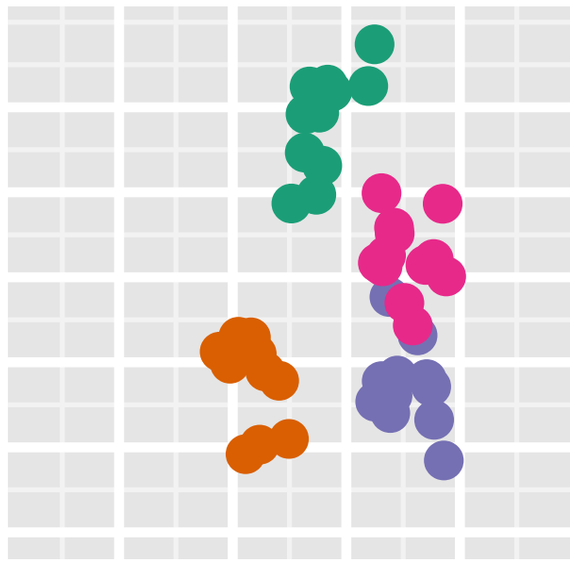
What do you see?



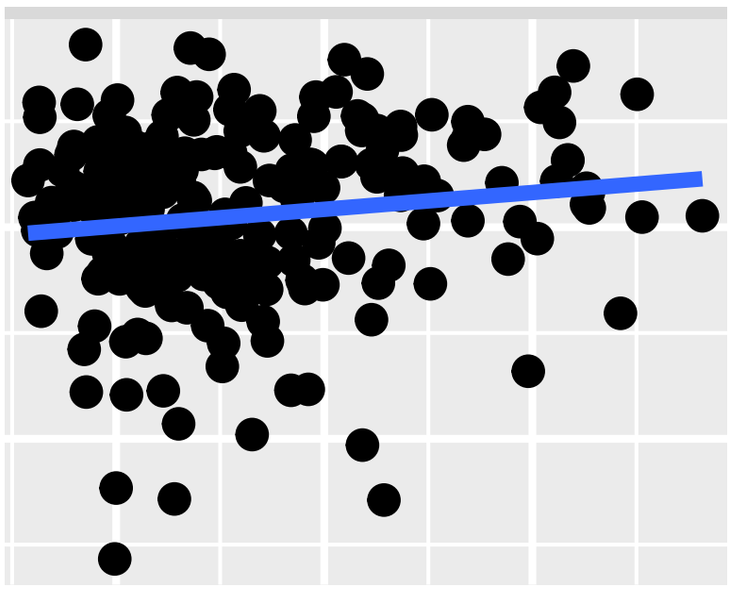


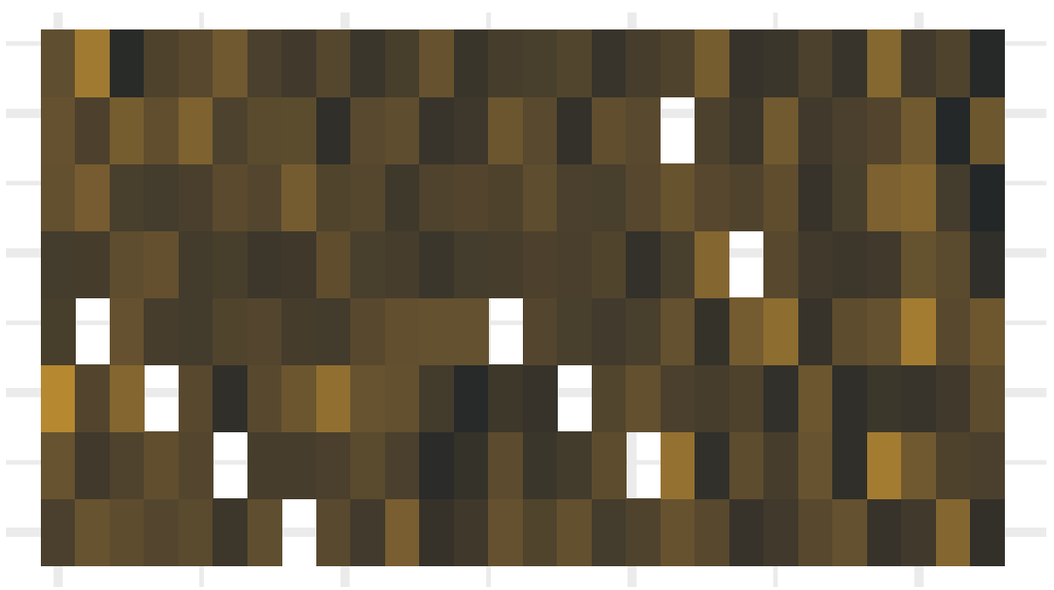


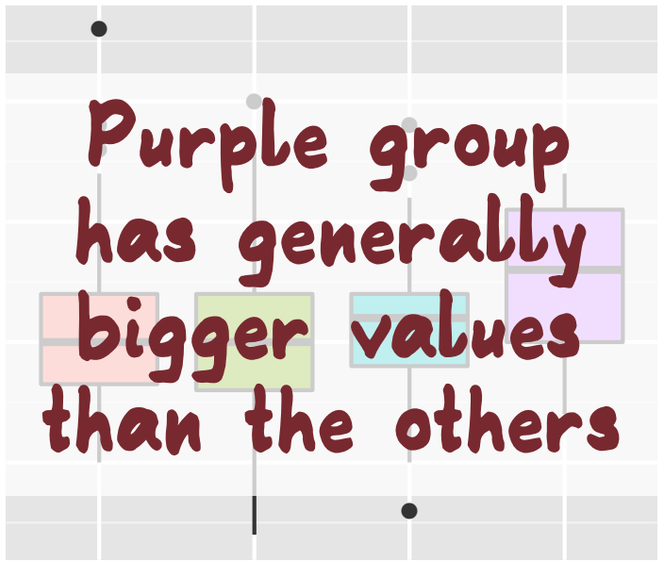


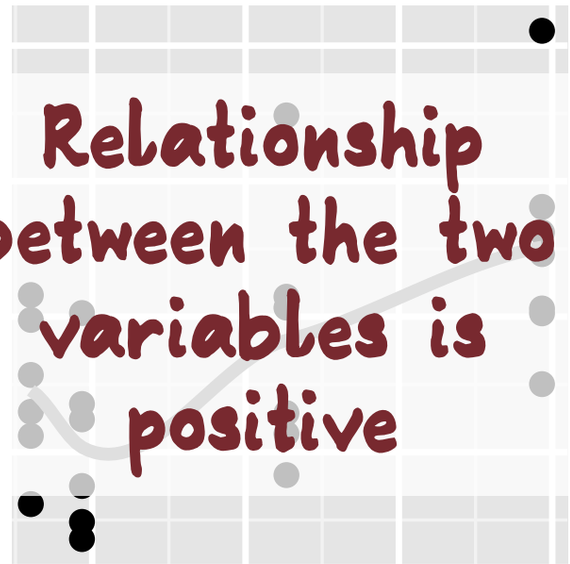


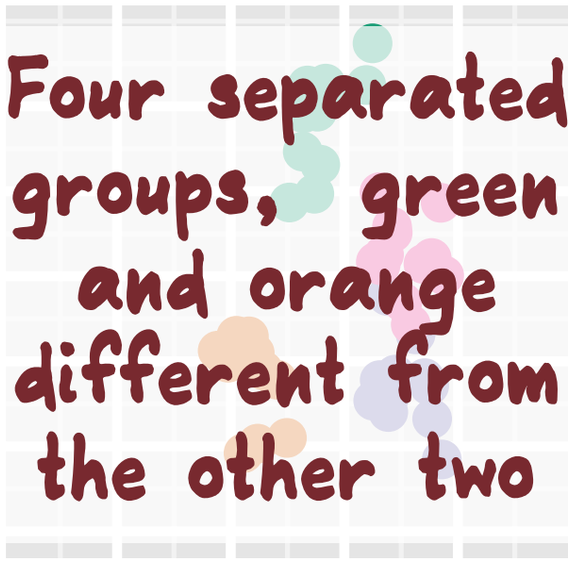
What do you see?



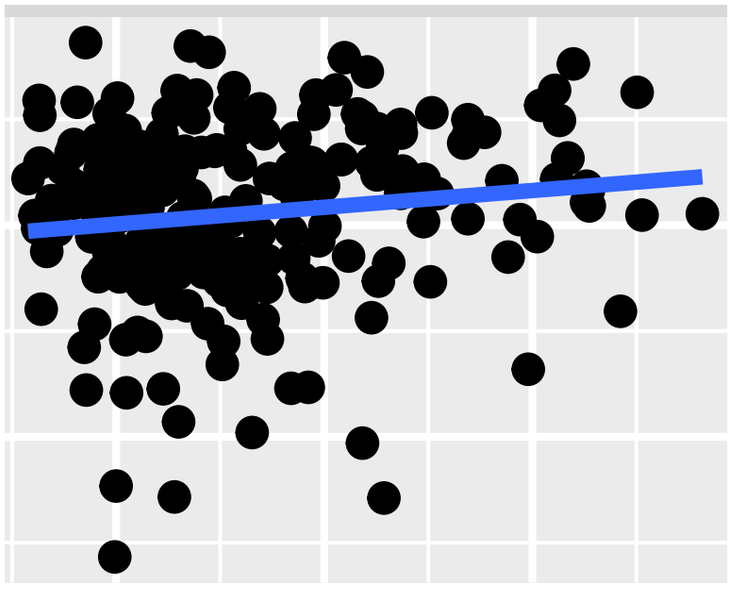


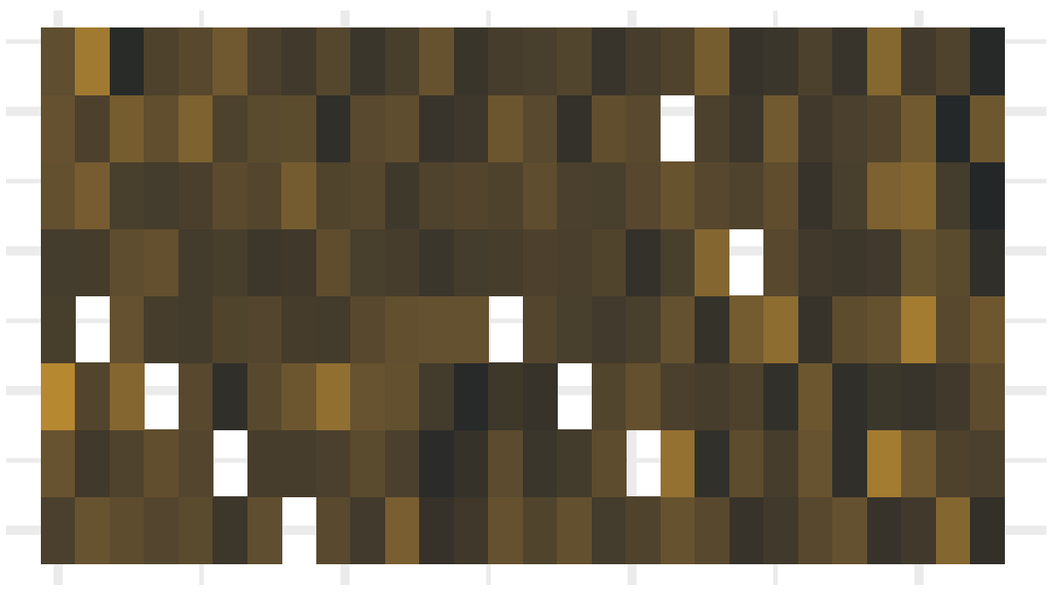


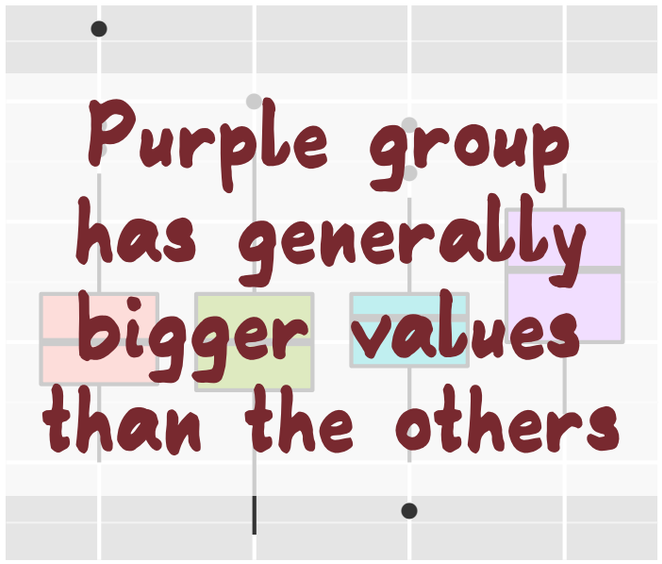


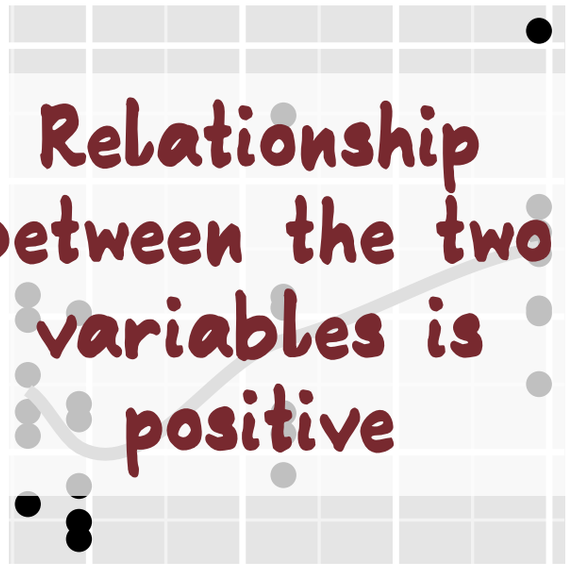


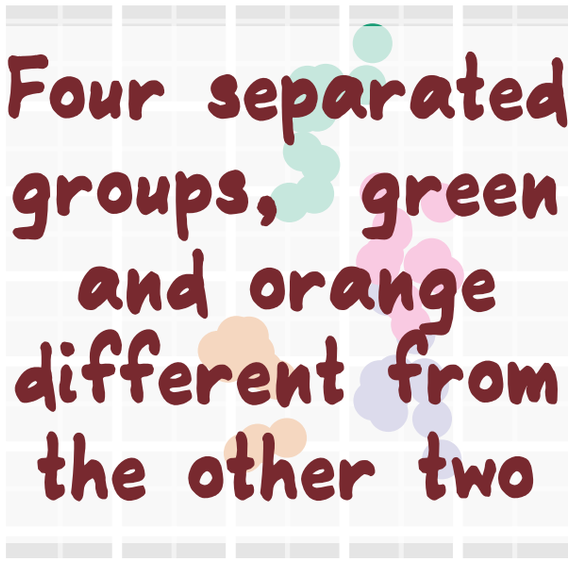
What do you see?



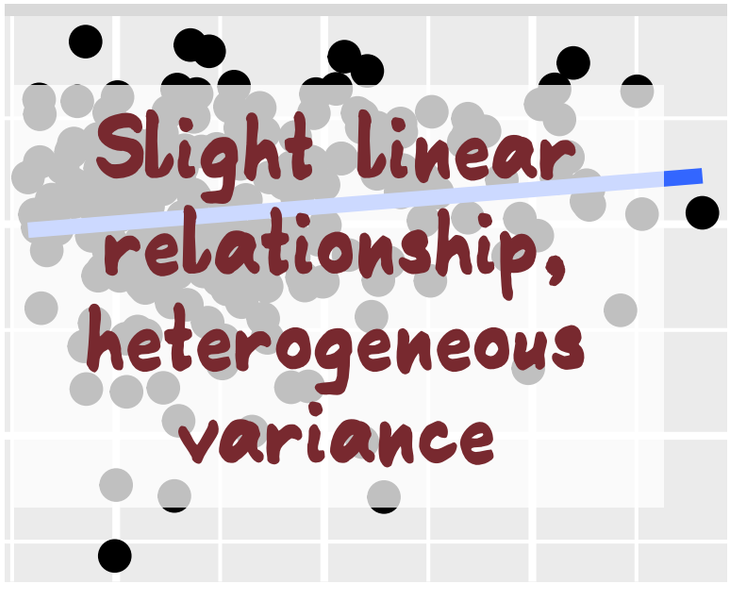


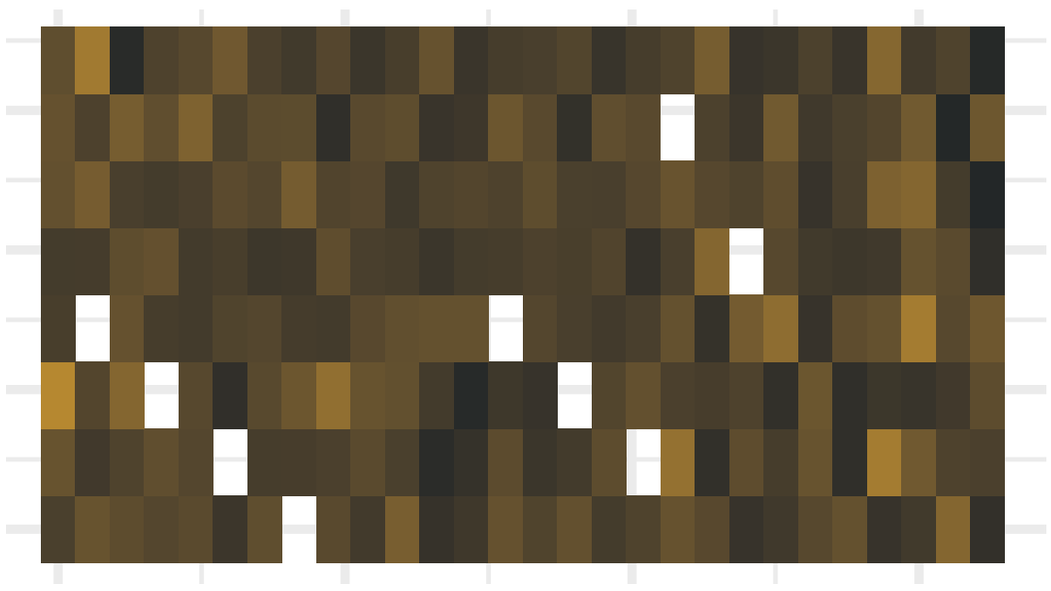


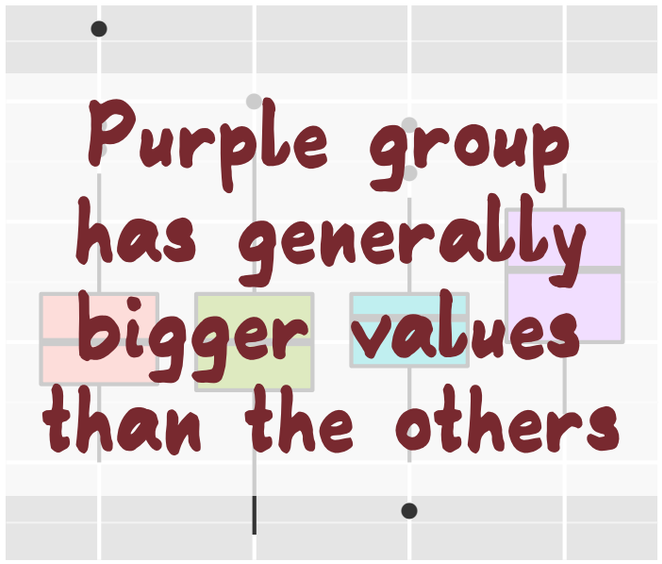


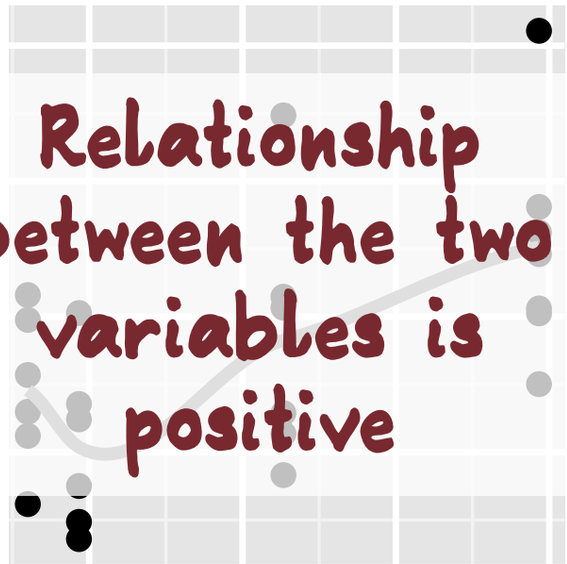


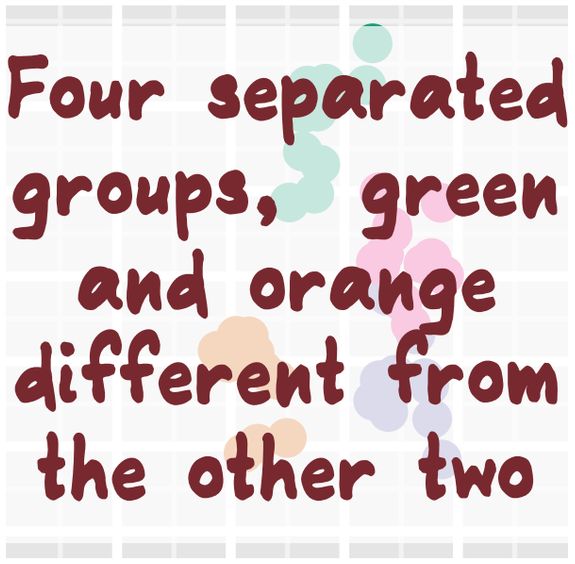
What do you see?



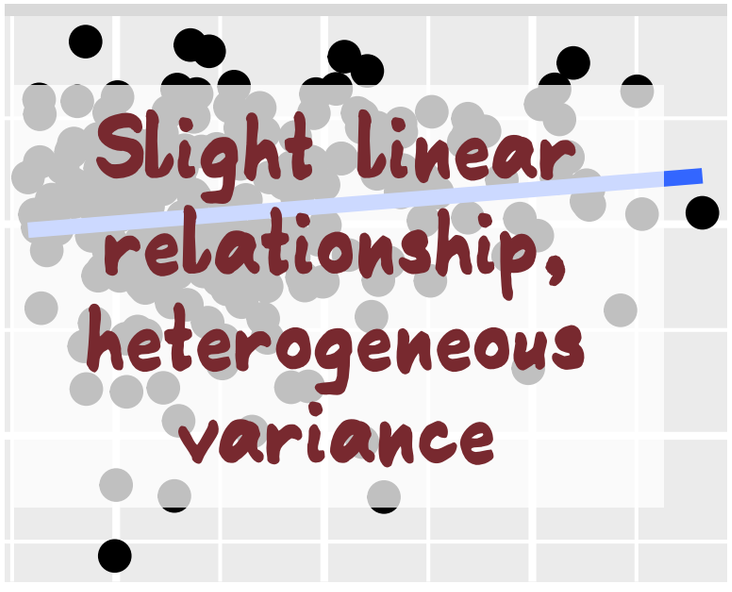








What do you see?



Statistical graphics

- Data plots are integral to our field
- Plots encourage curiosity and discovery when most of statistics is about being skeptical.
- This work bridges the gap between inferential and exploratory statistics, permitting curiosity with skepticism.

Can we believe what we see?
How can we see what we don't see?

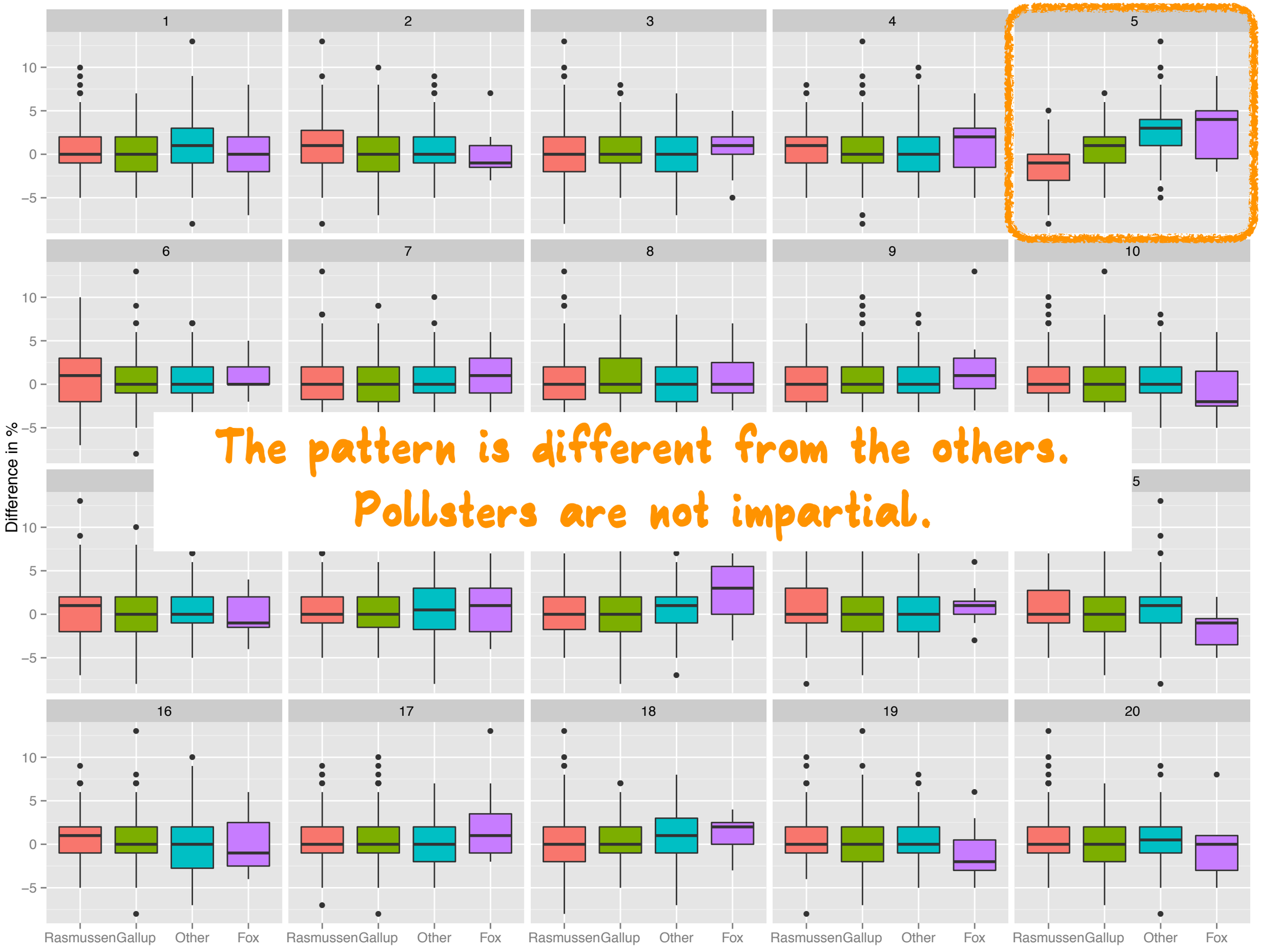
Idea

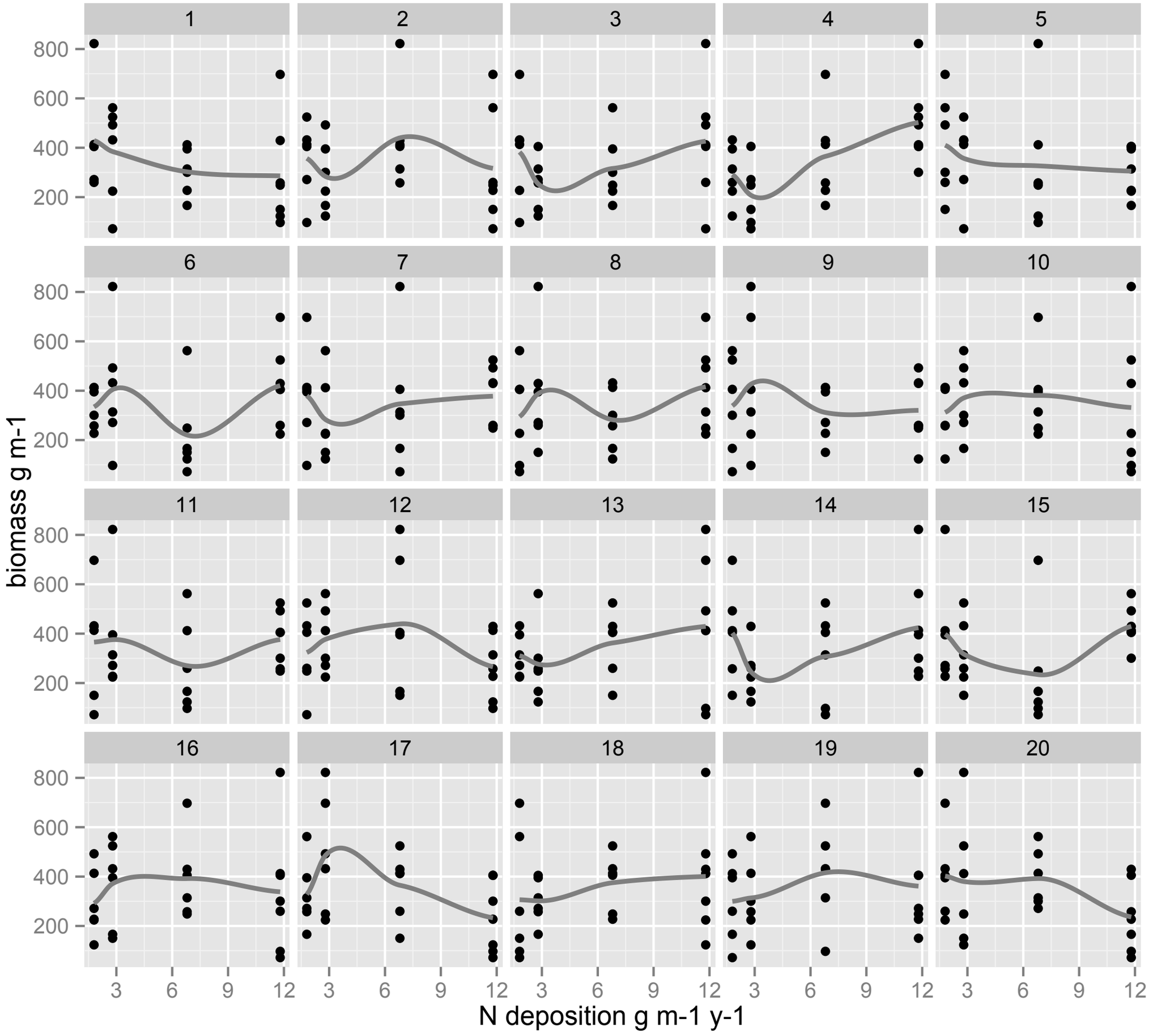
Instead of looking at the data plot in isolation, place it in the context of plots of data where there is nothing happening.

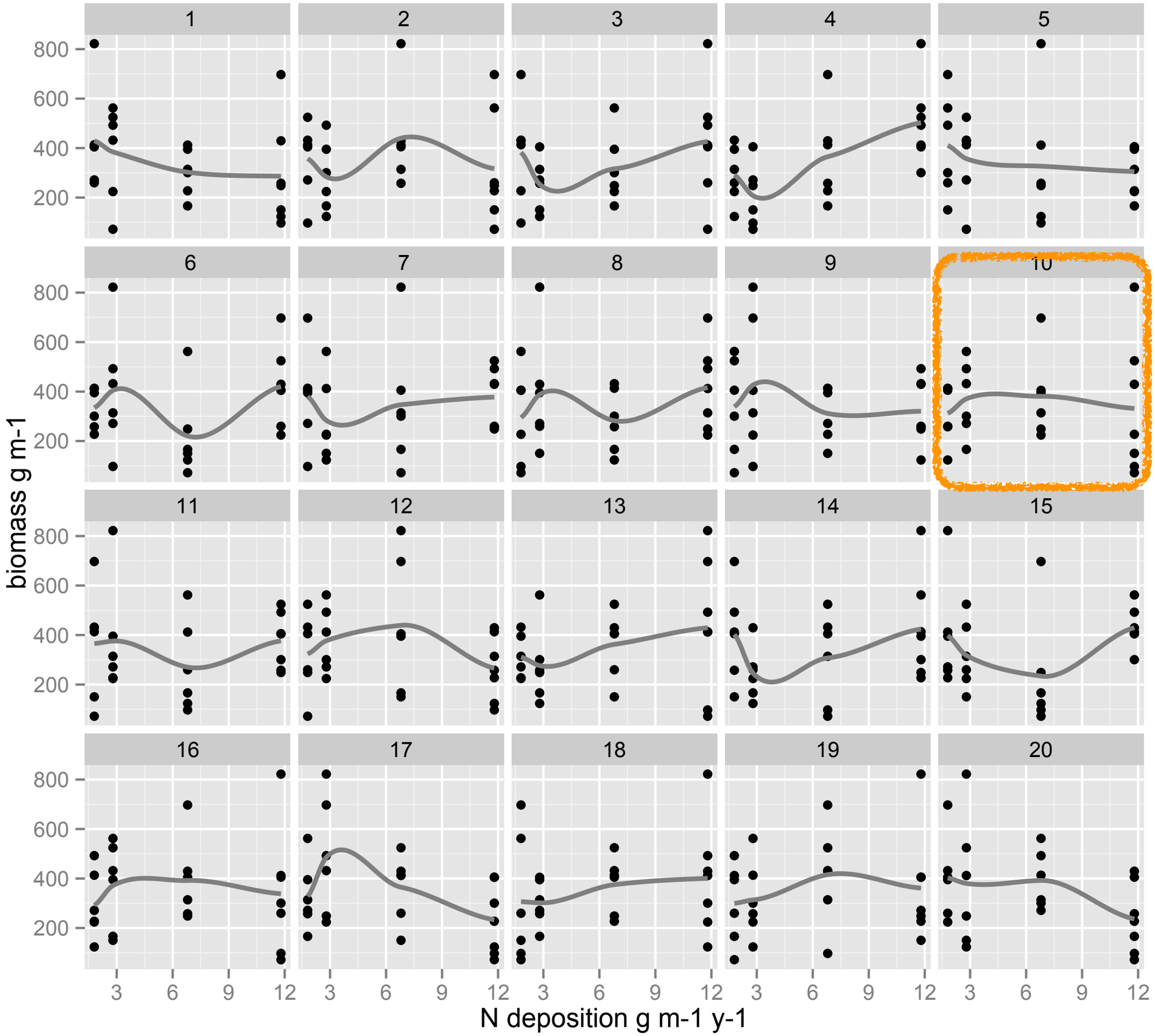
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
10
5
0
−5
Difference in %
Rasmussen
Gallup
Other
Fox

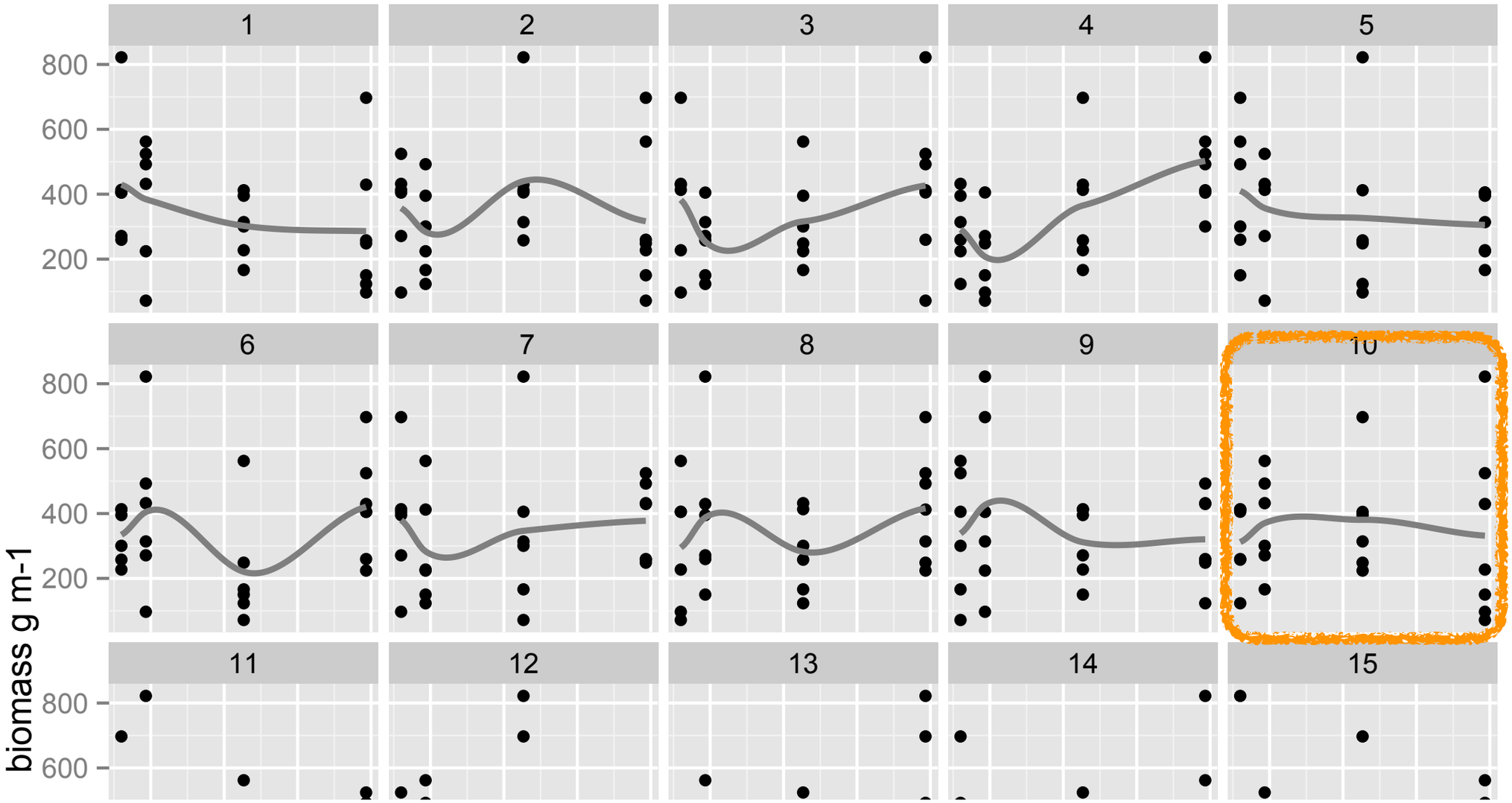
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
10
5
0
−5
Difference in %
Rasmussen
Gallup
Other
Fox

The pattern is different from the others.
Pollsters are not impartial.

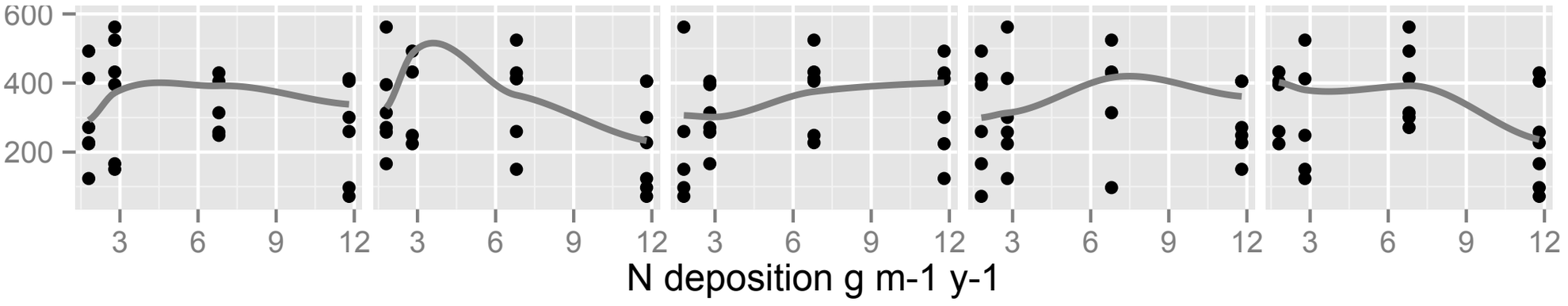


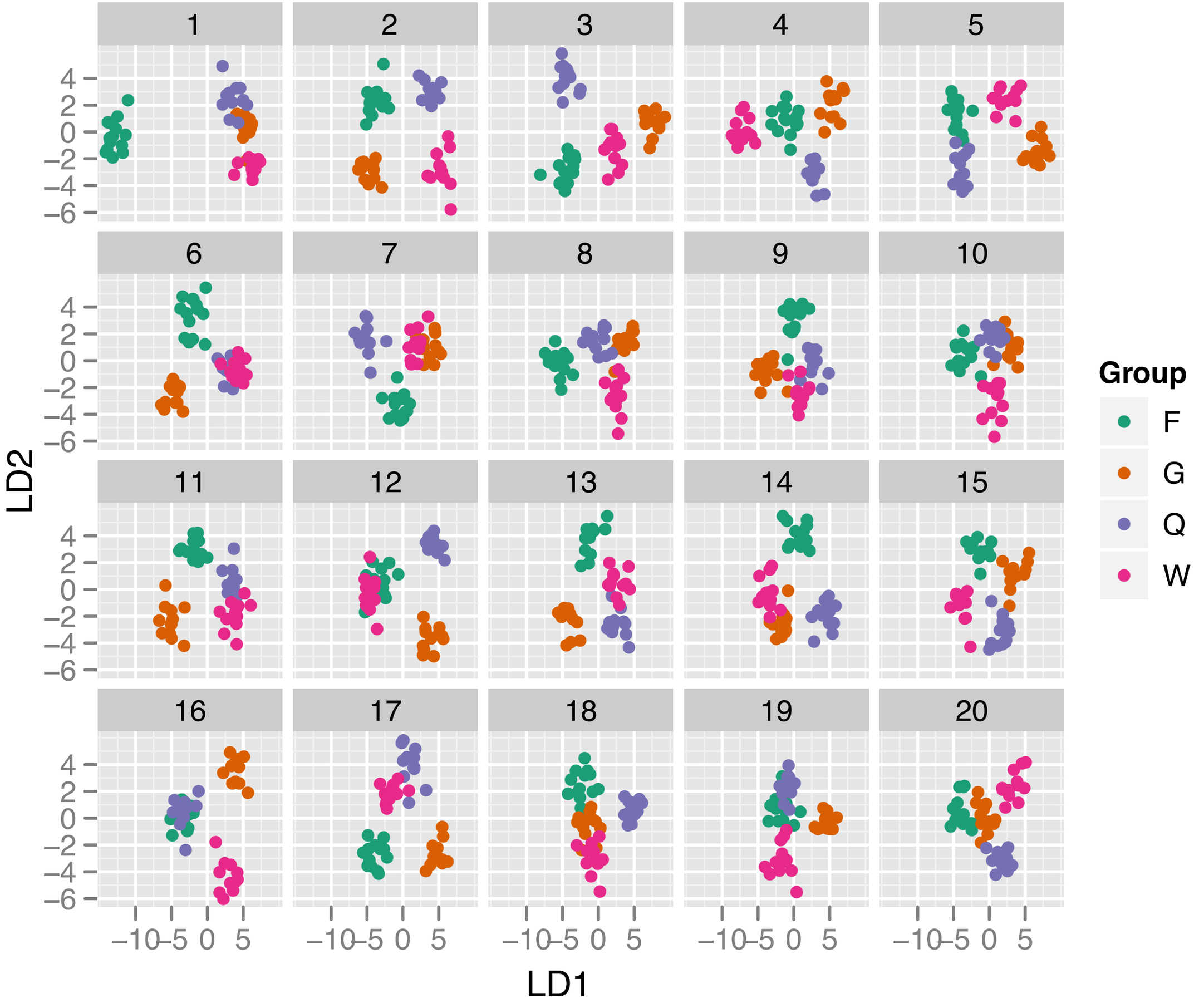

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
800
600
400
200
3
6
9
12
biomass g m-1
N deposition g m-1 y-1

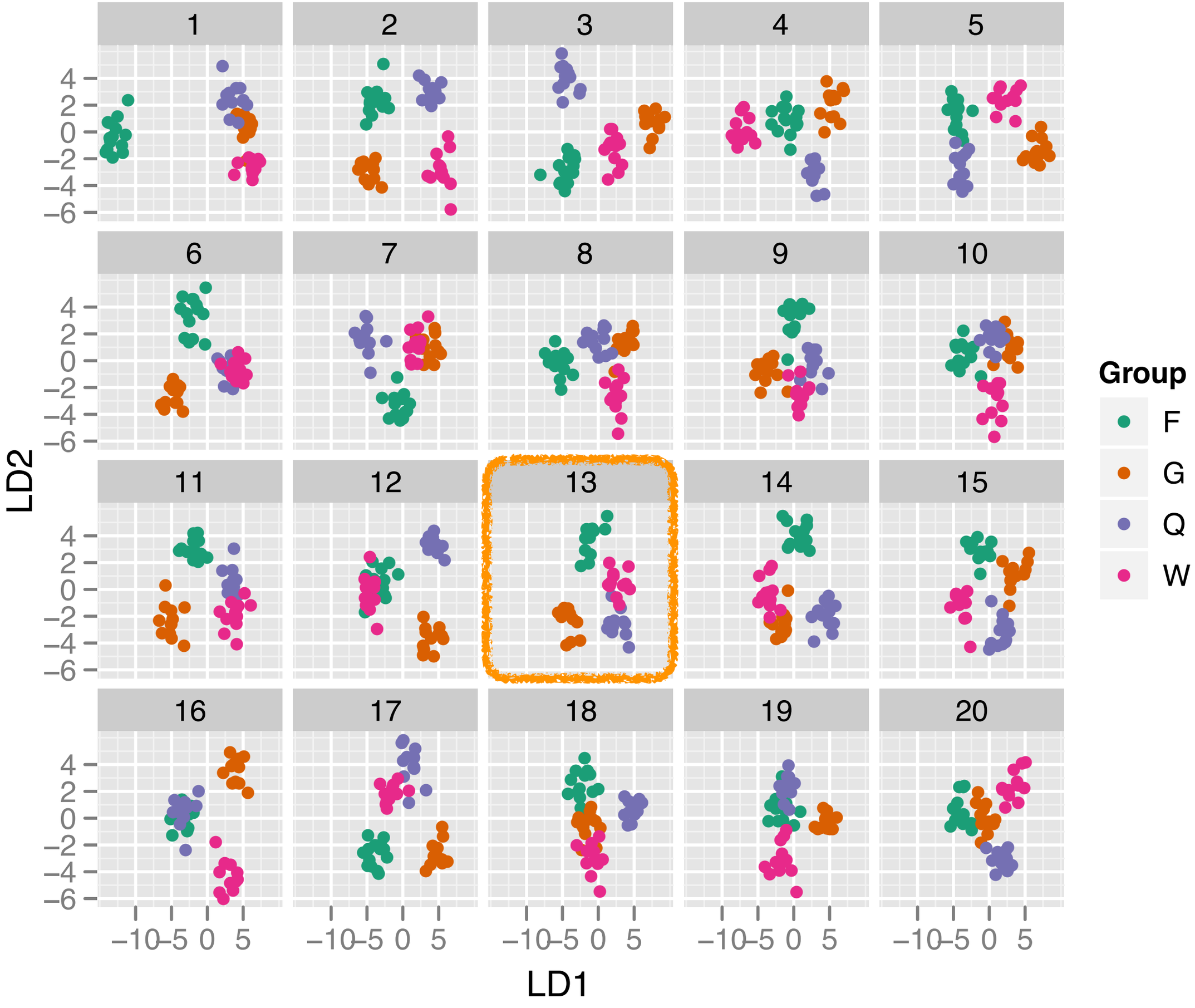

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
800
600
400
200
biomass g m-1
3
6
9
12
N deposition g m-1 y-1



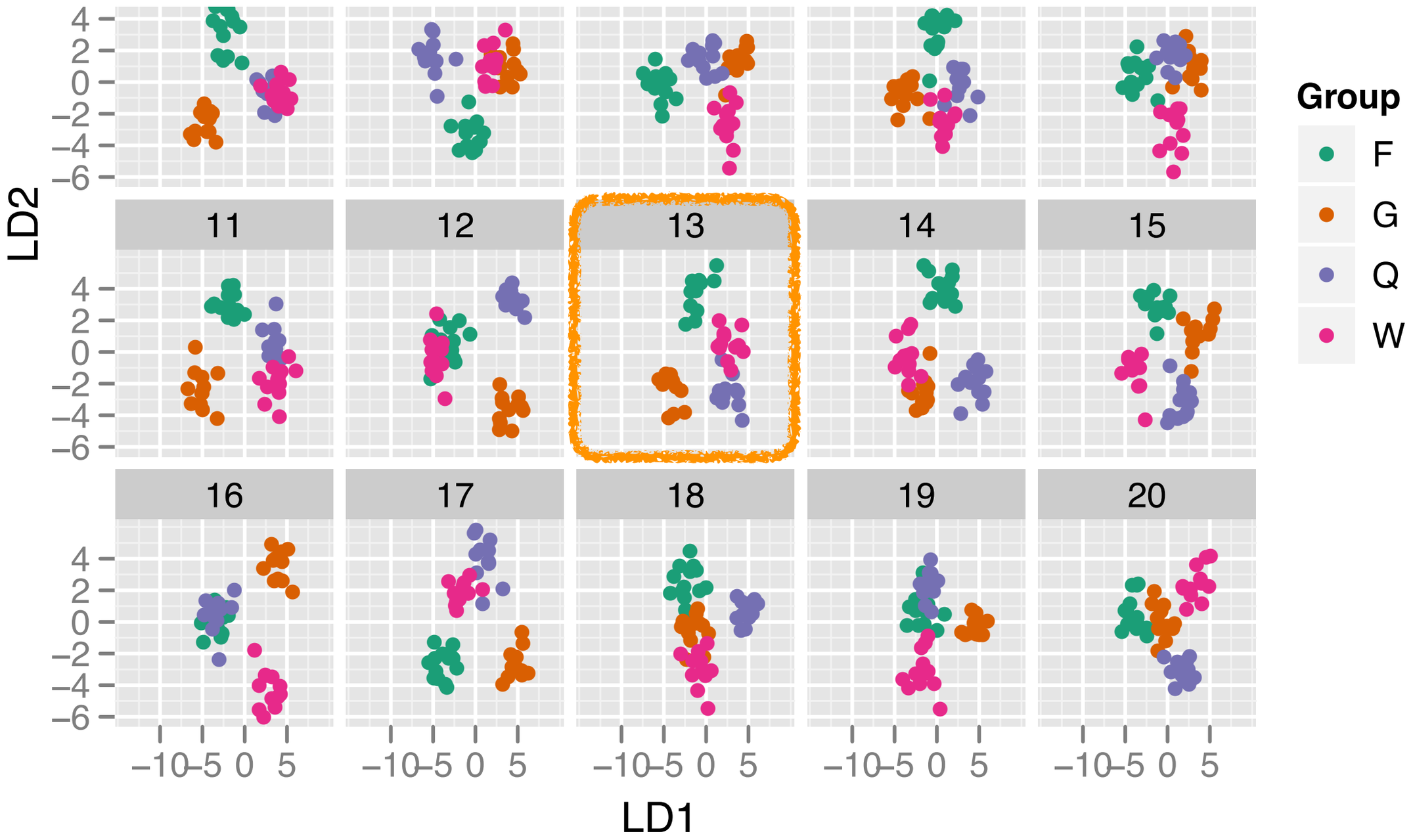
The pattern is not any more different than in other plots. The plot I showed earlier is actually a NULL plot, where I know the pattern is due to randomness.

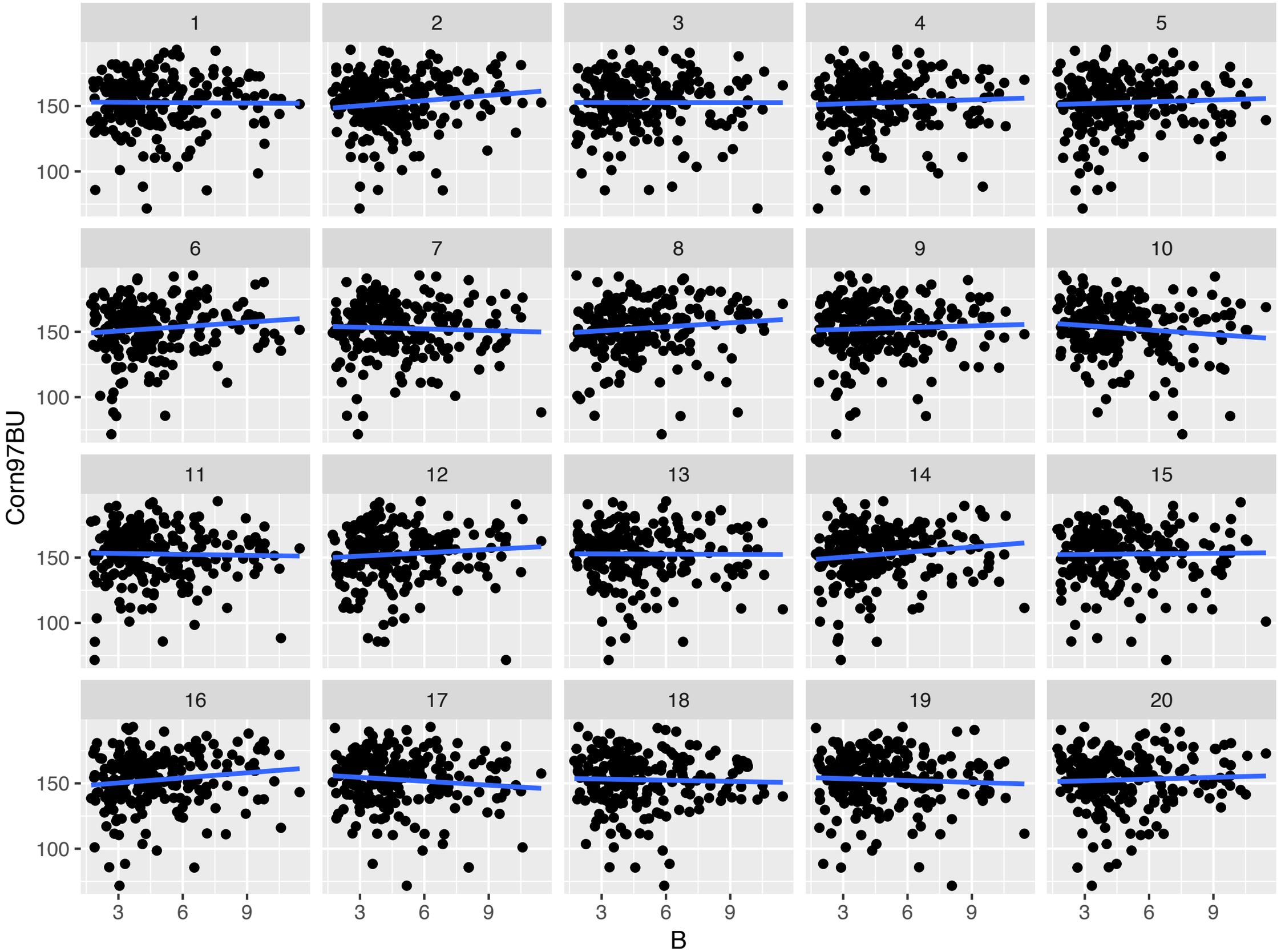


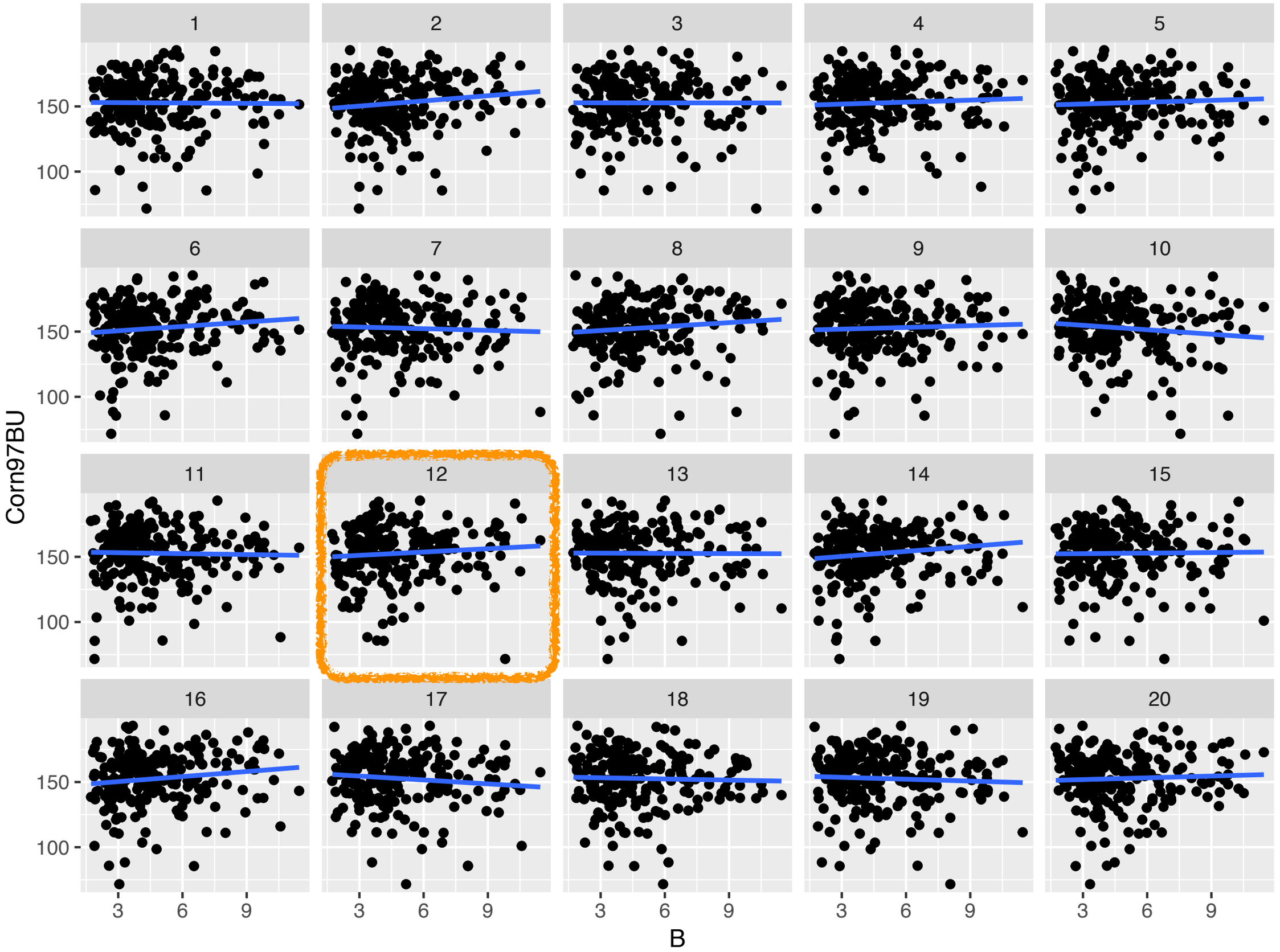

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
4
2
0
−2
−4
−6
−10
−5
0
5
LD2
LD1
Group
F
G
Q
W


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
4
2
0
−2
−4
−6
−10
−5
0
5
LD2
LD1
Group
F
G
Q
W

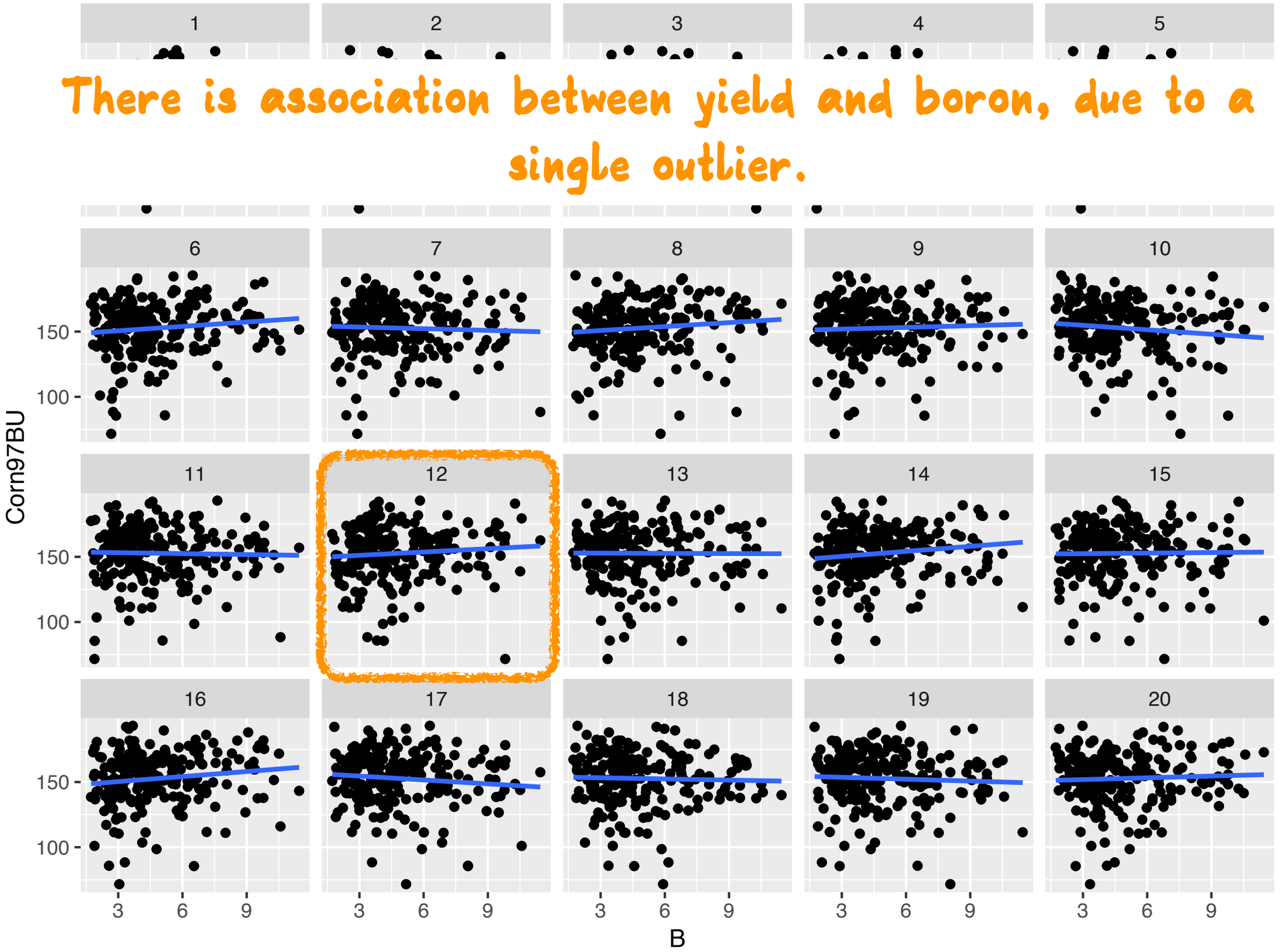
The groups are no more separated than in other plots. The separation in the data is purely by chance, actually because it is a projection from EMPTY high-dimensional space, there are many ways to find gaps.

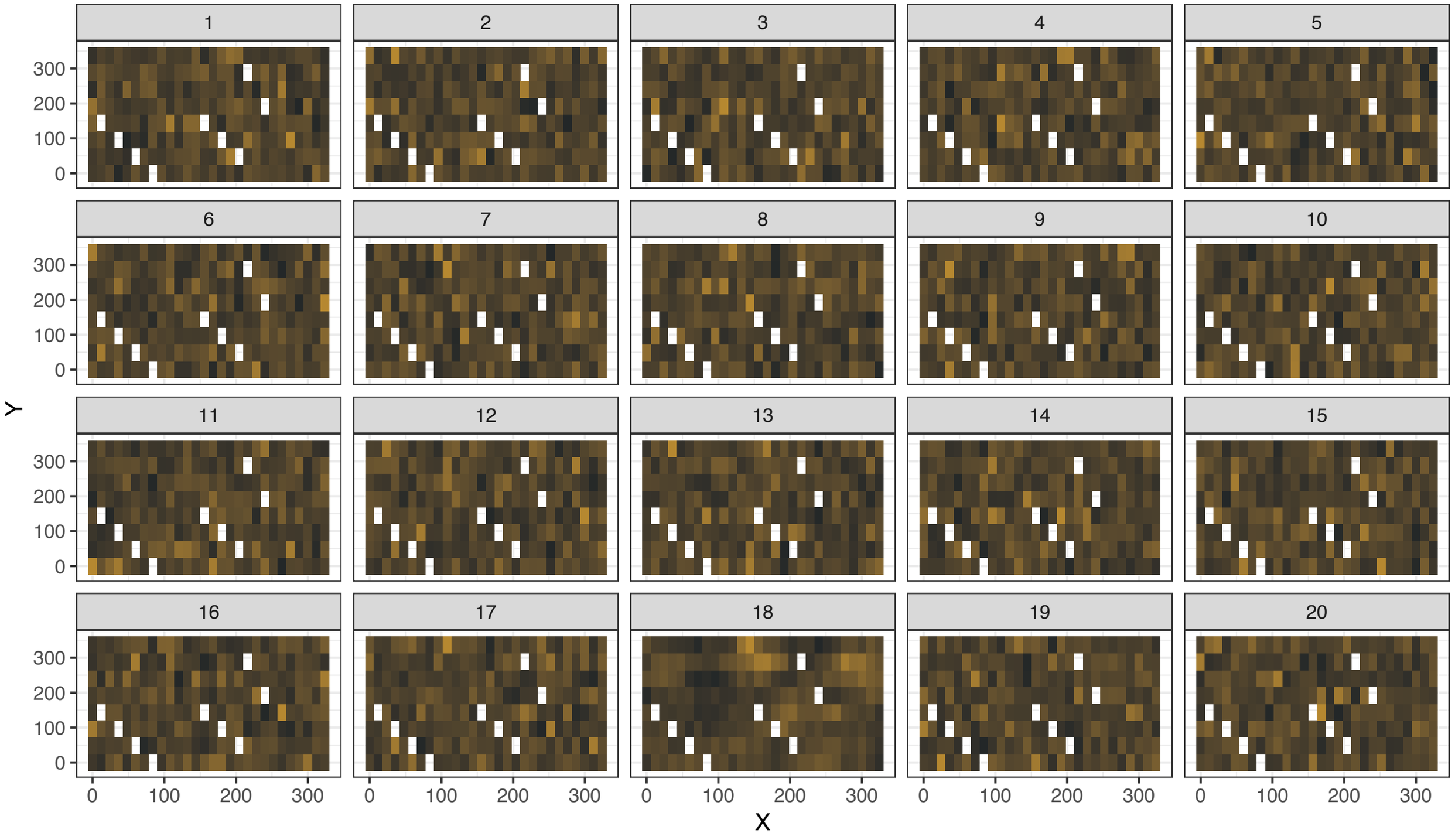


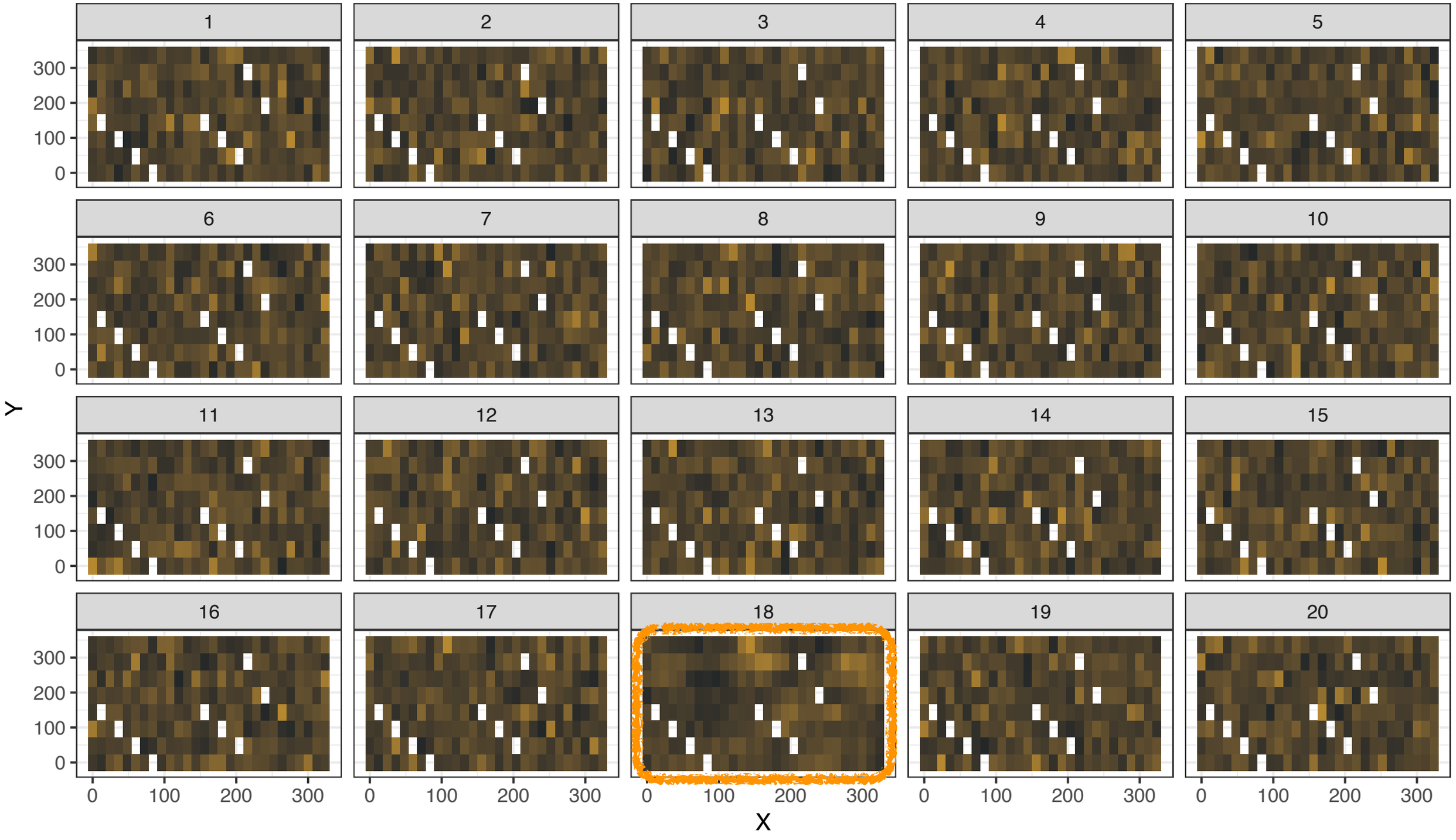

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
150
100
Cor
B
3
6
9

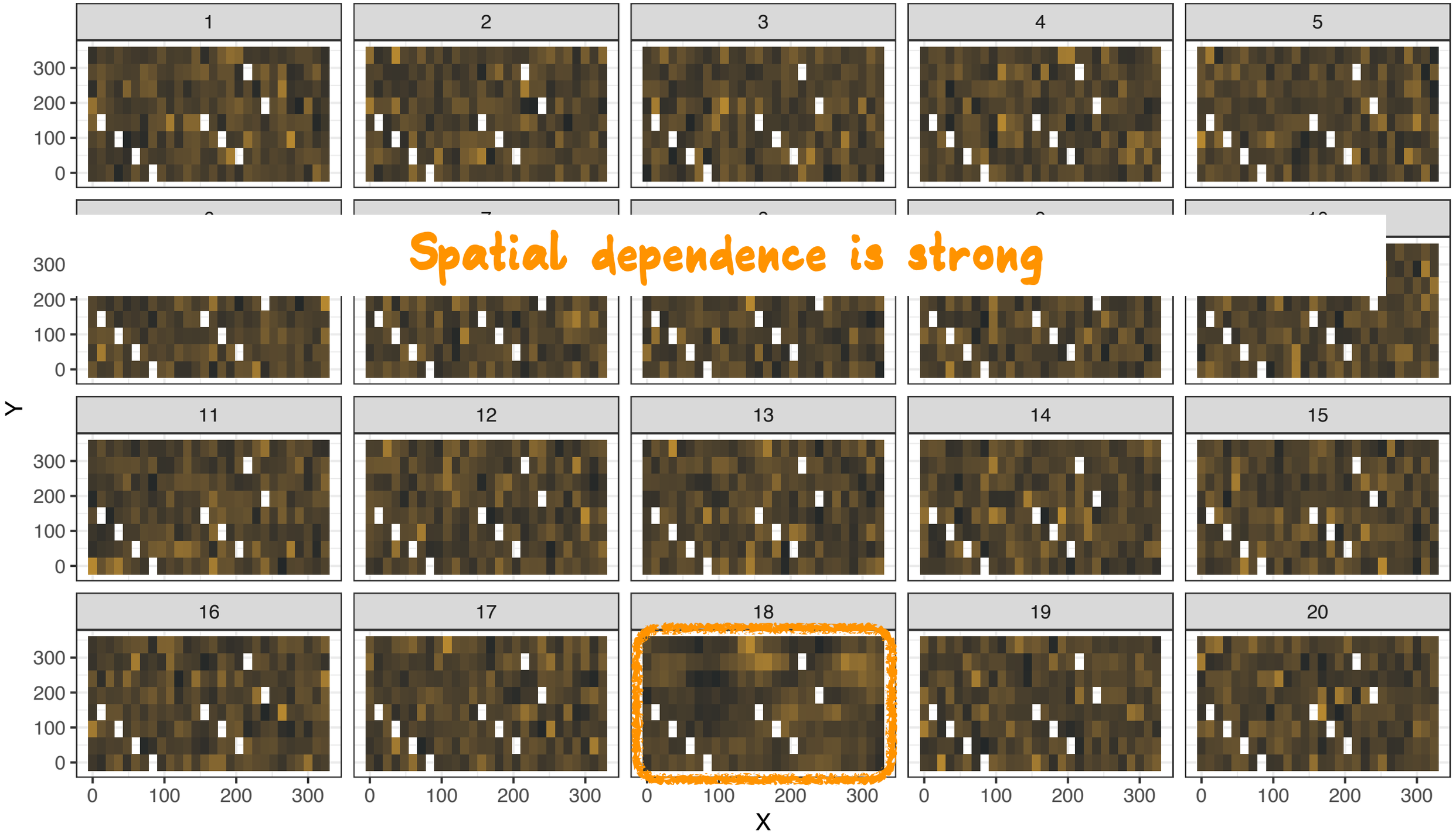

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
150
100
Co
rn97BU
3
6
9
B

There is association between yield and boron, due to a single outlier.




1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
0
100
200
300
Y
X


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
0
100
200
300
Y
X


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
0
100
200
300
Y
X
Spatial dependence is strong

Outline

- Protocols, computing p-values and power
- Fitting into the tidy verse
- Validation studies
- Applications
- How you can use these tools

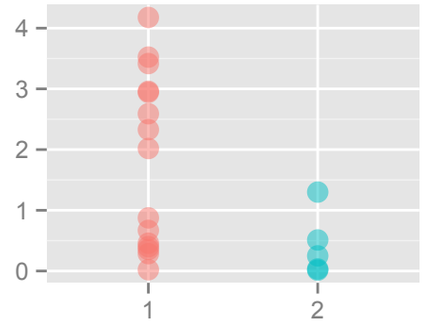
Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution
- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

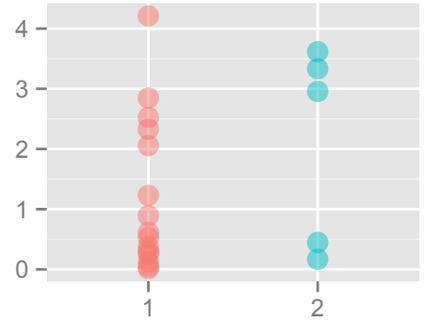


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

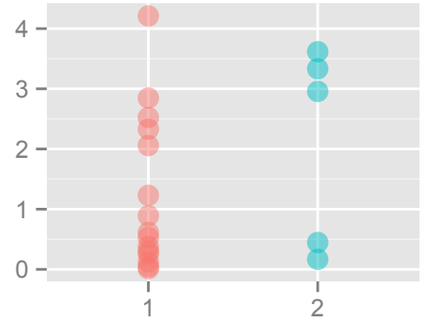


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

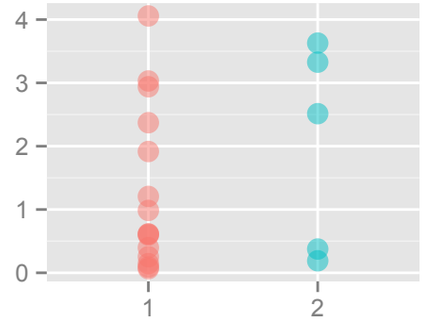


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

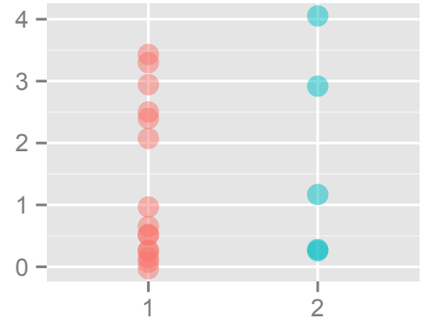


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

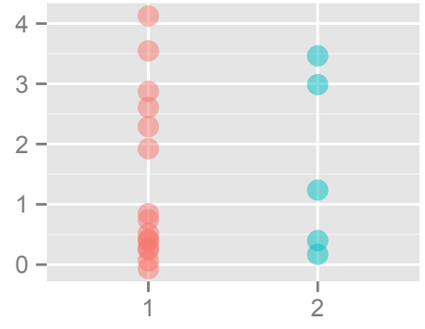


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

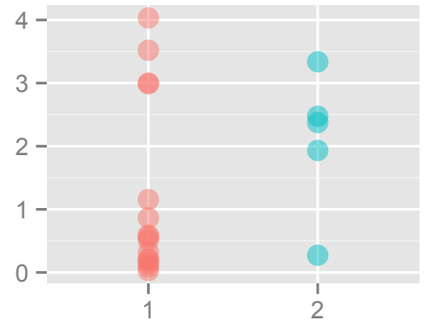


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

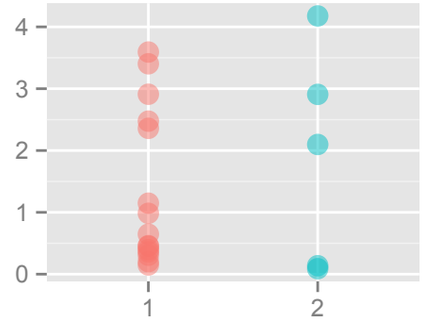


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

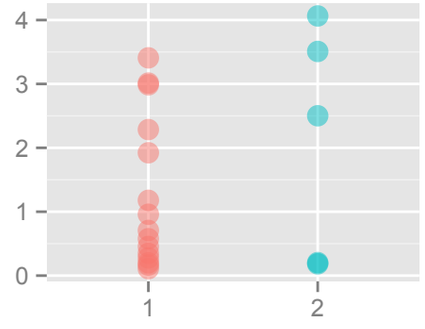


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

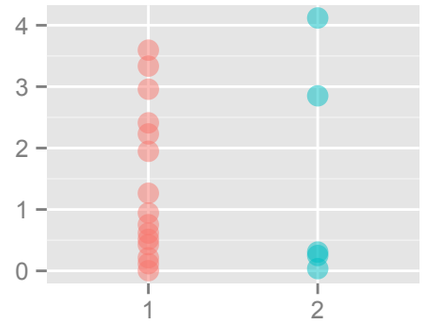


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

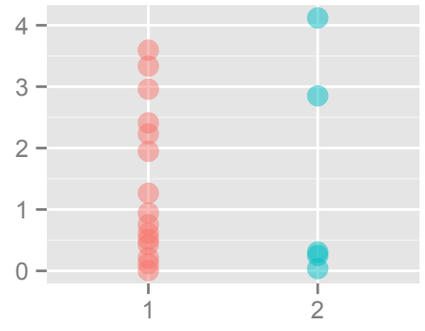


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

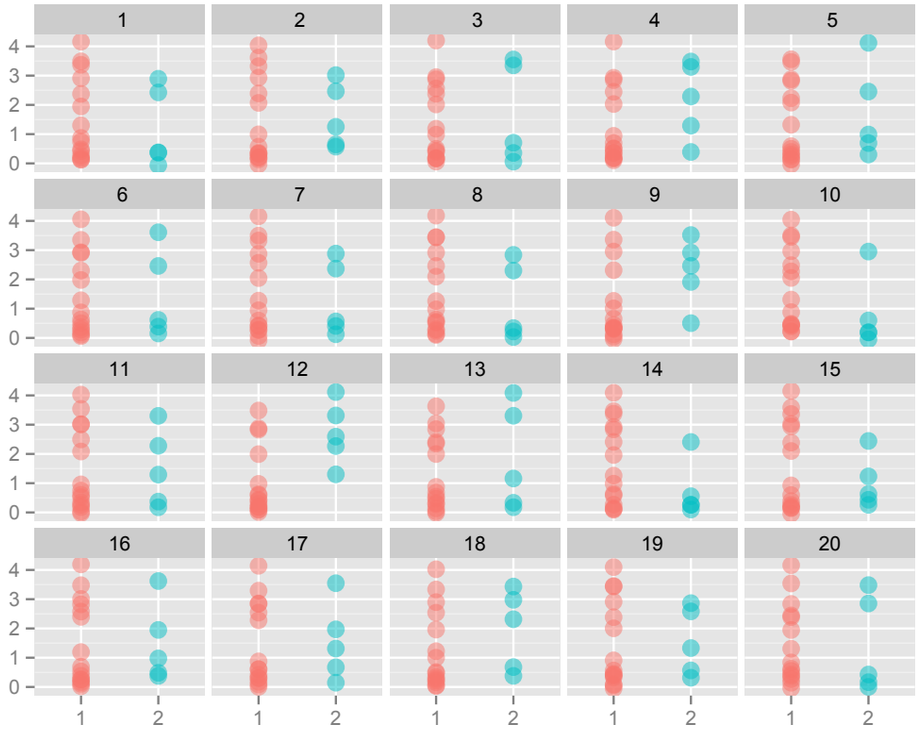
Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution



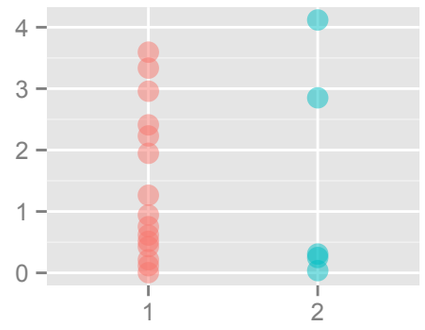
- Lineup: Embed the plot of the data among plots of data generated from the null distribution



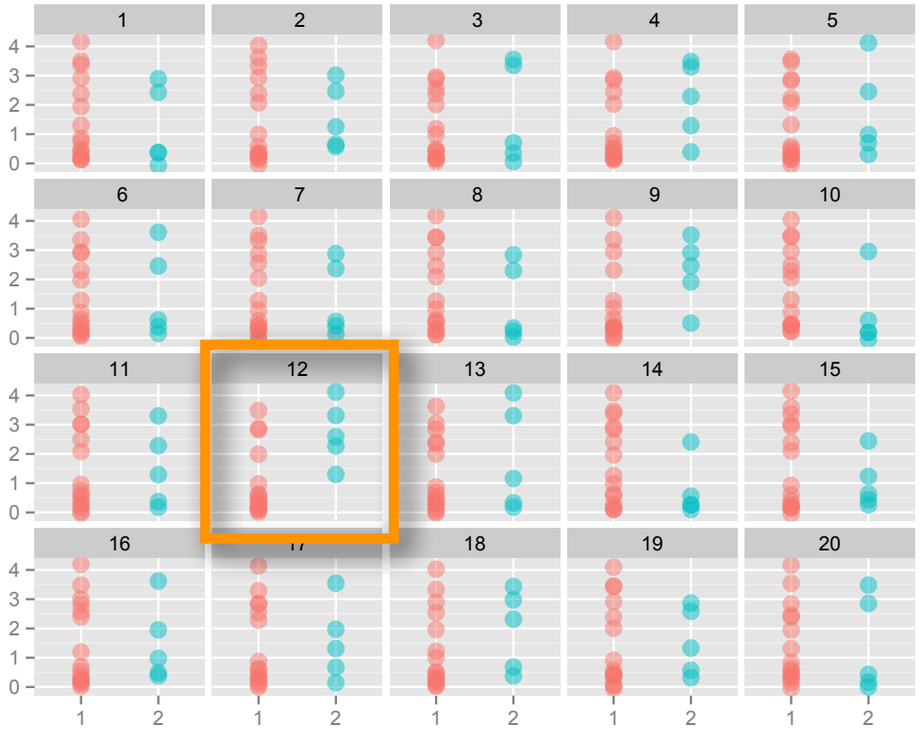
Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution



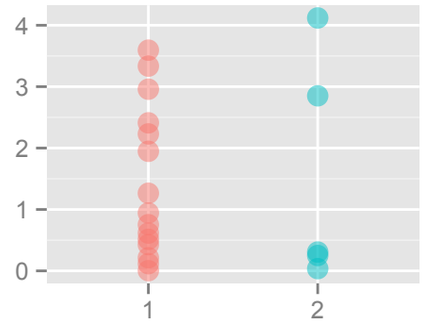
- Lineup: Embed the plot of the data among plots of data generated from the null distribution



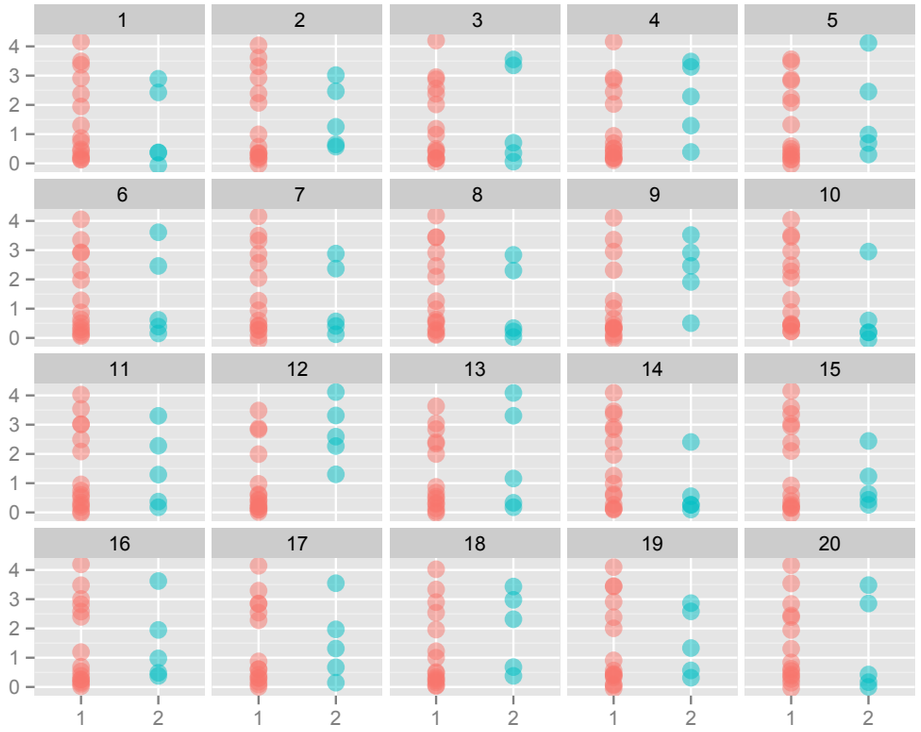
Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution



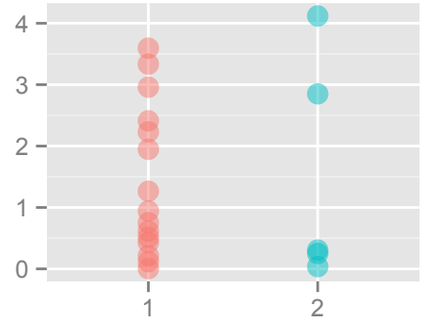
- Lineup: Embed the plot of the data among plots of data generated from the null distribution



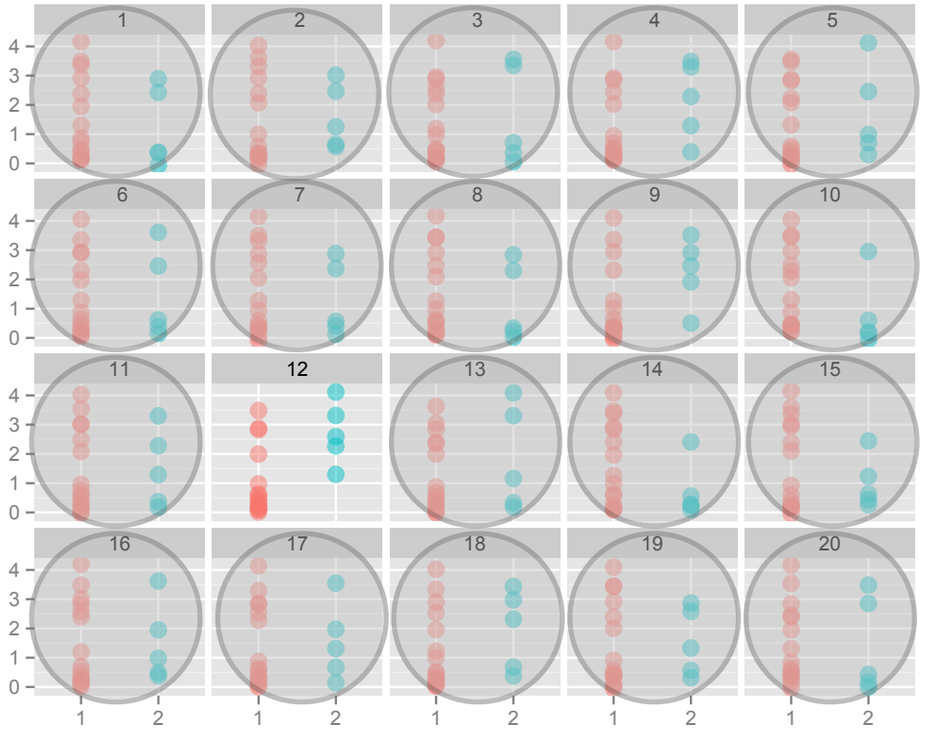
Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution



- Lineup: Embed the plot of the data among plots of data generated from the null distribution

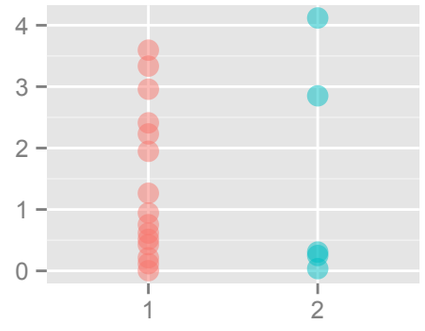


Null plots

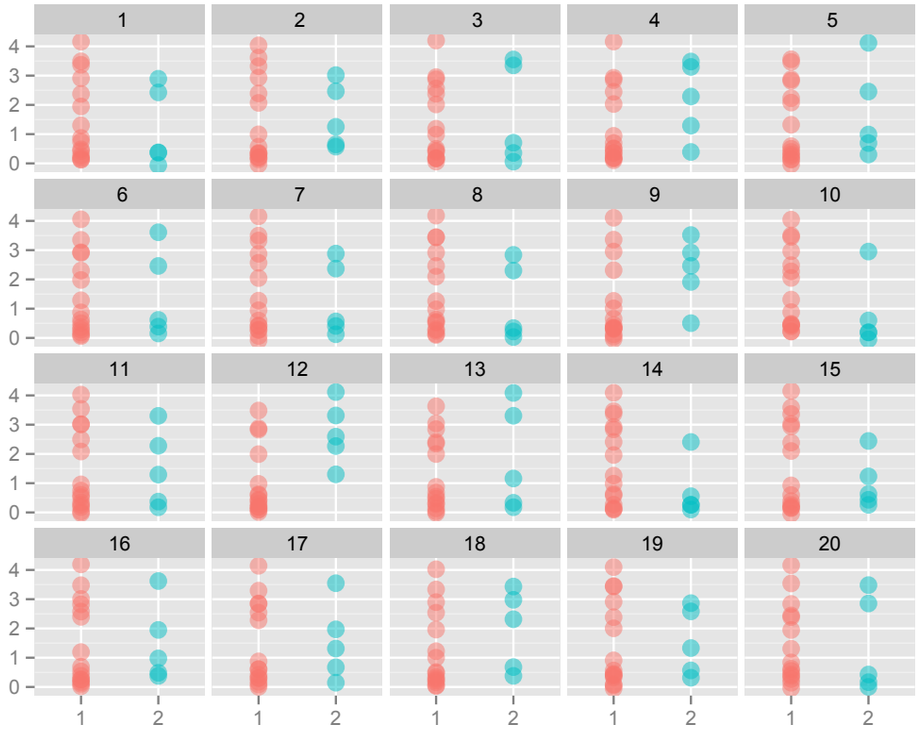
Source: Buja et al (2009) RSPT(A)

Protocols

- Rorschach: Show many pictures of data with "nothing" happening, pictures from a null distribution

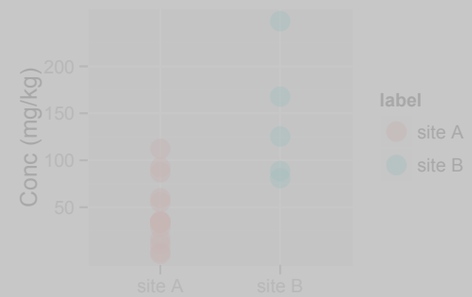
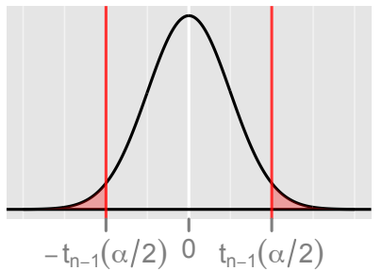
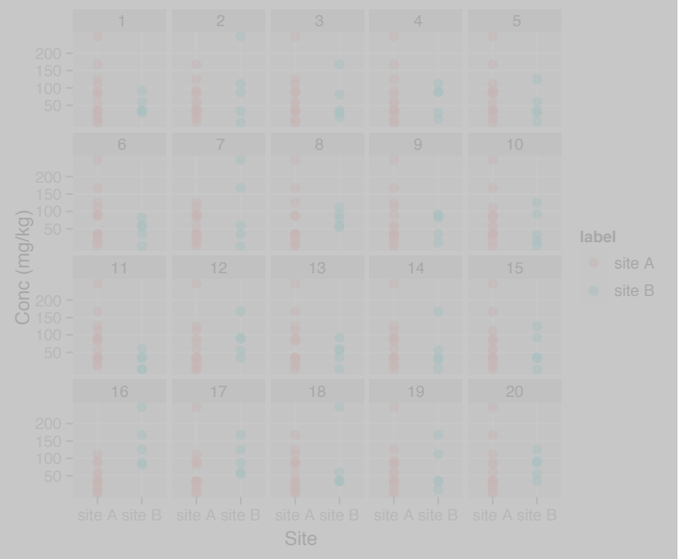


- Lineup: Embed the plot of the data among plots of data generated from the null distribution

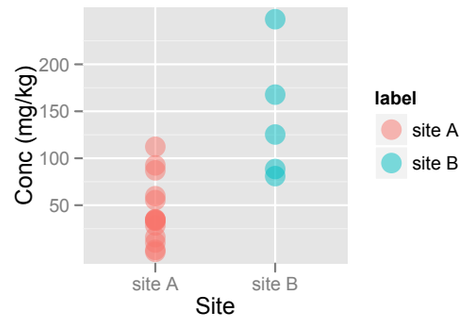
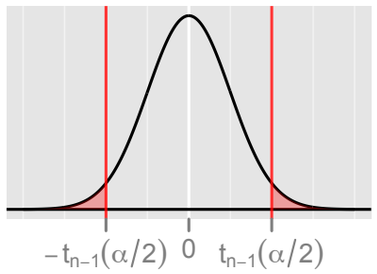
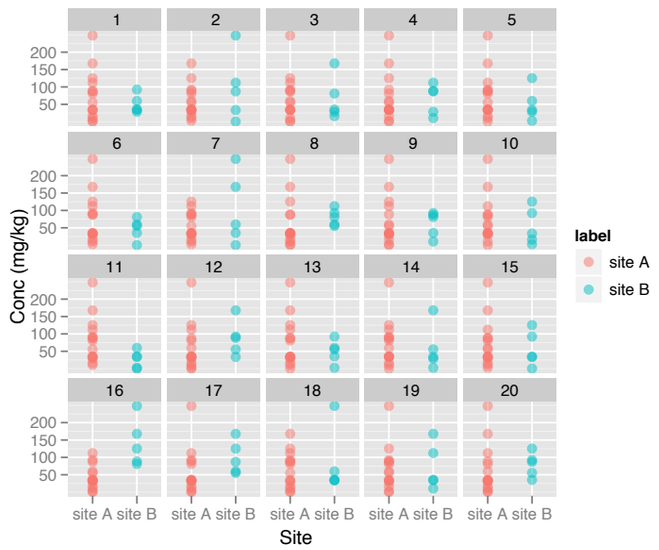


Source: Buja et al (2009) RSPT(A)

Hypothesis testing

	Mathematical Inference	Visual Inference
Hypothesis	$H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 \neq \mu_2$	$H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 \neq \mu_2$
	↓	↓
Test Statistic	$T(y) = \frac{\bar{y}_1 - \bar{y}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$T(y) =$ 
	↓	↓
Sampling Distribution	$f_{T(y)}(t)$; 	$f_{T(y)}(t)$; 
	↓	↓
Reject H_0 if	observed T is extreme	observed plot is identifiable

Hypothesis testing

	Mathematical Inference	Visual Inference
Hypothesis	$H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 \neq \mu_2$	$H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 \neq \mu_2$
	↓	↓
Test Statistic	$T(y) = \frac{\bar{y}_1 - \bar{y}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$T(y) =$ 
	↓	↓
Sampling Distribution	$f_{T(y)}(t);$ 	$f_{T(y)}(t);$ 
	↓	↓
Reject H_0 if	observed T is extreme	observed plot is identifiable

Plots as statistics

```
ggplot(data = <DATA>) +
  <GEOM_FUNCTION>(
     mapping = aes(<MAPPINGS>),
     stat = <STAT>,
     position = <POSITION>
  ) +
  <COORDINATE_FUNCTION> +
  <FACET_FUNCTION>
```

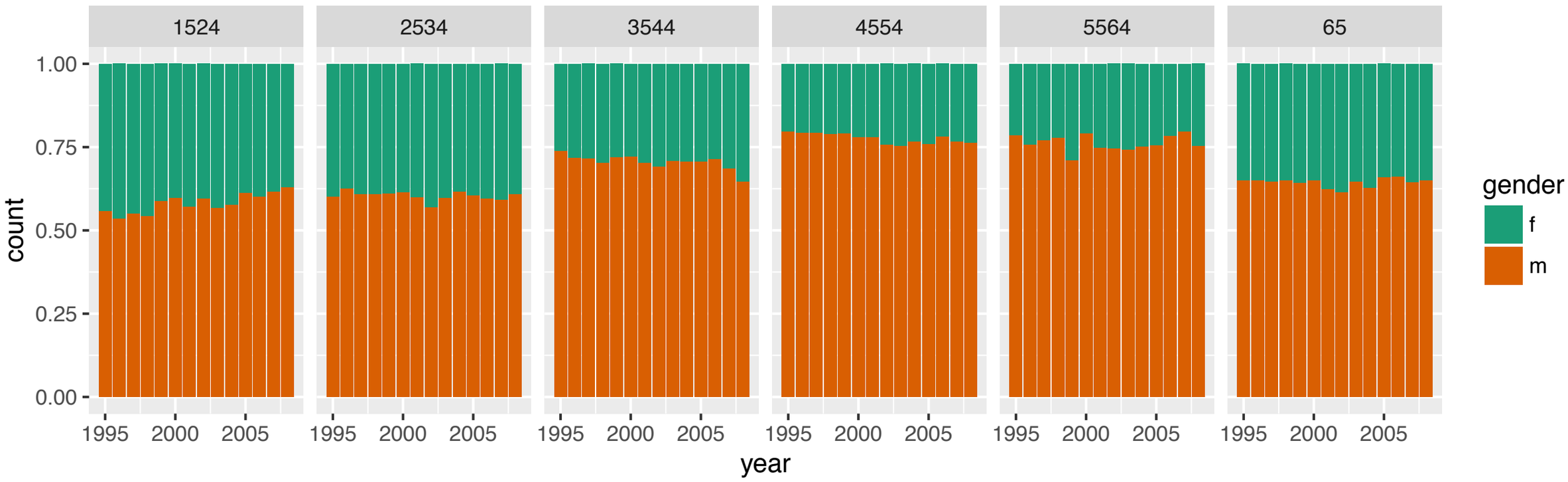
Plots as statistics

```
ggplot(data = <DATA>) +
  <GEOM_FUNCTION>(
     mapping = aes(<MAPPINGS>),
     stat = <STAT>,
     position = <POSITION>
  ) +
  <COORDINATE_FUNCTION> +
  <FACET_FUNCTION>
```

Connection between variables and graphical elements

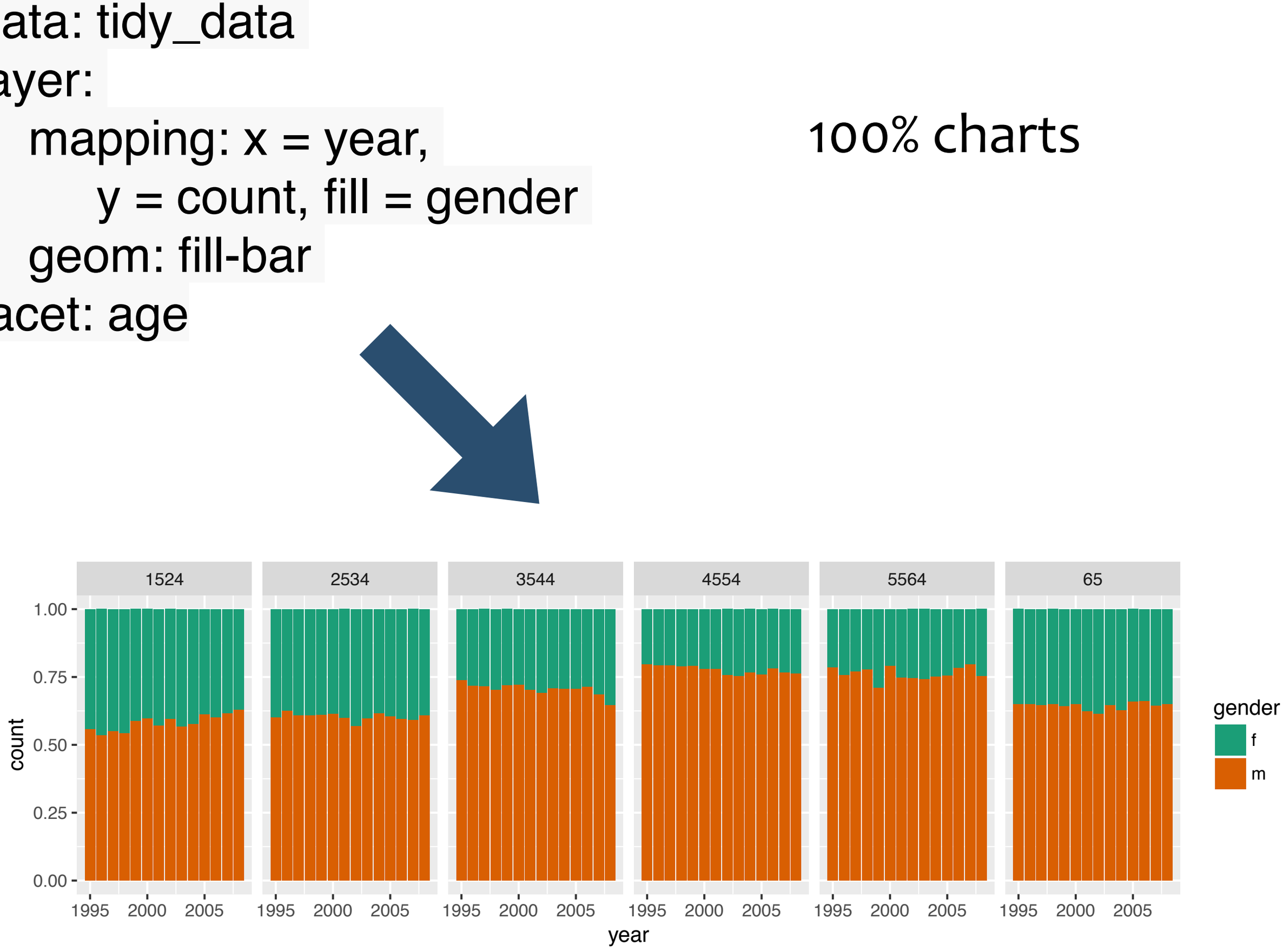
```
data: tidy_data
layer:
    mapping: x = year,
         y = count, fill = gender
    geom: fill-bar
facet: age
```

100% charts



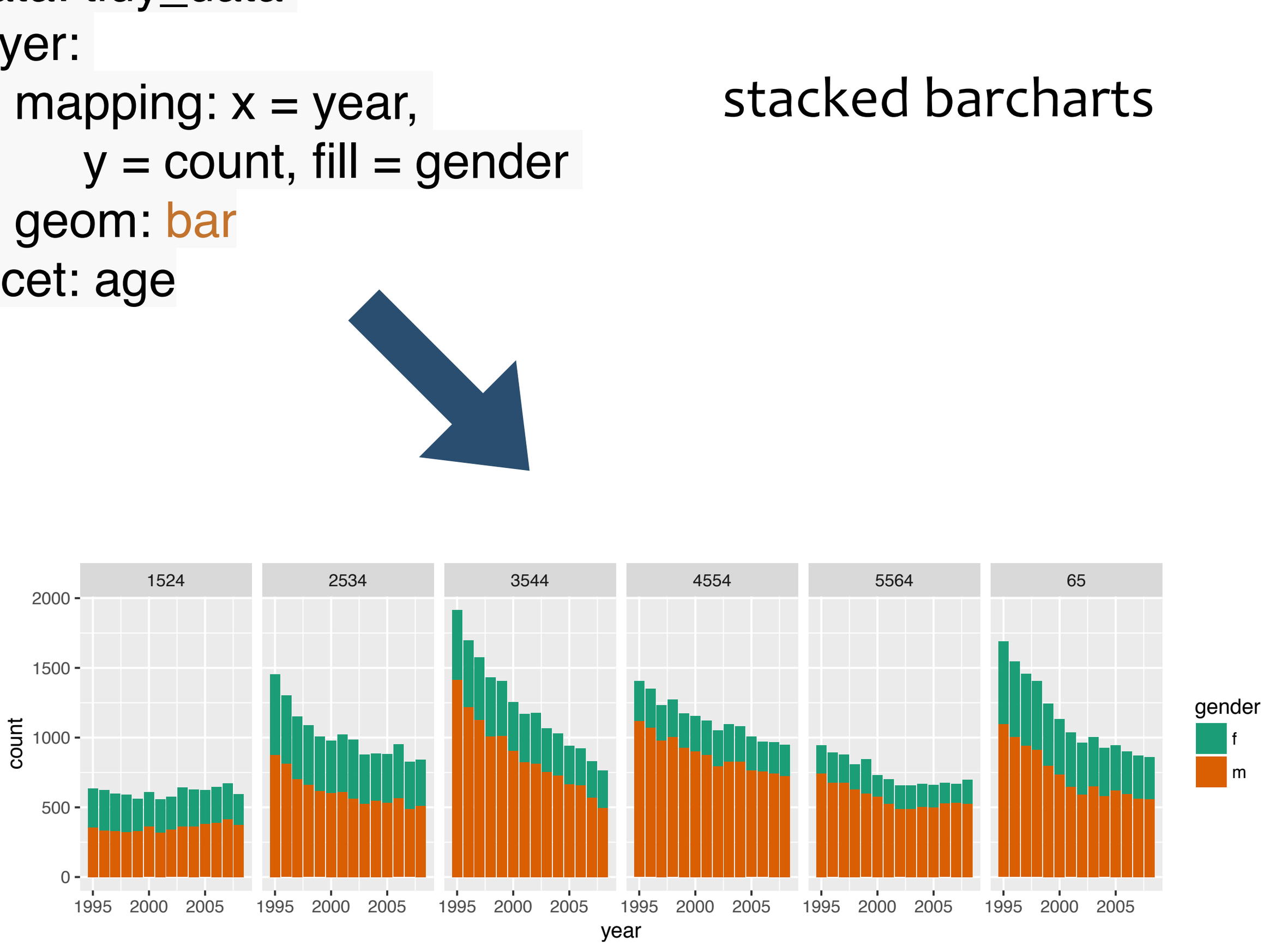
```
data: tidy_data
layer:
    mapping: x = year,
         y = count, fill = gender
    geom: fill-bar
facet: age
```

100% charts



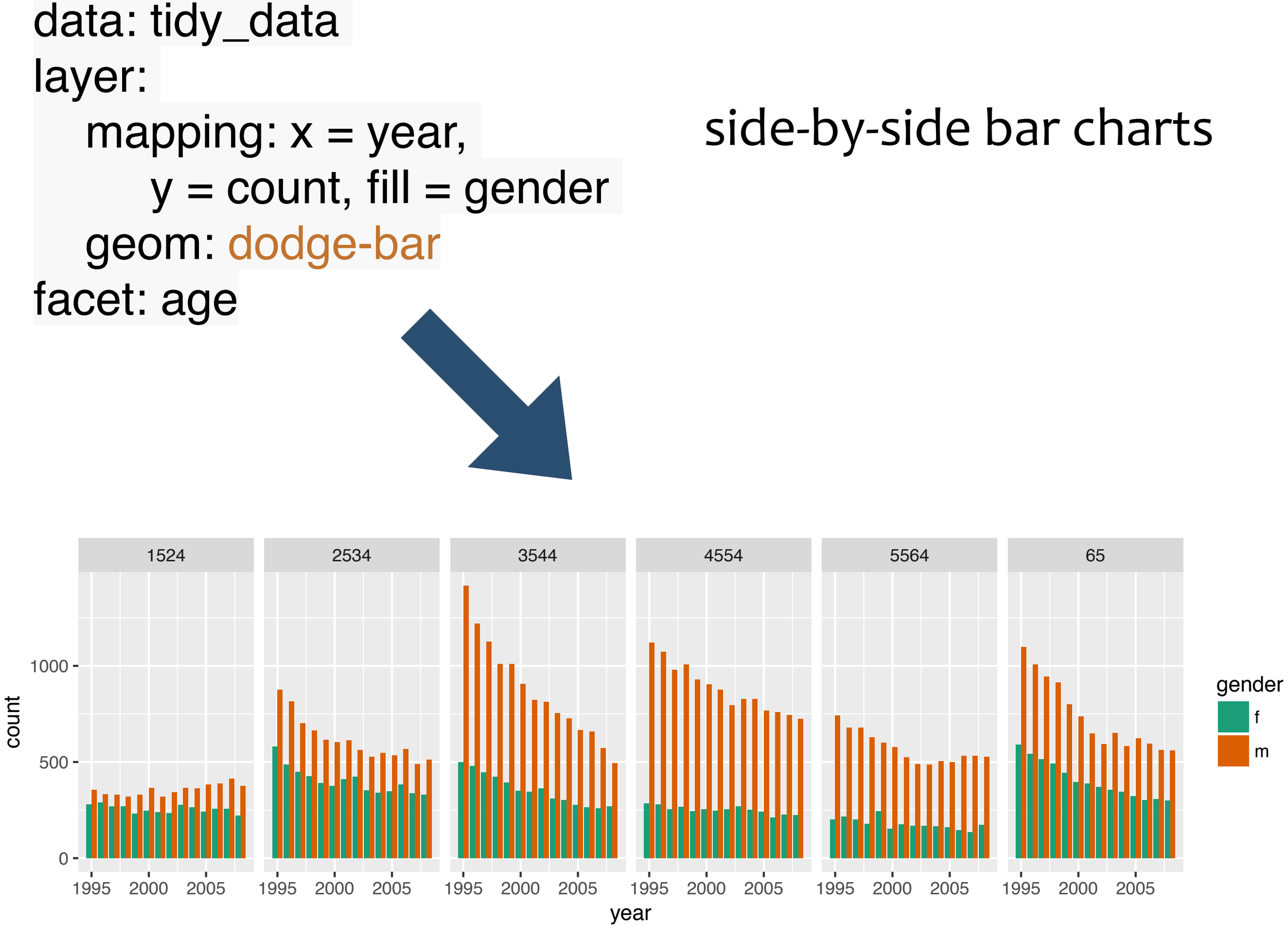
```
data: tidy_data
layer:
    mapping: x = year,
        y = count, fill = gender
    geom: bar
facet: age
```

stacked barcharts



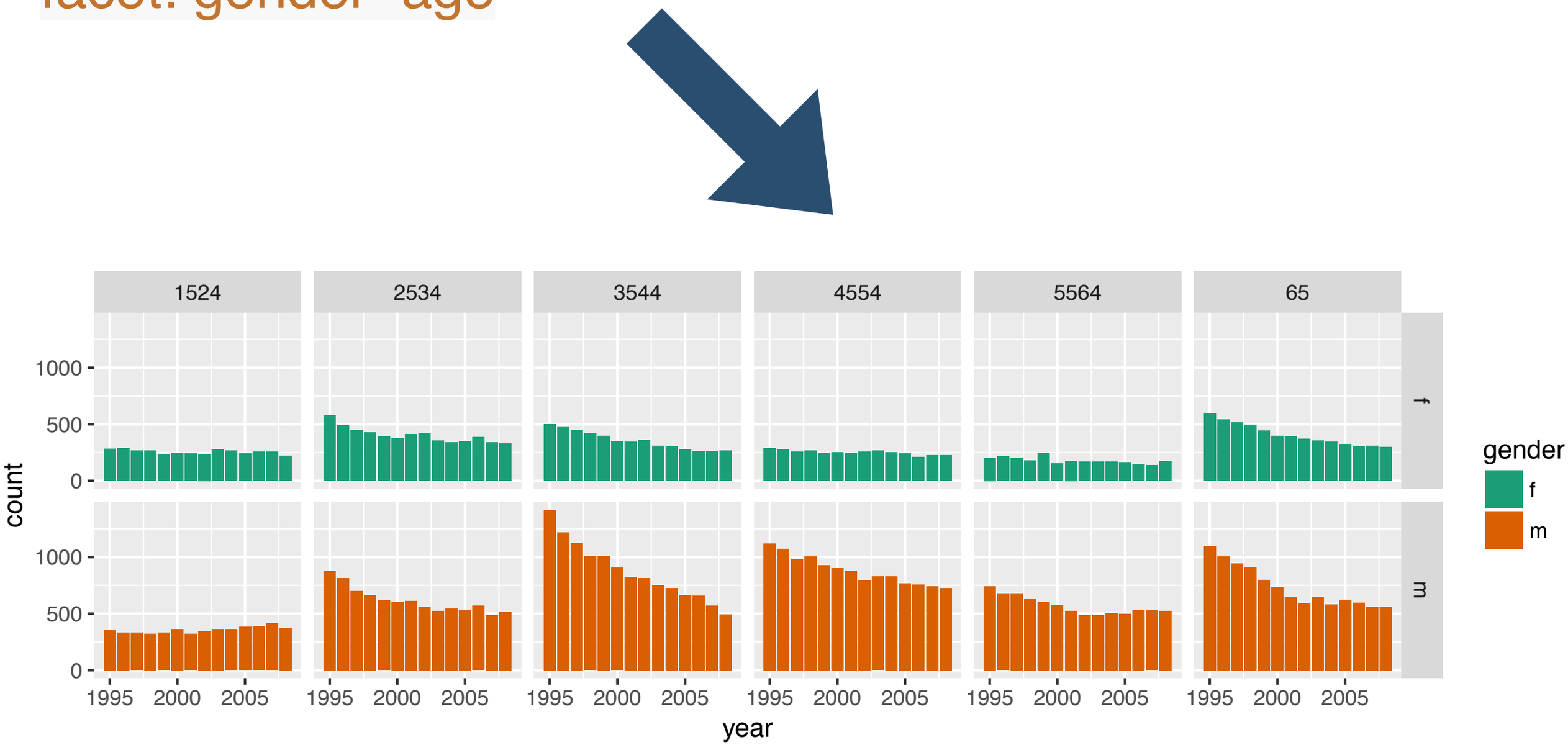
```
data: tidy_data
layer:
    mapping: x = year,
        y = count, fill = gender
    geom: dodge-bar
facet: age
```

side-by-side bar charts



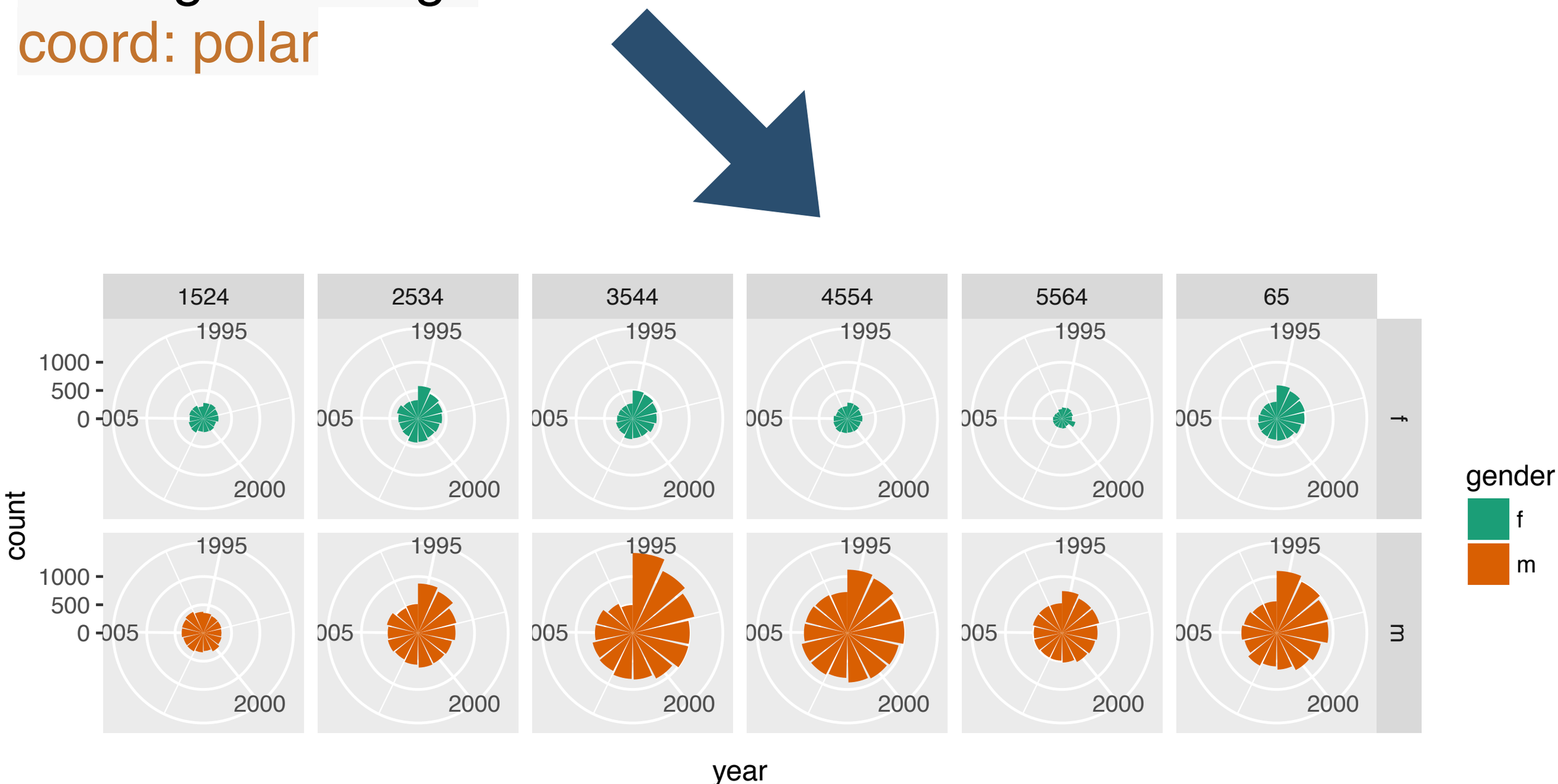
```
data: tidy_data
layer:
    mapping: x = year,
         y = count, fill = gender
    geom: bar
facet: gender~age
```

bar charts



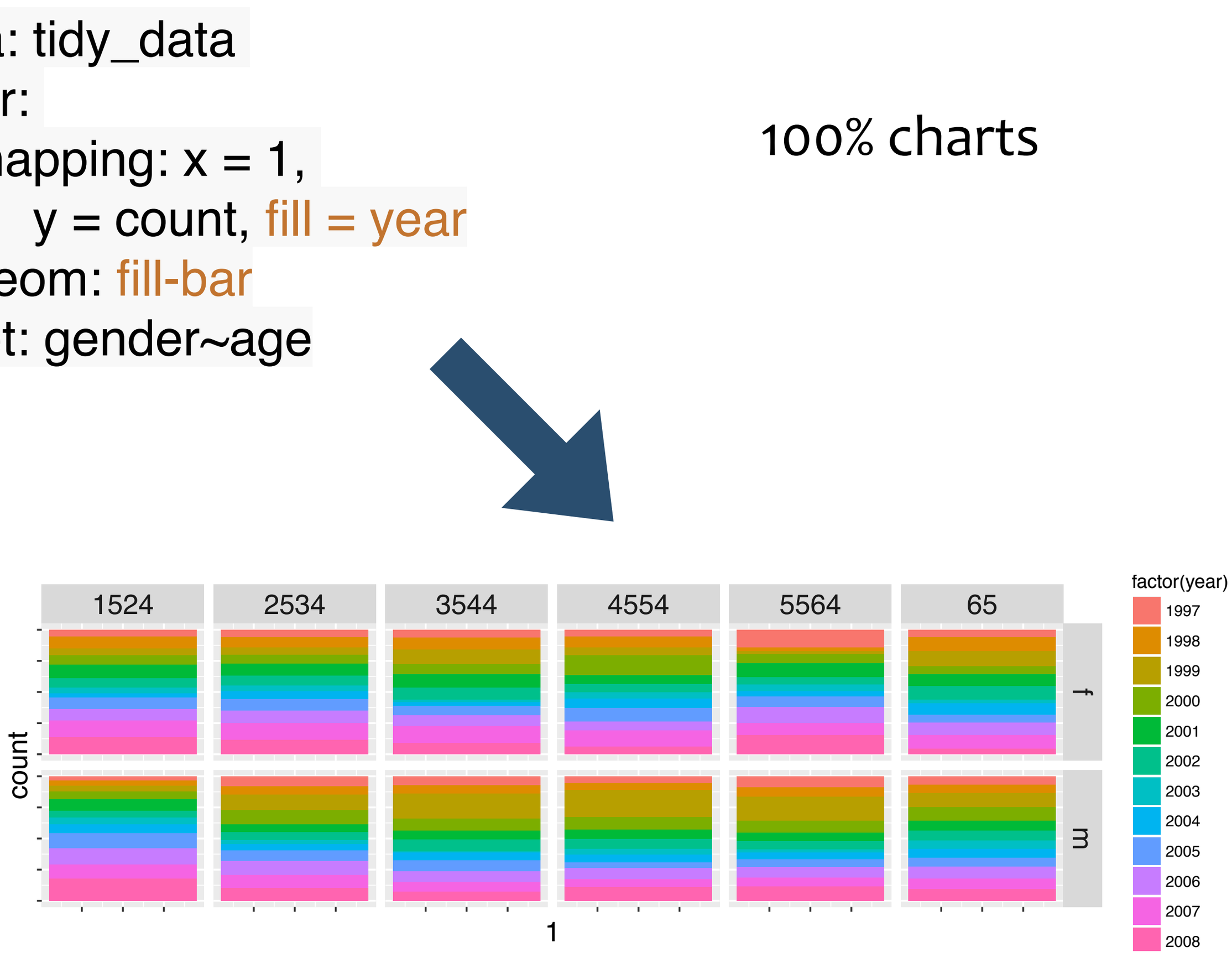
```
data: tidy_data
layer:
    mapping: x = year,
        y = count, fill = gender
    geom: bar
facet: gender~age
coord: polar
```

rose plots



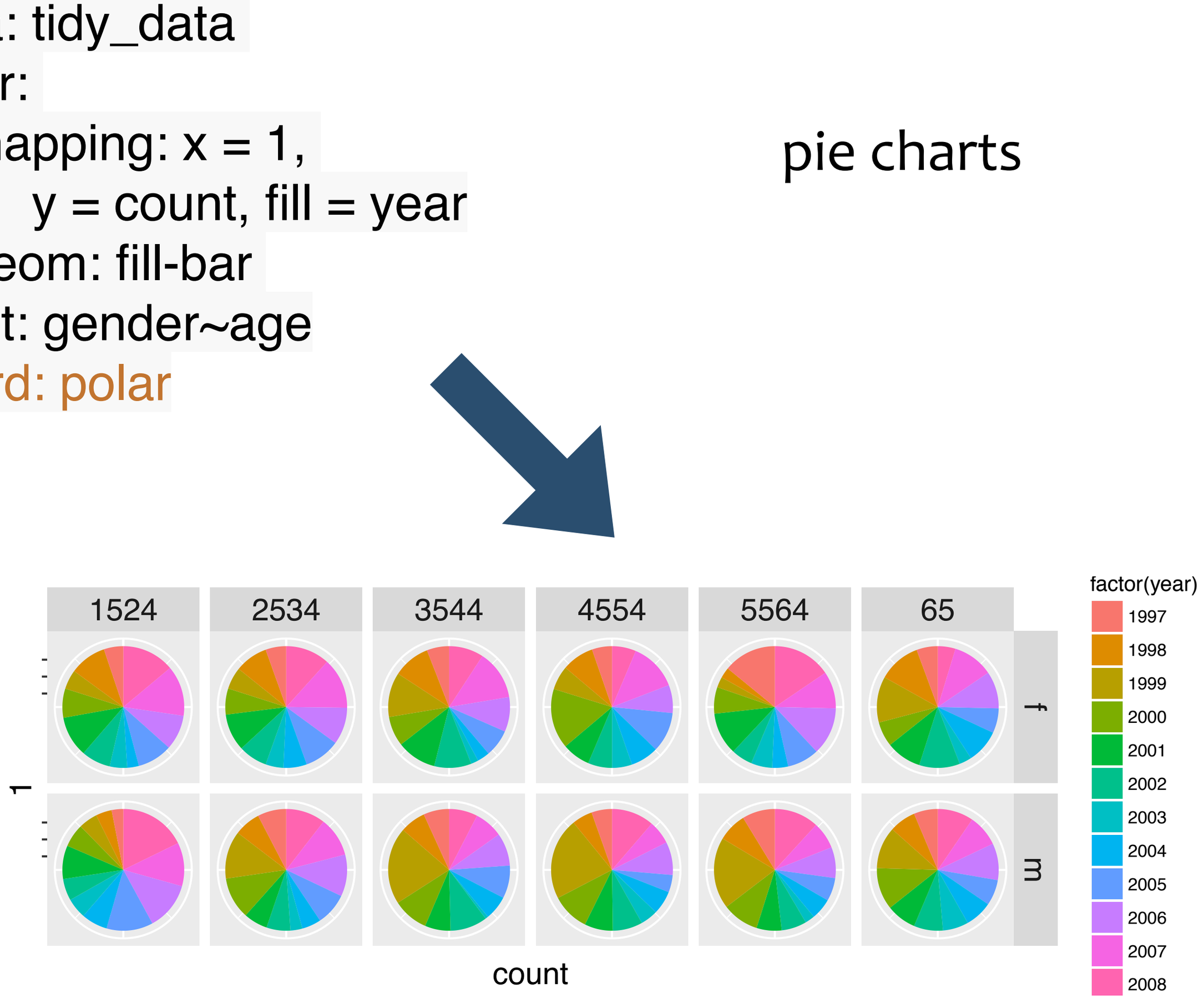
```
data: tidy_data
layer:
    mapping: x = 1,
        y = count, fill = year
    geom: fill-bar
facet: gender~age
```

100% charts



```
data: tidy_data
layer:
    mapping: x = 1,
         y = count, fill = year
    geom: fill-bar
facet: gender~age
coord: polar
```

pie charts



http://knowyourmeme.com/memes/morphing/photos/sort/favorites

http://knowyourmeme.com/memes/morphing/photos/sort/favorites

Key components

- A plot is a test statistic, adding the data is the observed test statistic
- Type of plot used indicates null/alternative hypothesis, eg scatterplot suggests null hypothesis "no association between two variables"
- Null hypothesis determines null generating mechanism
- Generate draws from the null, plot, compare with data plot using an uninvolved observer
- Data plot detected equivalent to rejection of null, it is extreme relative to the sampling distribution

Significance

- What is the *p*-value?
- For one observer, the probability of randomly selecting the data plot is $^1/_m$, where m is the number of plots in the lineup.
- With multiple observers, the *p*-value is estimated by

$$P(X \geq x) = 1 - \textit{Binom}_{K,1/m}(x-1) = \sum_{i=x}^{K} \binom{K}{i} \left(\frac{1}{m}\right)^i \left(\frac{m-1}{m}\right)^{K-i}$$

Source: Majumder et al (2013) JASA

Significance

- What is the *p*-value?
- For one observer, the probability of randomly selecting the data plot is $1/m$, where m is the number of plots in the lineup.
- With multiple observers, the *p*-value is estimated by

Number of independent observers

$$P(X \geq x) = 1 - \textit{Binom}_{K,1/m}(x-1) = \sum_{i=x}^{K} \binom{K}{i} \left(\frac{1}{m}\right)^{i} \left(\frac{m-1}{m}\right)^{K-i}$$

Source: Majumder et al (2013) JASA

Significance

- What is the *p*-value?
- For one observer, the probability of randomly selecting the data plot is $^1/_m$, where m is the number of plots in the lineup.
- With multiple observers, the *p*-value is estimated by

$$P(X \geq x) = 1 - \textit{Binom}_{K,1/m}(x-1) = \sum_{i=x}^{K} \binom{K}{i} \left(\frac{1}{m}\right)^i \left(\frac{m-1}{m}\right)^{K-i}$$

Source: Majumder et al (2013) JASA

Significance

- What is the *p*-value?
- For one observer, the probability of randomly selecting the data plot is $^1/_m$, where m is the number of plots in the lineup.
- With multiple observers, the *p*-value is estimated by

$$P(X \geq x) = 1 - Binom_{K,1/m}(x-1) = \sum_{i=x}^{K} \binom{K}{i} \left(\frac{1}{m}\right)^i \left(\frac{m-1}{m}\right)^{K-i}$$

Number of observers choosing data plot

Source: Majumder et al (2013) JASA

Significance

- What is the *p*-value?
- For one observer, the probability of randomly selecting the data plot is $^1/_m$, where m is the number of plots in the lineup.
- With multiple observers, the *p*-value is estimated by

$$P(X \geq x) = 1 - \textit{Binom}_{K,1/m}(x-1) = \sum_{i=x}^{K} \binom{K}{i} \left(\frac{1}{m}\right)^i \left(\frac{m-1}{m}\right)^{K-i}$$

Source: Majumder et al (2013) JASA

Computing *p*-values

Scenario: **same data plot and null plots, multiple observers**

Simulation:

1. Sample $p_{\text{data}} \sim U[0,1]$ (why uniform? Murdoch et al (2008) P-Values are Random Variables, TAS 62, 242-245)
2. *S*ample *(m-1)* null *p*-values, $p_j \sim U[0,1], j=1, \ldots, m$
3. *C*ompute 'strength' of data plot relative to other plots*:*
   $q = (1-p_{\text{data}}) / \sum j(1-p_j)$
4. Simulate number of data plot picks $X \sim Bin(K, q)$
5. Repeat steps 1-4, N times to obtain distribution of X
6. Examine where observed X is placed in the distribution of X*'s* generated assuming the null hypothesis is true.

Hofmann et al, to be published

Computing *p*-values

Scenario: **same data plot and null plots, multiple observers**

Simulation:

1. Sample $p_{data} \sim U[0,1]$ (why uniform? Murdoch et al (2008) P-Values are Random Variables, TAS 62, 242-245)
2. *S*ample *(m-1)* null *p*-values, $p_j \sim U[0,1], j=1, ..., m$
3. *C*ompute 'strength' of data plot relative to other plots:
   $q = (1-p_{data})/\sum j(1-p_j)$
4. Simulate number of data plot picks $X \sim Bin(K, q)$
5. Repeat steps 1-4, N times to obtain distribution of X
6. Examine where observed X is placed in the distribution of X's generated assuming the null hypothesis is true.

Hofmann et al, to be published

Computing *p*-values

Scenario: same data plot and null plots, multiple observers

Simulate picks from Binomial, with K observers all looking at the same lineup

Power of the Test

- Power is used to compare the performance of tests in statistics.

Definition 2.6. The *power* of a visual test, V_θ, is defined as the probability to reject the null hypothesis for a given parameter value θ:

$$\text{Power}_V(\theta) = \Pr(\text{Reject } H_0 \mid \theta).$$

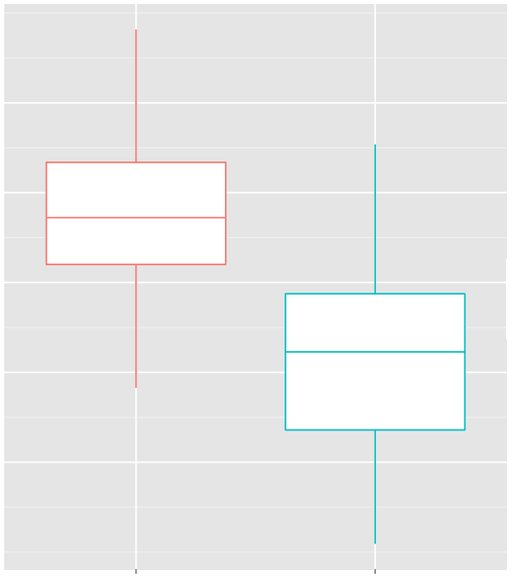
- Estimate by the proportion of observers picking the data plot.

Source: Majumder et al (2013) JASA

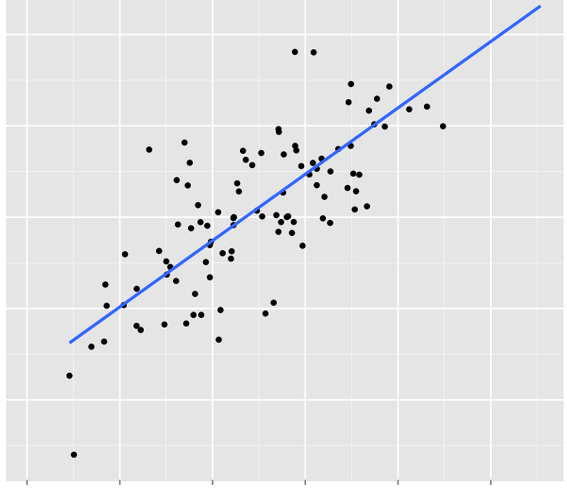
Validation experiments

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_1 X_{i2} + \beta_3 X_{i1} X_{i2} + \cdots + \epsilon_i$$

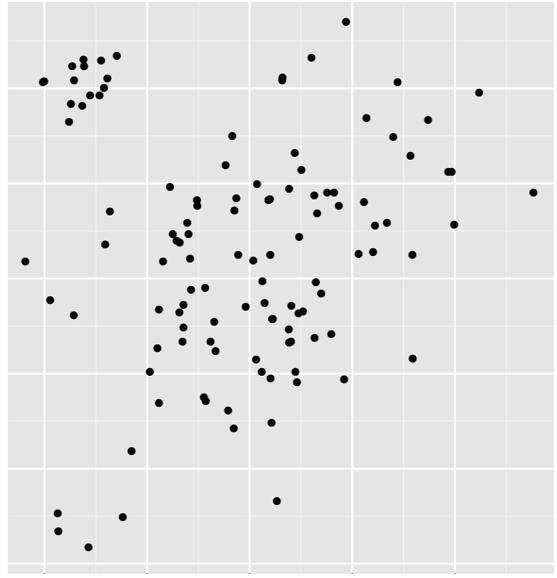
$$H_0 : \beta_k = 0 \text{ vs } H_1 : \beta_k \neq 0$$



Categorical X



Quantitative X

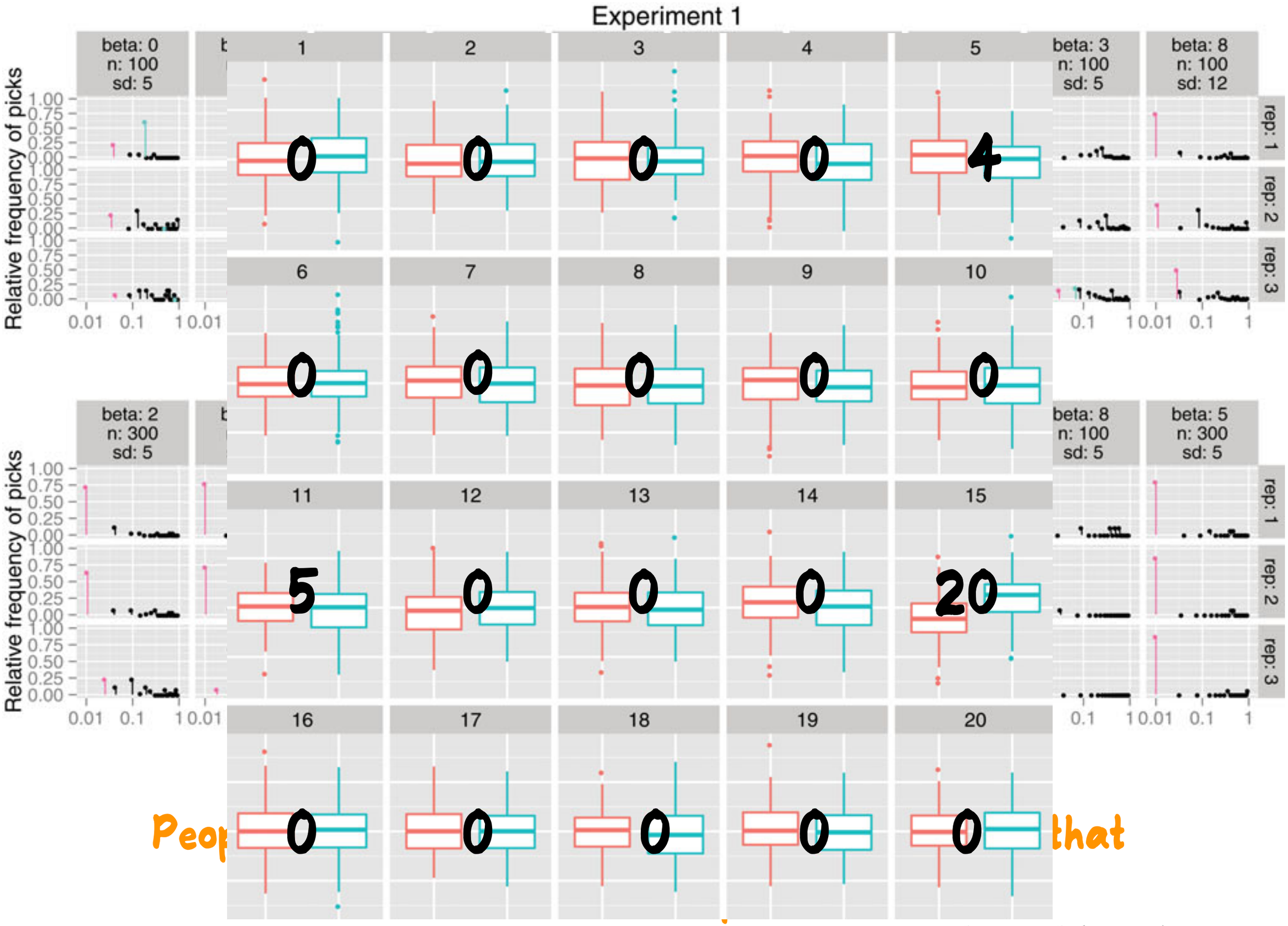


Contamination

`http://mamajumder.github.io/html/experiments.html`

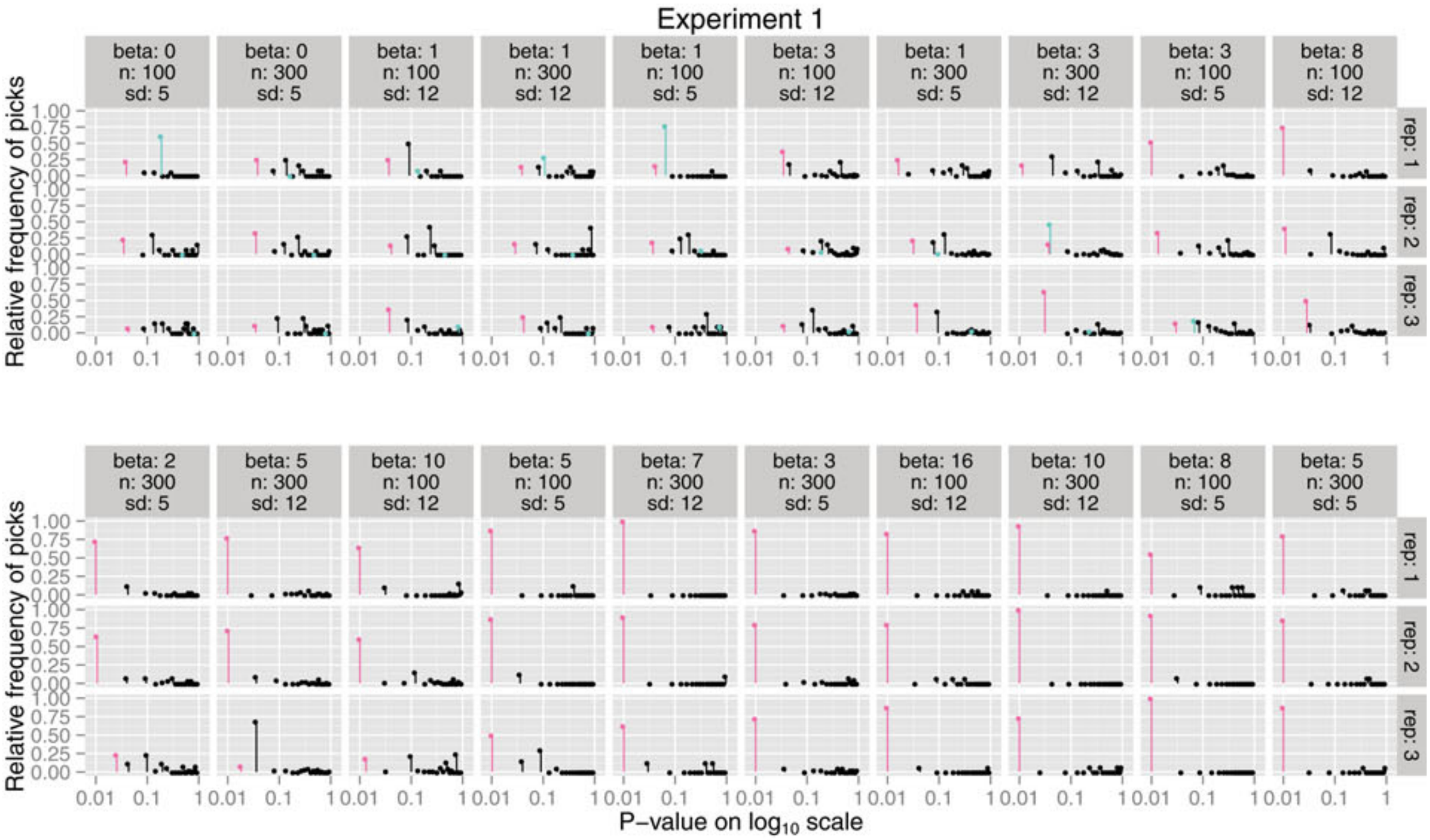
Results

- People's eyes operate similarly to classical test statistics.
- The detection of data plot is related to effect size, yielding results closer to practical significance. Less sensitive to sample size. (NOTE: We do not expect people to see statistical significance.)
- The more observers the more powerful the procedures.
- Education, gender and age do not affect detection rate



Peop that

Source: Majumder et al (2013) JASA



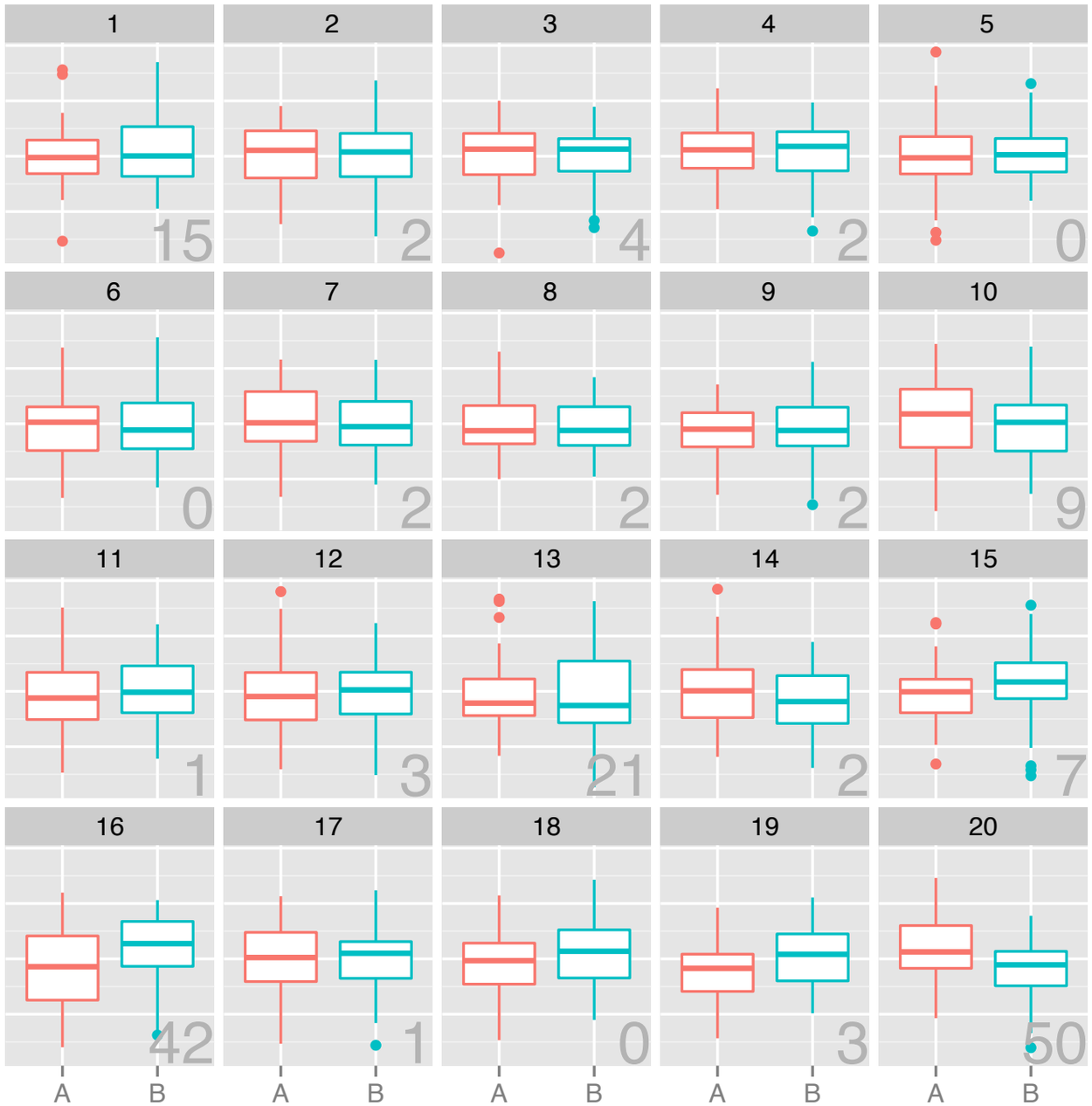
People tend to pick the plot from the lineup that has the lowest p-value

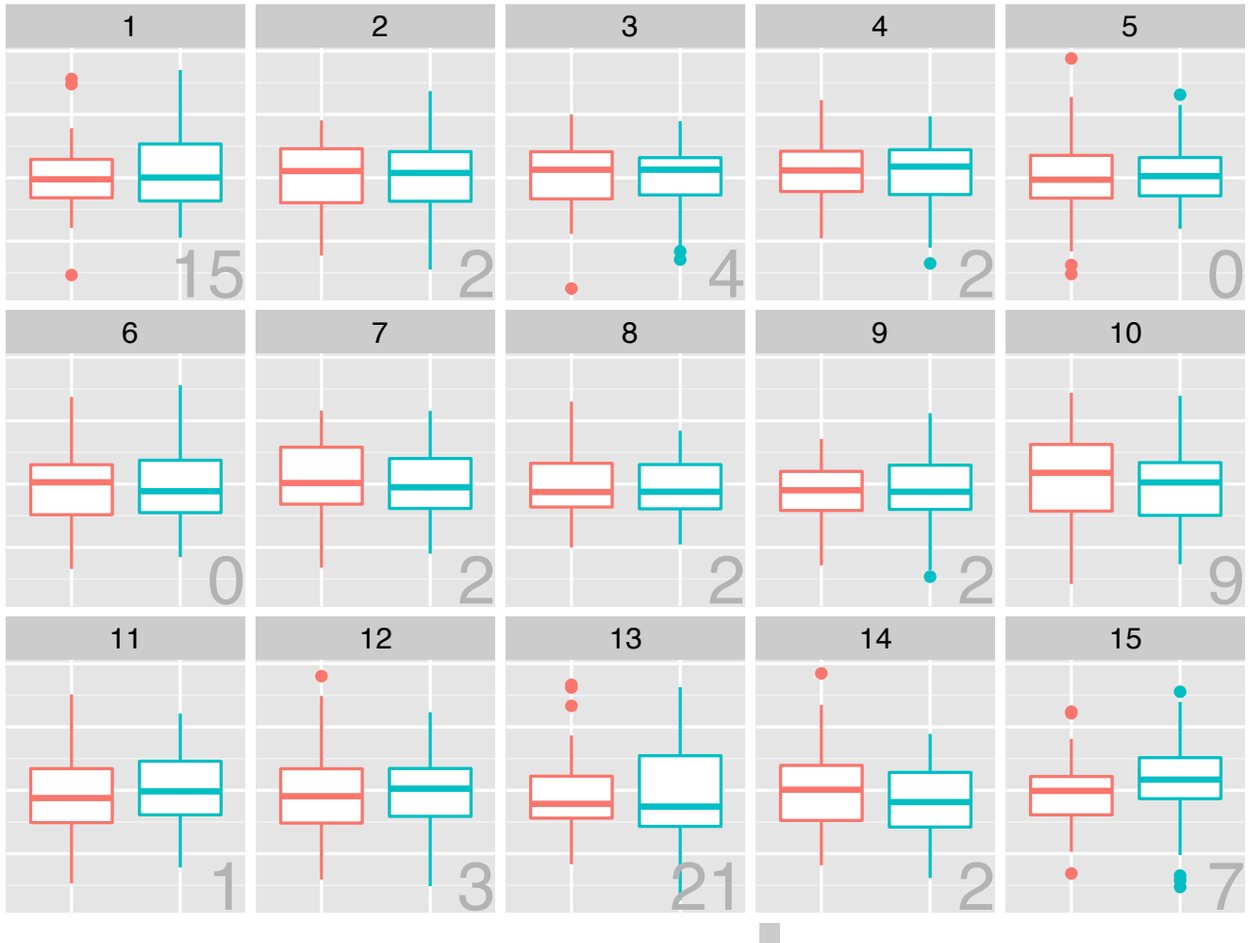
Source: Majumder et al (2013) JASA

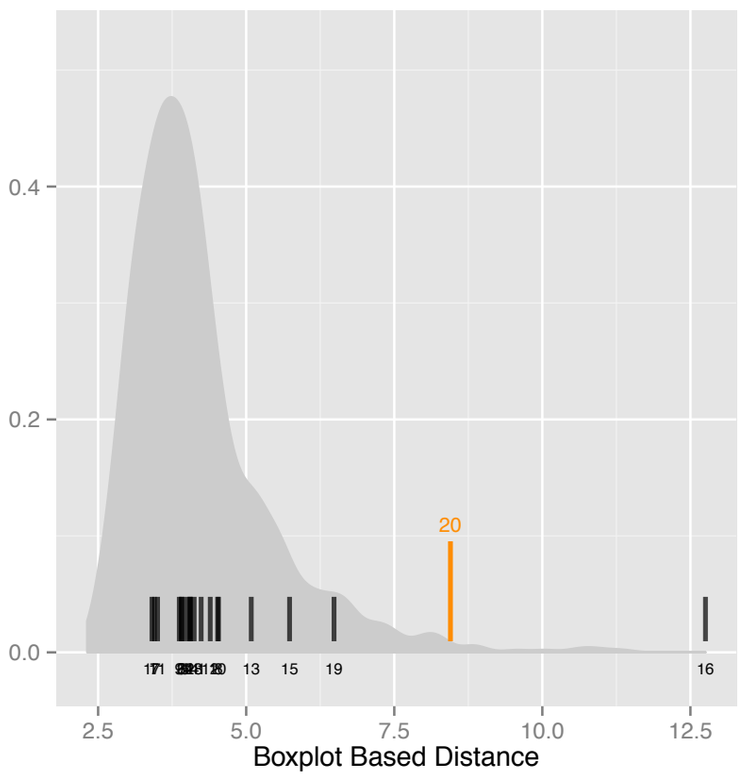
Quantify patterns?

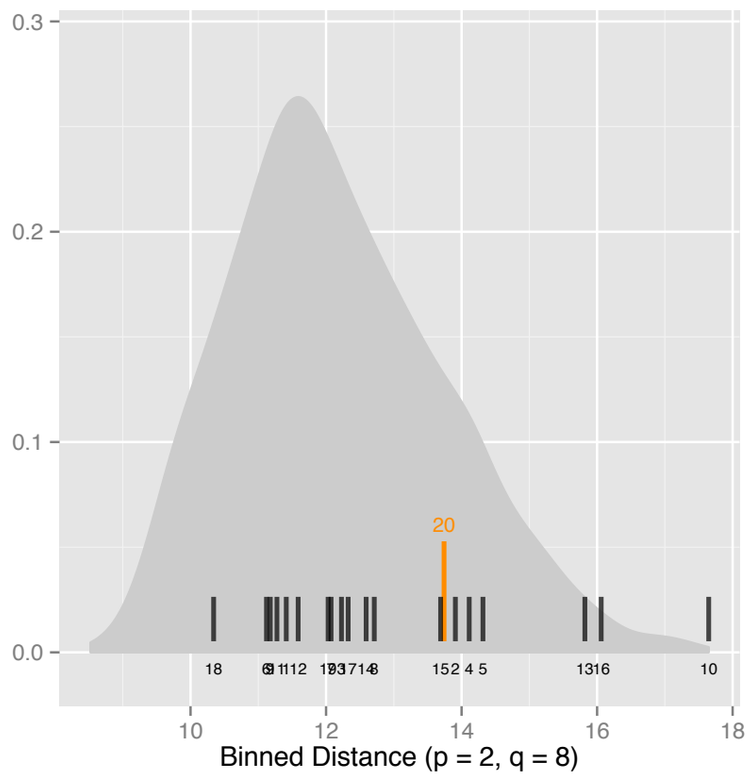
- Can we use this process to design metrics that capture structure, so that we can automate pattern recognition?
- Measure difference between data plot and null plots
- Match these to what people choose

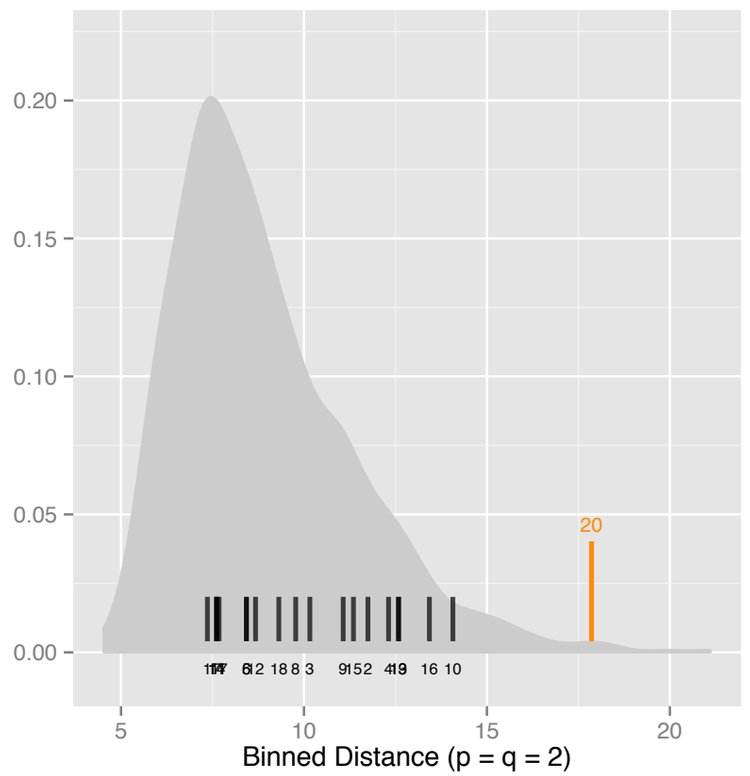
Source: Roy Chowdhury et al (2017) JCGS

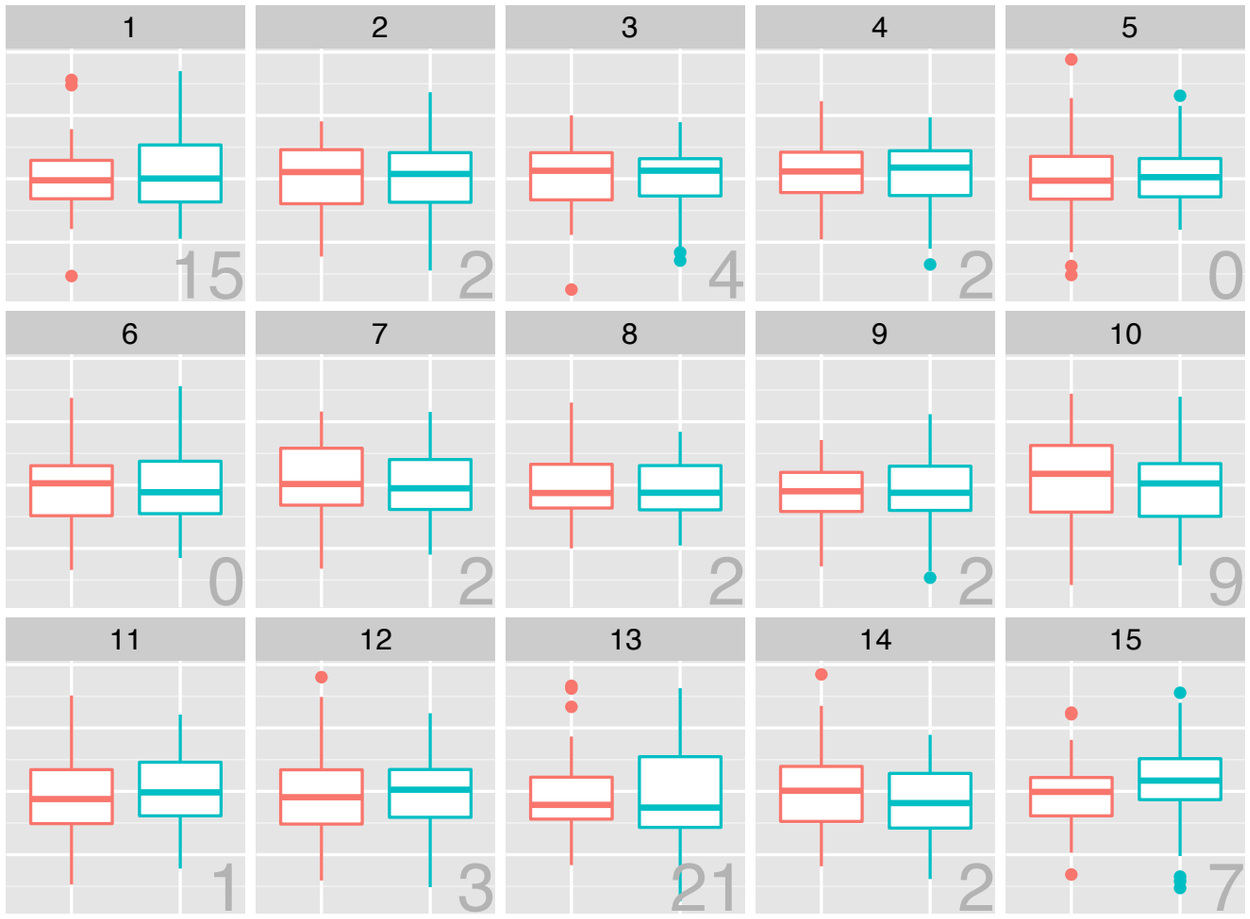

1
15
2
2
3
4
4
2
5
0
6
0
7
2
8
2
9
2
10
9
11
1
12
3
13
21
14
2
15
7
16
42
17
1
18
0
19
3
20
50
A
B


1
15
2
2
3
4
4
2
5
0
6
0
7
2
8
2
9
2
10
9
11
1
12
3
13
21
14
2
15
7

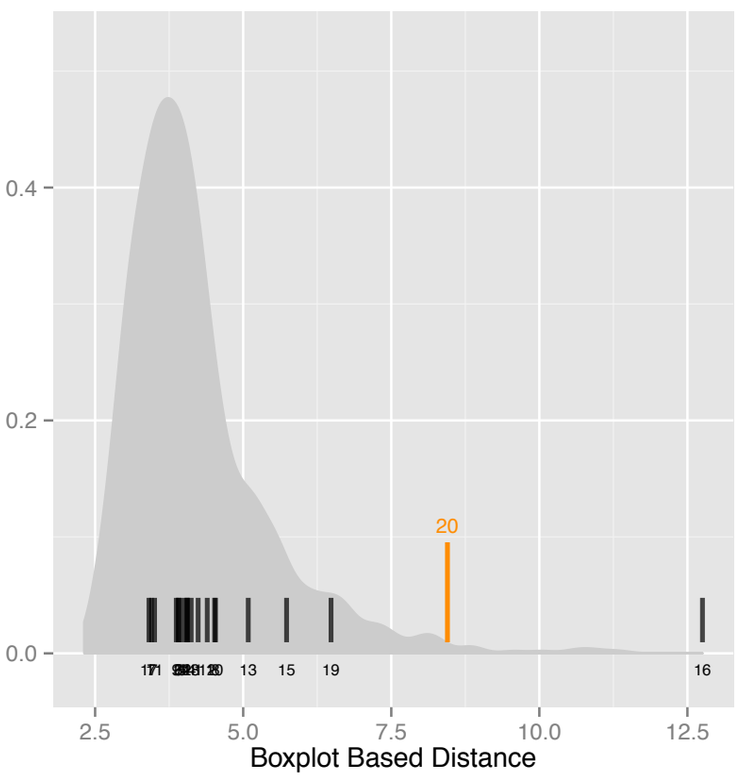

0.4
0.2
0.0
20
13
15
19
16
2.5
5.0
7.5
10.0
12.5
Boxplot Based Distance

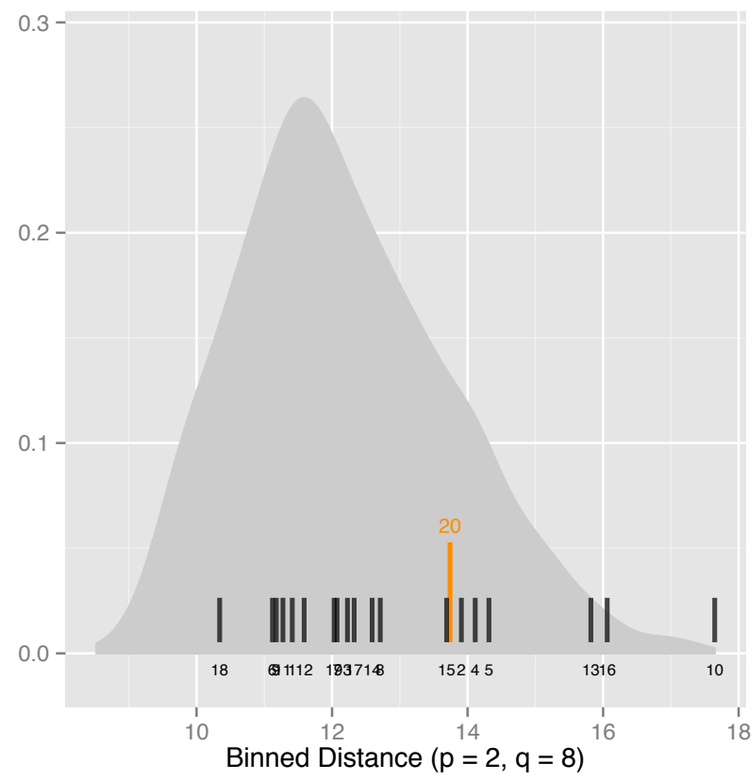

0.3
0.2
0.1
0.0
20
18
10
10
12
14
16
18
Binned Distance (p = 2, q = 8)

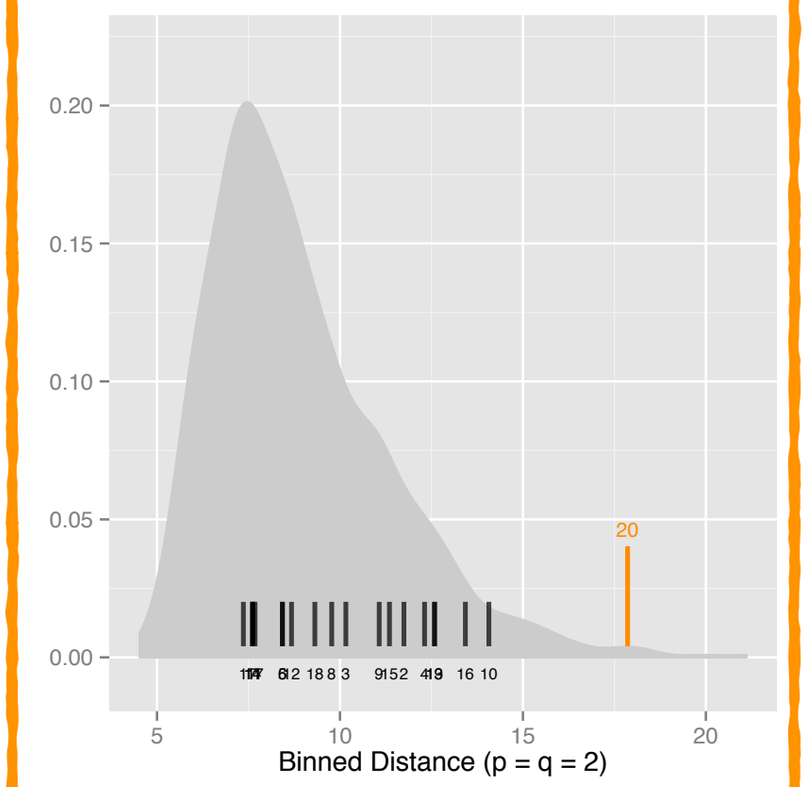

0.20
0.15
0.10
0.05
0.00
20
5
10
15
20
Binned Distance (p = q = 2)



This metric best captures what people select



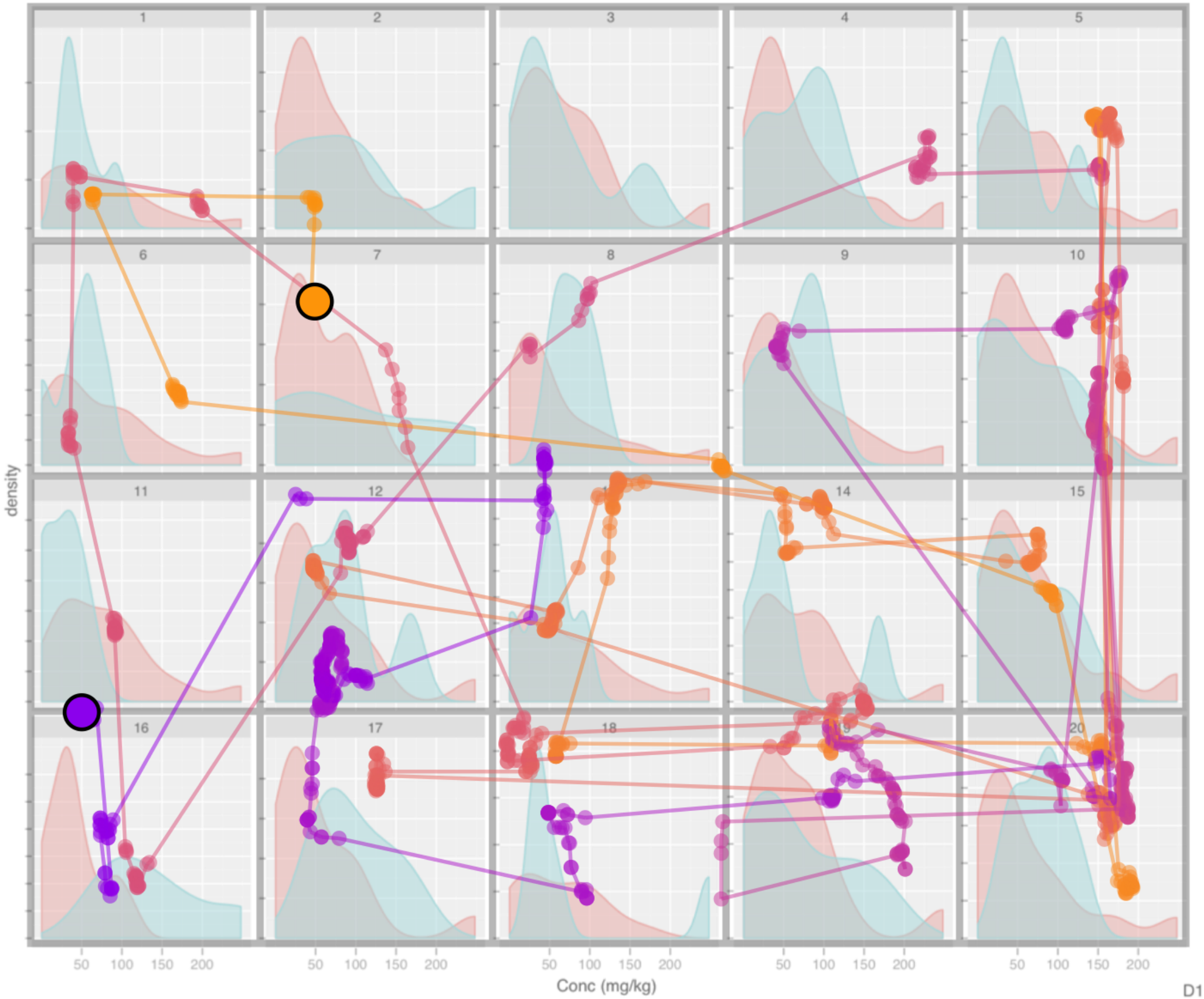




Follow-up eye tracking

- 12 lineups shown to 24 subjects
- Examined where there eyes focused in the page of plots while answering the same question as the Turk study

Source: Zhao et al (2013) IJITAS

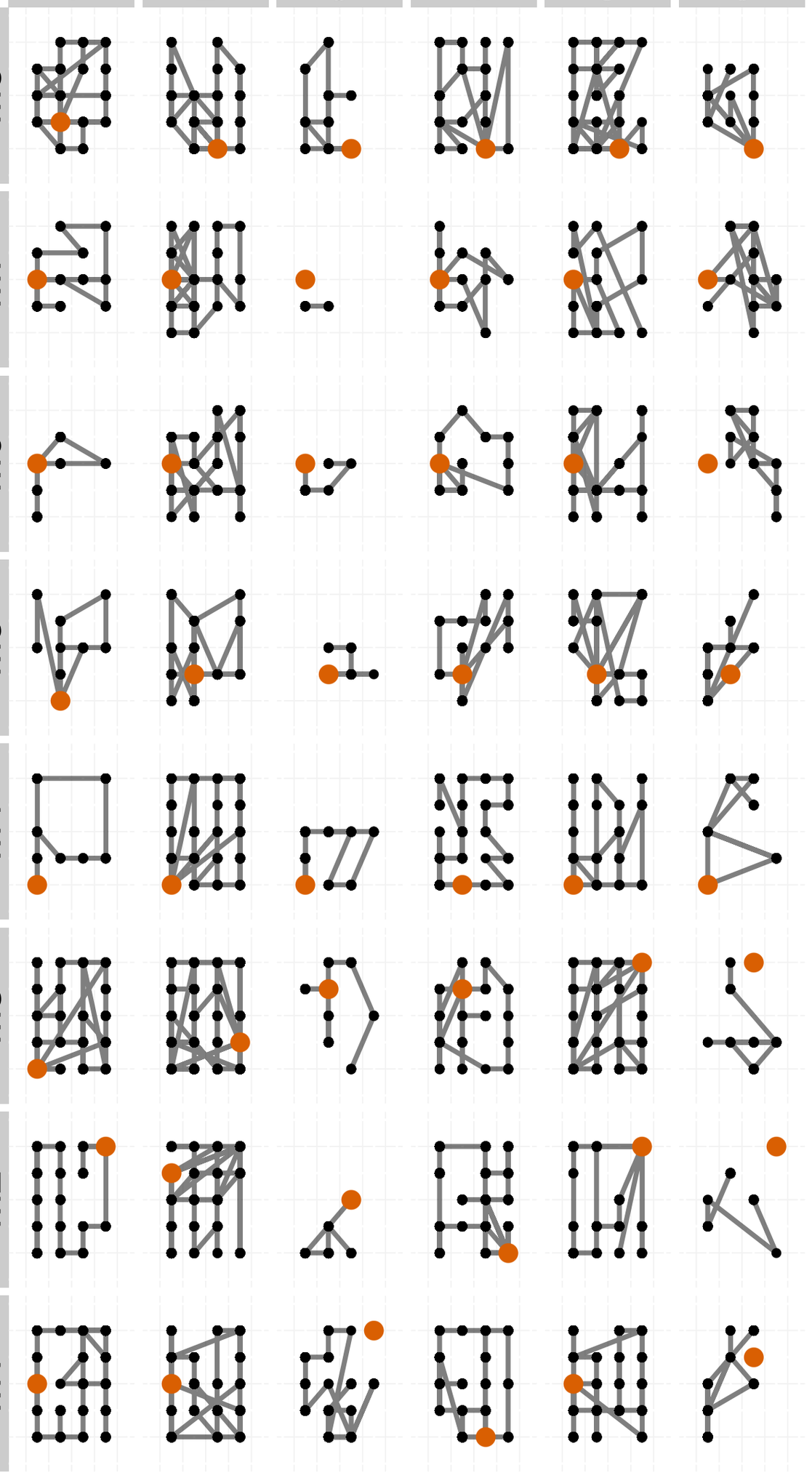

1
2
3
4
5
6
7
8
9
10
11
12
14
15
16
17
18
20
density
50
100
150
200
Conc (mg/kg)
D1

People

1 2 3 4 5 6

Lines

M8 M7 M6 M5 M4 M3 M2 M1



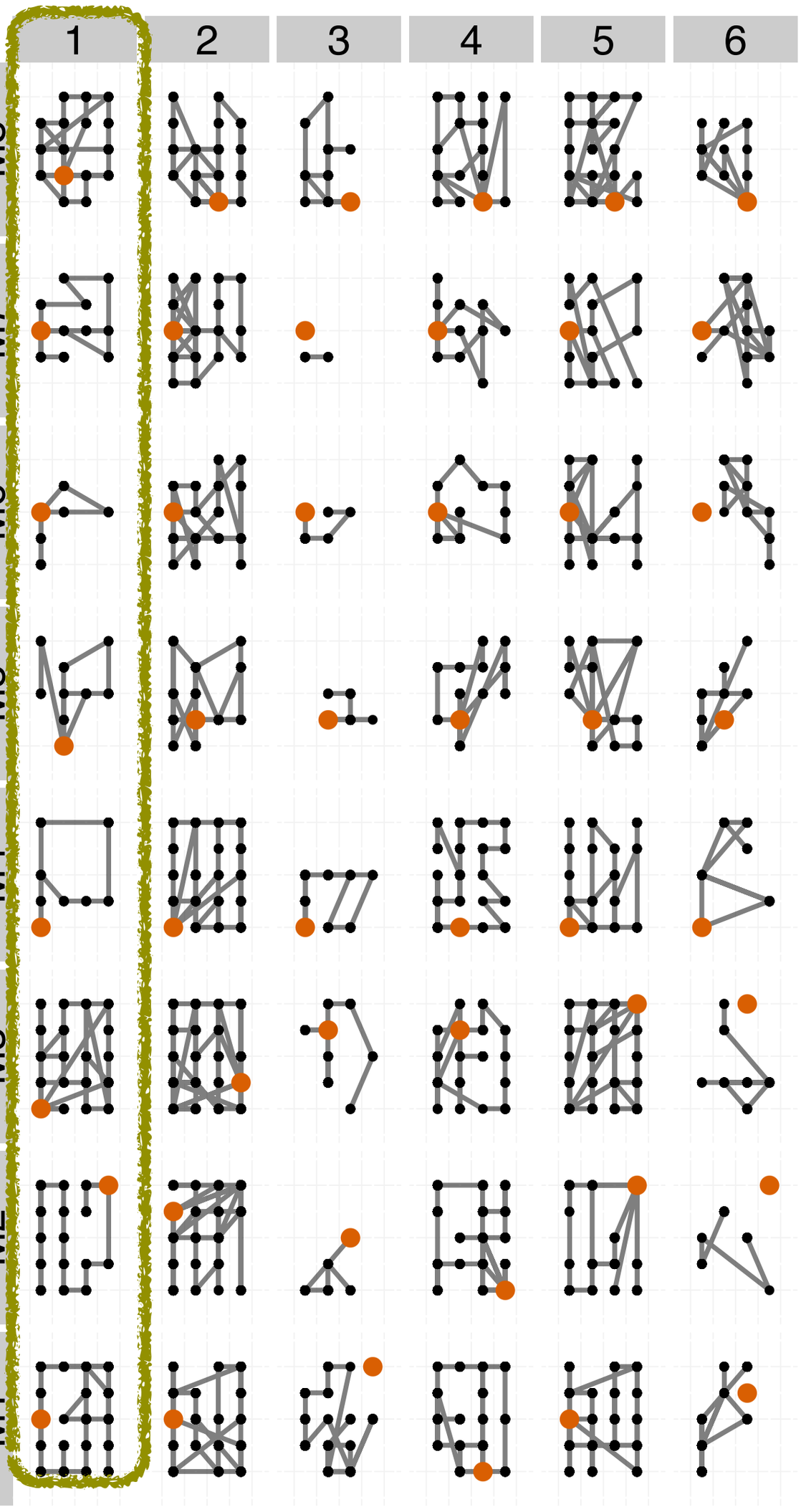
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lineups

M8 M7 M6 M5 M4 M3 M2 M1



Source: Zhao et al (2013) IJITAS

Lineups

M1 M2 M3 M4 M5 M6 M7 M8

People

1 2 3 4 5 6

Thorough

Source: Zhao et al (2013) IJITAS

Lineups

M1 M2 M3 M4 M5 M6 M7 M8

People

1 2 3 4 5 6

Thorough

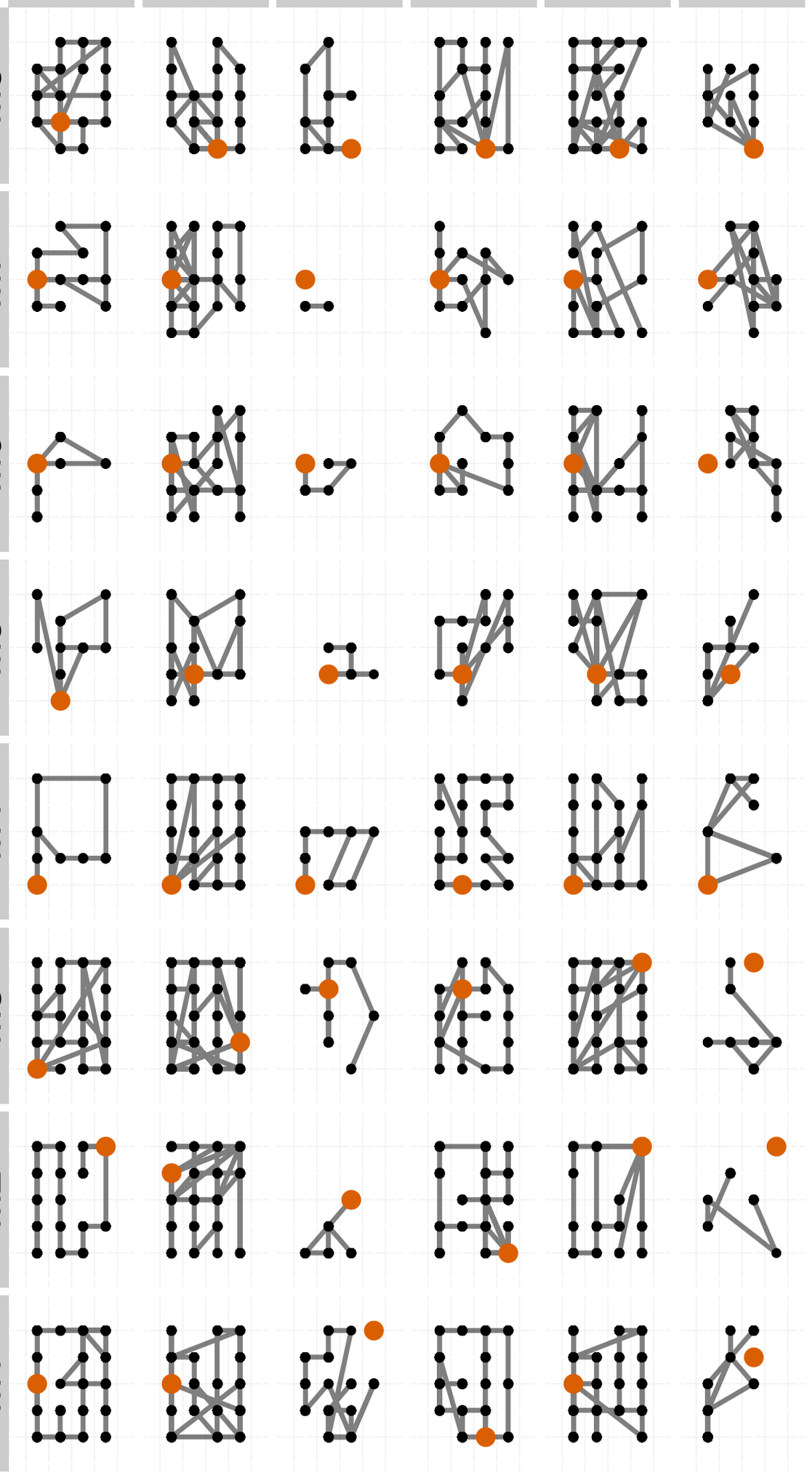
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lines

M8 M7 M6 M5 M4 M3 M2 M1



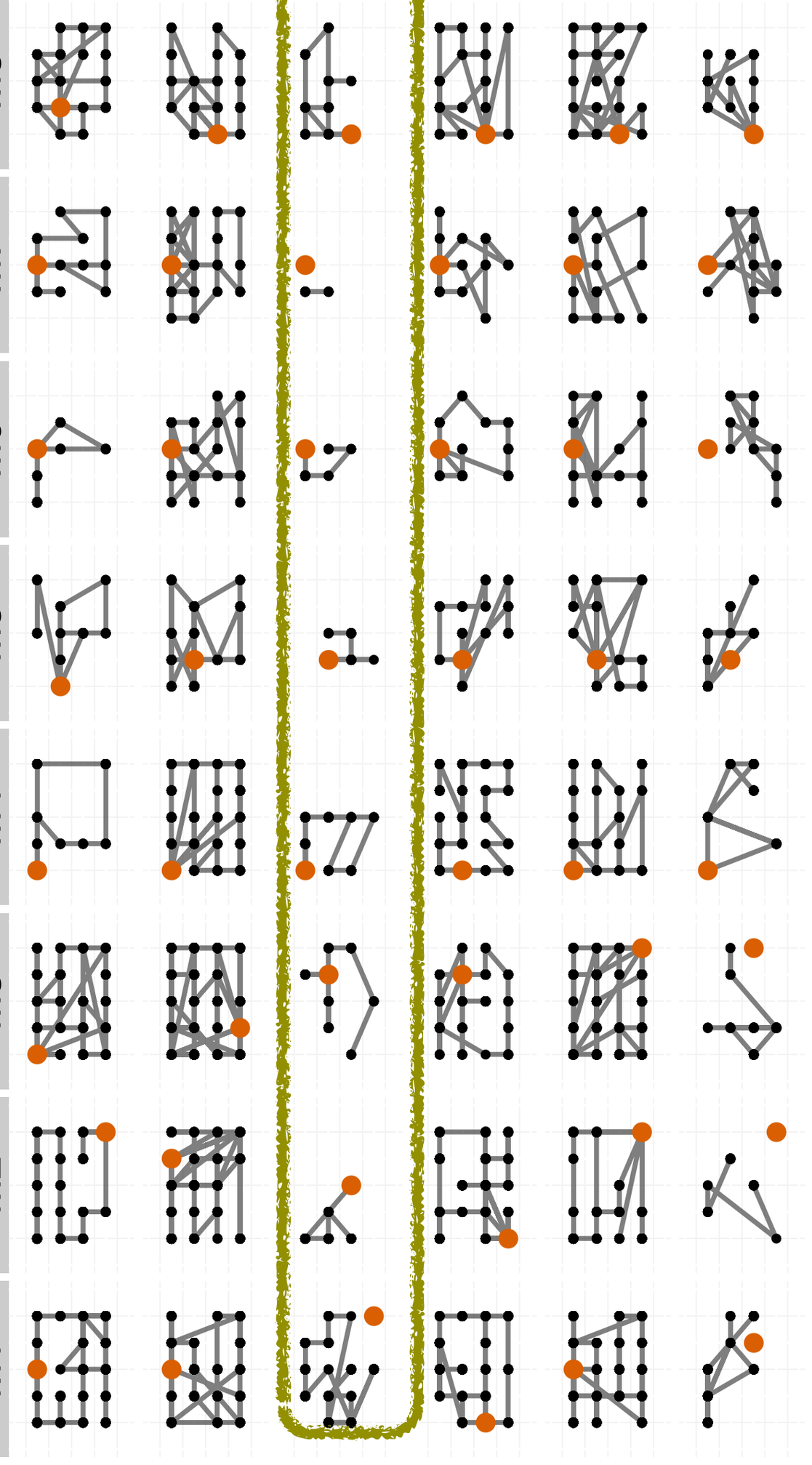
Source: Zhao et al (2013) IJITAS

People

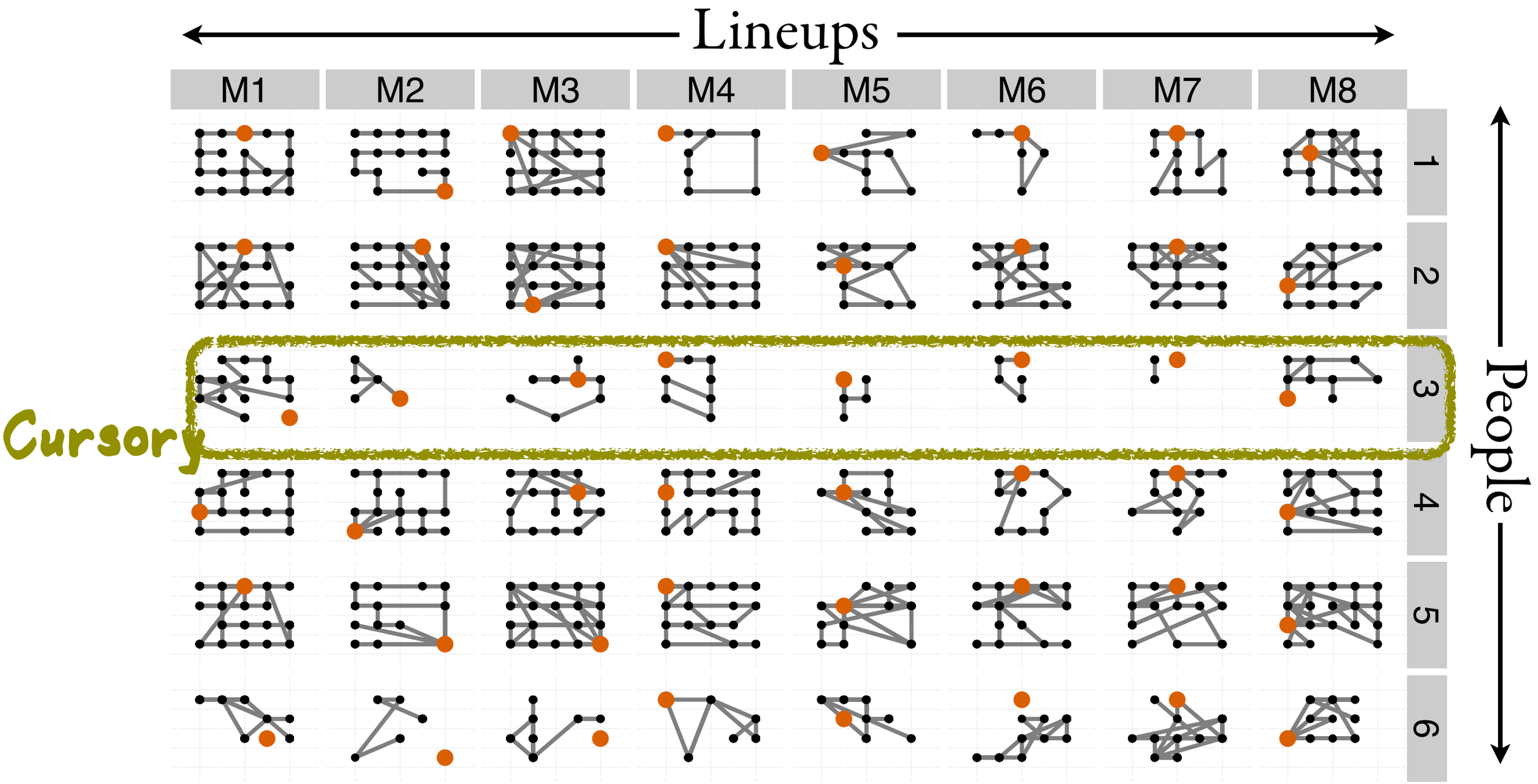
1 2 3 4 5 6

Lineups

M8 M7 M6 M5 M4 M3 M2 M1



Source: Zhao et al (2013) IJITAS



Source: Zhao et al (2013) IJITAS

Lineups

M1 M2 M3 M4 M5 M6 M7 M8

1 2 3 4 5 6

People

Cursory

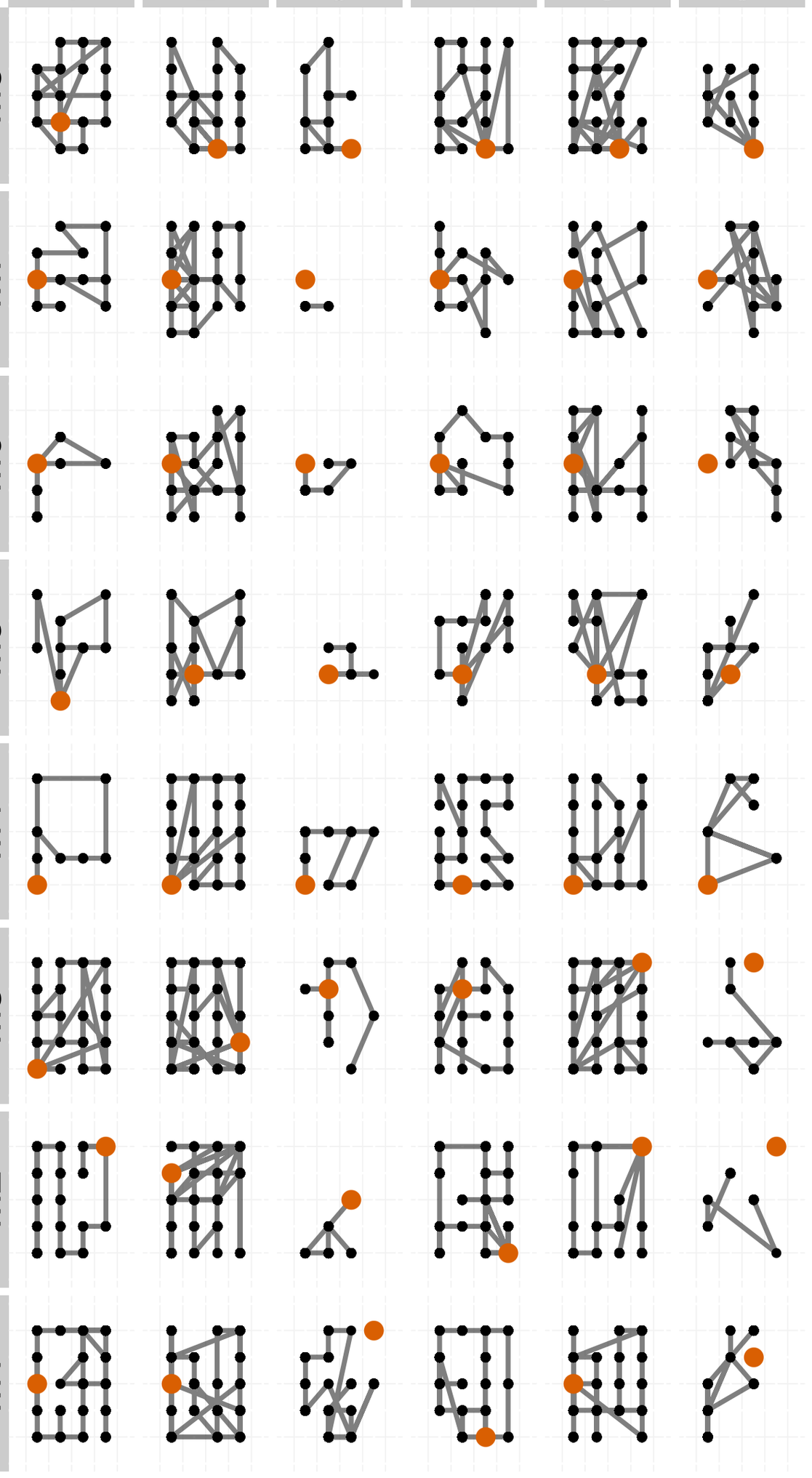
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lines

M8 M7 M6 M5 M4 M3 M2 M1



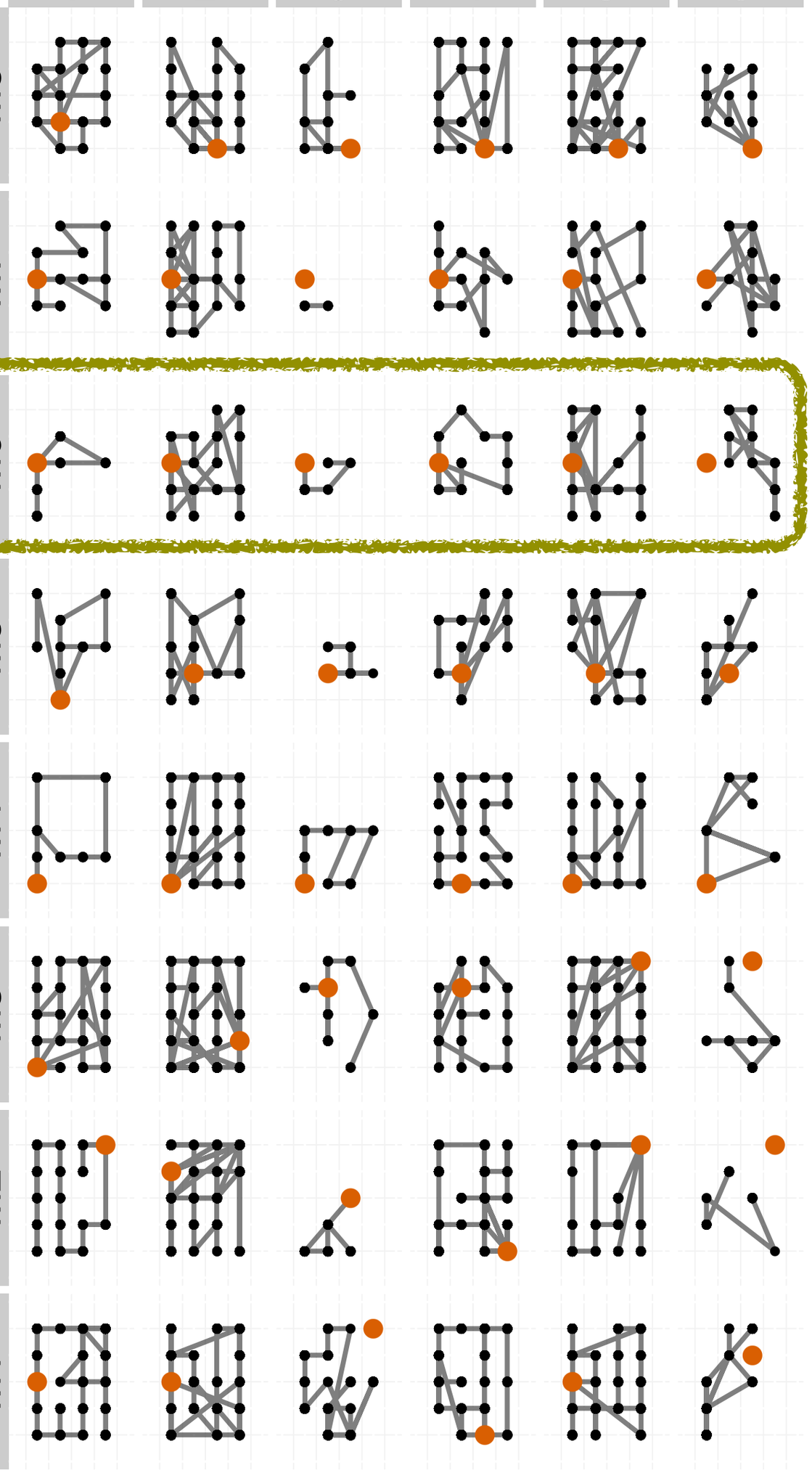
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lineups

M8 M7 M6 M5 M4 M3 M2 M1



Source: Zhao et al (2013) IJITAS

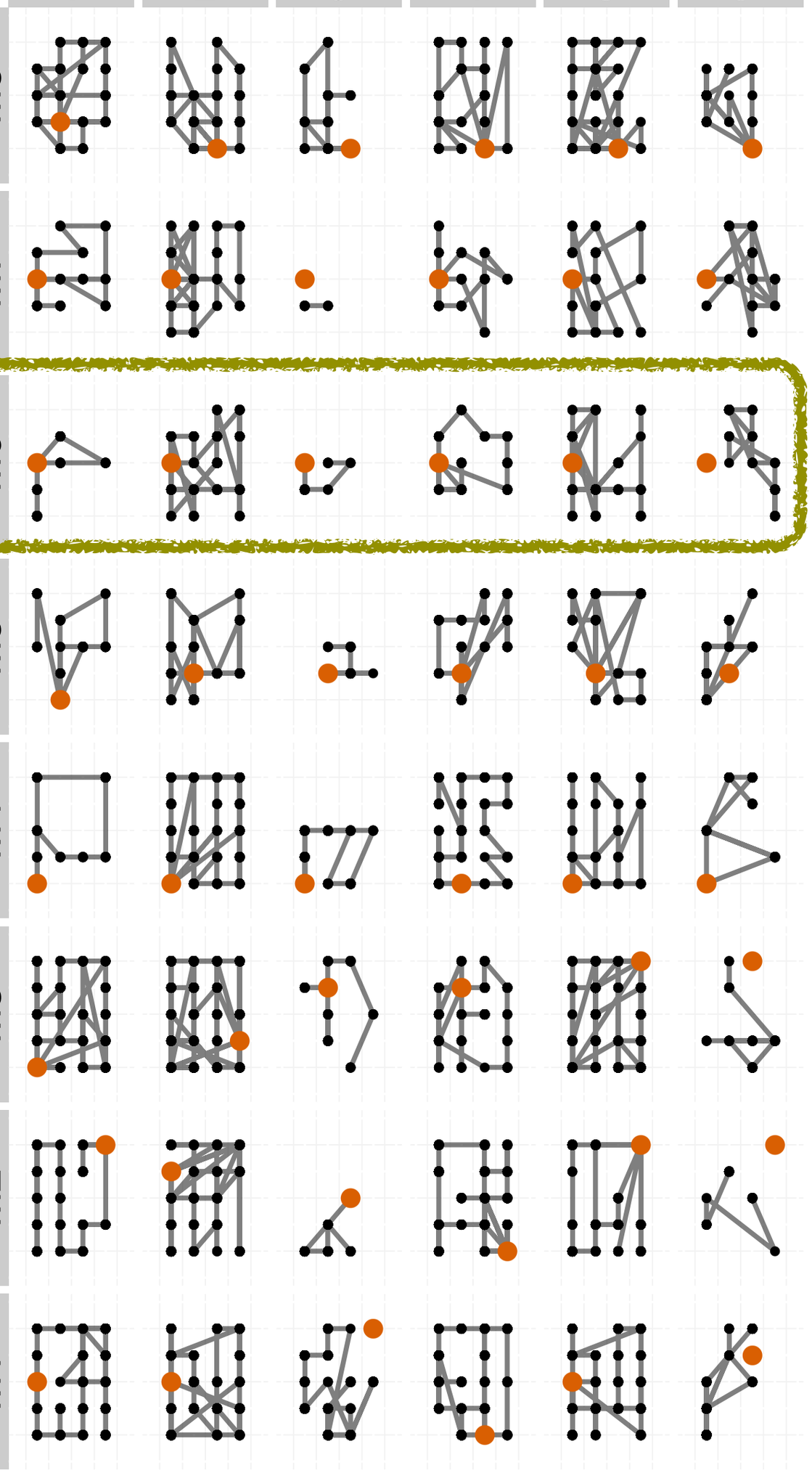
People

1 2 3 4 5 6

Easy

Lines

M8 M7 M6 M5 M4 M3 M2 M1



Source: Zhao et al (2013) IJITAS

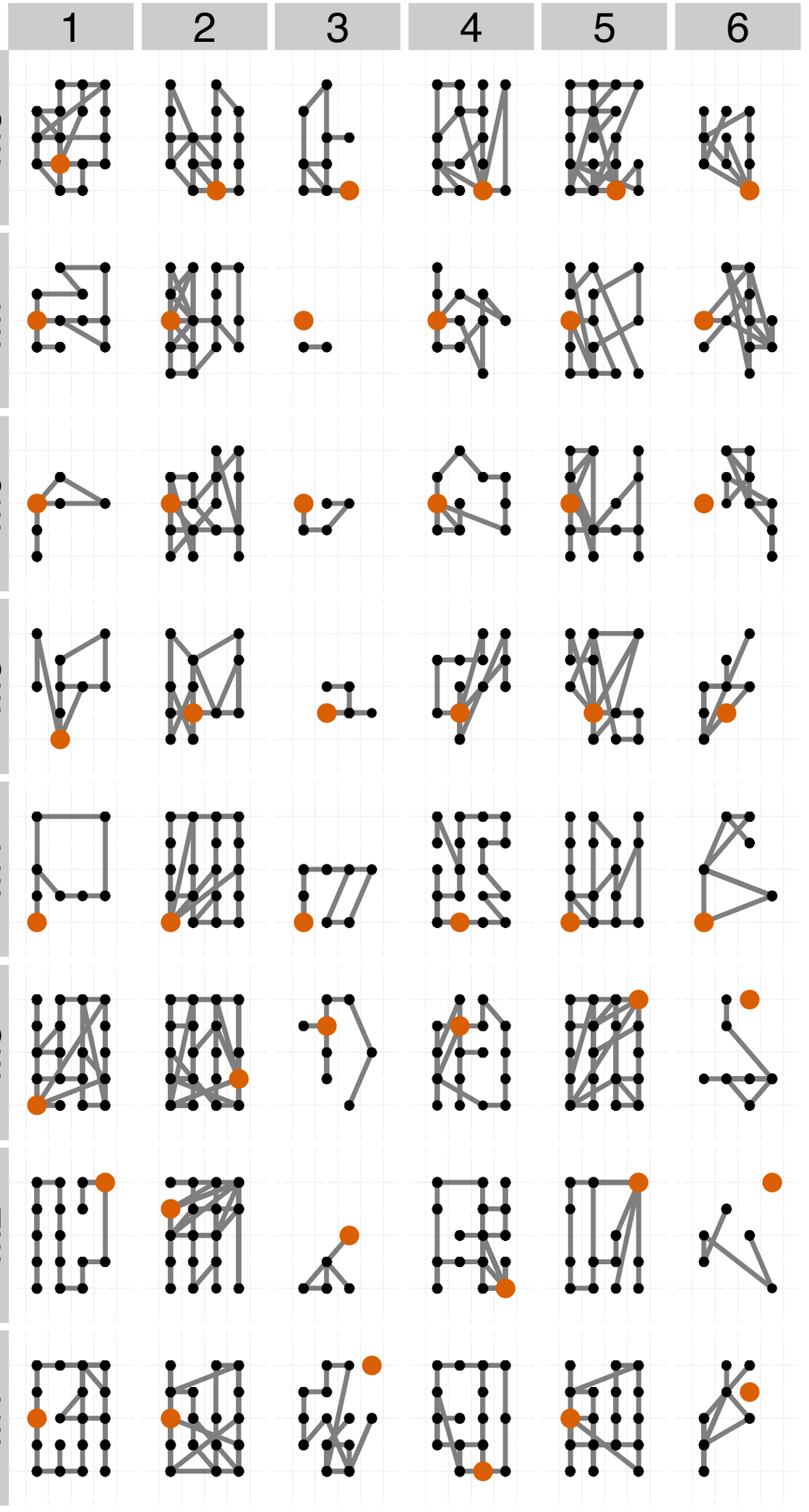
People

1 2 3 4 5 6

Easy

Line ups

M8 M7 M6 M5 M4 M3 M2 M1



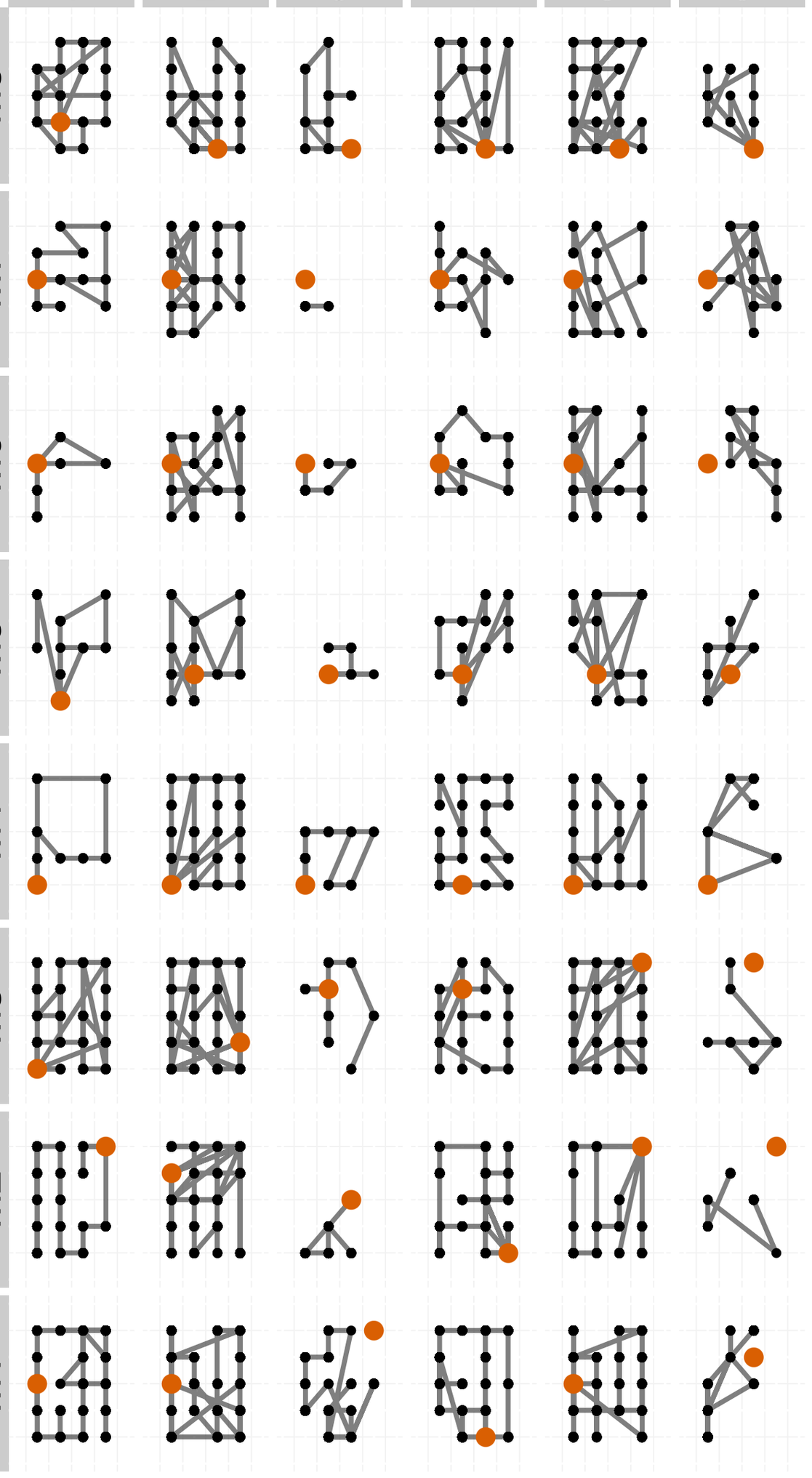
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lines

M8 M7 M6 M5 M4 M3 M2 M1



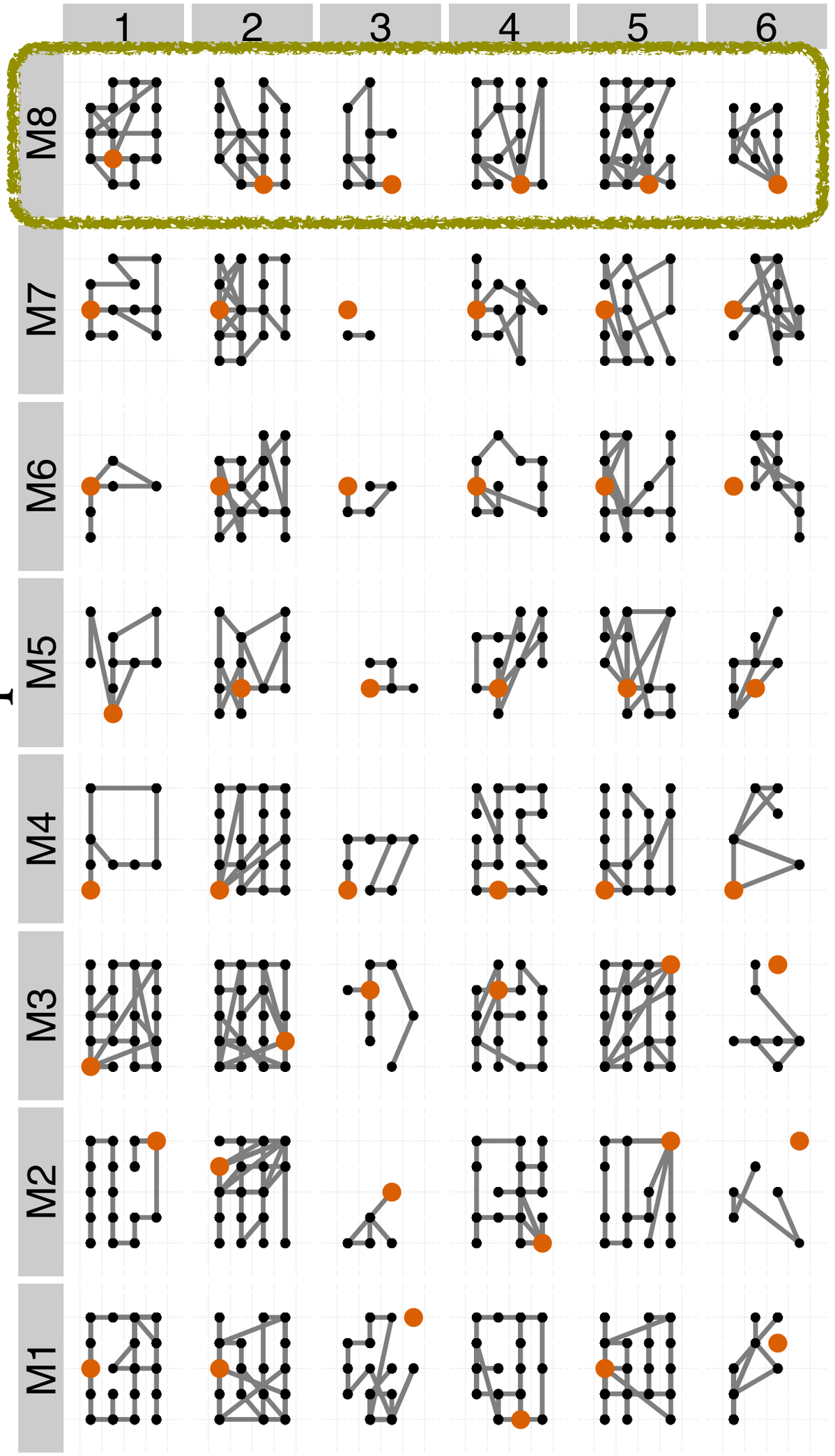
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Line ups

M8 M7 M6 M5 M4 M3 M2 M1



Source: Zhao et al (2013) IJITAS

Lineups

Hard

M1 M2 M3 M4 M5 M6 M7 M8

People

1 2 3 4 5 6

Source: Zhao et al (2013) IJITAS

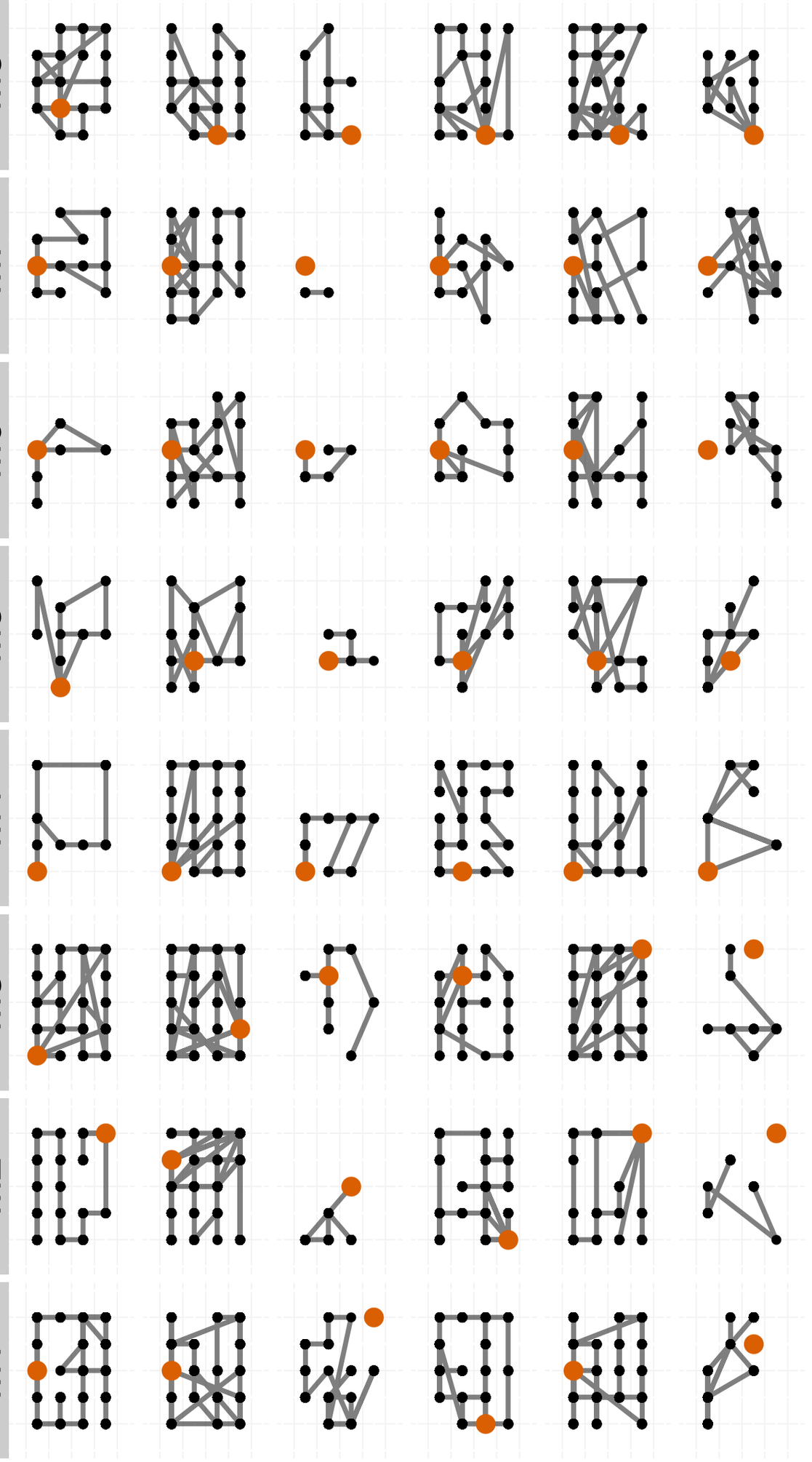
People

1 2 3 4 5 6

Hard

Lines

M8 M7 M6 M5 M4 M3 M2 M1



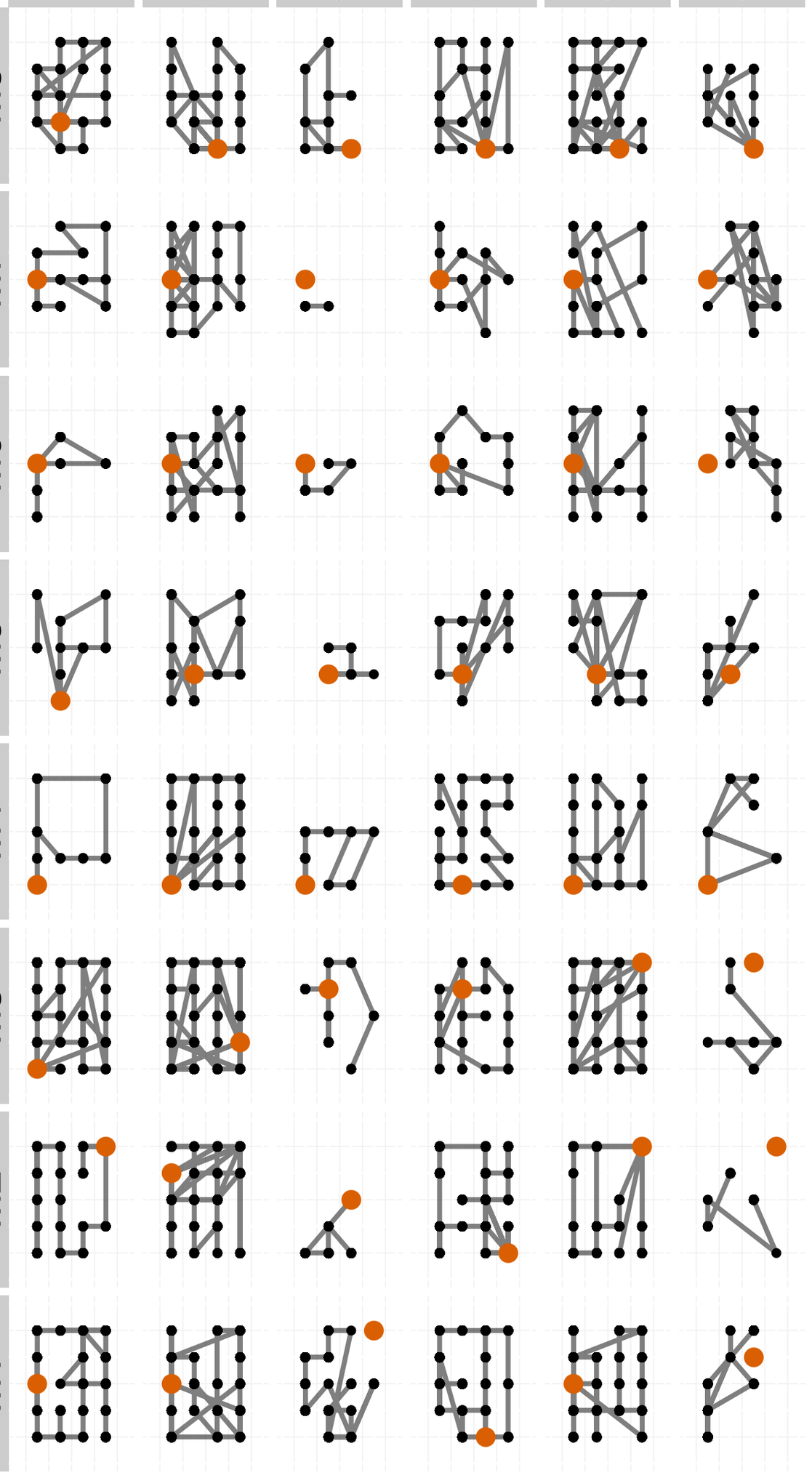
Source: Zhao et al (2013) IJITAS

People

1 2 3 4 5 6

Lines

M8 M7 M6 M5 M4 M3 M2 M1

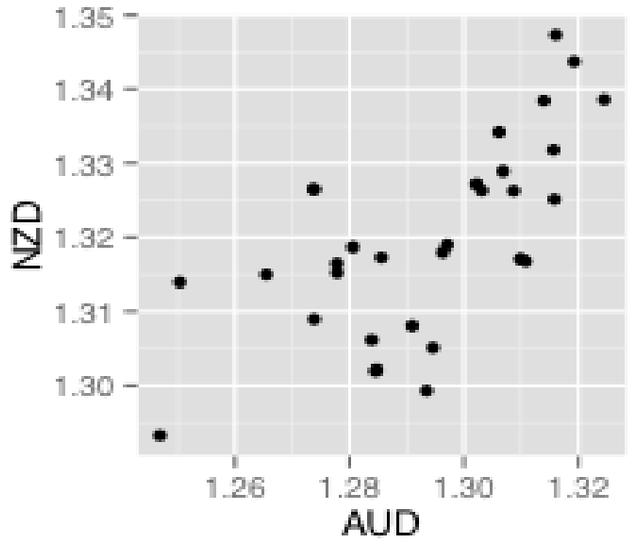


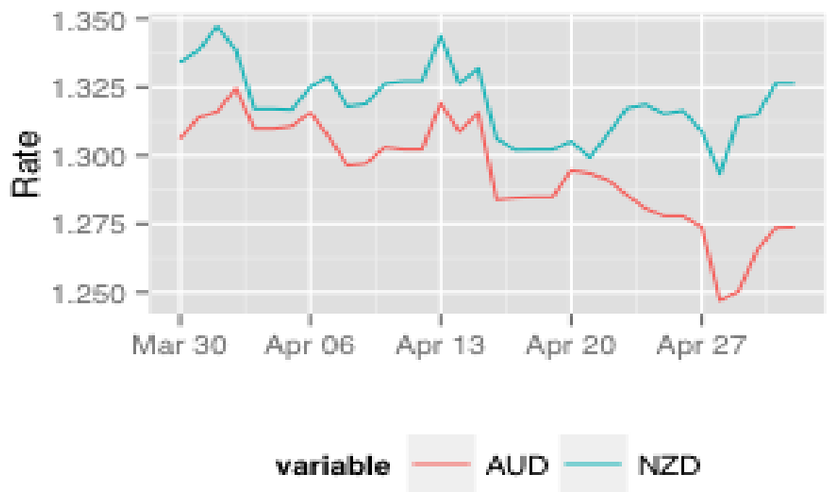
Source: Zhao et al (2013) IJITAS

Eye-tracking

- People tended to methodically look around the lineup to find the most unusual plot
- For more difficult lineups all looked harder, and for easier lineups the looking was fast
- Some people were much faster than others in looking
- There are two stages, initial scan, and a final comparison between a few plots

Testing plot design

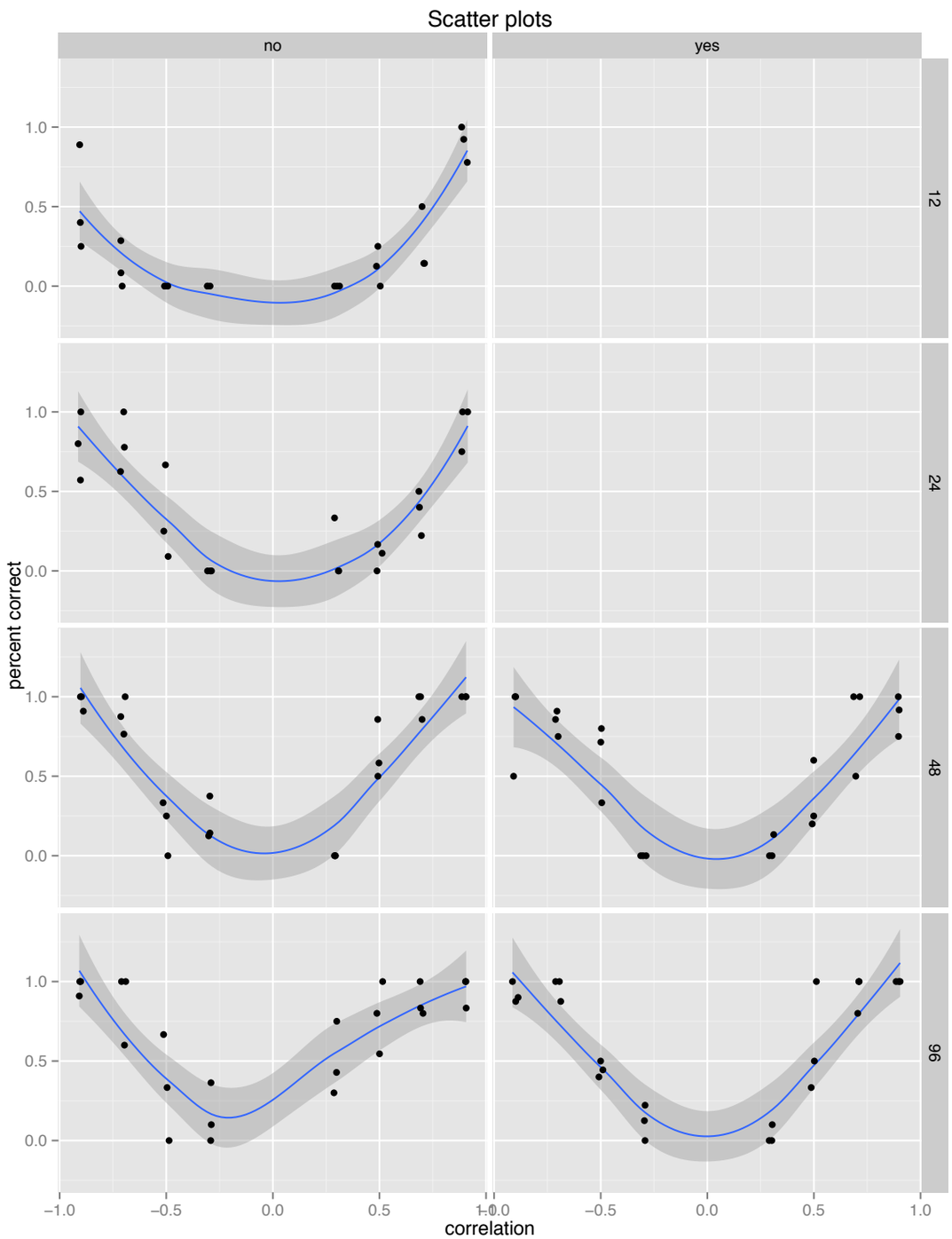


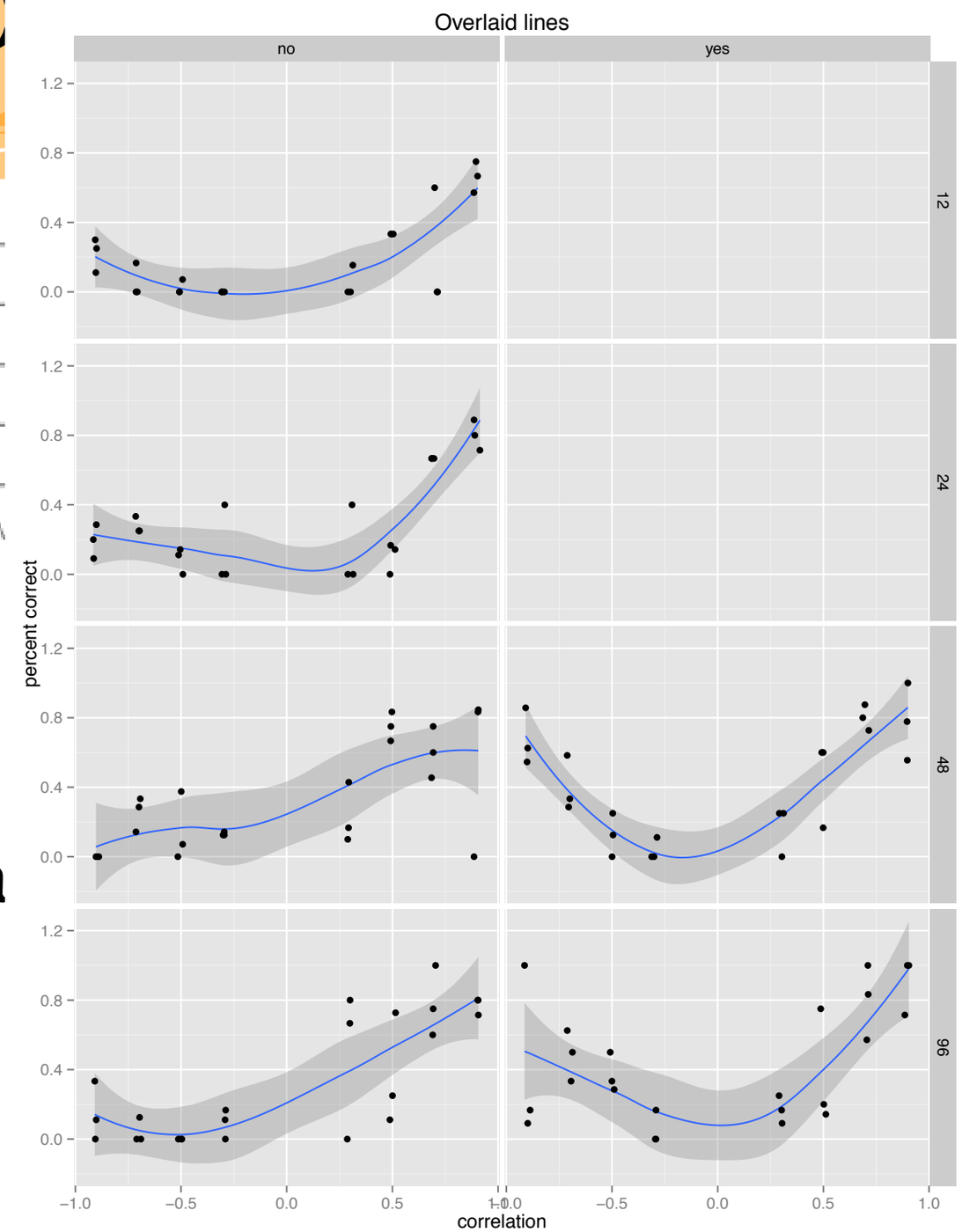


Which is the best display? Scatterplots universally better for reading association.

Methods in Hofmann et al (2012) InfoVis; Example from Tomasetti honours thesis, 2015

Testing plot design



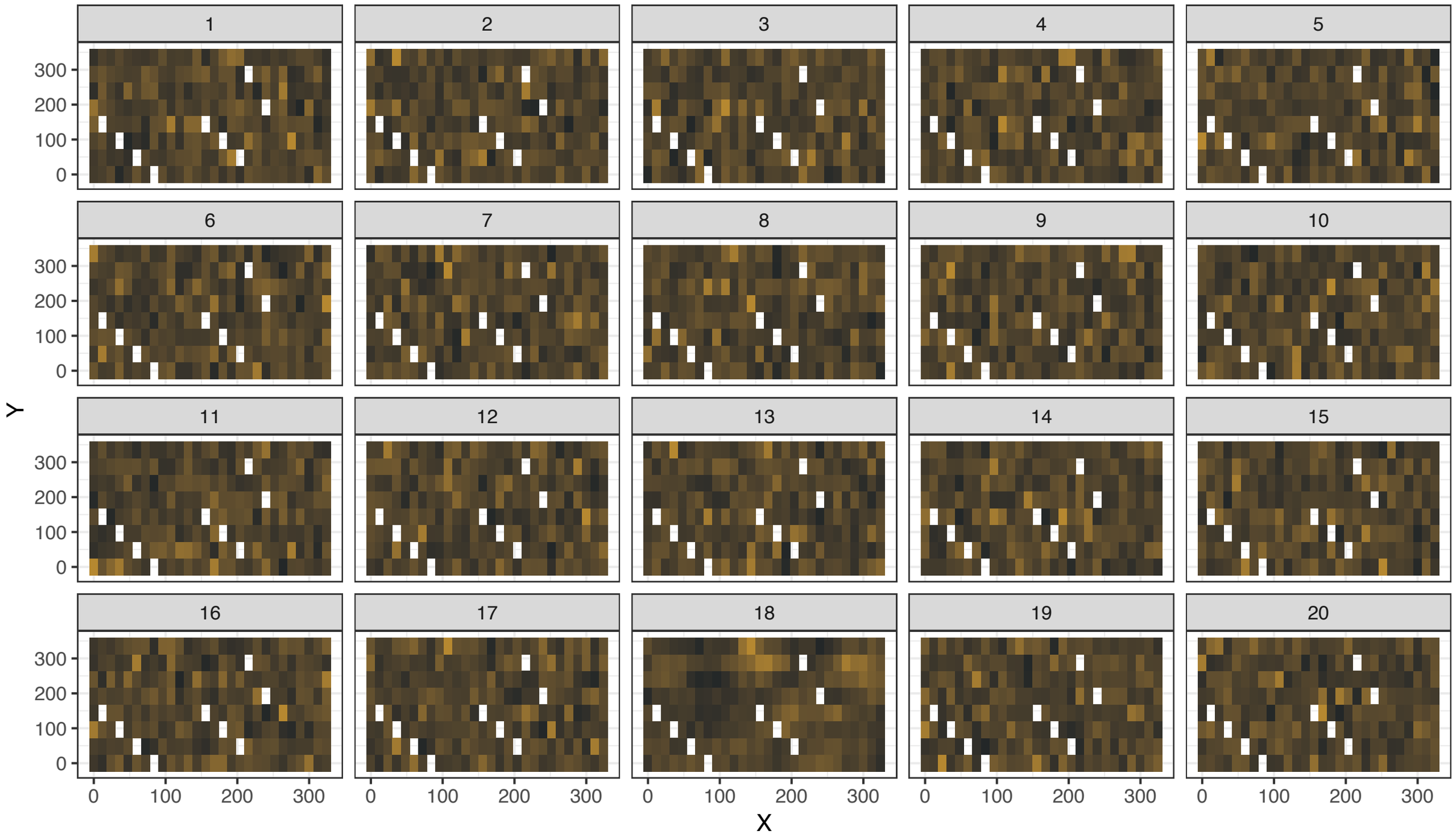


-1 <— 0 —> 1

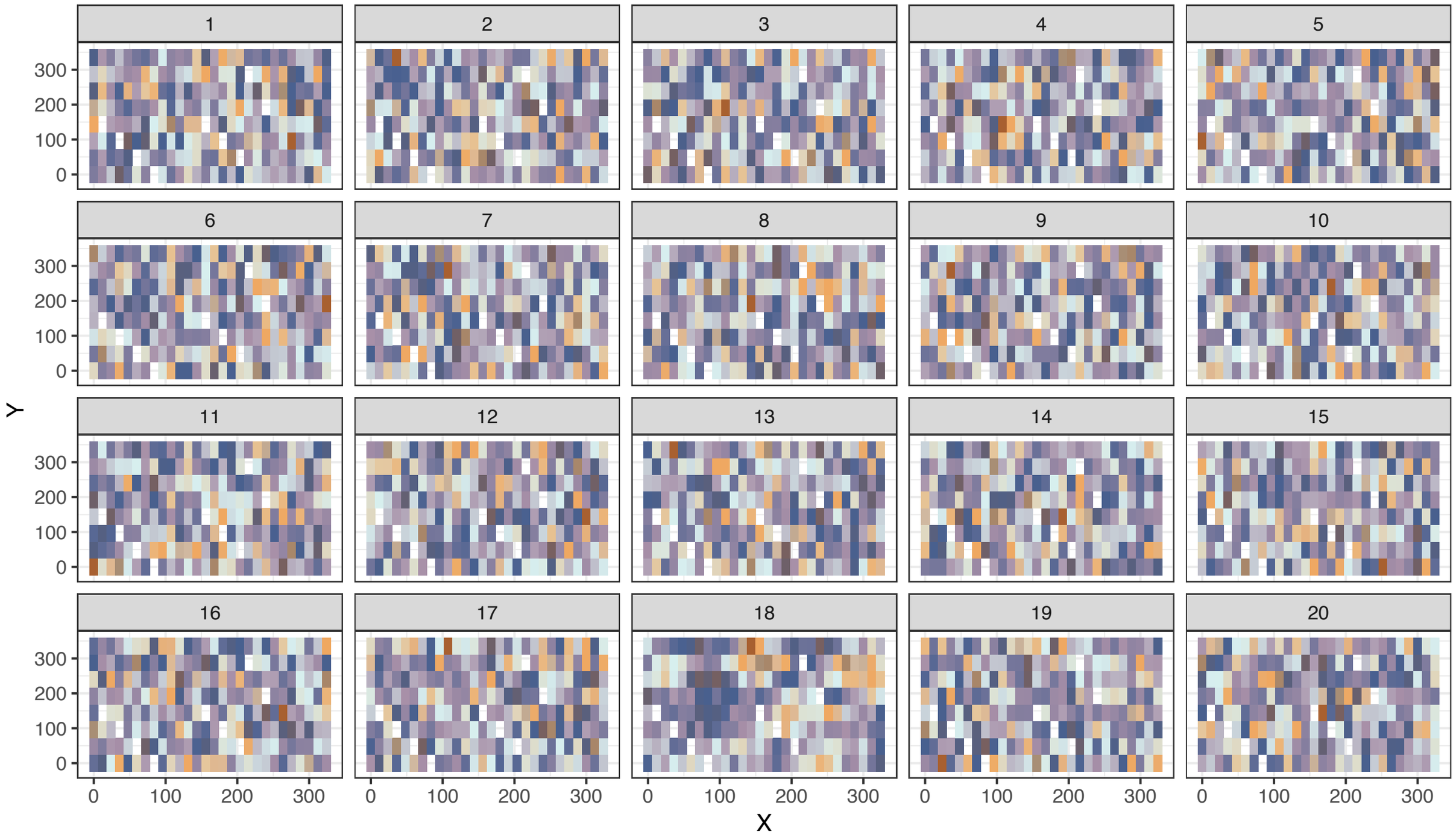
-1 <— 0 —> 1

-1 <— 0 —> 1

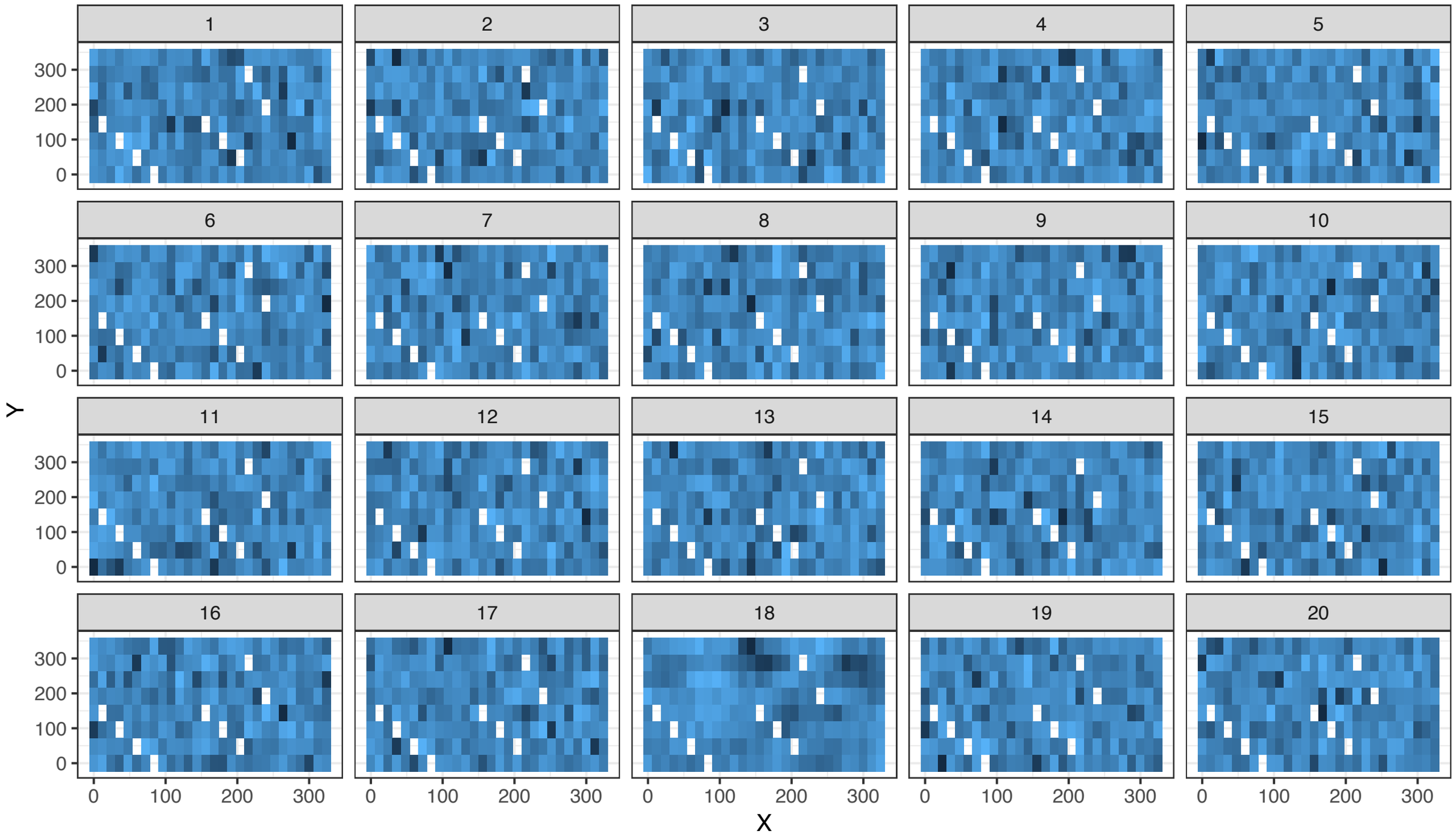
-1 <— 0 —> 1



Does colour affect the ease of detection of the data plot?



Does colour affect the ease of detection of the data plot?



Does colour affect the ease of detection of the data plot?

Some foundations...

- Scott et al (1954): Generated synthetic plates to compare with real astronomical plates, acknowledged in Brillinger's (2005) Neyman lecture.
- Daniel (1976) had 40 pages of null plots for industrial applications.
- Diaconis (1983) describes 'magical thinking'.
- Buja et al (1988) describe 'Informal Statistical Inference' in association with the software Dataviewer.
- Gelman (2004) simulate data from statistical models.
- Davies (2008) suggest viewing null data sets.

R Package

`nullabor` package on CRAN and `https://github.com/dicook/nullabor`

When you plot your data, plot it first in a lineup, so you can be the unbiased observer

Several null generating procedures included, *p*-value and power calculations

R Package

`nullabor` package on CRAN and `https://github.com/dicook/nullabor`

When you plot your data, plot it first in a lineup, so you can be the unbiased observer

```
> lineup(null_permute("Obama.Romney"),
tracking.polls[,c(9,11)])
> decrypt("fg0t DARA up iYzuRuYp Q")
[1] "True data in position 5"
```

Several null generating procedures included, *p*-value and power calculations

Procedure

1. Organise your data into tidy form
2. Define your plots using the grammar of graphics, where variables are clearly mapped to elements of the plot
3. Specify the null-generating mechanism
4. Make your plot using a lineup function that shows the data plot among null plots
5. Is the data plot distinguishable from the nulls? If so, you can say "YOU CAN SEE ..."

Summary and future work

- This work melds inferential and exploratory statistics.
- It has been used to provide new visual diagnostics for hierarchical linear models, and an objective way to compare plot designs.
- Develop new metrics on data plots, with a plan to automate some data plot reading.
- Lots more work on p-value and power estimation needed.
- Web service to evaluate lineups

Acknowledgements

Plots produced using R package **`ggplot2`** by Hadley Wickham

Lineups made with R package **`nullabor`**

Projection pursuit (experiment 7) done using R package **`tourr`** by Wickham, Cook, with PDA index from Lee

National Science Foundation grant DMS 1007697

These slides are available at: `https://github.com/dicook/Biometric2017`

Heike

Andreas

Debby

Hadley

Eun-kyung

Eric

Mahbub

Tengfei

Susan

Adam

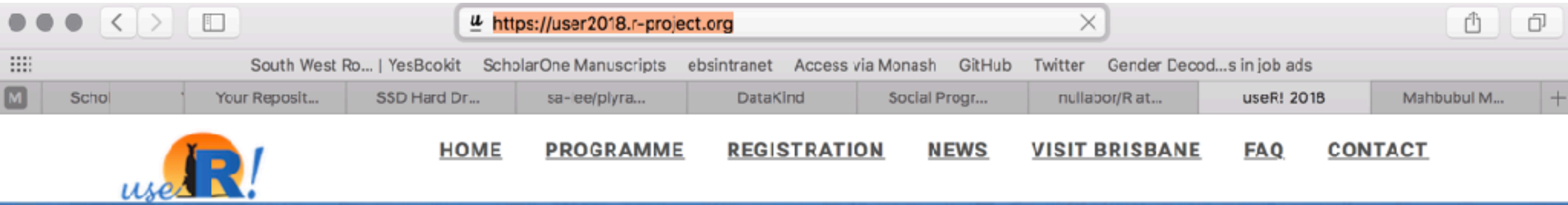
Niladri

Nathaniel

useR! 2018 is in Australia next year!!!!

First time ever it has been held outside Europe and USA. Starting a trend!!!

If you want a postcard sent to you, email me, dicook@monash.edu



Australia has the best place for conferences!

Thanks to the Biometrics by the Border organisers for changing the natural order of the world, and the great location!