Topics in Data Visualization

Perception Of Groups

Jul | | 2014

Charlotte Wickham



How many 6's are there?

How many 6's are there?

6078165891098515805655349090099011 7297800395994427458150494855524330 15220029552727741482039704437699

6078165891098515805655349090099011 7297800395994427458150494855524330 15220029552727741482039704437699

2 2 1 4 5 0 1 4 **6** 5 5 9 9 5 0 1 9 3 1 9 2 4 1 9 0 9 0 8 5 8 **6** 0 7 3 3 1 9 1 9 0 4 7 7 0 1 9 2 8 0 7 7 5 3 3 0 9 7 1 9 1 9 3 3 9 3 0 5 5 8 8 5 3 5 7 9 2 1 3 7 7 2 4 3 7 1 9 9 2 5 4 7 1 4 9 5 8 0 4 3 8 3 0 2 9 7 3 0 1 8 9 0 9 4 3 9 5 2 4 9 7 8 4 4 7 8 **6** 1 5 7 9 4 2 8 0 5 1 0 2 9 7 8 4 4 2 8 2 5 4 0 3 2 7 4 4 5 9 1 3 1 5 1 7 3 2 5 1 9 8 1 3 7 4 **6** 4 0 4 7 3 4 1 7 0 3 5 0 7 4 5 5 5 4 9 7 2 3 4 4 4 8 9

Preattentive processing

- Some things just seem to "pop out"
- A visual search consists of a series of eye fixations at "likely" locations.
- A pattern is preattentive if the likely locations are very quickly mapped, and the time it takes doesn't depend on how many distractors there are.
- For graphics: if you want something to stand out use a preattentive property.







C. Ware. Information visualization: perception for design. Elsevier, 2013.

not





Find points in group "a".

Distinction is easier when:

the difference between target and non-target is large the difference of non-target's from each other is small



How many colours/shapes/orientations/ sizes can I use before processing stops being preattentiive? How could I design an experiment to

find out?

Task: find a group, count members in the group, how many groups are there? Encoding: how are we choosing colours/ size

Design variable: number of groups

Response: time to complete task, accuracy.

Advice: Not more than 7, probably better if it's less than 5.

Exploratory analysis: interaction can help.

Combining features

Redundant coding is almost always beneficial.





Which two are most similar?



Which two are most similar?





Figure 5.19 (a) The width and height of an ellipse are perceived integrally, so the ellipses are seen as more similar to each other (because they have the same shape) than the pair having the same width. The color and height of a shape are perceived separably, so the two green shapes are seen as most similar. (b, c) Space plots of the two examples.



Figure 5.23 Examples of glyphs coded according to two display attributes. At the top are more integral coding pairs. At the bottom are more separable coding pairs.



"...robust rules that describe how we see patterns in visual displays"



Figure 6.2 Spatial proximity is a powerful cue for perceptual organization. A matrix of dots is perceived as rows on the left (a) and columns on the right (b). In (c) we perceive two groups of dots because of proximity relationships.



Figure 6.5 Connectedness is a powerful grouping principle that is stronger than (a) proximity, (b) color, (c) size, or (d) shape.

For graphics: perceived groupings should match real groupings.

http://www.nytimes.com/interactive/2014/07/07/upshot/how-england-italy-and-germany-are-dominating-the-world-cup.html?_r=0

How England, Italy and Germany Are Dominating the World Cup

By GREGOR AISCH and KEVIN QUEALY JULY 7, 2014

Top Ten Soccer Leagues by Share of Players Remaining in the World Cup

Group Stage		Round of 16		Quarterfinals		Semifinals		F	Finals	
Premier League (England)	15%	Premier League (England)	16%	Premier League (England)	22%	Bundesliga (Germany)	23% -	E	Bundesliga (Germany)	35%
Serie A (Italy)	11%	Serie A (Italy)	12%	Bundesliga (Germany)	14%	Premier League (England)	21%	, s	Serie A (Italy)	20%
Bundesliga (Germany)	10%	Bundesliga (Germany)	11%	Serie A (Italy)	11%	Serie A (Italy)	14%	F	Premier League (England)	15%
La Liga (Spain)	8%	La Liga (Spain)	10%	Ligue 1 (France)	9% •	Eredivisie (Netherlands)	11%	> L	La Liga (Spain)	11%
Ligue 1 (France)	6%	Ligue 1 (France)	6%	La Liga (Spain)	8%	La Liga (Spain)	9% -	7 F	Primeira Liga (Portugal)	7%
Premier League (Russia)	5%	Liga MX (Mexico)	5%	Eredivisie (Netherlands)	7%	Serie A (Brazil)	4%	• P	Primera Division (Argentina)) 7%
Liga MX (Mexico)	3%	Eredivisie (Netherlands)	4%	Primera Division (Costa Rica) 5%	Ligue 1 (France)	4%	/• L	igue 1 (France)	4%
Super Lig (Turkey)	3%	Major League Soccer (U.S.)	4%	Primera Division (Argentina)	4%	Primera Division (Argentina)	3%	L	liga MX (Mexico)	2%
Primeira Liga (Portugal)	3%	Primeira Liga (Portugal)	4%	Primeira Liga (Portugal)	4%	Primeira Liga (Portugal)	3%			
Major League Soccer (U.S.)	3%	Superleague (Greece)	3%	Serie A (Brazil)	2%	Premier League (Ukraine)	2%			
		Color	ad by socios	. Furana Nauth Amaria	Cau	th America				

What groupings are there? How are they indicated? What tasks are easy? Is it successful?

Conclusions

If you want something to stand out (or be easily found) use a preattentive property.

Don't expect people to distinguish more than seven levels of a property.

Perceived groupings should match real groupings.

	1 -	Premier League (England) 15%	Premier League (England) 16%	Premier League (England) 22%	Bundesliga (Germany) 23%	Bundesliga (Germany) 35%		
	2 -	Serie A (Italy) 11%	Serie A (Italy) 12%	Bundesliga (Germany) 14%	Premier League (England) 21%	Serie A (Italy) 20%		
	3 -	Bundesliga (Germany) 10%	Bundesliga (Germany) 11%	Serie A (Italy) 11%	Serie A (Italy) 14%	Premier League (England) 15%		
	4 -	La Liga (Spain) 8%	La Liga (Spain) 10%	Ligue 1 (France) 9%	Eredivisie (Netherlands) 11%	La Liga (Spain) 11%	regio	n
¥	5 -	Ligue 1 (France) 6%	Ligue 1 (France) 6%	La Liga (Spain) 8%	La Liga (Spain) 9%	Primeira Liga (Portugal) 7%		Europe
-a	6 -	Premier League (Russia) 5%	Liga MX (Mexico) 5%	Eredivisie (Netherlands) 7%	Serie A (Brazil) 4%	Primera Division (Argentina) 7%		North America
	7 -	Liga MX (Mexico) 3%	Eredivisie (Netherlands) 4%	Primera Division (Costa Rica) 5%	Ligue 1 (France) 4%	Ligue 1 (France) 4%		South America
	8 -	Super Lig (Turkey) 3%	Major League Soccer (U.S.) 4%	Primera Division (Argentina) 4%	Primera Division (Argentina) 3%	Liga MX (Mexico) 2%		
	9 -	Primeira Liga (Portugal) 3%	Primeira Liga (Portugal) 4%	Primeira Liga (Portugal) 4%	Primeira Liga (Portugal) 3%			
1	0 -	Major League Soccer (U.S.) 3%	Superleague (Greece) 3%	Serie A (Brazil) 2%	Premier League (Ukraine) 2%			
		Group Stage	Round of 16	Quarterfinals stage	Semifinals	Finals		

```
ggplot(leagues, aes(stage, rank)) +
geom_tile(aes(fill = region), width = 0.8, height = 0.9) +
geom_text(aes(label = labels), size = 3) +
scale_fill_manual(values = c("#E3ECE1", "#D0DFE9", "#F5E8C8")) +
theme_minimal()
```

Premier Lea	gue (England) 15	% ••	Premier League (England)	16% 🔍	Premier League (England)	22%	Bundesliga (Germany)	23%		Bundesliga (Germany)	35%
Serie A (Ital)	.) 11	%	Serie A (Italy)	12%	Bundesliga (Germany)	14%	Premier League (England)	21%	\smile	Serie A (Italy)	20%
Bundesliga (Germany) 10	%	Bundesliga (Germany)	115	Serie A (Italy)	11%	Serie A (Italy)	14%	\frown	Premier League (England)	15%
La Liga (Spa	in) 8	%	La Liga (Spain)	10%	Ligue 1 (France)	9%	Eredivisie (Netherlands)	11%		La Liga (Spain)	11%
Ligue 1 (Fra	nce) 6	%	Ligue 1 (France)	6%	🖲 La Liga (Spain)	8%	La Liga (Spain)	9%	< 1	Primeira Liga (Portugal)	7%
Premier Lea	gue (Russia) 5	× _	Liga MX (Mexico)	5%	Eredivisie (Netherlands)	7%	Serie A (Brazil)	4%	1	Primera Division (Argentina)	7%
Liga MX (Me	xico) 3	%	Eredivisie (Netherlands)	4%	Primera Division (Costa Rica)	5%	Ligue 1 (France)	4%	+	Ligue 1 (France)	4%
Super Lig (T	urkey) 3	× /	Major League Soccer (U.S.)	4%	Primera Division (Argentina)	4%	Primera Division (Argentina)	3%	1	Liga MX (Mexico)	2%
Primeira Liga	a (Portugal) 3	% ——	Primeira Liga (Portugal)	4%	Primeira Liga (Portugal)	4%	Primeira Liga (Portugal)	3%	(
Major Leagu	e Soccer (U.S.) 3	%	Superleague (Greece)	3%	Serie A (Brazil)	2%	Premier League (Ukraine)	2%			
	Group Stage		Round of 16		Quarterfinals		Semifinals			Finals	

region ---- Europe ---- North America ---- South America

```
ggplot(leagues, aes(stage, rank)) +
  geom_tile(aes(fill = region), width = 0.8, height = 0.9) +
  geom_text(aes(x = as.numeric(stage) - 0.38, label = league), size = 3, hjust = 0) +
  geom_text(aes(x = as.numeric(stage) + 0.38,
          label = paste(percent, "%", sep = "")),
     size = 3, hjust = 1) +
 geom_segment(aes(x = as.numeric(x) + 0.41, xend = as.numeric(xend) - 0.41,
          y = y, yend = yend, colour = region),
     data = connections, inherit.aes = FALSE, size = 1) +
  geom_point(aes(x = as.numeric(x) + 0.41, y = y, colour = region),
     data = connections, inherit.aes = FALSE) +
  geom_point(aes(x = as.numeric(xend) - 0.41, y = yend, colour = region),
     data = connections,
     inherit.aes = FALSE) +
  scale_fill_manual(values = c("#E3ECE1", "#D0DFE9", "#F5E8C8")) +
  scale_colour_manual(values = c("#C8DAC4", "#B1CADA", "#EED8A2")) +
  theme_minimal() +
  theme(axis.ticks = element_blank(), axis.text.y = element_blank(),
    panel.grid = element_blank(), axis.title = element_blank(),
    legend.position = "bottom")
```