

```
1 # Pino Trogu
2 # School of Design -- San Francisco State University
3 # DATA? TA-DA! 2 -- More Tips for Better Data Viz
4 # Workshop 2022-02-22
5 # HUM 404 -- 4:00 - 6:00 PM
6 # ALL CODE PROVIDED "AS IS"
7
8 # from environment pane, import two separate datasets:
9 # can18_1975_2017.csv (breast cancer rates from 1975 to 2017)
10 # shorten name during import to: can7517
11 # can18_2018.csv (breast cancer rates in 2018 by state)
12 # shorten name during import to: can18
13
14 # plot matrix of possible bivariate plots for can7517
15 plot(can7517)
16 # plot only selected variables in dataset
17 plot(can7517[1:3])
18
19 # plot matrix of possible bivariate plots for can18
20 plot(can18)
21 # plot only selected variables in dataset
22 plot(can18[1:5])
23 plot(can18[3:5])
24 plot(can18[5:7])
25
26 # plot (can18) white mortality on X axis and black mortality on Y axis
27 plot(can18$whiteMortRate, can18$blackMortRate)
28
29 # add state names
30 text(can18$whiteMortRate,
31       can18$blackMortRate,
32       can18$state,
33       cex=0.5)
34
35 # summary for can18 dataset (info appears in console)
36 summary(can18)
37 summary(can7517)
38
39 # line chart tests
40 plot(can18$blackMortRate, type="l", lty = 1)
41 plot(can18$blackMortRate, type="l", lty = 6)
42 plot(can18$whiteMortRate, type="l", lty = 1)
43 plot(can7517$black, type="l", lty = 1)
44
45 # *****
46 # START BAR CHART 1 WHITE INCIDENCE RATE
47
48 # barplot white breast cancer incidence rate per 100K F pop. by state 2018
49 barplot(can18$whiteIncRate, names.arg=can18$state)
50
51 #sort by incidence, not alpha by state name
52 sortByWhiteInc <- can18[order(can18$whiteIncRate, decreasing = TRUE), ]
53
54 # define anchor points for states labels & add title
```

```

55 midpts <- barplot(sortByWhiteInc$whiteIncRate, main="1. WHITE breast cancer INCIDENCE per
... 100K F pop. in 2018")
56
57 # add state names (vertical)
58 # text(x=midpts+.5, y=-2, sortByWhiteInc$state, cex=0.5, srt=90, xpd=TRUE, pos=2)
59
60 # add state names (45 degree angle -- run from line 35 but skip 44)
61 text(x=midpts+.5, y=-2, sortByWhiteInc$state, cex=0.5, srt=45, xpd=TRUE, pos=2)
62
63 # export plot (save as PDF)
64
65 # END BAR CHART 1 WHITE INCIDENCE RATE
66 # *****
67
68 # *****
69 # START BAR CHART 2 WHITE MORTALITY RATE
70
71 # barplot white breast cancer death rate per 100K F pop. by state 2018
72 barplot(can18$whiteMortRate, names.arg=can18$state)
73
74 #sort by incidence, not alpha by state name
75 sortByWhiteMort <- can18[order(can18$whiteMortRate, decreasing = TRUE), ]
76
77 # define anchor points for states labels & add title
78 midpts <- barplot(sortByWhiteMort$whiteMortRate, main="2. WHITE breast cancer MORTALITY per
... 100K F pop. in 2018")
79
80 # add state names (vertical)
81 # text(x=midpts+.5, y=-1, sortByWhiteMort$state, cex=0.5, srt=90, xpd=TRUE, pos=2)
82
83 # add state names (45 degree angle -- run from line 35 but skip 44)
84 text(x=midpts+.5, y=-1, sortByWhiteMort$state, cex=0.5, srt=45, xpd=TRUE, pos=2)
85
86 # export plot (save as PDF)
87
88 # END BAR CHART 2 WHITE MORTALITY RATE
89 # *****
90
91
92 # *****
93 # START BAR CHART 3 BLACK INCIDENCE RATE
94
95 # barplot black breast cancer incidence rate per 100K F pop. by state 2018
96 barplot(can18$blackIncRate, names.arg=can18$state)
97
98 #sort by incidence, not alpha by state name
99 sortByBlackInc <- can18[order(can18$blackIncRate, decreasing = TRUE), ]
100
101 # define anchor points for states labels & add title
102 midpts <- barplot(sortByBlackInc$blackIncRate, main="3. BLACK breast cancer INCIDENCE per
... 100K F pop. in 2018")
103
104 # add state names (vertical)
105 # text(x=midpts+.5, y=-2, sortByBlackInc$state, cex=0.5, srt=90, xpd=TRUE, pos=2)

```

```
106
107 # add state names (45 degree angle -- run from line 35 but skip 44)
108 text(x=midpts+.5, y=-2, sortByBlackInc$state, cex=0.5, srt=45, xpd=TRUE, pos=2)
109
110 # export plot (save as PDF)
111
112 # END BAR CHART 3 BLACK INCIDENCE RATE
113 # *****
114
115 # *****
116 # START BAR CHART 4 BLACK MORTALITY RATE
117
118 # barplot black breast cancer death rate per 100K F pop. by state 2018
119 barplot(can18$blackMortRate, names.arg=can18$state)
120
121 #sort by incidence, not alpha by state name
122 sortByBlackMort <- can18[order(can18$blackMortRate, decreasing = TRUE), ]
123
124 # define anchor points for states labels & add title
125 midpts <- barplot(sortByBlackMort$blackMortRate, main="4. BLACK breast cancer MORTALITY per
... 100K F pop. in 2018")
126
127 # add state names (vertical)
128 # text(x=midpts+.5, y=-1, sortByBlackMort$state, cex=0.5, srt=90, xpd=TRUE, pos=2)
129
130 # add state names (45 degree angle -- run from line 35 but skip 44)
131 text(x=midpts+.5, y=-1, sortByBlackMort$state, cex=0.5, srt=45, xpd=TRUE, pos=2)
132
133 # export plot (save as PDF)
134
135 # END BAR CHART 4 BLACK MORTALITY RATE
136 # *****
137
138
139 # *****
140 # START SIDE-BY-SIDE BAR CHART COMPARISON (various scales)
141 # 5.WHITE INC. -- 6.BLACK INC -- 7.WHITE MORT. -- 8.BLACK MORT.
142
143 # sort by white incidence (no plot)
144 sortByWhiteInc <- can18[order(can18$whiteIncRate, decreasing = TRUE), ]
145
146 # sort by white mortality (no plot)
147 sortByWhiteMort <- can18[order(can18$whiteMortRate, decreasing = TRUE), ]
148
149 # sort by black incidence (no plot)
150 sortByBlackInc <- can18[order(can18$blackIncRate, decreasing = TRUE), ]
151
152 # sort by black mortality (no plot)
153 sortByBlackMort <- can18[order(can18$blackMortRate, decreasing = TRUE), ]
154
155 # 5. White incidence
156 barplot(sortByWhiteInc$whiteIncRate,
157         names.arg=can18$state,
158         main="5. WHITE incidence")
```

```
159
160 # 6. Black incidence
161 barplot(sortByBlackInc$blackIncRate,
162         names.arg=can18$state,
163         main="6. BLACK incidence")
164
165 # 7. White mortality
166 barplot(sortByWhiteMort$whiteMortRate,
167         names.arg=can18$state,
168         main="7. WHITE mortality")
169
170 # 8. Black mortality
171 barplot(sortByBlackMort$blackMortRate,
172         names.arg=can18$state,
173         main="8. BLACK mortality")
174
175 # export plot (save as PDF)
176
177 # END SIDE-BY-SIDE BAR CHART COMPARISON (various scales)
178 # *****
179
180 # *****
181 # START SIDE-BY-SIDE BAR CHARTS (synchronized scales)
182 # 9. 10. 11. 12. WHITE & BLACK MORT.
183
184 # 9. White mortality -- default
185 barplot(sortByWhiteMort$whiteMortRate,
186         names.arg=can18$state,
187         main="9. WHITE mortality -- default")
188
189 # 10. Black mortality -- default
190 barplot(sortByBlackMort$blackMortRate,
191         names.arg=can18$state,
192         main="10. BLACK mortality -- default")
193
194 # 11. White mortality -- sync. axes
195 # define Y range (ylim)
196 barplot(sortByWhiteMort$whiteMortRate,
197         names.arg=can18$state,
198         ylim = c(0,35),
199         main="11. WHITE mortality -- sync. axes")
200
201 # 12. Black mortality -- sync. axes
202 # define Y range (ylim)
203 barplot(sortByBlackMort$blackMortRate,
204         names.arg=can18$state,
205         ylim = c(0,35),
206         main="12. BLACK mortality -- sync. axes")
207
208 # export plot (save as PDF)
209
210 # END SIDE-BY-SIDE BAR CHARTS (synchronized scales)
211 # *****
212
```

```
213
214 # *****
215 # START SCATTERPLOTS INCIDENCE & MORTALITY
216 # 13-14-15-16
217
218 plot(can18)
219
220 # plot only selected variables in dataset
221 plot(can18[1:5])
222 plot(can18[2:5])
223
224 summary(can18)
225
226 # ***** (default axes)
227
228 # 13 white incidence and mortality (default)
229 plot(can18$whiteIncRate,
230      can18$whiteMortRate,
231      main="13. White incidence and mortality 2018 -- default")
232
233 # 14. black incidence and mortality (default)
234 plot(can18$blackIncRate,
235      can18$blackMortRate,
236      main="14. Black incidence and mortality 2018 -- default")
237
238
239 # ***** (fixed axes)
240
241 # 15. white incidence and mortality
242 # define X and Y range (xlim and ylim)
243 plot(can18$whiteIncRate,
244      can18$whiteMortRate,
245      xlim = c(75,150),
246      ylim = c(15,35),
247      main="15. White incidence and mortality 2018 -- fixed")
248
249 # 16. black incidence and mortality
250 # define X and Y range (xlim and ylim)
251 plot(can18$blackIncRate,
252      can18$blackMortRate,
253      xlim = c(75,150),
254      ylim = c(15,35),
255      main="16. Black incidence and mortality 2018 -- fixed")
256
257 # add state names
258 text(can18$blackIncRate,
259      can18$blackMortRate,
260      can18$state,
261      cex=0.5)
262
263 # export plot (save as PDF)
264
265 # END SCATTERPLOTS INCIDENCE & MORTALITY
266 # *****
```

```
267
268 # *****
269 # START SCATTERPLOT -- WHITE & BLACK MORTALITY
270 # 17-18
271
272 plot(can18)
273
274 # plot only selected variables in dataset
275 plot(can18[1:5])
276 plot(can18[2:5])
277
278 summary(can18)
279
280 # *****
281
282 # 17 white & black mortality (default)
283 plot(can18$whiteMortRate,
284      can18$blackMortRate,
285      main="17. White & black mortality 2018 -- default")
286
287 # 18 white & black mortality (fixed)
288 # define X and Y range (xlim and ylim)
289 plot(can18$whiteMortRate,
290      can18$blackMortRate,
291      xlim = c(15,35),
292      ylim = c(15,35),
293      main="18. White & black mortality 2018 -- fixed")
294
295 # add state names
296 text(can18$whiteMortRate,
297      can18$blackMortRate,
298      can18$state,
299      cex=0.5)
300
301 # export plot (save as PDF)
302
303 # END SCATTERPLOT -- WHITE & BLACK MORTALITY
304 # *****
305
306 # *****
307 # START SCATTERPLOT -- ALL RACES INCIDENCE & MORTALITY
308 # 19. ALL
309 # 20. WHITE
310 # 21. BLACK
311 # 22. ALL
312 # 23. ALL CROPPED
313
314 plot(can18)
315
316 # plot only selected variables in dataset
317 plot(can18[1:5])
318 plot(can18[2:5])
319
320 summary(can18)
```

```
321
322 # 19. ALL races incidence and mortality 2018
323 plot(can18$all_IncRate,
324       can18$all_MortRate,
325       xlim = c(75,150),
326       ylim = c(15,35),
327       main="19. ALL races incidence and mortality 2018")
328
329 # 20. WHITE incidence and mortality 2018
330 plot(can18$whiteIncRate,
331       can18$whiteMortRate,
332       xlim = c(75,150),
333       ylim = c(15,35),
334       main="20. WHITE incidence and mortality 2018")
335
336 # 21. BLACK incidence and mortality 2018
337 plot(can18$blackIncRate,
338       can18$blackMortRate,
339       xlim = c(75,150),
340       ylim = c(15,35),
341       main="21. BLACK incidence and mortality 2018")
342
343 # 22. ALL AGAIN incidence and mortality 2018
344 plot(can18$all_IncRate,
345       can18$all_MortRate,
346       xlim = c(75,150),
347       ylim = c(15,35),
348       main="21. ALL incidence and mortality 2018")
349
350 # add state names
351 text(can18$all_IncRate,
352       can18$all_MortRate-.3,
353       can18$state,
354       cex=0.5)
355
356 # 23. ALL AGAIN CROPPED -- 2018
357 plot(can18$all_IncRate,
358       can18$all_MortRate,
359       xlim = c(110,140),
360       ylim = c(15,25),
361       main="21. ALL incidence and mortality 2018")
362
363 # add state names
364 text(can18$all_IncRate,
365       can18$all_MortRate-.3,
366       can18$state,
367       cex=0.5)
368
369 # export plot (save as PDF)
370
371 # END SCATTERPLOT -- ALL RACES INCIDENCE & MORTALITY
372 # *****
373
```