bivariate panel

2022-08-01

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.3.6 ✔ purrr 0.3.4  
## ✔ tibble 3.1.8 ✔ dplyr 1.0.9  
## ✔ tidyr 1.2.0 ✔ stringr 1.4.0  
## ✔ readr 2.1.2 ✔ forcats 0.5.1  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(readxl)  
library(extrafont)

## Registering fonts with R

library(rstatix)

##   
## Attaching package: 'rstatix'  
##   
## The following object is masked from 'package:stats':  
##   
## filter

headscan\_full1<-read\_excel("C:\\Users\\19177\\OneDrive - Colostate\\Desktop\\Dissertation\\headscan\_dissertation\\headscan\_full1.xlsx")  
headscan\_full1$coder <- as.factor(headscan\_full1$coder)  
headscan\_full1$gender <- as.factor(headscan\_full1$gender)  
headscan\_full1$race\_eth <- as.factor(headscan\_full1$race\_eth)  
headscan\_full1$age\_group <- as.factor(headscan\_full1$age\_group)  
str(headscan\_full1)

## tibble [2,016 × 33] (S3: tbl\_df/tbl/data.frame)  
## $ ID : chr [1:2016] "400-20201012-002" "400-20201012-003" "400-20201012-004" "400-20201012-005" ...  
## $ AA\_C : num [1:2016] 65 55 70 58 67 60 59 59 65 65 ...  
## $ BGl\_C : num [1:2016] 315 289 293 313 288 306 320 NA 300 277 ...  
## $ BiW\_C : num [1:2016] 130 127 143 140 137 130 141 138 143 150 ...  
## $ BiW\_L : num [1:2016] 115 108 121 109 104 106 109 111 113 116 ...  
## $ ChCh\_C : num [1:2016] 62 64 68 70 70 70 67 69 67 63 ...  
## $ GoSub\_C : num [1:2016] 93 93 115 93 103 100 79 106 85 102 ...  
## $ NRB\_L : num [1:2016] 17 18 19 21 19 14 17 18 16 17 ...  
## $ ProA\_L : num [1:2016] 28 25 31 23 28 28 26 27 32 28 ...  
## $ ProA\_C : num [1:2016] 31 27 33 27 31 29 27 29 34 31 ...  
## $ ProS\_C : num [1:2016] 18 20 14 13 22 22 19 14 26 24 ...  
## $ ProS\_L : num [1:2016] 17 18 14 13 20 20 18 12 24 22 ...  
## $ SelP\_C : num [1:2016] 42 41 51 45 47 48 46 42 47 44 ...  
## $ SelP\_L : num [1:2016] 42 41 51 44 47 48 46 41 46 44 ...  
## $ SelDH\_C : num [1:2016] 15 9 9 11 13 15 9 9 12 14 ...  
## $ SelM\_L : num [1:2016] 122 99 130 115 119 126 117 112 117 117 ...  
## $ SnasM\_C : num [1:2016] 82 55 84 74 73 80 78 76 64 75 ...  
## $ SmanM\_C : num [1:2016] 59 51 45 43 33 34 55 37 61 41 ...  
## $ SmanM\_L : num [1:2016] 55 50 45 42 33 34 50 36 59 40 ...  
## $ SnasM\_L : num [1:2016] 75 53 78 69 67 76 69 71 62 69 ...  
## $ TrHO\_C : num [1:2016] 179 163 169 166 159 162 169 NA 167 166 ...  
## $ TrEJ\_C : num [1:2016] 40 32 39 29 46 42 29 32 29 33 ...  
## $ TrGo\_C : num [1:2016] 84 57 70 61 68 70 75 61 67 64 ...  
## $ TrSel\_C : num [1:2016] 149 138 150 133 140 151 140 138 156 143 ...  
## $ TrSman\_C : num [1:2016] 177 145 178 147 157 164 149 159 151 160 ...  
## $ TrSnas\_C : num [1:2016] 163 142 167 145 152 157 148 149 157 NA ...  
## $ TrTr\_C : num [1:2016] 296 276 292 273 279 300 283 275 307 286 ...  
## $ TrTr\_L : num [1:2016] 155 141 156 149 146 146 147 151 157 144 ...  
## $ coder : Factor w/ 4 levels "Chandler","Isabel",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ age : num [1:2016] 31 49 49 34 49 55 26 18 25 27 ...  
## $ gender : Factor w/ 4 levels "Female","Male",..: 2 1 2 2 2 2 2 2 2 2 ...  
## $ race\_eth : Factor w/ 8 levels "AIAN","Asian",..: 3 8 8 8 8 8 3 3 8 8 ...  
## $ age\_group: Factor w/ 3 levels "18-36","37-54",..: 1 2 2 1 2 3 1 1 1 1 ...

headscan\_fullmm <- headscan\_full1  
  
#headscan\_fullmm$AA\_C <- headscan\_fullmm$AA\_C \*10  
#headscan\_fullmm$BGl\_C <- headscan\_fullmm$BGl\_C \* 10  
#headscan\_fullmm$BiW\_C <- headscan\_fullmm$BiW\_C \*10  
#headscan\_fullmm$BiW\_L <- headscan\_fullmm$BiW\_L \*10  
#headscan\_fullmm$ChCh\_C <- headscan\_fullmm$ChCh\_C \*10  
#headscan\_fullmm$GoSub\_C <- headscan\_fullmm$GoSub\_C \*10  
#headscan\_fullmm$NRB\_L <- headscan\_fullmm$NRB\_L \*10  
#headscan\_fullmm$ProA\_L <- headscan\_fullmm$ProA\_L \*10  
#headscan\_fullmm$ProA\_C <- headscan\_fullmm$ProA\_C \*10  
#headscan\_fullmm$ProS\_C <- headscan\_fullmm$ProS\_C \*10  
#headscan\_fullmm$ProS\_L <- headscan\_fullmm$ProS\_L \*10  
#headscan\_fullmm$SelP\_C <- headscan\_fullmm$SelP\_C \*10  
#headscan\_fullmm$SelP\_L <- headscan\_fullmm$SelP\_L \*10  
#headscan\_fullmm$SelDH\_C <- headscan\_fullmm$SelDH\_C \*10  
#headscan\_fullmm$SelM\_L <- headscan\_fullmm$SelM\_L \*10  
#headscan\_fullmm$SnasM\_C <- headscan\_fullmm$SnasM\_C \*10  
#headscan\_fullmm$SmanM\_C <- headscan\_fullmm$SmanM\_C \*10  
#headscan\_fullmm$SmanM\_L <- headscan\_fullmm$SmanM\_L \*10  
#headscan\_fullmm$SnasM\_L <- headscan\_fullmm$SnasM\_L \*10  
#headscan\_fullmm$TrHO\_C <- headscan\_fullmm$TrHO\_C \*10  
#headscan\_fullmm$TrEJ\_C <- headscan\_fullmm$TrEJ\_C \*10  
#headscan\_fullmm$TrGo\_C <- headscan\_fullmm$TrGo\_C \*10  
#headscan\_fullmm$TrSel\_C <- headscan\_fullmm$TrSel\_C \*10  
#headscan\_fullmm$TrSman\_C <- headscan\_fullmm$TrSman\_C \*10  
#headscan\_fullmm$TrSnas\_C <- headscan\_fullmm$TrSnas\_C \*10  
#headscan\_fullmm$TrTr\_C <- headscan\_fullmm$TrTr\_C \*10  
#headscan\_fullmm$TrTr\_L <- headscan\_fullmm$TrTr\_L \*10  
  
#str(headscan\_fullmm)

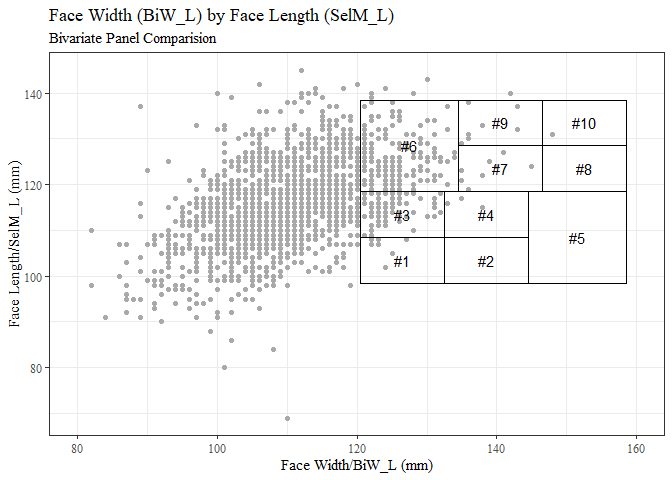
panel\_numbers <- data.frame(cell\_num = c("#1", "#2", "#3", "#4", "#5", "#6", "#7", "#8", "#9", "#10"),  
 SelM\_L = c(103.5, 103.5, 113.5, 113.5, 108.5, 128.5, 123.5, 123.5, 133.5, 133.5),  
 BiW\_L = c(126.5, 138.5, 126.5, 138.5, 151.5, 127.5, 140.5, 152.5, 140.5, 152.5),  
 gender = NA,  
 age\_group = NA,  
 race\_eth = NA)  
  
str(panel\_numbers)

## 'data.frame': 10 obs. of 6 variables:  
## $ cell\_num : chr "#1" "#2" "#3" "#4" ...  
## $ SelM\_L : num 104 104 114 114 108 ...  
## $ BiW\_L : num 126 138 126 138 152 ...  
## $ gender : logi NA NA NA NA NA NA ...  
## $ age\_group: logi NA NA NA NA NA NA ...  
## $ race\_eth : logi NA NA NA NA NA NA ...

From Zhaung 2007: menton-nasal root depression length is face length, bizygomatic breath is face width

biv1 <- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L))+  
 geom\_point(color="darkgray")+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width (BiW\_L) by Face Length (SelM\_L)",  
 subtitle = "Bivariate Panel Comparision",  
 y="Face Length/SelM\_L (mm)",  
 x="Face Width/BiW\_L (mm)")  
  
biv1 + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 226 rows containing missing values (geom\_point).



facelen\_outl<-headscan\_fullmm %>%   
 identify\_outliers(SelM\_L)  
  
tibble(facelen\_outl$SelM\_L)

## # A tibble: 8 × 1  
## `facelen\_outl$SelM\_L`  
## <dbl>  
## 1 86  
## 2 145  
## 3 143  
## 4 69  
## 5 84  
## 6 90  
## 7 88  
## 8 80

min(facelen\_outl$SelM\_L, na.rm = TRUE)

## [1] 69

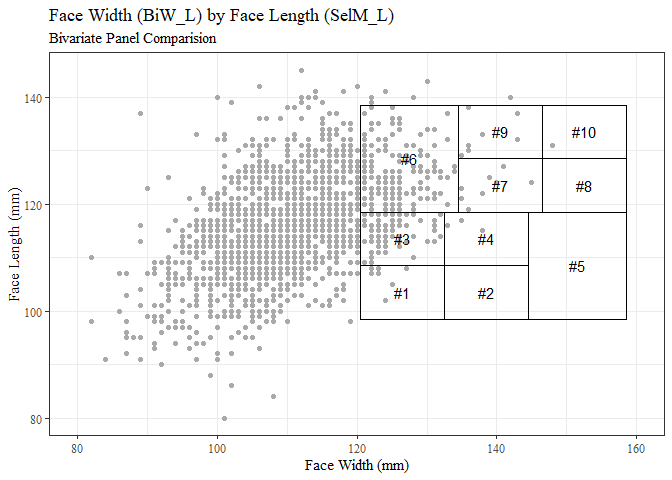
mean(headscan\_fullmm$SelM\_L, na.rm = TRUE)

## [1] 116.2444

headscan\_fullmm <- headscan\_fullmm %>% mutate(SelM\_L = replace(SelM\_L, SelM\_L<70, NA))

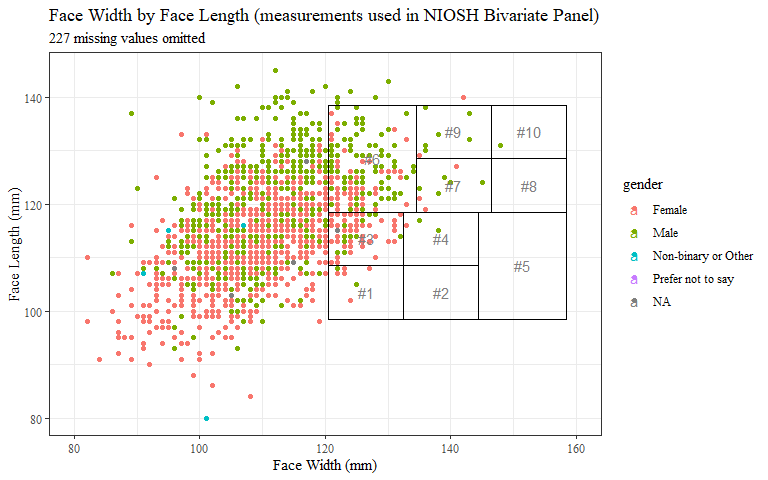
biv1 <- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L))+  
 geom\_point(color="darkgray")+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width (BiW\_L) by Face Length (SelM\_L)",  
 subtitle = "Bivariate Panel Comparision",  
 y="Face Length (mm)",  
 x="Face Width (mm)")  
  
biv1 + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 227 rows containing missing values (geom\_point).



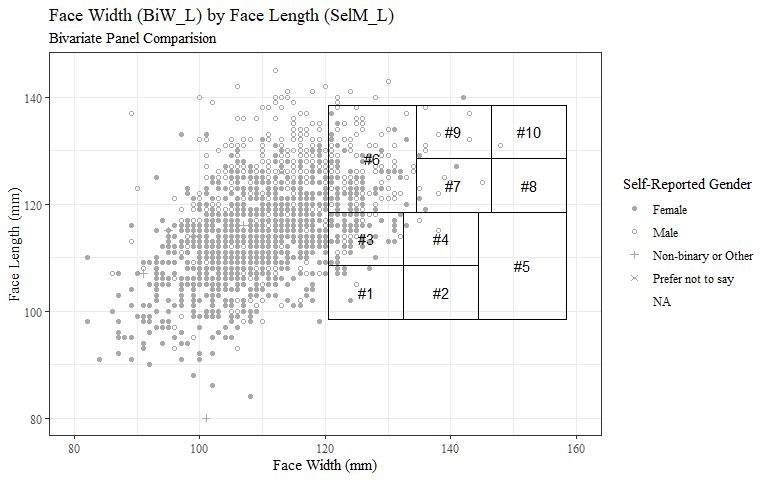
biv2 <- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L, color=gender))+  
 geom\_point()+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width by Face Length (measurements used in NIOSH Bivariate Panel)",  
 subtitle = "227 missing values omitted",  
 y="Face Length (mm)",  
 x="Face Width (mm)")  
  
biv2 + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 227 rows containing missing values (geom\_point).



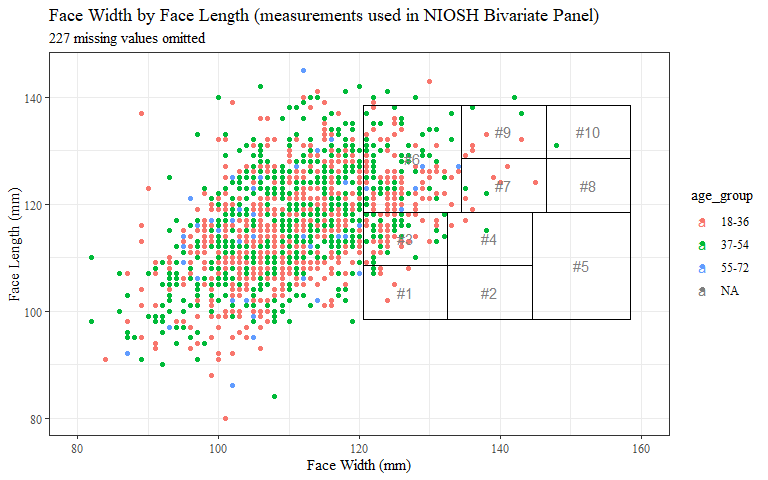
bivx <- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L))+  
 geom\_point(aes(shape=gender), color="darkgray")+  
 scale\_shape\_manual(values=c(16, 1, 3, 4))+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width (BiW\_L) by Face Length (SelM\_L)",  
 subtitle = "Bivariate Panel Comparision",  
 y="Face Length (mm)",  
 x="Face Width (mm)",   
 shape="Self-Reported Gender")  
  
bivx + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 235 rows containing missing values (geom\_point).



biv3<- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L, color=age\_group))+  
 geom\_point()+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width by Face Length (measurements used in NIOSH Bivariate Panel)",  
 subtitle = "227 missing values omitted",  
 y="Face Length (mm)",  
 x="Face Width (mm)")  
  
biv3 + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 227 rows containing missing values (geom\_point).



biv4 <- headscan\_fullmm %>%   
 ggplot(aes(x=BiW\_L, y=SelM\_L, color=race\_eth))+  
 geom\_point()+  
 theme\_bw()+theme(text=element\_text(family= "Times New Roman"))+  
 xlim(80,160)+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=98.5, ymax=108.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=144.5, xmax=158.5, ymin=98.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=132.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=132.5, xmax=144.5, ymin=108.5, ymax=118.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=120.5, xmax=134.5, ymin=118.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=134.5, xmax=146.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=118.5, ymax=128.5), fill= NA, alpha=0.4, colour="black")+  
 geom\_rect(aes(xmin=146.5, xmax=158.5, ymin=128.5, ymax=138.5), fill= NA, alpha=0.4, colour="black")+  
 labs(title="Face Width by Face Length (measurements used in NIOSH Bivariate Panel)",  
 subtitle = "227 missing values omitted",  
 y="Face Length (mm)",  
 x="Face Width (mm)")  
  
biv4 + geom\_text(data=panel\_numbers, aes(x=BiW\_L, y=SelM\_L, label=cell\_num))

## Warning: Removed 227 rows containing missing values (geom\_point).

