Beeswarm practice

Kate Huebner

11/1/2017

Modelling dependent~independent variables to determine variables associated with drug efficacy against TB. Dependent Variables are ELU (efficacy in the lungs) and ESP (efficacy in the spleen). Cmax and Trough indicate different drug concentrations and values are averaged across many mice at different drug doseages. Dependent variables include biologic and chemical (in vitro) measures.

Here I will use beeswarm to plot scatterplots to show variation of different variables. ggbeeswarm provides two methods of plotting categorical scatter plots such that the arrangement of points within a category reflects the density of data at that region, and avoids over-plotting

knitr::opts\_chunk$set(echo = TRUE, message = FALSE,  
 warning = FALSE, error = FALSE)

#Reading in data  
cmax <- read\_excel("../data/DRUG\_CLASS\_I\_Mean\_Cmax\_Trough\_Efficacy\_R\_DATA\_ANALYSIS.xlsx",  
 sheet = "Mean\_PK\_Efficacy\_In vitro", skip = 1, n\_max = 12)  
  
trough <- read\_excel("../data/DRUG\_CLASS\_I\_Mean\_Cmax\_Trough\_Efficacy\_R\_DATA\_ANALYSIS.xlsx",  
 sheet = "Mean\_PK\_Efficacy\_In vitro", skip = 16, n\_max = 12)  
  
head(cmax)

## # A tibble: 6 x 19  
## X\_\_1 Dose PLA ULU RIM OCS ICS SLU SLE ELU ESP cLogP  
## <chr> <chr> <dbl> <chr> <chr> <chr> <chr> <dbl> <dbl> <chr> <chr> <dbl>  
## 1 DRUG1 50BID 20700 21900 20300 23800 23300 10900 17500 0.8 2 0.30  
## 2 DRUG2 50BID 21433 36800 29900 45050 35950 20000 28200 2.33 2.18 2.38  
## 3 DRUG3 50BID 13295 9480 10400 8760 2480 8170 7370 1.65 0.63 1.70  
## 4 DRUG4 50BID 3203 11185 9185 1452 332 12450 5985 2.29 0.78 1.62  
## 5 DRUG5 50BID 2420 24950 17200 6535 612 20400 15313 1.56 1.31 0.71  
## 6 DRUG6 50BID 832 17100 11085 8080 2275 14700 9420 0.78 -0.2 0.69  
## # ... with 7 more variables: huPPB <chr>, muPPB <dbl>, MIC\_Erdman <dbl>,  
## # MICserumErd <dbl>, MIC\_Rv <chr>, Caseum\_binding <dbl>, MacUptake <chr>

head(trough)

## # A tibble: 6 x 19  
## X\_\_1 Dose PLA ULU RIM OCS ICS SLU SLE ELU ESP  
## <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <chr> <dbl> <dbl>  
## 1 DRUG1 50BID 21.5 49.7 49 117 196 59 63 0.80 2.00  
## 2 DRUG2 50BID 1550.0 3380.0 2290 8510 16800 NA NA 2.33 2.18  
## 3 DRUG3 50BID 19.0 37.0 68 283 511 NA NA 1.65 0.63  
## 4 DRUG4 50BID 79.0 187.0 347 551 953 577 651 2.29 0.78  
## 5 DRUG5 50BID 180.0 1740.0 1710 1900 837 257 617 1.56 1.31  
## 6 DRUG6 50BID 58.0 1475.0 1518 1362 305 831 1039 0.78 -0.20  
## # ... with 8 more variables: cLogP <dbl>, huPPB <chr>, muPPB <dbl>,  
## # MIC\_Erdman <dbl>, MICserumErd <dbl>, MIC\_Rv <dbl>,  
## # Caseum\_binding <dbl>, MacUptake <dbl>

#data cleaning  
cmax <- cmax %>%   
 rename(Drug = X\_\_1)  
   
  
cmax

## # A tibble: 11 x 19  
## Drug Dose PLA ULU RIM OCS ICS SLU SLE  
## <chr> <chr> <dbl> <chr> <chr> <chr> <chr> <dbl> <dbl>  
## 1 DRUG1 50BID 20700.00 21900 20300 23800 23300 10900.00 17500.00  
## 2 DRUG2 50BID 21433.00 36800 29900 45050 35950 20000.00 28200.00  
## 3 DRUG3 50BID 13295.00 9480 10400 8760 2480 8170.00 7370.00  
## 4 DRUG4 50BID 3203.00 11185 9185 1452 332 12450.00 5985.00  
## 5 DRUG5 50BID 2420.00 24950 17200 6535 612 20400.00 15313.00  
## 6 DRUG6 50BID 832.00 17100 11085 8080 2275 14700.00 9420.00  
## 7 DRUG7 50BID 4170.00 NA NA NA NA 3715.00 2960.00  
## 8 DRUG1 100QD 15619.62 24100 22400 22933.33 23233.33 15400.00 14900.00  
## 9 DRUG9 100QD 38675.00 49350 39100 37650 45200 26200.00 27675.00  
## 10 DRUG10 100QD 19600.00 14150 13925 12200 11322.5 20266.67 14233.33  
## 11 DRUG11 100QD 13275.00 11700 8335 9260 10950 11710.00 11650.00  
## # ... with 10 more variables: ELU <chr>, ESP <chr>, cLogP <dbl>,  
## # huPPB <chr>, muPPB <dbl>, MIC\_Erdman <dbl>, MICserumErd <dbl>,  
## # MIC\_Rv <chr>, Caseum\_binding <dbl>, MacUptake <chr>