Lab 4

Names

# Rules

In groups of 2 to 4 from the last lab session, complete the following.

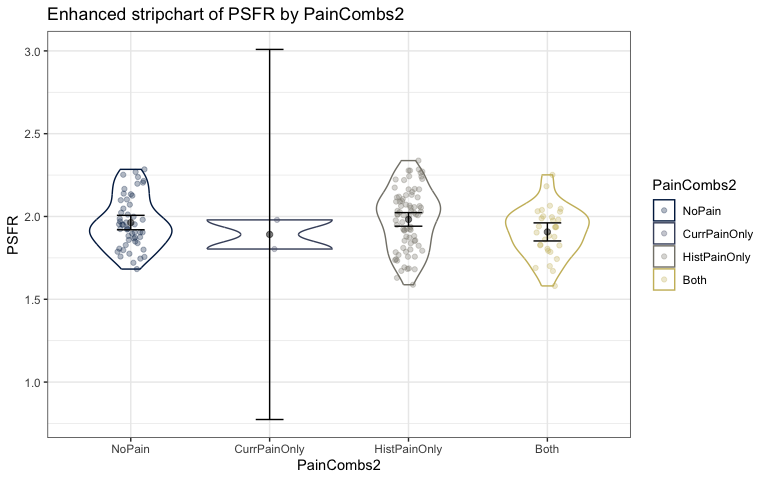
# Code/work from Lab 3:

data(TendonData)  
tendon <- TendonData  
  
 tendonN <- tendon %>%  
 rename(SubjectID = 'Subject ID',  
 Sex = 'Sex (M=1)',  
 CurrPain = 'CurrPain (Y=1)',  
 WaisttoHip = 'Waist to Hip',  
 VISAA = 'VISA-A',  
 Neovascularization = 'Neovascularization (Doppler) (Y=1)',  
 HistPain = 'Hx Pain (Y=1)')  
   
 tendonF <- tendonN %>%   
 mutate(across(where(~ is.character(.x)), as.factor))  
   
   
 tendonF <- tendonF %>%  
 mutate(CurrPain = factor(CurrPain),  
 HistPain = factor(HistPain),  
 Neovascularization = factor(Neovascularization),  
 Sex = factor(Sex),  
 Sex = fct\_recode(Sex,  
 "Female" = "0",  
 "Male" = "1"),  
 Location = factor(substr(SubjectID, 1, 3)),  
 PainCombs = as.numeric(str\_c(HistPain, CurrPain))  
 )  
tendonna1 <- tendonF %>% drop\_na(CSA)  
tendonna2 <- tendonF %>% drop\_na()

#Names may differ from your version of Lab 3 - can edit these or use our version:  
tendonna2 <- tendonna2 %>% mutate(PainCombs2 = factor(PainCombs),  
 PainCombs2 = fct\_recode(PainCombs2,  
 "NoPain" = "0", # 00   
 "CurrPainOnly" = "1", # 01   
 "HistPainOnly" = "10",  
 "Both" = "11"))  
levels(tendonna2$PainCombs2)

## [1] "NoPain" "CurrPainOnly" "HistPainOnly" "Both"

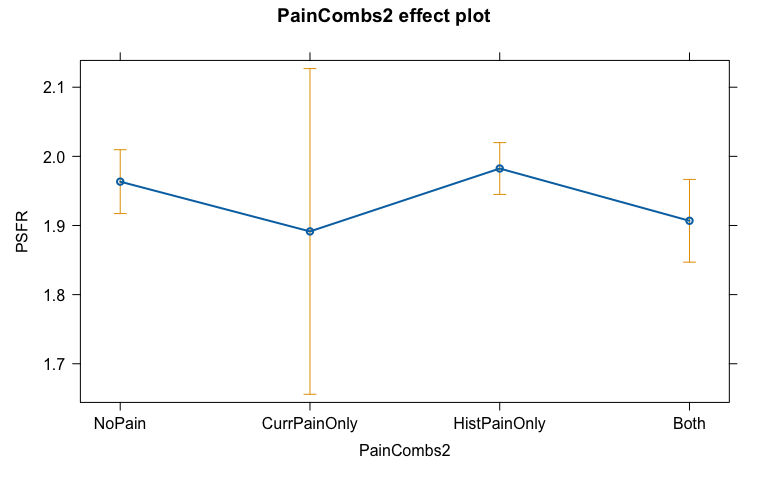
enhanced\_stripchart(PSFR ~ PainCombs2, data = tendonna2)



lm\_PC2 <- lm(PSFR ~ PainCombs2, data = tendonna2)  
summary(lm\_PC2)

##   
## Call:  
## lm(formula = PSFR ~ PainCombs2, data = tendonna2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.39465 -0.11606 -0.00146 0.12211 0.35510   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1.96342 0.02339 83.938 <2e-16  
## PainCombs2CurrPainOnly -0.07201 0.12154 -0.592 0.554  
## PainCombs2HistPainOnly 0.01898 0.03012 0.630 0.530  
## PainCombs2Both -0.05662 0.03827 -1.479 0.141  
##   
## Residual standard error: 0.1687 on 160 degrees of freedom  
## Multiple R-squared: 0.02926, Adjusted R-squared: 0.01106   
## F-statistic: 1.608 on 3 and 160 DF, p-value: 0.1897

plot(allEffects(lm\_PC2))



Q1 (Lab 3 number 6): Use favstats to obtain the estimated means for PSFR for each of the levels of the new PainCombs variable. No discussion.

favstats(PSFR ~ PainCombs2, data = tendonna2)

## PainCombs2 min Q1 median Q3 max mean sd  
## 1 NoPain 1.682439 1.856104 1.940357 2.078183 2.284604 1.963421 0.1566721  
## 2 CurrPainOnly 1.803465 1.847439 1.891414 1.935388 1.979363 1.891414 0.1243789  
## 3 HistPainOnly 1.587747 1.845178 2.011611 2.115495 2.337499 1.982401 0.1833098  
## 4 Both 1.580619 1.815095 1.935324 1.998517 2.251339 1.906802 0.1487741  
## n missing  
## 1 52 0  
## 2 2 0  
## 3 79 0  
## 4 31 0

Q2 (Lab 3 number 7): Generate an estimated mean from the previously estimated linear model for subjects with neither history of pain nor current pain. Show your work or explain how you found the result. Make sure you check your work versus the results from favstats in the previous question.

The estimated mean for subjects with neither history of pain nor current pain would be 1.96 as the indicator variables for the other beta coefficients would be set to 0, leaving just the baseline statistic.

Q3 (Lab 3 number 8): Generate an estimated mean from the model for subjects with current pain but no history of pain. Show your work.

Estimated mean from the model for subjects with current pain but no history of pain is 1.89.

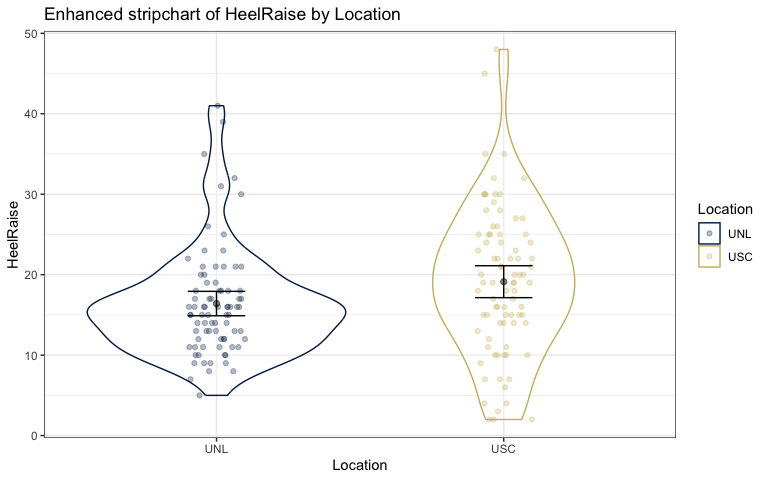
1.96 - 0.07201

## [1] 1.88799

## Heel Raise as a response variable

Q4 (Lab 3 number 9): Make an enhanced stripchart of the HeelRaise variable based on the Location (university where the measurements were made) variable. No discussion and you do not need to modify the axes or title.

enhanced\_stripchart(HeelRaise ~ Location, data = tendonna2)



Q5 (Lab 3 number 10): From their reference number 29, I found “The number of standing unilateral heel-raises above a level 5 cm from the floor was counted. The starting side for the test was the same as in the isokinetic tests. Each test was performed on both sides, followed by a 4 min rest… The leg that was being tested had to be kept straight. The test was stopped when the subject was unable to continue performing proper heel-raises due to fatigue.” Assuming that the testing protocol was the same in this study, what are the units of the measurements in HeelRaise? What is the smallest and largest observed value in the data set?

* The smallest measurement recorded was 2.0 heel raise repetitions and the largest was 48.0 heel raise repetitions.

summary(tendonna2$HeelRaise)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 2.0 12.0 16.0 17.8 22.0 48.0

# Part II: Model for heel raises:

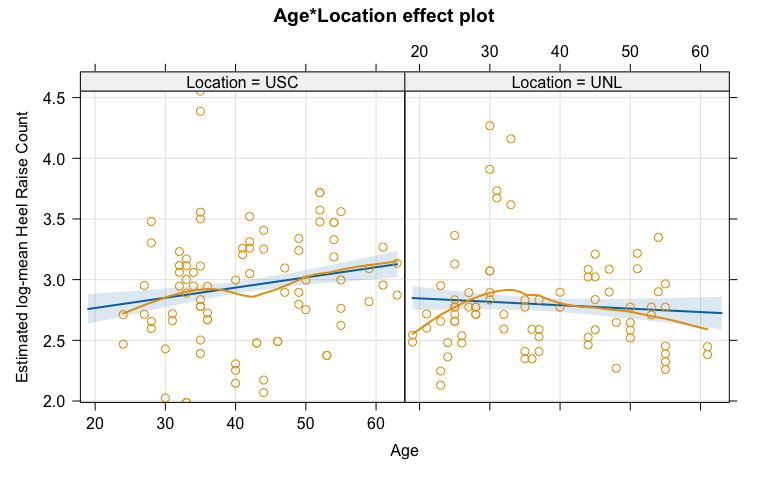
Q6) Fit an appropriate GLM for heal raises based on Age and Location and their interaction. Generate a model summary and effects plots on the link and response scales with improved y-axis labels (see the lecture notes).

* Notes: (i) The provided code changes the baseline level of the Location variable, use the modified version. (ii) If your two effects plots look the same, you should check whether you actually fit a Poisson GLM and not a Gaussian glm.

tendonna2 <- tendonna2 %>% mutate(Location = relevel(Location, "USC"))  
  
glm1 <- glm(HeelRaise ~ Age\*Location, tendonna2, family = poisson)  
summary(glm1)

##   
## Call:  
## glm(formula = HeelRaise ~ Age \* Location, family = poisson, data = tendonna2)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 2.598627 0.106610 24.375 < 2e-16  
## Age 0.008392 0.002441 3.438 0.000586  
## LocationUNL 0.303186 0.140640 2.156 0.031103  
## Age:LocationUNL -0.011207 0.003413 -3.284 0.001025  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 609.80 on 163 degrees of freedom  
## Residual deviance: 579.63 on 160 degrees of freedom  
## AIC: 1343.4  
##   
## Number of Fisher Scoring iterations: 4

plot(allEffects(glm1, residuals = T), type = "link", grid = T, ylab = "Estimated log-mean Heel Raise Count")



plot(allEffects(glm1, residuals = T), type = "response", grid = T, ylab = "Estimated mean Heel Raise Count")



Q7) Write out the estimated model, defining any indicator variable(s) you use.

* Heel Raise Poisson() #defines random component in model
* where is 1 when location is UNL and 0 if USC

Q8) Simplify the estimated model for the log-mean on the link scale for a USC subject.

Q9) Interpret the Age predictor “size” on the response (mean count) scale. You can use modelname %>% tbl\_regression(intercept = T, exponentiate = T) to get the numbers for the interpretation and related 95% CI (no exponentiation needed since it does that for you).

**For two otherwise similar subjects, but that differ by 1 year in age, we estimate that the mean number of heel raises until failure will change by 1.01 times that of the younger individual at USC (95% CI: exp(1.00) to exp(1.01)).**

glm1 %>% tbl\_regression(intercept = T, exponentiate = T)

| **Characteristic** | **IRR***1* | **95% CI***1* | **p-value** |
| --- | --- | --- | --- |
| (Intercept) | 13.4 | 10.9, 16.6 | <0.001 |
| Age | 1.01 | 1.00, 1.01 | <0.001 |
| Location |  |  |  |
| USC | — | — |  |
| UNL | 1.35 | 1.03, 1.78 | 0.031 |
| Age \* Location |  |  |  |
| Age \* UNL | 0.99 | 0.98, 1.00 | 0.001 |
| *1*IRR = Incidence Rate Ratio, CI = Confidence Interval | | | |

Q10) Cite any external resources used. If none, report that.

* NONE