Lab 11

Preparation for Lab 11:

- Please review the materials provided in
  https://rconnect.math.montana.edu/Data_Wrangling/#section-relational-data-with-dplyr

- Pay particular attention to the discussion of `left_join` and think about how we might use it to join your repeated measures of grip strength with the demographics data set.

Part 1: Grip strength potential interaction:

```r
GripStrengthMeasurements <- read_excel("GripStrengthMeasurementsF24.xlsx")
GS <- GripStrengthMeasurements %>% drop_na(GripStrength)
tally(Hand~Arm, data = GS)
```

```
##      Arm
## Hand 90 down Down up Up <NA>
##   D  24    5   19  8 16    0
##   ND 24    6   17  8 16    1
```

```r
GS <- GS %>% drop_na(Arm)
GS <- GS %>% mutate(Arm = forcats::fct_collapse(Arm,
                                       Up = c("Up", "up"),
                                       Down = c("Down", "down"),
                                       "90" = c("90", "90 degree")),
                SubjectID = forcats::fct_collapse(SubjectID,
                                      MDEarl = c("MDEarl","MDEar")),
                Arm = factor(Arm),
                Hand = factor(Hand),
                OrderF = factor(Order)
)
tally(~Arm, data = GS)
```
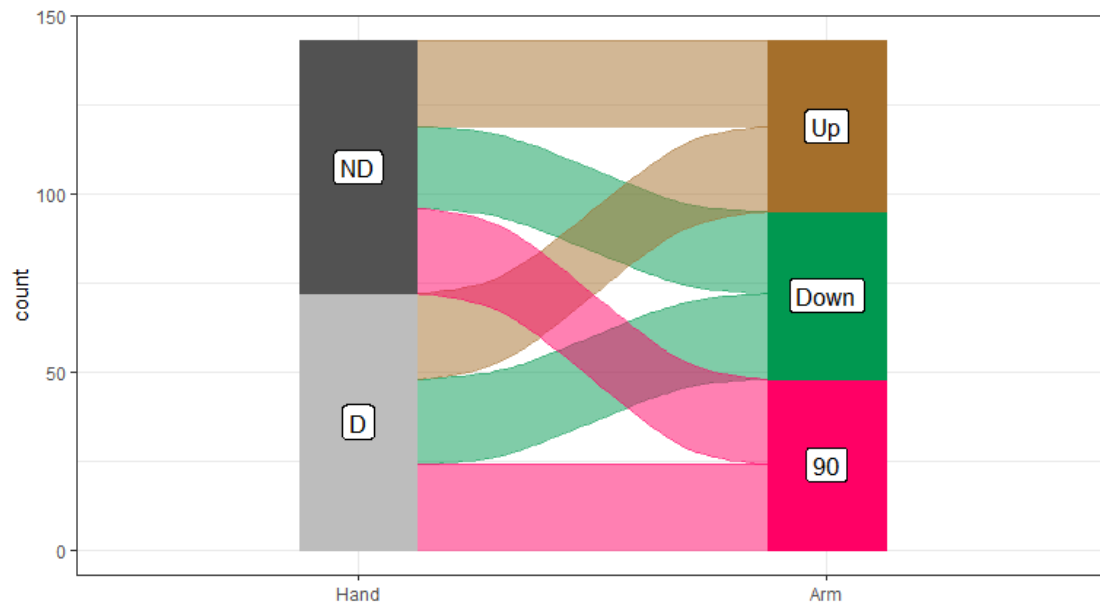
```
## Arm
##   90 Down   Up
##   48   47   48
```

```r
tally(Arm ~ Hand, data = GS)
```

```
##        Hand
## Arm      D ND
##   90    24 24
##   Down  24 23
##   Up    24 24
```

```
GS %>% dplyr::select(Hand, Arm) %>% alluvial_wide(bin_labels = "mean",
fill_by = "last_variable", bins = 6)
```
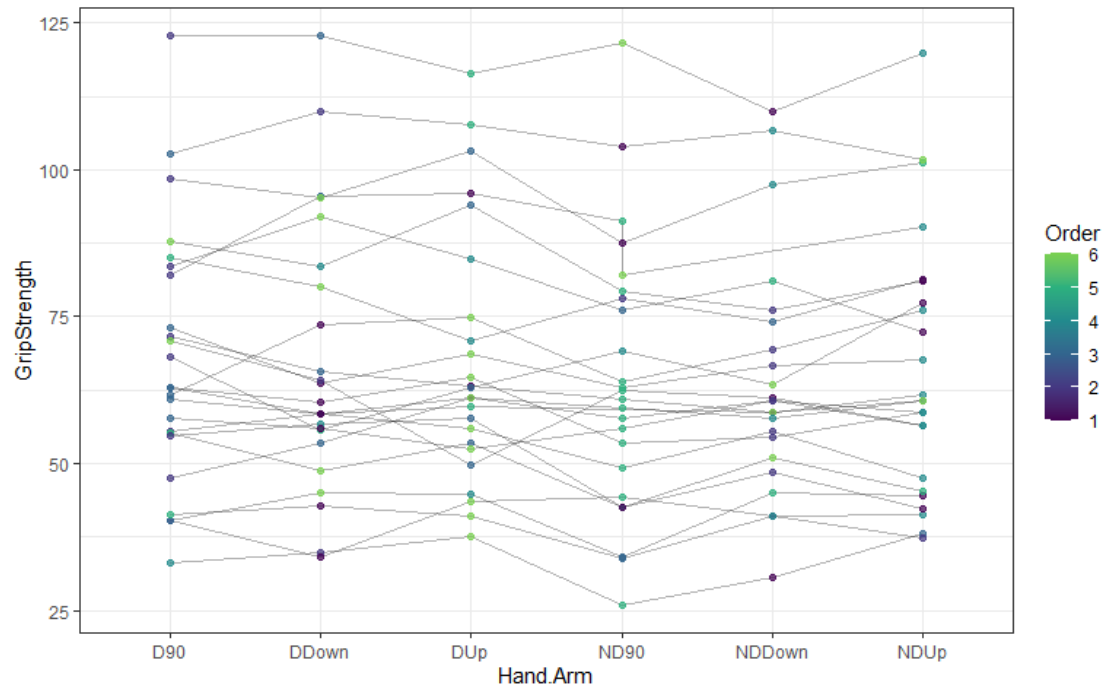

Number of flows: 6
Original Dataframe reduced to 4.2 %
Maximum weight of a single flow 16.8 %

```
GS <- GS %>% mutate(Hand.Arm = factor(str_c(Hand, Arm)))

Spaghetti_GS <- GS %>% ggplot(mapping = aes(x = Hand.Arm, y = GripStrength,
group = SubjectID)) +
  geom_line(alpha = 0.3) +
  geom_point(aes(color = Order), alpha = 0.7) +
  scale_color_viridis_c(end = 0.8)

Spaghetti_GS
```

```
Spaghetti_GS %>% ggplotly()
```
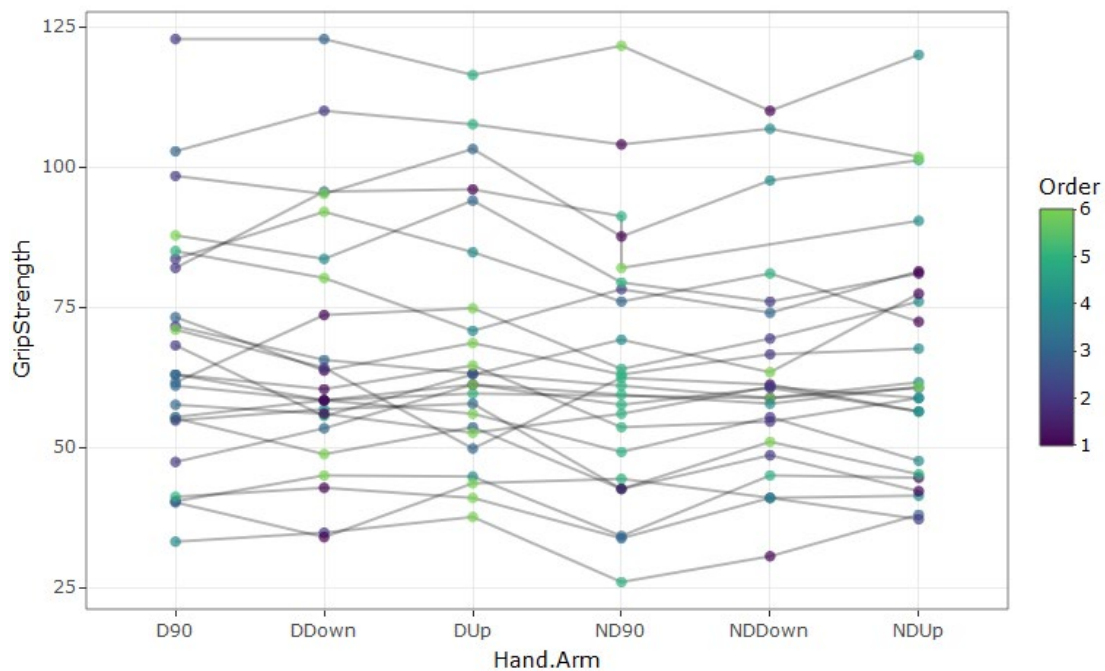

1) What did the `fct_collapse` function do and why was it needed in the previous code?

- Combines the upper and lowercase values of the same variable. E.g., up and Up to one column.

2) The addition of the line of code of `webshot::install_phantomjs(force = T)` should allow you to knit with interactive documents, but uncomment in the `ggplotly` lines if that is problematic. Find someone in your group or an observation that looks interesting and track that value across the treatment combinations. Does the subject tend to do better or worse on dominant or non-dominant hand across arm positions? Is there an optimal arm position for either or both hands for that subject? Now, look at the whole plot. What can you say about the differences in grip strength across hands (dominant/non-dominant), arm positions, and subjects?

- Following PATaller. NDDown had the lowest value, but the others were relatively similar. Down and 90 for the dominant hand were the highest, and the exact same at 122.8 lbs. 90 degree for both hands had high values, and differed by less than a pound (0.8 lbs.). There is not much variation for either arm position of hand. There is a clear difference in subjects across the whole data set, but difference in dominant and non-dominant appear to be minimal. There are exceptions to this. For example, IAGarthBrooks shows a clear decline in strength between dominant and non-dominant.

3) Is this a balanced design relative to the Hand and Arm variables?

- It is not balanced. There are 24 observations for most measures but only 23 observations for the non-dominant down category, so it is not completely balanced.

4) Make an interaction plot and discuss the potential for an interaction of Arm and Hand on the Grip Strength responses. If you think there is a potential interaction, discuss the pattern. If you don't think there is a potential interaction, discuss the possible pattern of results in the two enhanced stripcharts that provide a hint about the results for each of the main effects.

```
ggintplot(data = GS, response = 'GripStrength', groupvars = c('Hand', 'Arm'))
```

Two-way interaction of Hand and Arm
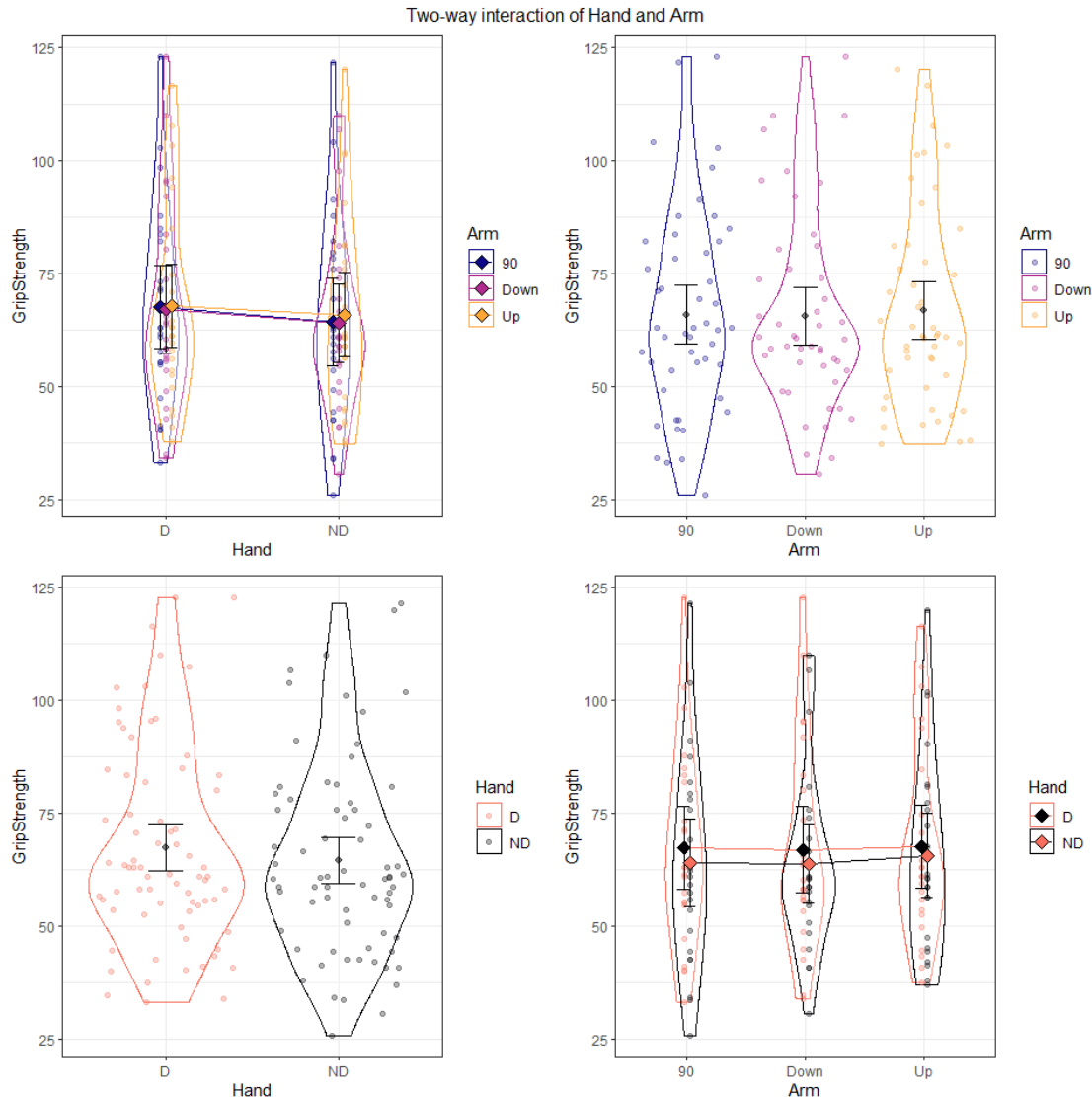
- We do not see much evidence of an interaction. The means of hand and arm position are similar. Additionally, the distributions are similar among all positions and both hands, respective of the variable. Based on the distributions, it does not appear there will be a main effect of either hand or position.

5) We will dive more deeply into linear mixed models in the coming weeks, but the following code uses the `lmer` function from the `lme4` package (with some modifications from the `lmerTest` package so you can see p-values) to fit a mixed model that incorporates subject as a "random" effect. You can interpret the results for the "fixed" effect (starting with arm and hand and their interaction on grip strength) as "conditional on" or "adjusted for" subject much like we did if it was included as a fixed effect. Report an evidence sentence for the interaction test from the following results and what it suggests for the interaction in the model.

```
lmer1 <- lmer(GripStrength ~ Hand*Arm + (1|SubjectID), data = GS)
Anova(lmer1, test.statistic = "F")
```

```
## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: GripStrength
##                F Df Df.res      Pr(>F)
## Hand     12.9542  1 114.00 0.0004737
## Arm       1.2829  2 114.01 0.2812085
## Hand:Arm  1.2018  2 114.01 0.3044415
```

- There is little to no evidence against the null hypothesis of no interaction $F(2,114.01) = 1.2018$, p = 0.304, controlling for the random effect of subject, and therefore we would not include the interaction in the model.

6) An interactive residual panel is available below. Which subject(s) had the largest and smallest observed residuals (so were most above/below what the model predicted for them)? Generally how large/small was the fitted value for the two most extreme residuals?

```
resid_interact(lmer1, c("resid", "qq", "yvp", "hist"))
```
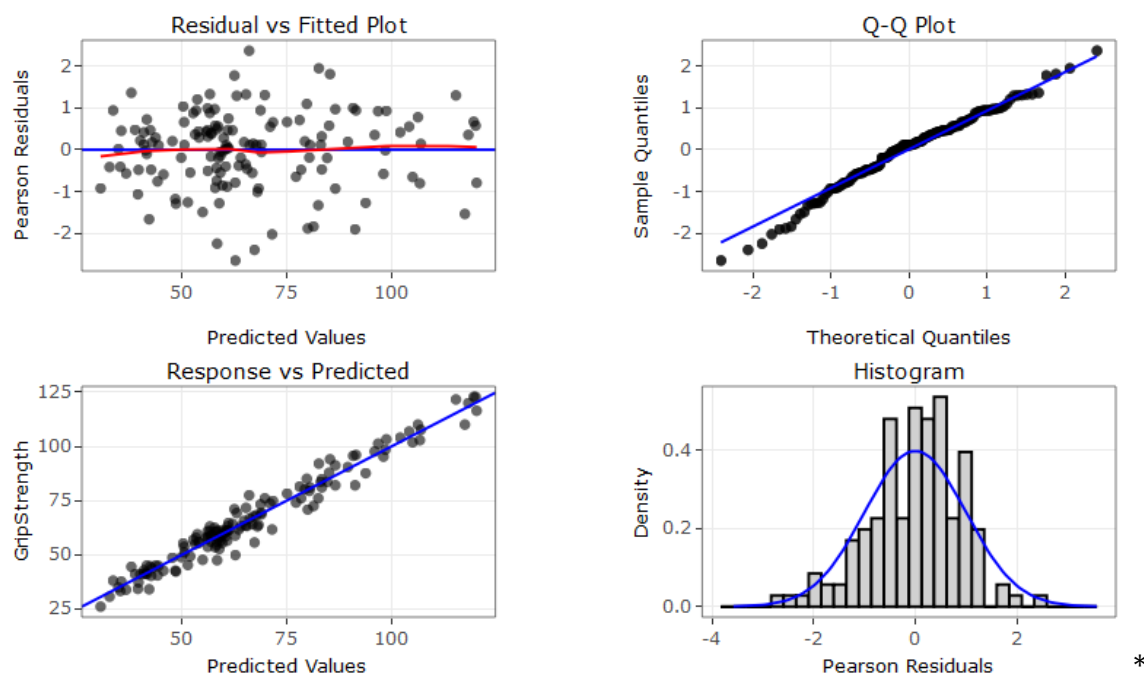

WATaller had the highest (2.355) residual values and WYJourney had the lowest (-2.64). Both fall in the middle of the distribution of the predicted values.

Part II: Merging (left_join-ing) the demographics data

The following code will read in and join the demographics data.

```
dim(GS)
```

```
## [1] 143   7
```

```
Demographics <- read_excel("DemographicsF24.xlsx")
data.frame(sort(unique(GS$SubjectID)), sort(unique(Demographics$SubjectID)))
```

```
##      sort.unique.GS.SubjectID.. sort.unique.Demographics.SubjectID..
## 1             CAJustinBieber                      CAJustinBieber
## 2             COMichealFranti                     COMichealFranti
## 3                      DCSNAP                              DCSNAP
## 4                    FLTaller                            FLTaller
## 5                      GASNAP                              GASNAP
## 6               IAGarthBrooks                       IAGarthBrooks
## 7                       IDAJR                               IDAJR
## 8                       IDPAD                               IDPAD
## 9                 LADeathGrips                            LADeath
## 10                     MDEarl                              MDEarl
## 11                  NDLawrence                          NDLawrence
## 12                  NMSnuggies                          NMSnuggies
## 13                       NVKP                                NVKP
## 14              ORMichelangelo                      ORMichelangelo
## 15                    PATaller                            PATaller
## 16                     UTSNAP                              UTSNAP
## 17                     VTPink                              VTPink
## 18                    WATaller                            WATaller
## 19                   WIWiggles                           WIWiggles
## 20                       WYIW                                WYIW
## 21                  WYJourney                           WYJourney
## 22                 WYSlipKnot                          WYSlipKnot
## 23                    WYSNAP                              WYSNAP
## 24                   WYTaller                            WYTaller
```

7) The following code highlights a potential issue with one of the subject IDs. Fix the problem in the Demographics data.frame using R code.

```
data.frame(sort(unique(GS$SubjectID)), sort(unique(Demographics$SubjectID)))
```

```
##      sort.unique.GS.SubjectID.. sort.unique.Demographics.SubjectID..
## 1             CAJustinBieber                      CAJustinBieber
## 2             COMichealFranti                     COMichealFranti
## 3                      DCSNAP                              DCSNAP
## 4                    FLTaller                            FLTaller
## 5                      GASNAP                              GASNAP
## 6               IAGarthBrooks                       IAGarthBrooks
## 7                       IDAJR                               IDAJR
## 8                       IDPAD                               IDPAD
## 9                 LADeathGrips                            LADeath
## 10                     MDEarl                              MDEarl
## 11                  NDLawrence                          NDLawrence
## 12                  NMSnuggies                          NMSnuggies
## 13                       NVKP                                NVKP
## 14              ORMichelangelo                      ORMichelangelo
## 15                    PATaller                            PATaller
## 16                     UTSNAP                              UTSNAP
```

```
## 17                      VTPink                      VTPink
## 18                     WATaller                     WATaller
## 19                    WIWiggles                    WIWiggles
## 20                         WYIW                         WYIW
## 21                    WYJourney                    WYJourney
## 22                    WYSlipKnot                   WYSlipKnot
## 23                       WYSNAP                       WYSNAP
## 24                     WYTaller                     WYTaller
```

```r
# Fix the problem:
Demographics <- Demographics %>% mutate(SubjectID =
forcats::fct_collapse(SubjectID,

LADeathGrips = 'LADeath' ))

#Check that problem was fixed:
data.frame(sort(unique(GS$SubjectID)), sort(unique(Demographics$SubjectID)))
```

```
##     sort.unique.GS.SubjectID.. sort.unique.Demographics.SubjectID..
## 1              CAJustinBieber                       CAJustinBieber
## 2              COMichealFranti                     COMichealFranti
## 3                       DCSNAP                              DCSNAP
## 4                      FLTaller                            FLTaller
## 5                       GASNAP                              GASNAP
## 6                  IAGarthBrooks                       IAGarthBrooks
## 7                        IDAJR                               IDAJR
## 8                        IDPAD                               IDPAD
## 9                 LADeathGrips                        LADeathGrips
## 10                      MDEarl                              MDEarl
## 11                   NDLawrence                          NDLawrence
## 12                   NMSnuggies                          NMSnuggies
## 13                        NVKP                                NVKP
## 14               ORMichelangelo                      ORMichelangelo
## 15                     PATaller                            PATaller
## 16                      UTSNAP                              UTSNAP
## 17                      VTPink                              VTPink
## 18                     WATaller                            WATaller
## 19                    WIWiggles                           WIWiggles
## 20                         WYIW                                WYIW
## 21                    WYJourney                           WYJourney
## 22                    WYSlipKnot                          WYSlipKnot
## 23                       WYSNAP                              WYSNAP
## 24                     WYTaller                            WYTaller
```

```r
Demographics <- Demographics %>% mutate(Weights = factor(Weights)) %>%
mutate(
                                        Weights =
forcats::fct_collapse(Weights,
                                                    no = c("no",
"No"),
```

```
                                                                 yes = c("yes",
"Yes")))

combined <- left_join(x = GS, y = Demographics, by = "SubjectID")

combinedR <- combined %>% drop_na()

combinedR <- combinedR %>% mutate(forearm_bins   = cut_number(ForearmLength,
n = 3),
                                   epworth_bins = cut_number(Epworth, n = 3),
                                   balance_bins = cut_number(BalanceTime, n =
3))
dim(combinedR)

## [1] 137  20

tally(~SubjectID, data = combinedR)

## SubjectID
##  CAJustinBieber COMichealFranti          DCSNAP          FLTaller
GASNAP
##              6               5               6               6
6
##   IAGarthBrooks           IDAJR           IDPAD     LADeathGrips
MDEarl
##              6               6               6               6
0
##      NDLawrence       NMSnuggies            NVKP   ORMichelangelo
PATaller
##              6               6               6               6
6
##          UTSNAP           VTPink         WATaller        WIWiggles
WYIW
##              6               6               6               6
6
##       WYJourney      WYSlipKnot          WYSNAP         WYTaller
##              6               6               6               6
```

8) What is the sample size before and after the `left_join`? How many subjects are in the data set as analyzed in `combinedR`?

```
dim(GS)

## [1] 143    7

dim(combinedR)

## [1] 137  20
```

- Before the left join the df 'GS' had 143 observations. The new df 'CombinedR' has 137 observations. The new df has 23 subjects because MDEarl had 0 observations was dropped from the df.

9) Now we can incorporate forearm length into the model using the `forearm_bins, which we will treat as a fixed effect, and consider whether the differences in hand between arm positions on grip strength might change based on the length of the arm. Report an evidence sentence for the hand by arm by forearm bin interaction test.

```
lmer2 <- lmer(GripStrength ~ Hand*Arm*forearm_bins + (1|SubjectID), data =
combinedR)
Anova(lmer2, test.statistic = "F")

## Analysis of Deviance Table (Type II Wald F tests with Kenward-Roger df)
##
## Response: GripStrength
##                            F Df Df.res      Pr(>F)
## Hand                 13.2026  1 99.006 0.0004453
## Arm                   0.8257  2 99.006 0.4409032
## forearm_bins         15.3843  2 19.999 9.004e-05
## Hand:Arm              1.8485  2 99.007 0.1628710
## Hand:forearm_bins     1.3798  2 99.006 0.2564255
## Arm:forearm_bins      0.7275  4 99.006 0.5752496
## Hand:Arm:forearm_bins 2.1716  4 99.007 0.0776882

plot(allEffects(lmer2), multiline = T, ci.style = "bars")
```


Hand*Arm*forearm_bins effect plot
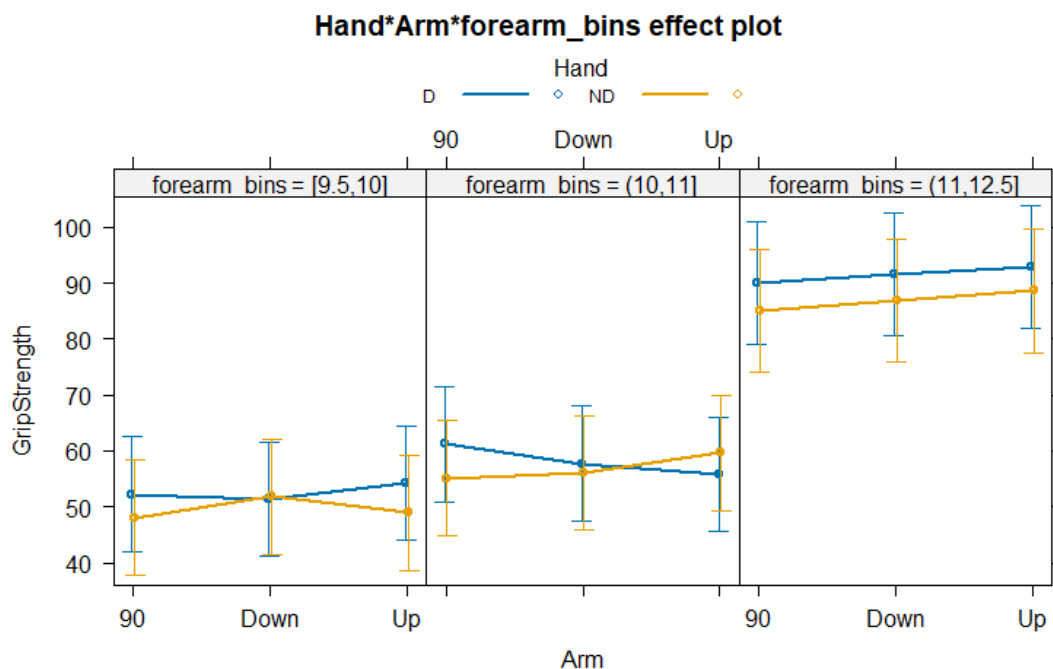
* There is some/ moderate evidence against the null that there of no interaction of hand:arm:forearm interaction $F(4, 99.007) = 2.1716$, $p = 0.077$, conditional of the random effect of subject, and we will keep this term in the model.

10) What combination of hand, arm, and forearm had the highest estimated mean grip strength based on the model (Hint: review the effects plot)? Based on the following results, what combinations of the three variables is that detectably different from (Hint: use the CLD)?

```
m1_em <- emmeans(lmer2, pairwise ~ Hand*Arm*forearm_bins, adjust = "tukey")
multcomp::cld(m1_em, Letters = LETTERS)
```

```
##  Hand Arm  forearm_bins emmean   SE   df lower.CL upper.CL .group
##  ND   90   [9.5,10]       48.0 5.24 24.7    37.3     58.8  A
##  ND   Up   [9.5,10]       48.9 5.19 23.9    38.2     59.6  A
##  D    Down [9.5,10]       51.4 5.19 23.9    40.7     62.1  A
##  ND   Down [9.5,10]       51.8 5.19 23.9    41.1     62.5  A
##  D    90   [9.5,10]       52.3 5.19 23.9    41.6     63.0  A
##  D    Up   [9.5,10]       54.2 5.19 23.9    43.5     65.0  A
##  ND   90   (10,11]        55.1 5.19 23.9    44.4     65.8  A
##  D    Up   (10,11]        55.8 5.19 23.9    45.1     66.5  AB
##  ND   Down (10,11]        56.0 5.19 23.9    45.3     66.8  AB
##  D    Down (10,11]        57.7 5.19 23.9    47.0     68.4  ABC
##  ND   Up   (10,11]        59.6 5.19 23.9    48.9     70.3  ABCD
##  D    90   (10,11]        61.2 5.19 23.9    50.5     72.0  ABCDE
##  ND   90   (11,12.5]      85.1 5.55 23.9    73.6     96.5   BCDEF
##  ND   Down (11,12.5]      86.8 5.55 23.9    75.3     98.2    CDEF
##  ND   Up   (11,12.5]      88.6 5.55 23.9    77.1    100.0     DEF
##  D    90   (11,12.5]      90.0 5.55 23.9    78.6    101.5      EF
##  D    Down (11,12.5]      91.6 5.55 23.9    80.1    103.0       F
##  D    Up   (11,12.5]      92.8 5.55 23.9    81.3    104.2       F
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 18 estimates
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
pwpm(m1_em)
```

```
##                 D 90 [9.5,10] ND 90 [9.5,10] D Down [9.5,10] ND Down
[9.5,10]
## D 90 [9.5,10]        [52.3]        0.9571         1.0000
1.0000
## ND 90 [9.5,10]        4.234        [48.0]         0.9962
0.9871
## D Down [9.5,10]       0.900        -3.334         [51.4]
1.0000
```

```
## ND Down [9.5,10]          0.500          -3.734          -0.400
[51.8]
## D Up [9.5,10]            -1.975          -6.209          -2.875          -
2.475
## ND Up [9.5,10]            3.350          -0.884           2.450
2.850
## D 90 (10,11]            -8.975         -13.209          -9.875          -
9.475
## ND 90 (10,11]           -2.800          -7.034          -3.700          -
3.300
## D Down (10,11]          -5.450          -9.684          -6.350          -
5.950
## ND Down (10,11]         -3.775          -8.009          -4.675          -
4.275
## D Up (10,11]            -3.525          -7.759          -4.425          -
4.025
## ND Up (10,11]           -7.350         -11.584          -8.250          -
7.850
## D 90 (11,12.5]         -37.754         -41.988         -38.654          -
38.254
## ND 90 (11,12.5]        -32.811         -37.045         -33.711          -
33.311
## D Down (11,12.5]       -39.282         -43.516         -40.182          -
39.782
## ND Down (11,12.5]      -34.496         -38.731         -35.396          -
34.996
## D Up (11,12.5]         -40.496         -44.731         -41.396          -
40.996
## ND Up (11,12.5]        -36.296         -40.531         -37.196          -
36.796
##                 D Up [9.5,10] ND Up [9.5,10] D 90 (10,11] ND 90 (10,11]
## D 90 [9.5,10]         1.0000         0.9937         0.9977         1.0000
## ND 90 [9.5,10]        0.5154         1.0000         0.9238         0.9999
## D Down [9.5,10]       0.9989         0.9999         0.9937         1.0000
## ND Down [9.5,10]      0.9998         0.9990         0.9959         1.0000
## D Up [9.5,10]         [54.2]         0.7094         0.9999         1.0000
## ND Up [9.5,10]         5.325         [48.9]         0.9531         1.0000
## D 90 (10,11]          -7.000        -12.325         [61.3]         0.4497
## ND 90 (10,11]         -0.825         -6.150          6.175         [55.1]
## D Down (10,11]        -3.475         -8.800          3.525         -2.650
## ND Down (10,11]       -1.800         -7.125          5.200         -0.975
## D Up (10,11]          -1.550         -6.875          5.450         -0.725
## ND Up (10,11]         -5.375        -10.700          1.625         -4.550
## D 90 (11,12.5]       -35.779        -41.104        -28.779        -34.954
## ND 90 (11,12.5]      -30.836        -36.161        -23.836        -30.011
## D Down (11,12.5]     -37.307        -42.632        -30.307        -36.482
## ND Down (11,12.5]    -32.521        -37.846        -25.521        -31.696
## D Up (11,12.5]       -38.521        -43.846        -31.521        -37.696
## ND Up (11,12.5]      -34.321        -39.646        -27.321        -33.496
##                 D Down (10,11] ND Down (10,11] D Up (10,11] ND Up
```

```
                        (10,11]
## D 90 [9.5,10]           1.0000          1.0000          1.0000
0.9998
## ND 90 [9.5,10]          0.9952          0.9994          0.9996
0.9731
## D Down [9.5,10]         1.0000          1.0000          1.0000
0.9991
## ND Down [9.5,10]        1.0000          1.0000          1.0000
0.9995
## D Up [9.5,10]           1.0000          1.0000          1.0000
1.0000
## ND Up [9.5,10]          0.9982          0.9999          0.9999
0.9862
## D 90 (10,11]            0.9891          0.7447          0.6727
1.0000
## ND 90 (10,11]           0.9996          1.0000          1.0000
0.8929
## D Down (10,11]          [57.7]          1.0000          1.0000
1.0000
## ND Down (10,11]          1.675          [56.1]          1.0000
0.9874
## D Up (10,11]             1.925           0.250          [55.8]
0.9754
## ND Up (10,11]           -1.900          -3.575          -3.825
[59.6]
## D 90 (11,12.5]         -32.304         -33.979         -34.229        -
30.404
## ND 90 (11,12.5]        -27.361         -29.036         -29.286        -
25.461
## D Down (11,12.5]       -33.832         -35.507         -35.757        -
31.932
## ND Down (11,12.5]      -29.046         -30.721         -30.971        -
27.146
## D Up (11,12.5]         -35.046         -36.721         -36.971        -
33.146
## ND Up (11,12.5]        -30.846         -32.521         -32.771        -
28.946
##                  D 90 (11,12.5] ND 90 (11,12.5] D Down (11,12.5]
## D 90 [9.5,10]         0.0043          0.0193          0.0027
## ND 90 [9.5,10]        0.0012          0.0054          0.0007
## D Down [9.5,10]       0.0032          0.0147          0.0020
## ND Down [9.5,10]      0.0037          0.0166          0.0023
## D Up [9.5,10]         0.0079          0.0344          0.0049
## ND Up [9.5,10]        0.0015          0.0070          0.0009
## D 90 (10,11]          0.0616          0.2169          0.0401
## ND 90 (10,11]         0.0101          0.0436          0.0063
## D Down (10,11]        0.0224          0.0904          0.0142
## ND Down (10,11]       0.0136          0.0574          0.0086
## D Up (10,11]          0.0126          0.0535          0.0079
## ND Up (10,11]         0.0390          0.1473          0.0250
```

```
## D 90 (11,12.5]                    [90.0]              0.8797              1.0000
## ND 90 (11,12.5]                   4.943               [85.1]              0.4865
## D Down (11,12.5]                 -1.529              -6.471               [91.6]
## ND Down (11,12.5]                 3.257              -1.686               4.786
## D Up (11,12.5]                   -2.743              -7.686              -1.214
## ND Up (11,12.5]                   1.457              -3.486               2.986
##                    ND Down (11,12.5] D Up (11,12.5] ND Up (11,12.5]
## D 90 [9.5,10]                      0.0116              0.0018              0.0067
## ND 90 [9.5,10]                     0.0032              0.0005              0.0018
## D Down [9.5,10]                    0.0088              0.0014              0.0051
## ND Down [9.5,10]                   0.0100              0.0016              0.0058
## D Up [9.5,10]                      0.0210              0.0034              0.0123
## ND Up [9.5,10]                     0.0042              0.0006              0.0024
## D 90 (10,11]                       0.1451              0.0282              0.0914
## ND 90 (10,11]                      0.0268              0.0044              0.0157
## D Down (10,11]                     0.0572              0.0098              0.0343
## ND Down (10,11]                    0.0356              0.0059              0.0210
## D Up (10,11]                       0.0331              0.0055              0.0195
## ND Up (10,11]                      0.0957              0.0175              0.0588
## D 90 (11,12.5]                     0.9978              0.9997              1.0000
## ND 90 (11,12.5]                    1.0000              0.1978              0.9953
## D Down (11,12.5]                   0.9052              1.0000              0.9993
## ND Down (11,12.5]                  [86.8]              0.6233              1.0000
## D Up (11,12.5]                    -6.000               [92.8]              0.9684
## ND Up (11,12.5]                   -1.800               4.200               [88.6]
##
## Row and column labels: Hand:Arm:forearm_bins
## Upper triangle: P values    adjust = "tukey"
## Diagonal: [Estimates] (emmean)
## Lower triangle: Comparisons (estimate)   earlier vs. later
```

- The combination of dominant, up, and length 11, 12.5 had the highest estimated mean of grip strength.
- This combination is different from the hand and arm combinations in the two lower forearm bins (i.e., [9.5,10] and [10,11]). In other words, all CLD that does not contain the letter F is different from this combination (i.e., dominant, up, and length bin 11, 12.5).

11) Note any additional resources used to complete this lab or NONE.

NONE