TMSTT_McRae

Yuchao Wang

2/12/2021

Matching McRae with MRC

Matching McRae 2005 ratings with additional ratings from MRC

```
CONCS_brm <- read.delim("~/Documents/GitHub/RStudio/TMSTT/McRae-BRM-InPress/CONCS_brm.txt")

McRae_MRC <- read.delim("~/Documents/GitHub/RStudio/TMSTT/McRae-BRM-InPress/McRae_MRC.txt")
McRae_MRC <- data.frame(lapply(McRae_MRC, function(x) {gsub("-", "NA", x)}))
McRae_MRC[, 2:7] <- sapply(McRae_MRC[, 2:7], as.numeric)
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
totalMRC <- nrow(McRae_MRC)
totalBRM <- nrow(CONCS_brm)
CombinedList <- cbind(CONCS_brm[FALSE, ], McRae_MRC[FALSE, 2:7]) #create empty dataframe with all relev

for (BRMindex in 1:totalBRM) {
  for (MRCindex in 1:BRMindex) {
    searchString <- paste("^",CONCS_brm$Concept[BRMindex],"$", sep = "") #this makes sure that the stri
    if (grepl(searchString, McRae_MRC$Word[MRCindex], ignore.case = TRUE)) {
      CombinedList <- rbind(CombinedList, (cbind(CONCS_brm[BRMindex, ], McRae_MRC[MRCindex, 2:7])))
    }
  }
}

CombinedList_short <- subset(CombinedList, select = c("Concept","ln.KF.", "Familiarity", "Length_Phoneme
CombinedList_short <- CombinedList_short[complete.cases(CombinedList_short), ]
write.csv(CombinedList_short, file = "~/Documents/GitHub/RStudio/TMSTT/McRae-BRM-InPress/CombinedList_sh
```

##Check difference in MRC and McRae ratings of familiarity

Highly correlated (r=0.7), using McRae 2005 since it was more recent than MRC and was done in the US.

##Check if possible to increase total number of features

Higher familiarity correlates with higher number of features produced

##Check group differences in number of taxonomic features

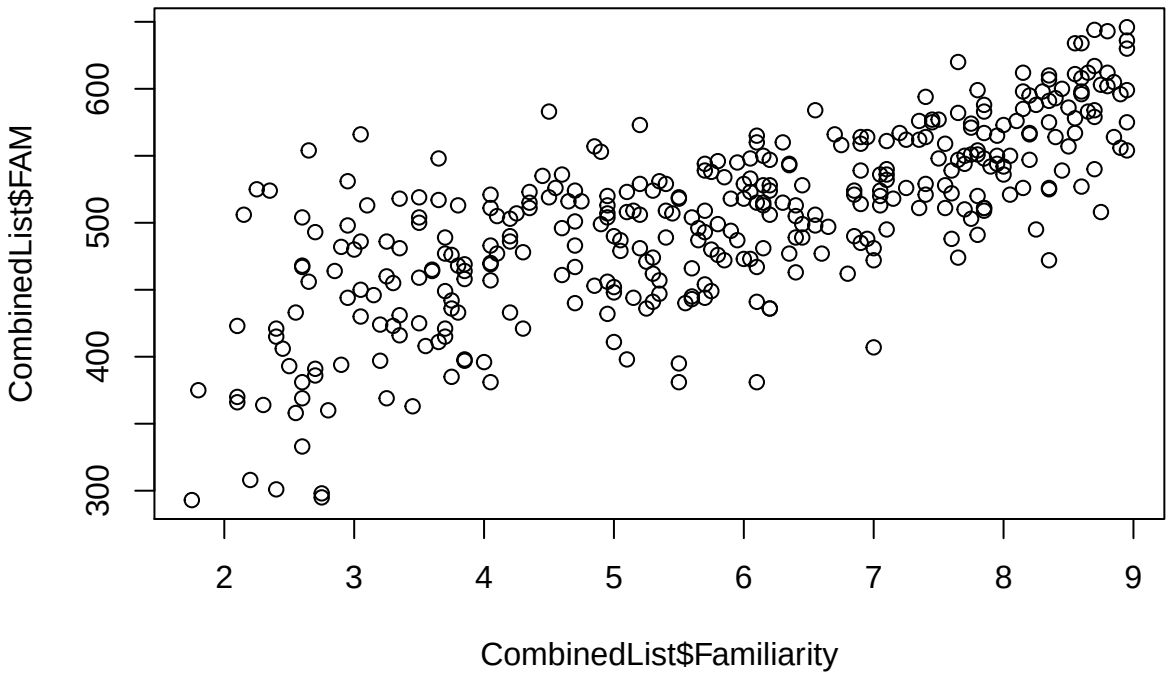
Animals tend to have more taxonomic features than household items and food

```
cor.test(CombinedList$Familiarity, CombinedList$FAM, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  CombinedList$Familiarity and CombinedList$FAM
## t = 20.571, df = 358, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6847418 0.7800321
## sample estimates:
##       cor
## 0.7360116
```

```
plot(CombinedList$Familiarity, CombinedList$FAM)

library(ggplot2)
```




```
cor.test(CombinedList$Num_Feats_Tax, CombinedList$Familiarity, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  CombinedList$Num_Feats_Tax and CombinedList$Familiarity
```

```
## t = 3.7223, df = 498, p-value = 0.00022
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.07796065 0.24862996
## sample estimates:
##       cor
## 0.1645265
```

```
Fam_NumFeats <- ggplot(data = CombinedList, aes(x = Num_Feats_Tax, y = Familiarity, label = Concept)) +
  geom_text(size = 1.5, alpha = 0.7, angle = 20, position = position_jitter(width = 1, height = 0.5)) +
  xlab("Total Number of Features") +
  ylab("Familiarity") +
  theme(text = element_text(size=15))
ggsave("Fam_NumFeats.png", plot = Fam_NumFeats, device = png(),
    scale = 1, width = 300, height = 180, units = c("mm"),
    dpi = 600, limitsize = TRUE)

library(readxl)
CombinedList_short_grouped <- read_excel("~/Documents/GitHub/RStudio/TMSTT/McRae-BRM-InPress/CombinedLis
ggplot(data = CombinedList_short_grouped, aes(x = Group, y = Num_Tax)) +
  geom_boxplot()
```

##Check agreement in 3 lists of stimuli

- Need 3 lists of 10 words each
- Within each list, 4 household items, 3 animals, 3 food
- Need to match at least on familiarity, concreteness, and imageability

```
PotentialList <- read.csv("~/Documents/GitHub/RStudio/TMSTT/McRae-BRM-InPress/Potential_Stim.csv")

summary(manova(cbind(Familiarity, CNC, IMG) ~ ListNum, data = PotentialList)) #one-way manova on famili
```

```
##           Df   Pillai approx F num Df den Df Pr(>F)
## ListNum    1 0.019572  0.17301      3     26 0.9137
## Residuals 28
```

```
summary(manova(cbind(ln.KF., Familiarity, Length_Phonemes, Length_Syllables, Num_Feats_Tax, Num_Feats_N
```

```
##           Df  Pillai approx F num Df den Df Pr(>F)
## ListNum    1 0.31226  0.86269     10     19 0.5803
## Residuals 28
```