Maxi

Claudia Falsetti

17 February 2018

R Markdown

This is an R Markdown document. Markdown is a formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com (http://rmarkdown.rstudio.com).

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

** Data Analysis **

Firstly, I set the working directory to point at the folder in which I saved the results.csv files and load the libraries that will be needed throughout the code.

```
library(tidyverse)
```

```
## ── Attaching packages ──────────────────────── tidyverse 1.2.1 ──
```

```
## ✔ ggplot2 2.2.1     ✔ purrr   0.2.4
## ✔ tibble  1.3.4     ✔ dplyr   0.7.4
## ✔ tidyr   0.7.2     ✔ stringr 1.2.0
## ✔ readr   1.1.1     ✔ forcats 0.2.0
```

```
## ── Conflicts ─────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following object is masked from 'package:tidyr':
##
##     expand
```

Now the .csv files are loaded in as a list and imported within the same dataframe named data_frame.

```
results = list.files(pattern = ".csv")
data = lapply(results, read.csv, header = FALSE, stringsAsFactors = FALSE)
data_frame <- do.call("rbind", data)
```

Data preprocessing

I eliminated unneeded columnes and rows and prepared the data for descriptive investigation and statistical analysis

```
data_frame <- data_frame[-1: -2, ] %>%  ## remove first 2 rows (completion code)
  select(V2, V7, V27, V28, V29, V33, V34)  ## and select relevant columns

colnames(data_frame) <- data_frame[1,]   ####### set the first row to be the header o
f the data frame
final_data <- data_frame[-1, ]

final_data <- filter(final_data, trialNo != "na", trialNo != "", trialNo != "trialNo"
)

final_data <- final_data %>%  # add a column for sujb_id, 275 represents n° of trials
 per participant
  mutate("subj_id" = rep(1:48,each=275))   ### change 5 with actual number of partici
pants
```

Descriptives.

Before proceding to the analysis, the dataset is explored. This includes measures of means, standard deviations and plots. A new dataframe called Ultimatum is created. It only includes results relative to the ultimatum game. The response to the ultimatum game is recoded as accepted = 1, rejected = 0. The descriptive statistics for the ultimatum game are summarised in the data_frame called ultim_summary and are visualised with a bar plot. The offers are coded as:

- 1: 90/10
- 2: 75/25
- 3: 60/40

```
final_data$RT <- as.numeric(final_data$RT)  ### set some variables to be as.numeric f
or future things
final_data$response <- as.numeric(final_data$response)

ultimatum <- final_data %>%   ## create a dataframe with only UltimatumGame results
  group_by(subj_id) %>%   #### recode response as acceptance -> accepted=1, rejected
 = 0
    filter(game == "ultimatum") %>%
      mutate(acceptance = if_else(response == 1, 0L, 1L)) %>%
        mutate(RT_S = RT / 1000) %>%
          select(subj_id, everything())

ultim_summary <- ultimatum %>%    #### descriptive statistics
  group_by(stim_type, offer) %>%
  summarise(p = mean(acceptance),
            sd = sd(acceptance))

accepted_only <- ultimatum %>%
  filter(acceptance == 1) %>%
    group_by(offer) %>%
      count(acceptance) %>%
        mutate(perc = n/3600*100)

accepted_only_2 <- ultimatum %>%
filter(acceptance == 1) %>%
group_by(offer, stim_type) %>%
count(acceptance)


united <- full_join(accepted_only, accepted_only_2, by = "offer") %>%
  select(offer, stim_type, n.x, perc, n.y) %>%
    group_by(offer, stim_type)

acceptance_rates <- vector("numeric", length = length(united$n.y))

for (i in 1:length(united$n.y)) {
    count <- (united[i, "n.y"])
    tot <- (united [i, "n.x"])


    acceptance_rates[i] <- count/tot*100

    }

rates <- do.call(rbind.data.frame, acceptance_rates)
colnames(rates) <- "acceptance_rates"

united["acceptance_rates"] <- rates$acceptance_rates


acceptance_overall <- ggplot(accepted_only, aes(x = offer, y = perc, fill = offer)) +
 geom_col() + coord_cartesian(ylim = c(0,100))


acceptance_stim <- ggplot(united, aes(x = stim_type, y = acceptance_rates, fill = sti
m_type)) + geom_col() +
  facet_grid(~offer)
```

```
col_plot <- ggplot(ultim_summary, aes(x = stim_type, y = p, fill = stim_type)) + geom
_col() +
  facet_grid(~offer)
```

Preparing the data for the analysis : **recoding** .

Due to the binary nature of the dependent variable and its binomial distribution, a mixed effects model was used to analyse the data. In the model, the acceptance to the ultimatum game was the response variable and offer and stimulus type were the predictors.

```
###### re- coding

ultimatum_dev <- ultimatum %>%
  mutate(offer_1 = ifelse(offer == "1", .5, -.5),
    offer_2 = ifelse(offer == "2", .5, -.5),
      brand = ifelse(stim_type == "brand", .5, -.5),
        human = ifelse(stim_type == "human", .5, -5))



model <- glmer(acceptance ~ (offer_1 + offer_2) * (brand + human) + ( (offer_1 + offe
r_2) * (brand + human)||subj_id), ultimatum_dev, binomial, control=glmerControl(optim
izer="bobyqa", optCtrl= list(maxfun=100000)))

summary(model)

##likelihood ratio tests

mod2 <- update(model, . ~ . -offer_1 -offer_2)
anova(model, mod2) # test main effect of type of offer

mod3 <- update(model, . ~ . -brand - human)
anova(model, mod3) # test main effect of source (stim_type)


mod4 <- update(model, . ~ . -offer_1:brand - offer_1:human - offer_2:brand - offer_2:
human)
anova(model, mod4) # test interaction
```

As expected, there was a main effect of offer type such that participants accepted the £40 offer 35% more often than the 25£ offer and 72% more often than the £10 offer: $\chi^2(2)$ = 102.57, p < .001

Means and standard deviations for each offer can be viewed in the table below. Offer 1 is the £10 offer, 2 is £25 and 3 is £40.

```
knitr::kable(ultim_summary)
```

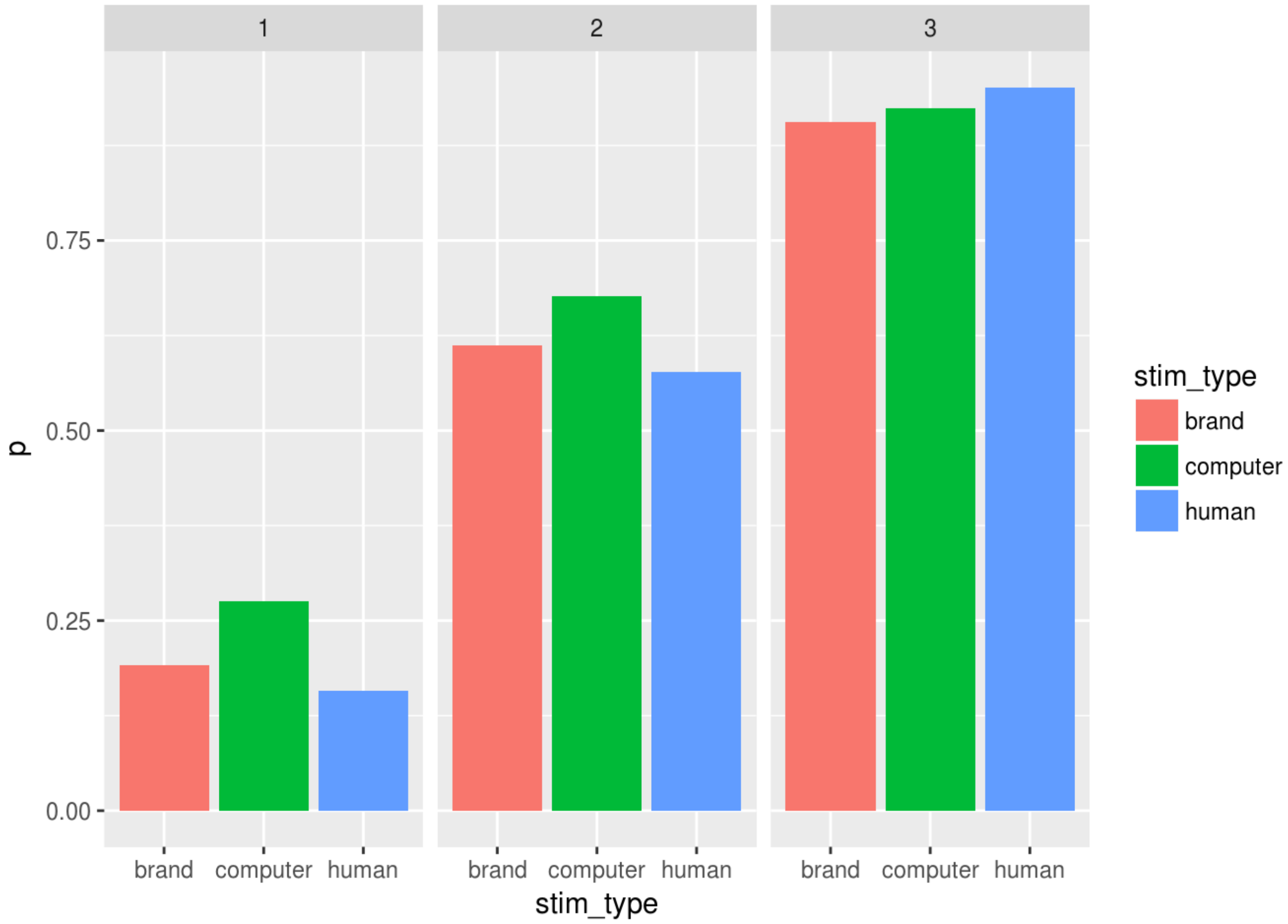
stim_type	offer	p	sd
brand	1	0.1908333	0.3931218
brand	2	0.6125000	0.4873825
brand	3	0.9058333	0.2921822
computer	1	0.2750000	0.4467004
computer	2	0.6766667	0.4679438

stim_type	offer	p	sd
computer	3	0.9241667	0.2648416
human	1	0.1583333	0.3652055
human	2	0.5766667	0.4942933
human	3	0.9516667	0.2145590

Moreover, the analysis revealed a significant main effect of the source (stim_type). As expected, unfair offers (£10) from the computer were accepted on average more often as compared to the same offers proposed by brands and humans : $\chi^2(2)$ = 15.761, p < .001

However, testing the interaction between offer type and proposer did not yielded significant results, $\chi^2(4)$ = 8.55, p = .073. This is attributable to the fact that the main effect of the source was observable only on one level of the offer type, namely the £10 offer (offer1). This is easily visualised in the bar chart below.

```
print(col_plot)
```



Code for the ANOVA, with the same variables

```
ultim_summary <- ultimatum %>%
  group_by(subj_id, stim_type, offer) %>%
   summarise(p = mean(acceptance), sd = sd(acceptance))

anova <- with(ultim_summary,
              aov(p ~ offer * stim_type +
                    Error(subj_id / (offer * stim_type)))
)

summary(anova)
```

```
##
## Error: subj_id
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  1 0.4767  0.4767
##
## Error: subj_id:offer
##       Df Sum Sq Mean Sq
## offer  2  29.59   14.79
##
## Error: subj_id:stim_type
##           Df Sum Sq Mean Sq
## stim_type  2 0.3349  0.1675
##
## Error: subj_id:offer:stim_type
##                 Df Sum Sq Mean Sq
## offer:stim_type  4 0.2692 0.06731
##
## Error: Within
##                  Df Sum Sq Mean Sq F value Pr(>F)
## offer             2   8.31   4.157  46.619 <2e-16 ***
## stim_type         2   0.03   0.015   0.165  0.848
## offer:stim_type   4   0.10   0.025   0.284  0.888
## Residuals       414  36.92   0.089
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ratings of Trustworthiness

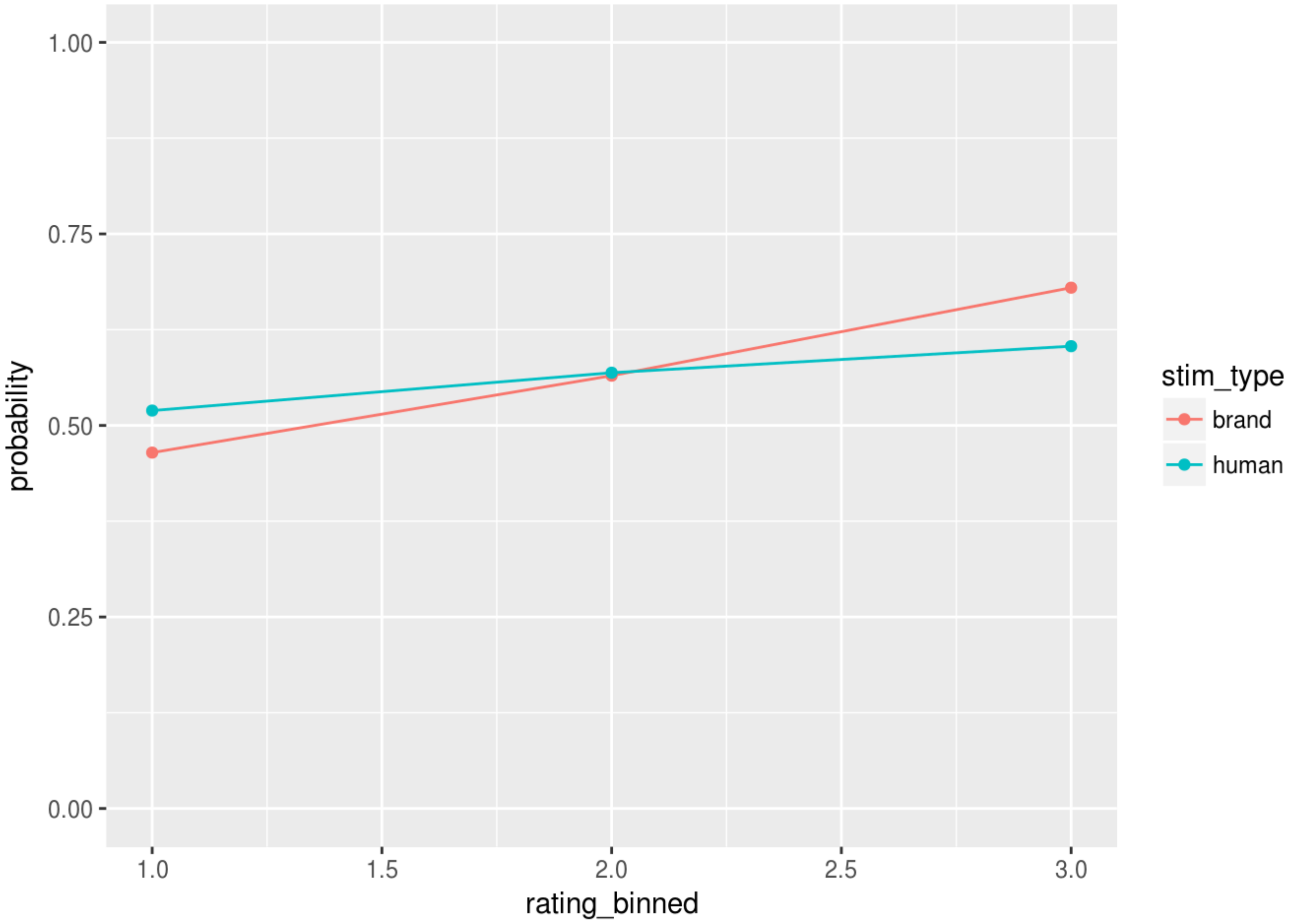
Working with the ratings of trustworthiness. Below, I created a dataframe with the trustworthiness ratings only. Ratings of trustworthiness are grouped into categories:

- rating1 = [0,10,20,30]
- rating2 = [4,50,60]
- rating3 = [70,80,90,100]

```
### recoding of variabels, binning the trustworthiness ratings into 3 categories.

trust <- final_data %>%
  group_by(subj_id) %>%
    filter(game == "trust") %>%
   mutate(rating_binned = ifelse(response <=4, 1,
          ifelse(response %in% 5:7, 2,
                 ifelse(response >= 8, 3, 0)))) %>%
        mutate(rating1 = ifelse(response <= 4, .5, -.5 ),
            rating2 = ifelse(response == 7 | response == 6| response ==5, .5, -.5), r
ating3 = ifelse(response >= 8, .5,-.5))
```

Now I created a merged data_frame including ratings of trustworthiness and ultimatum responses. In order to analyse the effect of brand and human only and account for the ratings of trustworthiness, the ultimatum game responses relative to computer offers were dropped from the following computations.

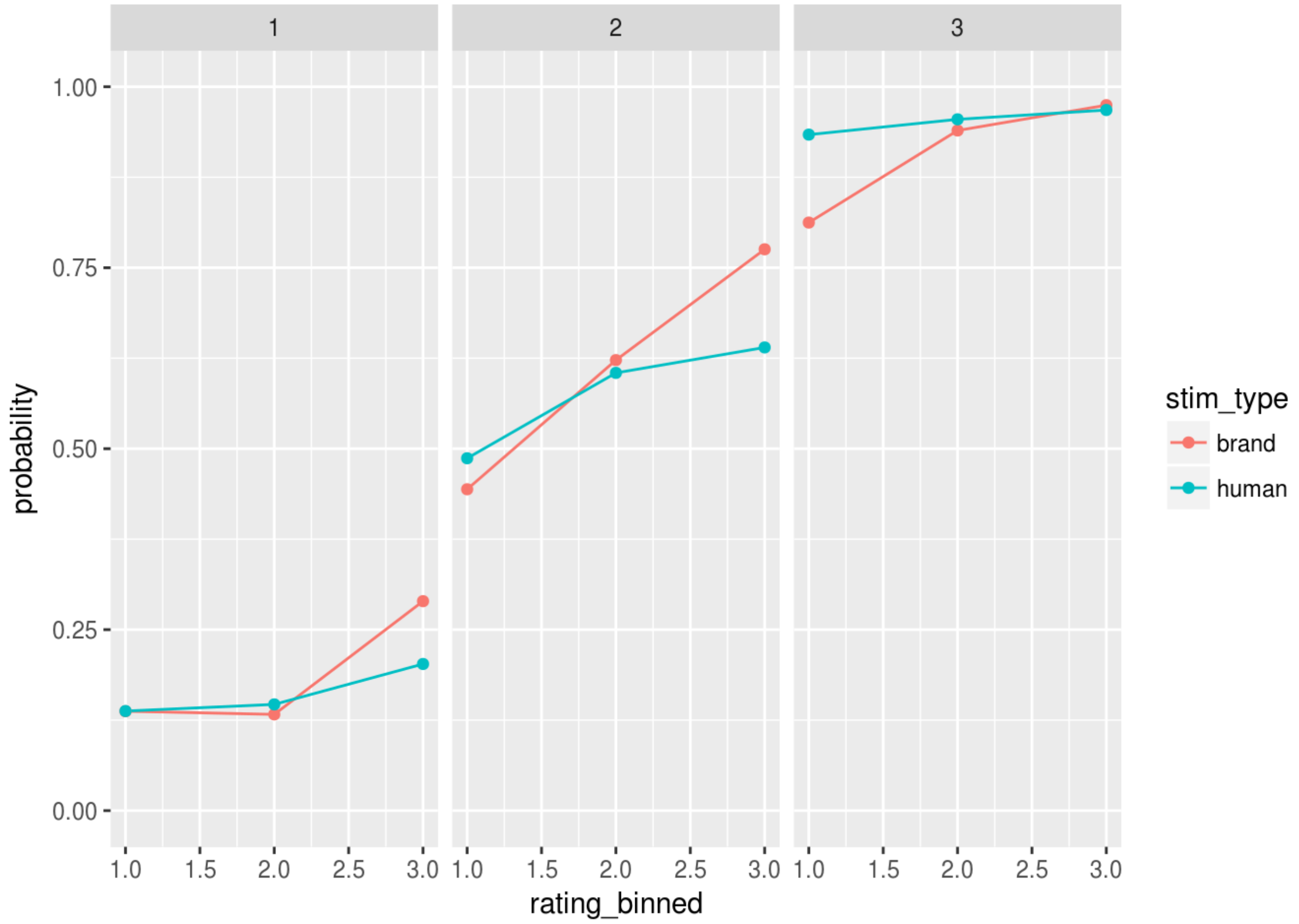
```
ultimatum_no_computer <- ultimatum %>%
  filter(stim_type != "computer") %>%
  select(stim_type, stim1, subj_id, acceptance, offer)

rating <- trust %>%
  group_by(stim1) %>%
select(rating_binned, subj_id, stim1, stim_type)

rating_mod <- trust %>%
  group_by(stim1) %>%
  select(rating1, rating2, rating3, stim1, subj_id, stim_type)

joined_mod <- merge(rating_mod, ultimatum_no_computer)

rat_mod_2 <- joined_mod %>%
  filter(offer < 3) %>%
  mutate(offer1 = ifelse(offer==1, .5, -.5),
        stim_type_b = ifelse(stim_type == "brand", .5, -.5)) %>%
  select(stim1, subj_id, stim_type_b, offer1, rating1, rating3, acceptance)

joined <- merge(rating, ultimatum_no_computer)

join_summary <- joined %>%
  group_by(rating_binned, stim_type, offer) %>%
    summarise(p = mean(acceptance))

join2 <- joined %>%
  group_by(rating_binned, stim_type) %>%
    summarise(p = mean(acceptance),
              sd = sd(acceptance))


plot2 <- ggplot(join2, aes(rating_binned, p, colour=stim_type)) +
        geom_line() +   geom_point() +   coord_cartesian(ylim = c(0, 1)) +
           labs(y = "probability")

print(plot2)
```



From the graph above it is possible to notice that at equal ratings of trustworthiness, participants accepted offers from brands and humans in a slight yet dissimilar way. Specifically, higher ratings of trustworthiness seemed to affect only acceptance of brands offer, which slightly increased.

In order to visualise the relationship in a more detailed manner, I added the different offer to the plot.

```
plot <- ggplot(join_summary, aes(rating_binned, p, colour=stim_type)) +
        geom_line() +   geom_point() +   coord_cartesian(ylim = c(0, 1)) +
           labs(y = "probability") + facet_grid(~offer)
print(plot)
```



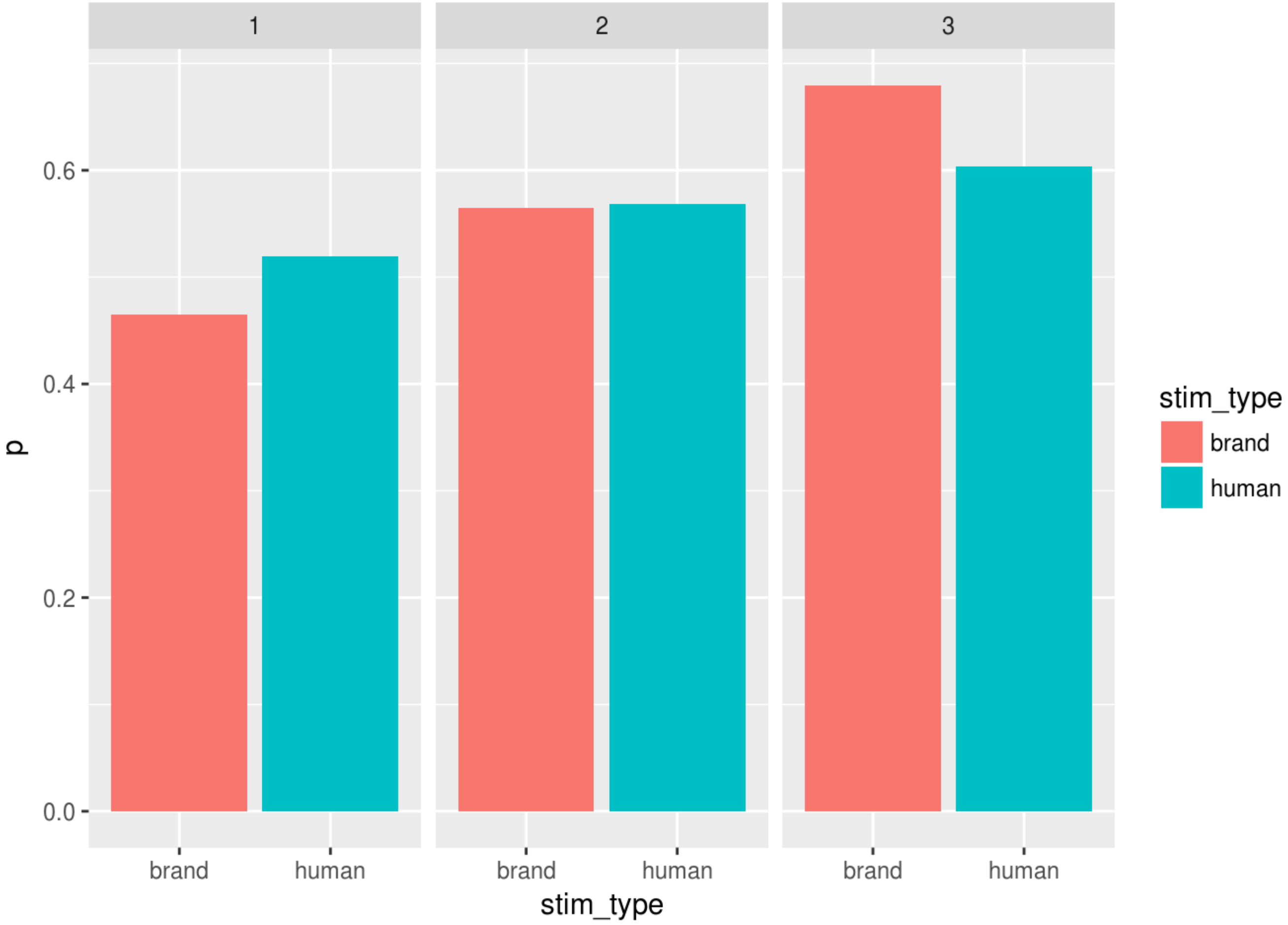
Althought no clear pattern or relationship seem to emerge from the graph, it is still noticeable that perceived trust has a bigger effect on acceptance of offers from brands as compared to offers from humans.

As the table and bar chart below show, differences in acceptance rates between brands and human offers were higher in the highest rating categories (rating3) and showed a preference to accept offers proposed by brands.

```
knitr::kable(join2)
```

rating_binned	stim_type	p	sd
1	brand	0.4645309	0.4989307
1	human	0.5194004	0.4998439
2	brand	0.5649547	0.4960128
2	human	0.5688193	0.4954029
3	brand	0.6797840	0.4667396
3	human	0.6034298	0.4894477

```
plot_join <- ggplot(join2, aes(x = stim_type, y = p, fill = stim_type)) + geom_col()
+
  facet_grid(~rating_binned)

print(plot_join)
```



ANALYSIS

Acceptance rates were analysed with a logistic mixed effects model in which offer type (£10 or 25£), rating category (binned into 3 categories as specified above) and source (brand or human) were coded as predictors. In order to account for the repeated measures design and allow generalization of results, subjects and items (stimuli, coded as stim1) were introduced as random factors.

```
########## mixed model

model_trust <- glmer(acceptance ~ offer1 * stim_type_b * (rating1 + rating3) + (offer
1 * stim_type_b ||subj_id) + (offer1 || stim1), rat_mod_2, binomial, control=glmerCon
trol(optimizer="bobyqa", optCtrl= list(maxfun=100000)))

summary(model_trust)

mod_trust2 <- update(model_trust, . ~ . -offer1)
anova(model_trust, mod_trust2) # test main effect of first factor

mod_trust3 <- update(model_trust, . ~ . -stim_type_b)
anova(model_trust, mod_trust3) # test main effect of second factor

mod_trust4 <- update(model_trust, .~ . -rating1 -rating3)
anova(model_trust, mod_trust4) ## test main effect of third factor

mod_trust_inter <- update(model_trust, . ~ . -offer1:stim_type_b - offer1:rating1 - o
ffer1:rating3 - stim_type_b:rating1 - stim_type_b:rating3)
anova(model_trust, mod_trust_inter) # test interaction


mod_new <- glmer(acceptance ~ (rating1 + rating3) * stim_type_b + (stim_type_b || sub
j_id), rat_mod_2, binomial, control=glmerControl(optimizer="bobyqa", optCtrl= list(ma
xfun=100000)))

summary(mod_new)
```

Interpretation

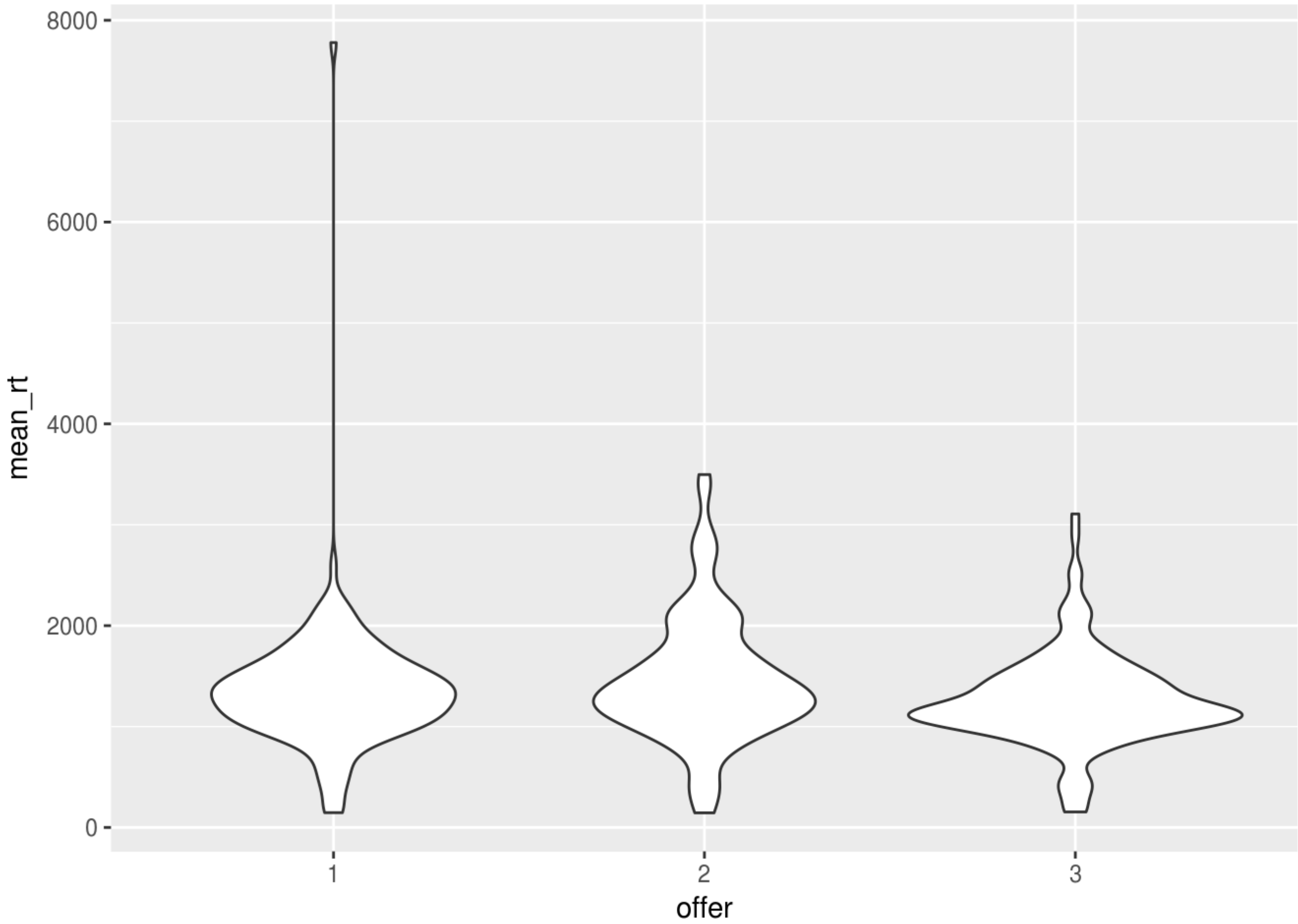
There was no main effect of source, $\chi^2(1)$ = 2.54, p = .110. As predicted from the graph, there was however a main effect of rating of trustworthiness, $\chi^2(2)$ = 12.031, p < .001, meaning that differences in ratings predicted differences in acceptance rates. Specifically, higher ratings were associated with increased acceptance.

Analysis of the interaction between ratings of trustworthiness and source type were non significant, suggesting that at equal ratings predict similar patterns of acceptance for both human and brands offer.

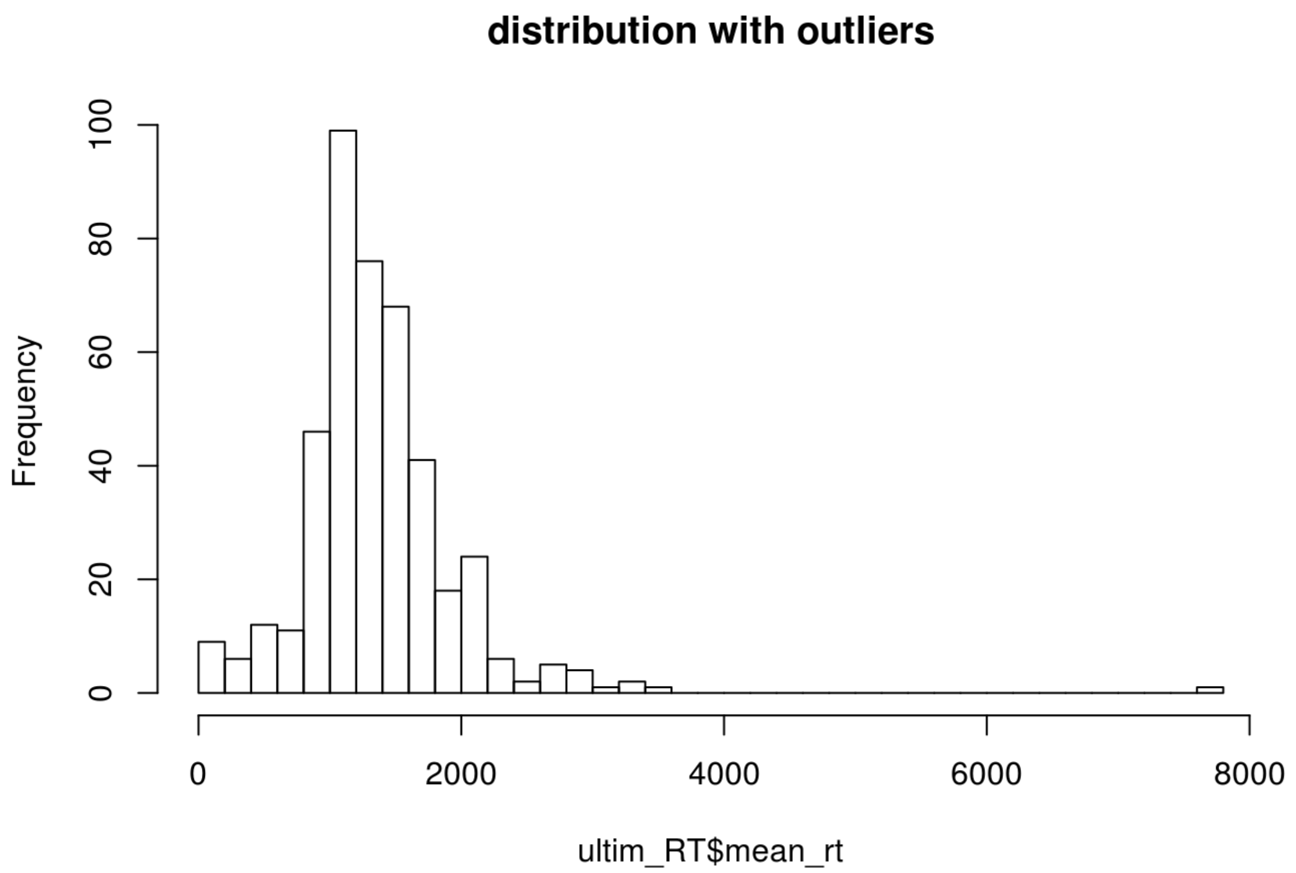
SECOND PART, **REACTION TIMES**.

I created a data_frame and summarised the reaction times grouping by subjects, stimulus type and kind of offer.

```
ultim_RT <- ultimatum %>%
  group_by(subj_id, stim_type, offer) %>%
  summarise(mean_rt = mean(RT), mean_p = mean(acceptance))

violinplot <- ggplot(ultim_RT, aes(offer, mean_rt)) + geom_violin()
print(violinplot)
```




```
col_plot_rt <- ggplot(ultim_RT, aes(x = offer, y = mean_rt, fill = offer)) + geom_col
() + facet_grid(~stim_type)

hist <- hist(ultim_RT$mean_rt, breaks = 30, main = "distribution with outliers")
```

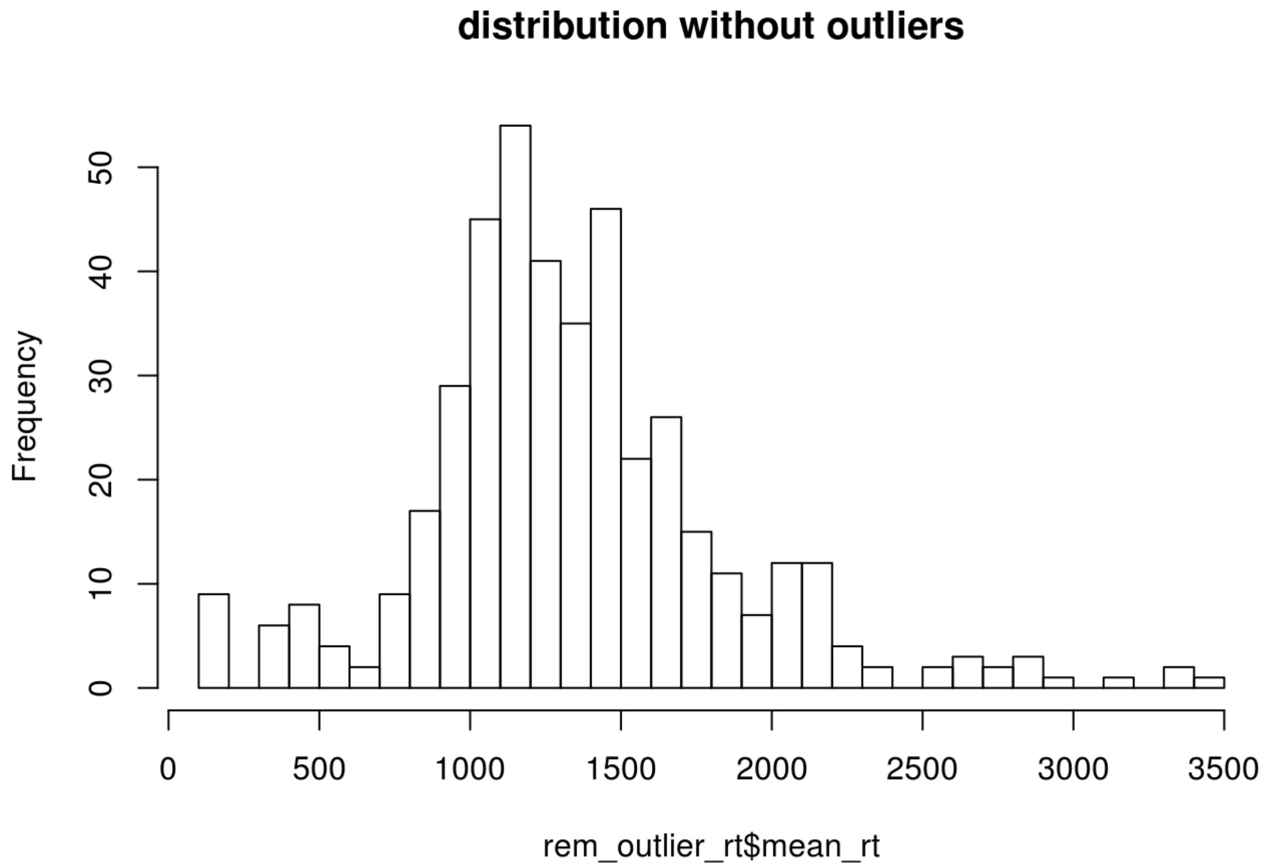


The distribution of the mean reaction times is visualised with a violin_plot split by offer type, while overall distribution is presented in the histogram below. It is possible to notice the presence of outliers which will be removed before continuing with the analysis.

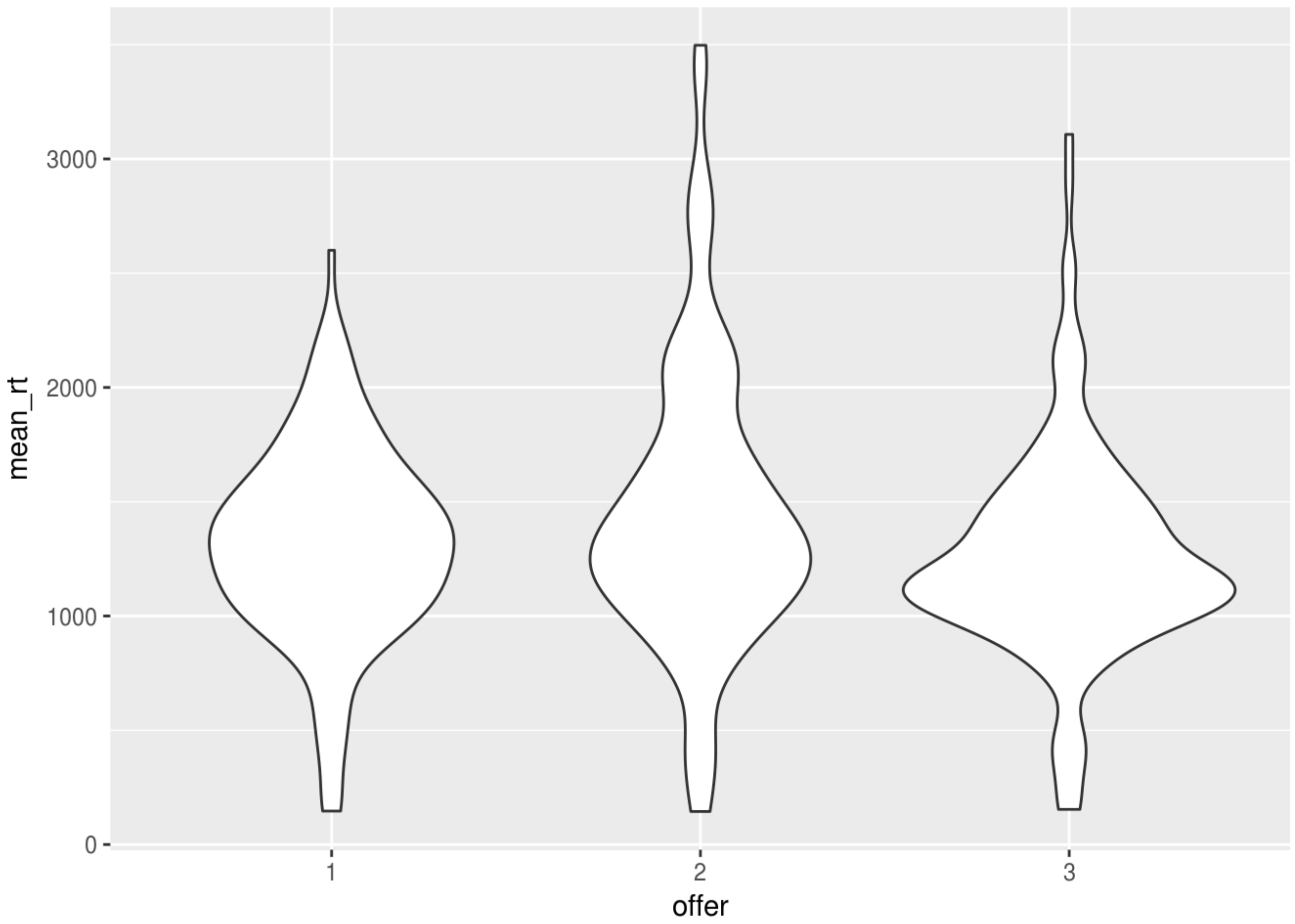
```
outliers <- ggplot(ultim_RT, aes(offer, mean_rt)) + geom_boxplot() + ggtitle("Mean Reaction Times With Outliers")

rem_outlier_rt <- ultim_RT %>%
  filter(mean_rt < 6000)

rem_outliers <- ggplot(rem_outlier_rt, aes(offer, mean_rt)) + geom_boxplot() + ggtitle("Mean Reaction Times Without Outliers")

###### extra graph

col_plot_no_out <- ggplot(rem_outlier_rt, aes(x = offer, y = mean_rt, fill = offer))
 + geom_col() + facet_grid(~stim_type)
```

The distribution without the outliers is presented in the figures below and looks closer to a normal distribution.

```
hist_no_out <- hist(rem_outlier_rt$mean_rt, breaks = 30, main ="distribution without
 outliers")
```




```
violinplot_no_out <- ggplot(rem_outlier_rt, aes(offer, mean_rt)) + geom_violin()
print(violinplot_no_out)
```

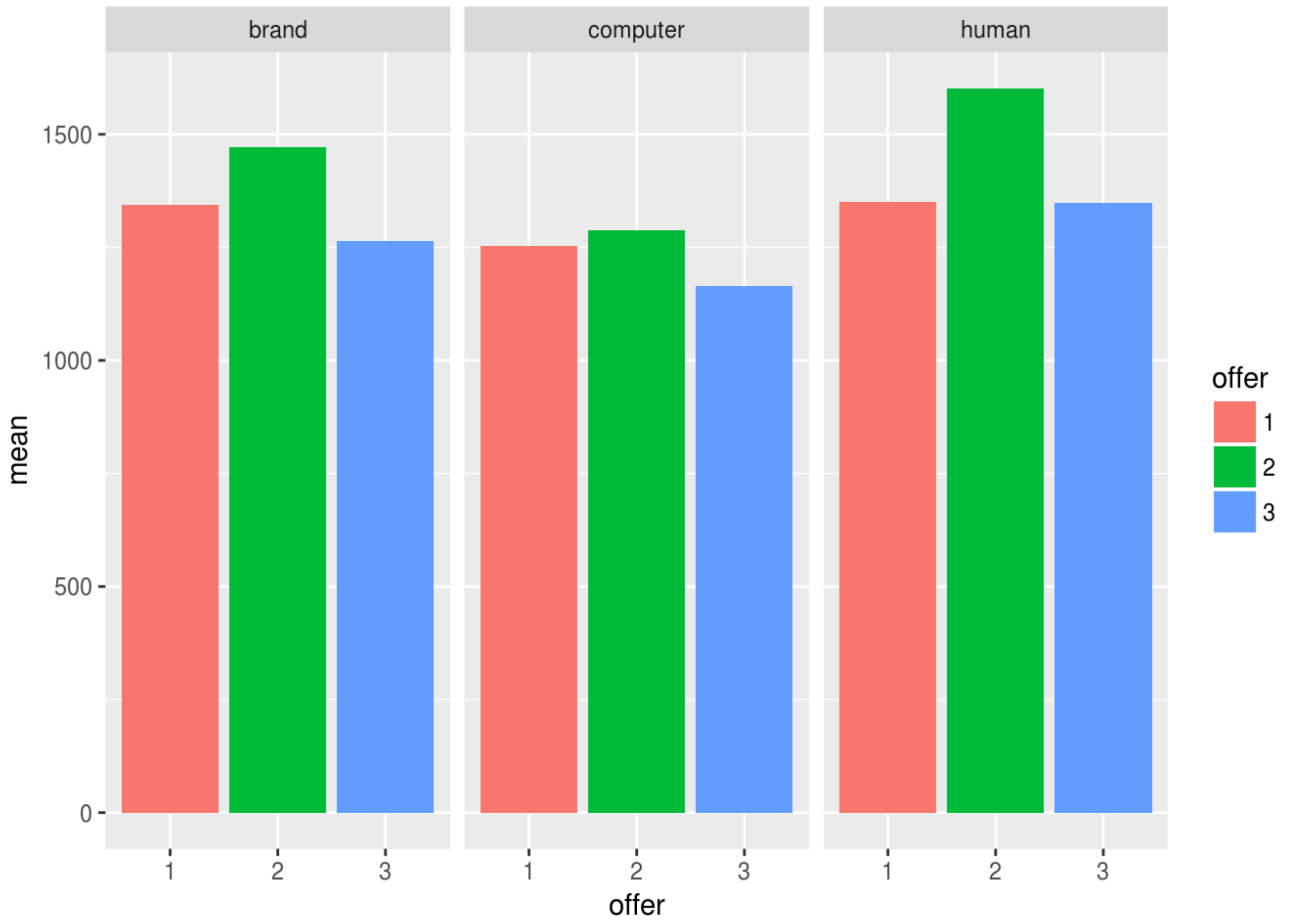


Descriptive statistics are provided in the table below and presented in the bar chart.

```
plot_rt <- rem_outlier_rt %>%
  group_by(offer, stim_type) %>%
  summarise(mean = mean(mean_rt),
         SD = sd(mean_rt)) %>%
  select(stim_type, offer, mean, SD) %>%
  print()
```

```
## # A tibble: 9 x 4
## # Groups:   offer [3]
##   stim_type offer     mean       SD
##       <chr> <chr>    <dbl>    <dbl>
## 1     brand     1 1344.566 450.7238
## 2  computer     1 1253.787 390.0796
## 3     human     1 1350.324 442.1020
## 4     brand     2 1472.200 619.9832
## 5  computer     2 1287.874 445.6001
## 6     human     2 1601.638 739.7303
## 7     brand     3 1264.629 432.0602
## 8  computer     3 1165.452 357.9413
## 9     human     3 1347.614 559.8645
```

```
plot__ <- ggplot(plot_rt, aes(x = offer, y = mean, fill = offer)) + geom_col() + face
t_grid(~stim_type)

print(plot__)
```



The bar plot shows that intensity of the offer (highest fairness and highess unfairness (offer 1 and 3) are associated as expected with shorter reaction times as compared to the highest uncertainty condition simulated with offer2 (25/75 split). This relationship was analysed with a repeated measures analysis of variance

```
#### recoding

ultim_RT_recode <- rem_outlier_rt %>%
    mutate(offer_1 = ifelse(offer == "1", .5, -.5),
    offer_2 = ifelse(offer == "2", .5, -.5),
      brand = ifelse(stim_type == "brand", .5, -.5),
        human = ifelse(stim_type == "human", .5, -5))

## mixed model formula

mod_rt <- lmer(mean_rt ~ (offer_1 + offer_2) * (brand + human) + ( (offer_1 + offer_
2) * (brand + human)||subj_id), ultim_RT_recode)

summary(mod_rt)

 ##likelihood ratio tests

rt_offer_mod <- update(mod_rt, . ~ . -offer_1 -offer_2)
anova(mod_rt, rt_offer_mod) # test main effect of first factor
```

```
## refitting model(s) with ML (instead of REML)
```

```
rt_stim_mod <- update(mod_rt, . ~ . -brand - human)
anova(mod_rt, rt_stim_mod) # test main effect of second factor
```

```
## refitting model(s) with ML (instead of REML)
```

```
rt_interaction_mod <- update(mod_rt, . ~ . -offer_1:brand - offer_1:human - offer_2:b
rand - offer_2:human)
anova(mod_rt, rt_interaction_mod) # test interaction
```

```
## refitting model(s) with ML (instead of REML)
```

The analysis showed a main effect of offer, such that offer1 and offer2 significantly differed in the acceptance reaction times as compared to offer3, set as baseline measure. Specifically, offers in the most uncertain condition (£25) were accepted slower than offers in the other categories.

However, no significant interaction was found between the factors.

```
options(contrasts = c("contr.sum","contr.poly"))

mod_rt <- aov(mean_rt ~ stim_type*offer + Error(subj_id), ultim_RT_recode)
summary(mod_rt)
```

```
##
## Error: subj_id
##           Df  Sum Sq Mean Sq
## stim_type  1 5112537 5112537
##
## Error: Within
##                  Df    Sum Sq Mean Sq F value  Pr(>F)
## stim_type         2   2874022 1437011   5.861 0.00308 **
## offer             2   2883654 1441827   5.881 0.00303 **
## stim_type:offer   4    594073  148518   0.606 0.65867
## Residuals       421 103214484  245165
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The anova yielded results similar to the mixed models regression by finding a main effect of source type and offer, and no effect of the interaction of the factors.