Tunes Together: Understanding Users' Attitudes and Behaviors Towards Collaborative Music Playlists

playlist

So Yeon Park & Alexa Siu

2019-03-21 21:10:37

Contents

1 Introduction

Music has traditionally prompted people to "engage in social cognition, participate in co-pathy, communicate, coordinate their actions, and cooperate with each other, leading to increased social cohesion" (Koelsch 2013). The importance of the social implications of music enjoyment have been underscored by research on how to better afford social interactions using systems with community-related activities (Brown, Sellen, and Geelhoed 2001), understanding music in cars (Axelsson and Östergren 2002), and designing social music players that aim to heighten the "extensive social functions" that music serves using physical devices (Lenz et al. 2012) and digital platforms (Liu and Reimer 2008). Furthermore, even research not focused on the social aspects highlight the importance of this. For example, in exploring music information needs and behaviors through a large-scale user survey, Lee et al. have found that "there is a strong social component to people's experience of and interaction with music, and music services that successfully incorporate such social features are well received," highlighting the importance of the "social aspect" in design implications found in their research (Lee, Cho, and Kim 2016).

Yet, investigations on the social aspect of music enjoyment is relatively lacking compared to other realms of music technology research such as AI recommendations and machine-learned music creation. And while there is multiple works on creating social playlist systems, whether tangible or digital, we lack an understanding of the current state of how users feel about and interact with collaborative music platforms. There have been multiple social platforms, such as Turntable.fm which provided such a service but failed to keep the original value proposition and became defunct (Danielson 2011). There are, on the other hand, more long-running and successful platforms with collaborative playlist (CP) functions, such as Spotify which allows users to co-create and co-modify a playlist since 2008 ("Collaborate on Playlists with Spotify's Collaboration Feature," n.d.). Despite the increasing popularity of music co-enjoyment platforms, there has yet to be published research looking explicitly into this phenomenon of collaborative playlisting in consideration of the importance of music's social qualities. There are some literature on this topic, but many ought to be revisited as the landscape of music co-enjoyment has evolved radically with music enjoyment nowadays is primarily through access-based consumption (Bardhi and Eckhardt 2012).

Moreover, looking at the landscape of collaborative tools, we see that investigations of those in music enjoyment is lacking compared to technologies in other fields. The social aspects of today's technologies are increasingly gaining importance, and technology's role in collaboration, in general, is a prevailing topic of our conversations today. Research in computer-mediated productivity and teamwork is particularly prolific, with meeting mediators discouraging overlapped speaking whilst encouraging more team interaction (Kim et al. 2008), robots acting as emotional regulators during team conflicts (Jung, Martelaro, and Hinds 2015), and digital systems facilitating organizational flexibility for "globally distributed online workforce to accomplish complex work" (Valentine et al. 2017). Conversely, there is a relative dearth of research in current collaborative technologies for our most intimate and social experience of music. These underpin the importance of technology's mediation in music and have led to our collaborative music technologies today. However, music's socially engaging traits are increasingly jeopardized by technologies that propagate individualized consumption of music (Frank 2009).

This investigation aims at gaining a better understanding of the phenomenon or act of engaging in CP, which is increasingly gaining traction. In preliminary work (Park and Kaneshiro 2017), we found that there are both advantages and disadvantages to engaging in such collaborative endeavors and wish to further explore these trends in detail. In particular we want to identify current behaviors and sentiments users have towards CP so that we can better support their needs from a user-centered perspective. A concrete understanding of designing HCI through music co-enjoyment will help build HCI principles that can influence the landscape of human collaborations.

1.1 Research Questions & Hypotheses

Two broad research questions drive this investigation:

1. Do people that participate in collaborative playlists have different music values and habits compared to people that don't participate in collaborative playlists? We hypothesize that those who participate in collaborative playlists:
   a. Value listening and sharing music with others.
   b. Frequently engage in sharing music with others.
   c. Frequently engage in discovering new music.
   d. Experiment more with new and unfamiliar music.
2. Is user's engagement in a collaborative playlist related to their sense of ownership? We hypothesize that for those who engage in collaborative playlists:
   a. Their sense of CP ownership is associated with perceived role and frequency of contributions.
   b. Their perceived role of adding new songs to collaborative playlist is most indicative of ownership.

Three exploratory questions were also considered as part of the scope of this investigation:

1. Do people use different music platform for different activities?
2. Is there a relationship between when users' first collaborative playlist was formed and when their favorite collaborative playlist was formed? (i.e., Are users' favorite collaborative playlist the first one they have created?)
3. How has the use of collaborative playlist changed users' music habits and preferences? Are these different from what non-users predict that collaborative playlist usage would engender?

2 Methods

2.1 Survey study

We conducted a survey that was built upon findings from preliminary work (Park and Kaneshiro 2017) that revealed differences amongst collaborative playlist users and non-users. We selected 57 questions asking about current music habits and experiences, whether participants use collaborative playlists or not, changes (or predicted changes) in behavior they believe engaging in the collaborative playlist would have on them, and whether collaborative playlist usage enhances social connectedness. Depending on survey answers, participants answered a minimum of 26 questions and up to 47 questions, with participants participating in collaborative playlists answering more questions, as we inquired about one of their favorite playlists.

The survey was created using the Qualtrics platform. Once we finalized the survey, we recruited participants through an introductory music class at Stanford University, a class predominantly consisting of undergraduate students. We screened the participants to meet the eligibility criteria of being 18 years and older and fluency in English. There was no compensation provided for participating in the survey study. The study was conducted under an approved IRB.

2.2 Statistical analyses

For the first part of the research analysis, we were interested in comparing differences between CP and non-CP users. To compare music values and habits between users of CP and non-CP users we looked at four aspects: (1) music as a means of social interaction, (2) music sharing, (3) music discovery, and (4) music taste. These metrics were associated with the following questions:

1. Social Interaction: How important are these activities to your social relationships? (Very Important, Somewhat Important, Neither Important nor Unimportant, Somewhat Unimportant, Very Unimportant)
   a. Experiencing musical events with others
   b. Performing or creating music with others
   c. Listening to recorded music with others
   d. Sharing music with others
   e. Discussing music with others
   f. Having non-musical interactions with others
2. Music Sharing: How often do you engage in sharing music with others? (At Least Once Per Day, At Least Once Per Week, At Least Once Per Month, A Few Times Per Year or Less, Never)
3. Music Discovery: How often do you engage in discovering music with others? (At Least Once Per Day, At Least Once Per Week, At Least Once Per Month, A Few Times Per Year or Less, Never)
4. Music Taste: I often enjoy listening to music that ... (Strongly Agree, Somewhat Agree, Neither Agree nor Disagree, Somewhat Disagree, Strongly Disagree)
   a. Is familiar.
   b. Is similar to other music that I listen to.

We used a logistic regression model to determine if these four factors were predictors for users or non-users of CP. The dependent variable, use of CP was binary. The independent variables, the Likert scale responses, were treated as continuous scales. For all tests, an alpha level of 0.05 was used to assess significance.

For the second part of the research analysis, we focused on questions answered by CP users only. Two metrics were hypothesized to quantify user engagement: user's perceived role and their frequency of contribution to the playlist. A collection of questions was used to determine these metrics.

For a user's perceived role, we used the following questions:

1. Please select your role of contribution to the collaborative playlist in terms of the following: (Lead/Primary, Equal to Others, Supporting, Minimal, None)
   a. Adding music
   b. Deleting music
   c. Reordering music
   d. Other

For frequency of contribution to a playlist, we used the following questions:

2. Please select the frequency with which you perform the following actions on the collaborative playlist: (Always, Often, Sometimes, Rarely, Never)
   a. Check (e.g. see if new songs were added)
   b. Listen/play alone
   c. Listen/play with others
   d. Contribute (e.g. modify, add, delete)
   e. Share (e.g. tell others about the playlist, post on social media)

Last, to determine sense of ownership, we used the following questions:

3. The songs that my collaborator(s) not including myself have contirbuted to the collaborative playlist are… (Strongly Agree, Somewhat Agree, Neither Agree nor Disagree, Somewhat Disagree, Strongly Disagree)
   a. my songs
   b. songs in my playlist
   c. songs in our collaborative playlist
   d. songs belonging equally to me and my collaborator(s)
   e. my collaborator(s)'s songs in our collaborative playlist

To assess significance, we compared two linear models, a compact model including only frequency and an augmented model including also user's role. The overall mean for the questions was used in the analysis. The models were compared using an anova and proportion reduction of error (PRE).

To further understand what actions might be most indicative of the sense of ownership CP users had, we ran a linear regression between their sense of ownership and their role of contribution to the collaborative playlist in terms. We also used an anova to measure how well the linear model generated fit the data.

3 Results

3.1 Data

3.1.1 Load

```
df.raw <- read.csv("../data/SY Music Playlist - V3_March 1, 2019_14.26.csv")
```

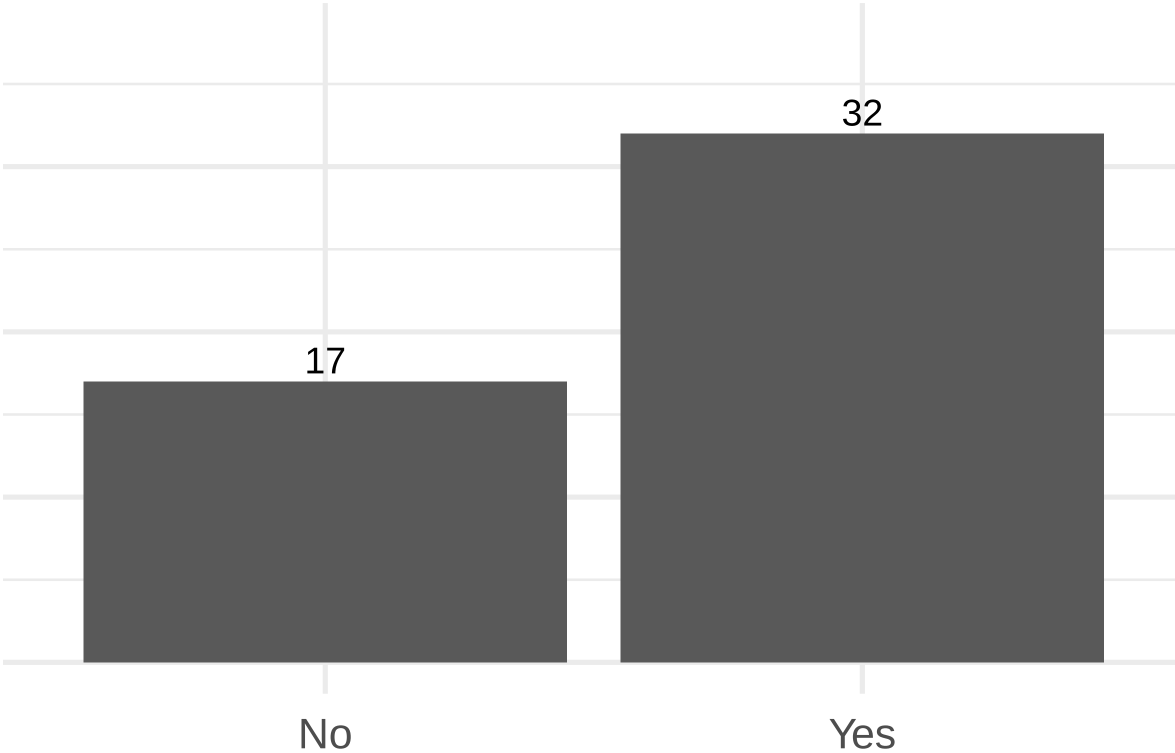
3.1.2 Filter

```
# remove test data
num_under18 <- df.raw %>%
  filter(Q1.3 == "No")

df.raw_all <- df.raw %>%
  filter(Progress == 100,
         Status == "IP Address",
         Finished == TRUE, Q3.1 != "asdf", Q1.3 != "No") %>%
         # Finished == "TRUE", Q3.1 != "asdf", Q1.3 != "No") %>%
    select(-StartDate, -EndDate, -IPAddress, -Progress, -Duration..in.seconds.,
           -Finished, -RecordedDate, -Status,
           -ResponseId, -RecipientLastName, -RecipientFirstName,
           -RecipientEmail, ExternalReference, -LocationLatitude,
           -LocationLongitude, -DistributionChannel, -UserLanguage )

df.raw_yes = df.raw_all %>%
    # only get people that use collaborative playlists
    filter(Q2.3 == "Yes")
```

```r
num_under18 <- df.raw %>%
  filter(Q1.3 == "No")

df.raw <- df.raw %>%
  filter(Progress == 100, Status == "IP Address",
         Finished == TRUE, Q3.1 != "asdf", Q1.3 != "No")
```

3.2 Survey Statistics

```r
totalAnswers = nrow(df.raw_all)

groupByCPUser = df.raw_all %>%
  group_by(Q2.3) %>%
  summarize(tally = n() )

totalCPUsers = groupByCPUser$tally[2]
totalNonCPUsers = groupByCPUser$tally[1]

df.raw_all %>%
  mutate(id = row_number()) %>%
  select(Q2.3) %>%
  group_by(Q2.3) %>%
  summarize(tally = n() ) %>%
  ggplot(aes(x = Q2.3, y = tally)) +
  geom_bar(stat="identity") +
  geom_text(aes(label=tally), vjust=-0.3, color="black", size=5.0) +
  ylim(0,38) +
  xlab("Do you use collaborative playlists?") +
  theme_minimal(base_size = 20) +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.y = element_blank(),
        axis.text.y = element_blank())

mean.age = df.raw %>%
  select(Q2.3, Q8.1_3) %>%
  mutate_if(is.factor, as.character) %>%
  mutate_if(is.character, as.numeric) %>%
  summarize(mean = round(mean(Q8.1_3),3))

df.gender = df.raw %>%
  select(Q2.3, Q8.1_4)
df.gender$gender[df.gender$Q8.1_4 == "Female" | df.gender$Q8.1_4 == "female"] = "F"
df.gender <- df.gender %>% group_by(gender) %>% tally() %>% mutate(percentage = n/nrow(df.raw))
```



Do you use collaborative playlists?

There were a total of 49 respondents to the survey. 32 reported using collaborative playlists and 17 reported never using a collaborative playlist. The mean age for survey respondents was 20.898 years, with 47% reported female and 53% reported male. The mean time taken to complete survey 1273 seconds (~ 21 minutes), a value we calculated after eliminating extreme values, which are caused by not completing the survey in one-go but keeping it open and taking days to complete it, as we allowed respondents to save and continue the survey later.

3.3 Confirmatory analysis

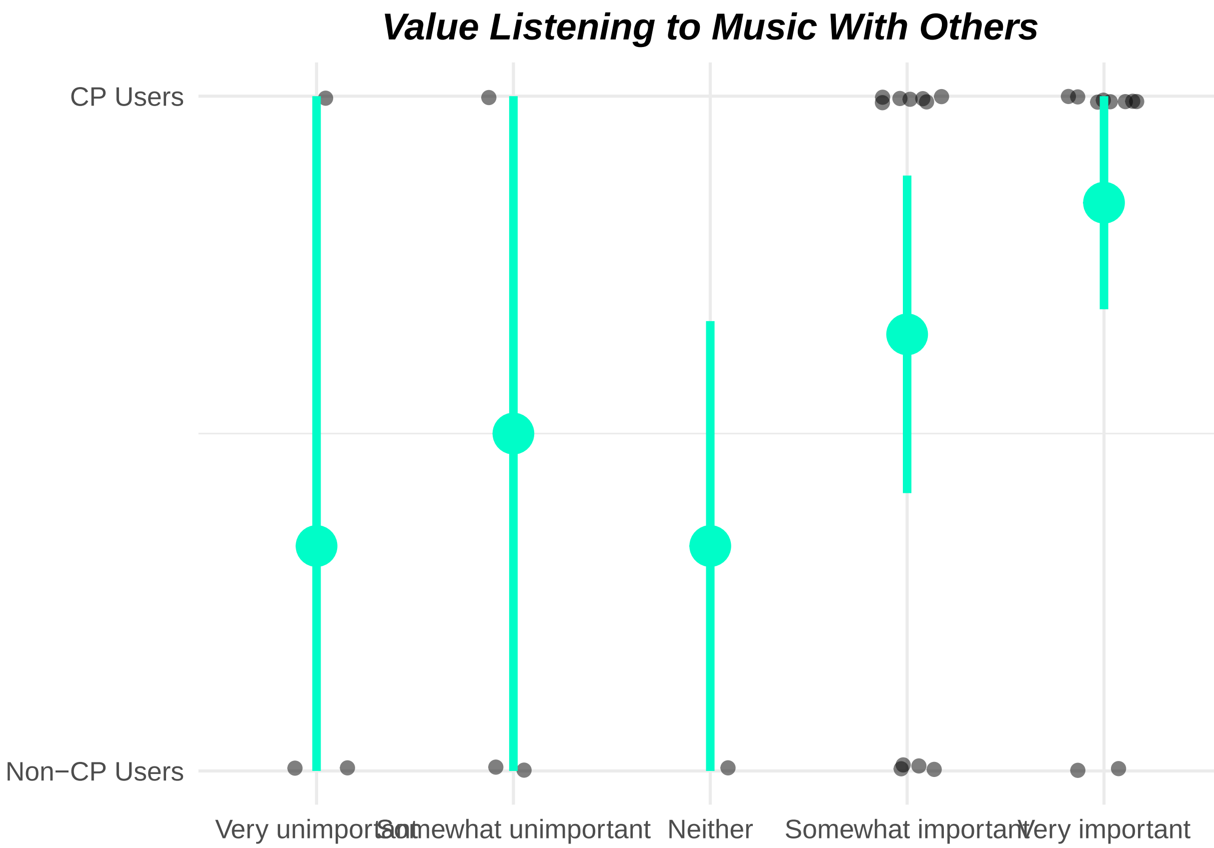
3.3.1 Q1 Do people that participate in collaborative playlists have different music values and habits compared to people that don't participate in collaborative playlists?

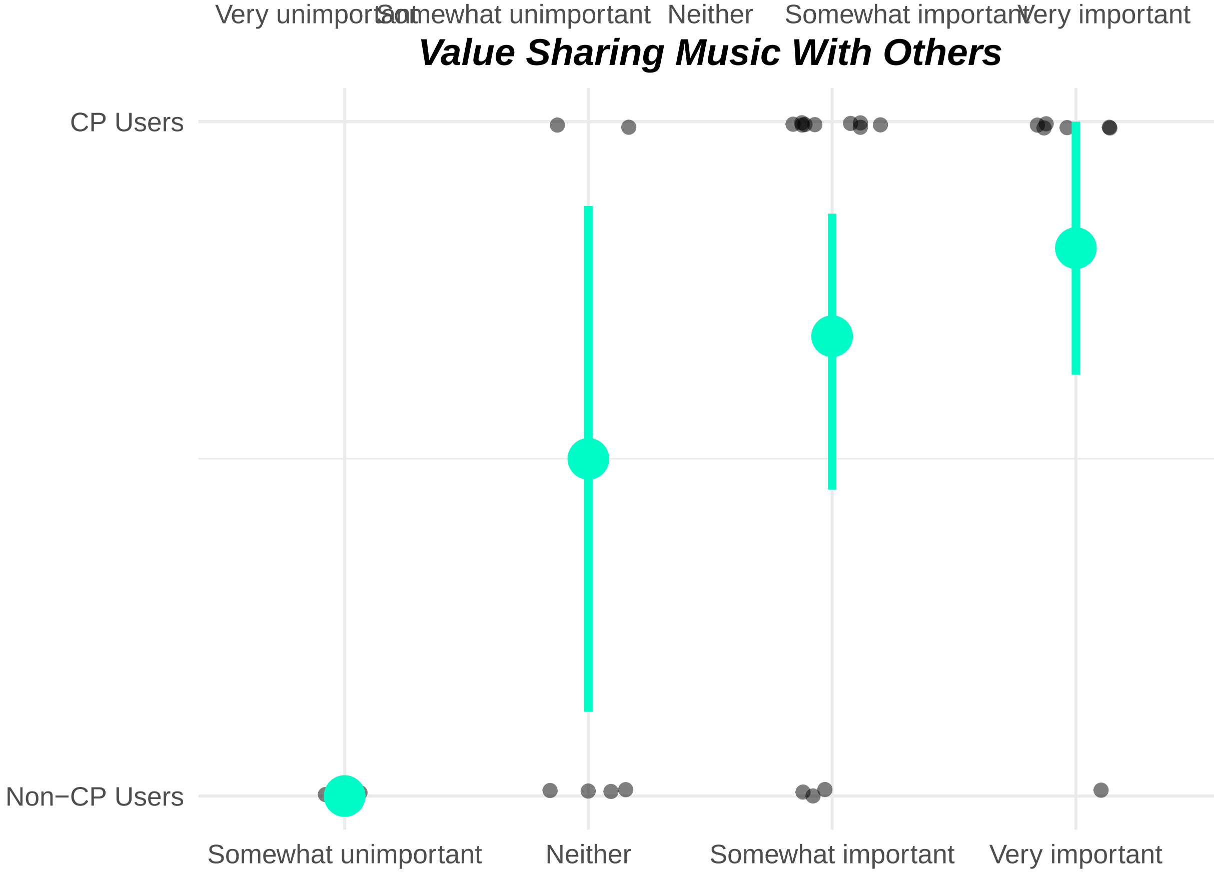
We hypothesize that users of CP have different values and habits around music compared to non-CP users; specifically around four areas: (1) social interaction, (2) music sharing, (3) music discovery, and (4) music taste.

3.3.1.1 H1a Value of Music for Social Interaction. Users of CP value listening and sharing music with others more than non-CP users.

```
df.impActivities <- df.raw %>%
  select(Q2.3, grep("Q7.4", names(df.raw))) %>%
  lapply(FUN = function(x) recode(x,
    "'Very important' = 2;
    'Somewhat important' = 1;
    'Neither important nor unimportant' = 0;
    'Somewhat unimportant' = -1;
```

```r
    'Very unimportant' = -2;
    'Yes' = 1;
    'No' = 0;")) %>%
  lapply(as.character) %>% lapply(as.numeric)

# plot value listening with others
ggplot(data = data.frame(df.impActivities),
       mapping = aes(y = (Q2.3),
                     x = as.factor(Q7.4_3))) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_y_continuous(limits = c(0, 1),labels = c("Non-CP Users",
                                  "CP Users"), breaks = c(0, 1)) +
  scale_x_discrete(breaks = c(-2, -1, 0, 1, 2),
                   labels = c("Very unimportant", "Somewhat unimportant", "Neither",
                              "Somewhat important", "Very important")) +
  theme_minimal(base_size = 12) +
  xlab("Value Listening to Music With Others") +
  ggtitle("Value Listening to Music With Others") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.x = element_blank(),
        axis.title.y = element_blank())

# plot value sharing with others
ggplot(data = data.frame(df.impActivities),
       mapping = aes(y = (Q2.3),
                     x = as.factor(Q7.4_4))) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_y_continuous(limits = c(0, 1),labels = c("Non-CP Users",
                                  "CP Users"), breaks = c(0, 1)) +
  scale_x_discrete(breaks = c(-2, -1, 0, 1, 2),
                   labels = c("Very unimportant", "Somewhat unimportant", "Neither",
                              "Somewhat important", "Very important")) +
  theme_minimal(base_size = 12) +
  xlab("Value Sharing Music With Others") +
  ggtitle("Value Sharing Music With Others") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.x = element_blank(),
        axis.title.y = element_blank())
```


Value Listening to Music With Others
CP Users
Non−CP Users
Very unimportant
Somewhat unimportant
Neither
Somewhat important
Very important


Value Sharing Music With Others
CP Users
Non−CP Users
Somewhat unimportant
Neither
Somewhat important
Very important

From plotting the data, we see that there is a slight trend associated with CP users and non-CP users. CP users seem to rate higher importance to sharing music with others. A trend is less clear when we look at the importance of listening to music with others.

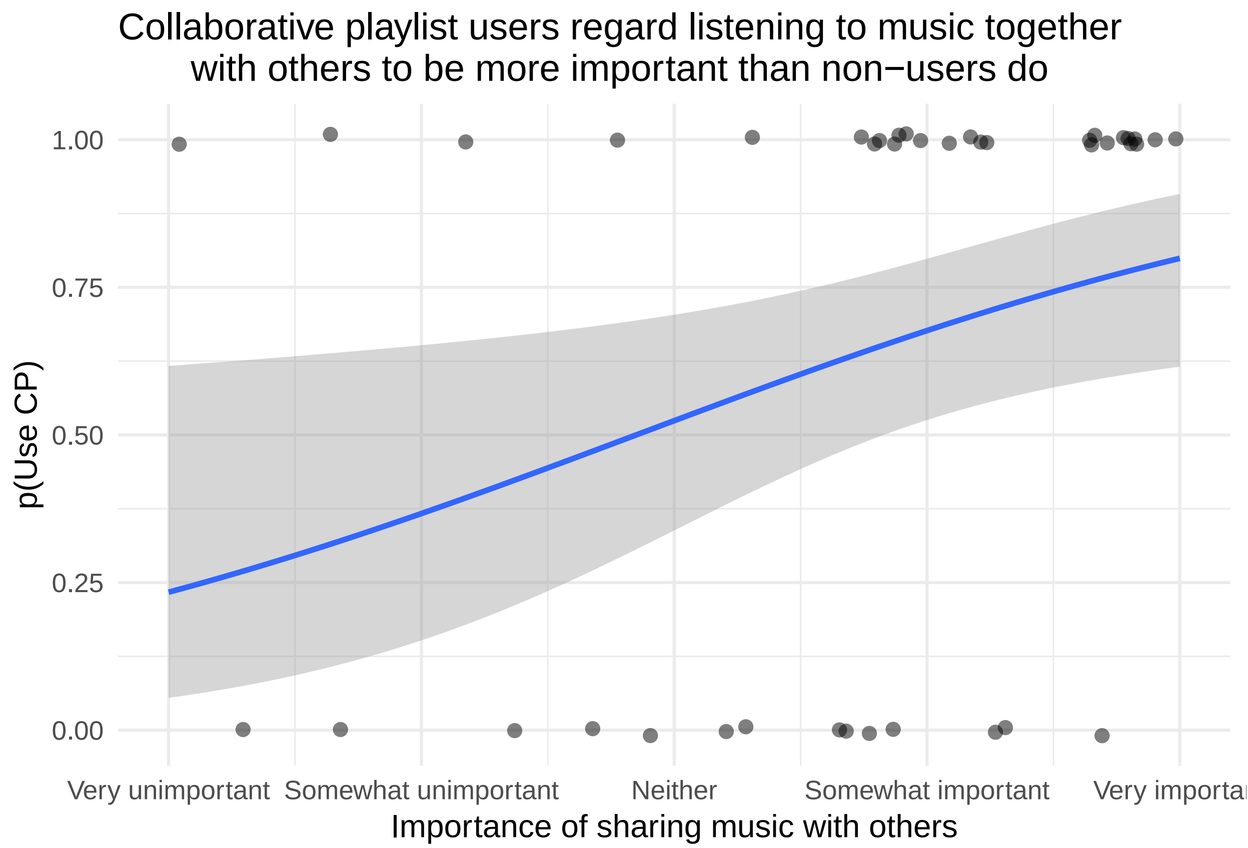
```
glm(Q2.3 ~ Q7.4_3,
    family = "binomial",
    data = df.impActivities) %>% summary()

glm(Q2.3 ~ Q7.4_4,
    family = "binomial",
    data = df.impActivities) %>% summary()

df.raw_binom <- df.raw %>%
  mutate_if(is.factor, as.character)
df.raw_binom[df.raw_binom == "Neither agree nor disagree"] <- "Neither"
df.raw_binom[df.raw_binom == "Strongly agree"] <- "Agree"
df.raw_binom[df.raw_binom == "Somewhat agree"] <- "Agree"
df.raw_binom[df.raw_binom == "Strongly disagree"] <- "Disagree"
df.raw_binom[df.raw_binom == "Somewhat disagree"] <- "Disagree"

df.raw_binom[df.raw_binom == "Neither important nor unimportant"] <- "Neither"
df.raw_binom[df.raw_binom == "Very important"] <- "Important"
df.raw_binom[df.raw_binom == "Somewhat important"] <- "Important"
df.raw_binom[df.raw_binom == "Very unimportant"] <- "Unimportant"
df.raw_binom[df.raw_binom == "Somewhat unimportant"] <- "Unimportant"

df.raw_binom[df.raw_binom == "At least once per day"] <- "Often"
df.raw_binom[df.raw_binom == "At least once per week"] <- "Often"
df.raw_binom[df.raw_binom == "At least once per month"] <- "Rare"
df.raw_binom[df.raw_binom == "A few times per year or less"] <- "Rare"
df.raw_binom[df.raw_binom == "Never"] <- "Never"

df.impActivities_binom <- df.raw_binom %>% select(Q2.3, grep("Q7.4", names(df.raw_binom)))
df.impActivities_binom[,1:7] <- lapply(df.impActivities_binom[,1:7],
                                       FUN = function(x) recode(x,
                                                                "'Neither' = 0;
                                                                'Important' = 1;
                                                                'Unimportant' = -1;
                                                                'Yes' = 1;
                                                                'No' = 0;"))

table.impActivities_binom_3 <- with(df.impActivities_binom, table(Q2.3, Q7.4_3))
table.impActivities_binom_4 <- with(df.impActivities_binom, table(Q2.3, Q7.4_4))
fisher.test (table.impActivities_binom_3)
fisher.test (table.impActivities_binom_4)
```

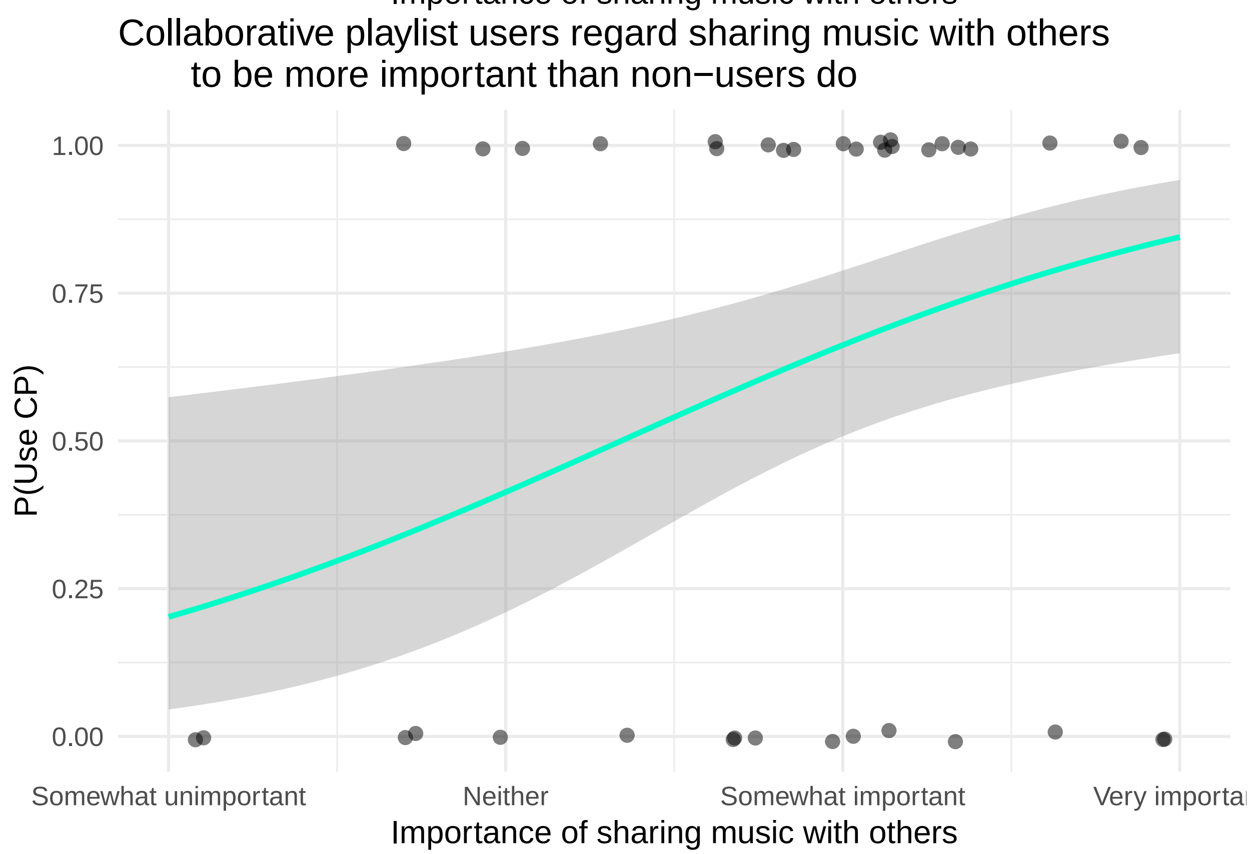
```
Call:
glm(formula = Q2.3 ~ Q7.4_3, family = "binomial", data = df.impActivities)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7916  -1.2187   0.6697   0.8838   1.7052
```

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.09652    0.39173   0.246    0.805
Q7.4_3       0.64209    0.28038   2.290    0.022 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 63.262  on 48  degrees of freedom
Residual deviance: 57.284  on 47  degrees of freedom
AIC: 61.284

Number of Fisher Scoring iterations: 4


Call:
glm(formula = Q2.3 ~ Q7.4_4, family = "binomial", data = df.impActivities)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9309  -1.0328   0.5804   0.9081   1.3292

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.3501     0.4975  -0.704    0.482
Q7.4_4        1.0229     0.4116   2.485    0.013 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 63.262  on 48  degrees of freedom
Residual deviance: 55.814  on 47  degrees of freedom
AIC: 59.814

Number of Fisher Scoring iterations: 4


    Fisher's Exact Test for Count Data

data:  table.impActivities_binom_3
p-value = 0.05247
alternative hypothesis: two.sided


    Fisher's Exact Test for Count Data

data:  table.impActivities_binom_4
p-value = 0.01616
alternative hypothesis: two.sided
```

```
# plot model value listening to music with others
ggplot(data = data.frame(df.impActivities),
```

```r
       mapping = aes(x = Q7.4_3,
                     y = Q2.3)) +
  geom_smooth(method = "glm",
              method.args = list(family = "binomial")) +
  geom_jitter(height = 0.01, size = 2, alpha = 0.5) +
  theme(axis.text.x = element_text(angle = 20)) +
  scale_x_continuous(limits = c(-2, 2),
                     breaks = c(-2, -1, 0, 1, 2),
                     labels = c("Very unimportant", "Somewhat unimportant", "Neither",
                                "Somewhat important", "Very important")) +
  theme_minimal(base_size = 12) +
  labs(title = "Collaborative playlist users regard listening to music together
       with others to be more important than non-users do",
       x = "Importance of sharing music with others",
       y = "p(Use CP)")

# plot model value sharing music with others
ggplot(data = data.frame(df.impActivities),
       mapping = aes(x = Q7.4_4,
                     y = Q2.3)) +
  geom_smooth(method = "glm",
              method.args = list(family = "binomial"), color = "#00fdc8") +
  geom_jitter(height = 0.01, size = 2, alpha = 0.5) +
  theme(axis.text.x = element_text(angle = 20)) +
  scale_x_continuous(limits = c(-1, 2),
                     breaks = c(-1, 0, 1, 2),
                     labels = c("Somewhat unimportant", "Neither",
                                "Somewhat important", "Very important")) +
  scale_color_manual(values="#00fdc8") +
  labs(title = "Collaborative playlist users regard sharing music with others
       to be more important than non-users do",
       x = "Importance of sharing music with others",
       y = "P(Use CP)") +
  theme_minimal(base_size = 12)
```


Collaborative playlist users regard listening to music together with others to be more important than non−users do
p(Use CP)
1.00
0.75
0.50
0.25
0.00
Very unimportant
Somewhat unimportant
Neither
Somewhat important
Very importa
Importance of sharing music with others

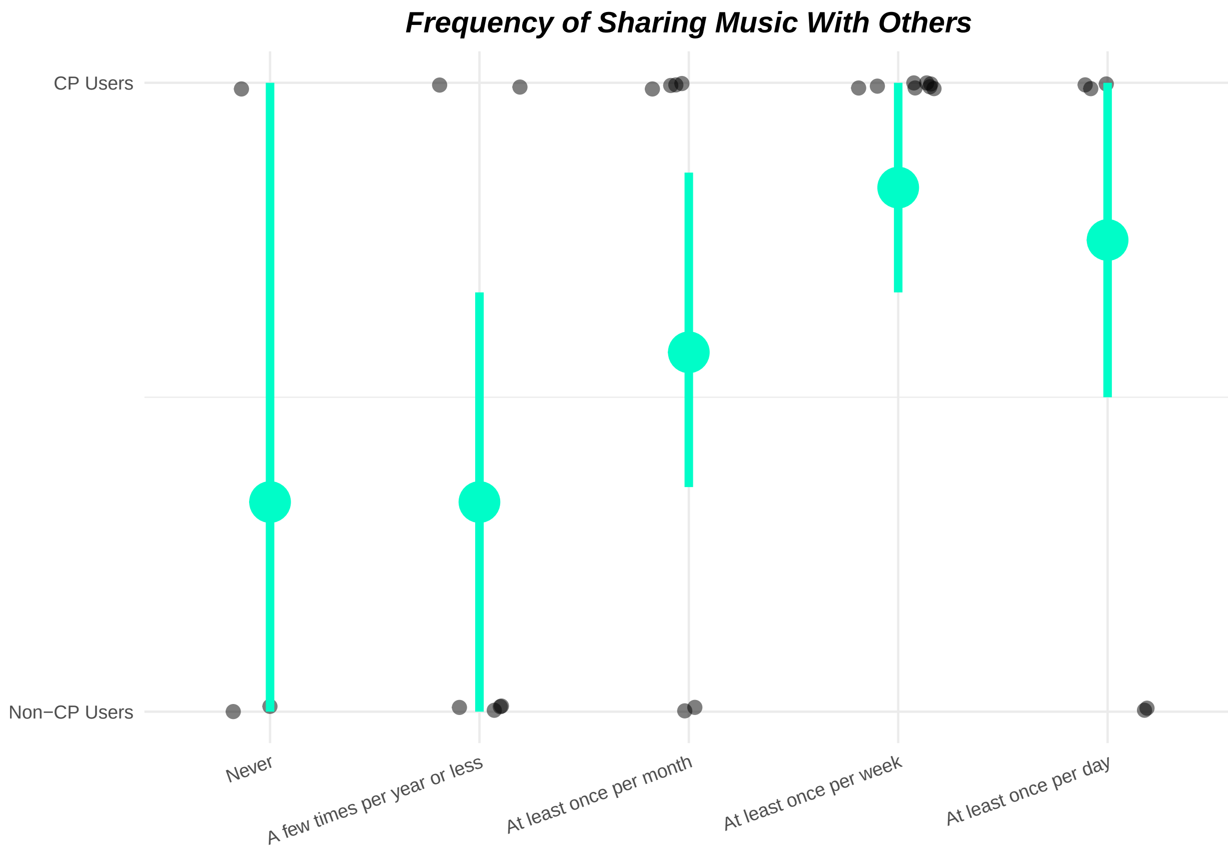

Collaborative playlist users regard sharing music with others to be more important than non−users do
P(Use CP)
1.00
0.75
0.50
0.25
0.00
Somewhat unimportant
Neither
Somewhat important
Very importa
Importance of sharing music with others

Logistic regressions on whether use of CP can be predicted by how much they value listening to and sharing music with others show that there is a significant association ($p < 0.05$ for both). Plots show that CP users rate higher importance to these two social activities. A Fisher's Exact tests for both activities also verifies the statistical significance ($p \sim 0.05$ and $p < 0.05$ respectively).

3.3.1.2 H1b Music Sharing: Users of CP frequently engage in sharing music with others.

Following H1a, we believed that those who value an activity would frequently engage in that activity more. In other words, value towards sharing music with others would translate to more frequent sharing of music.

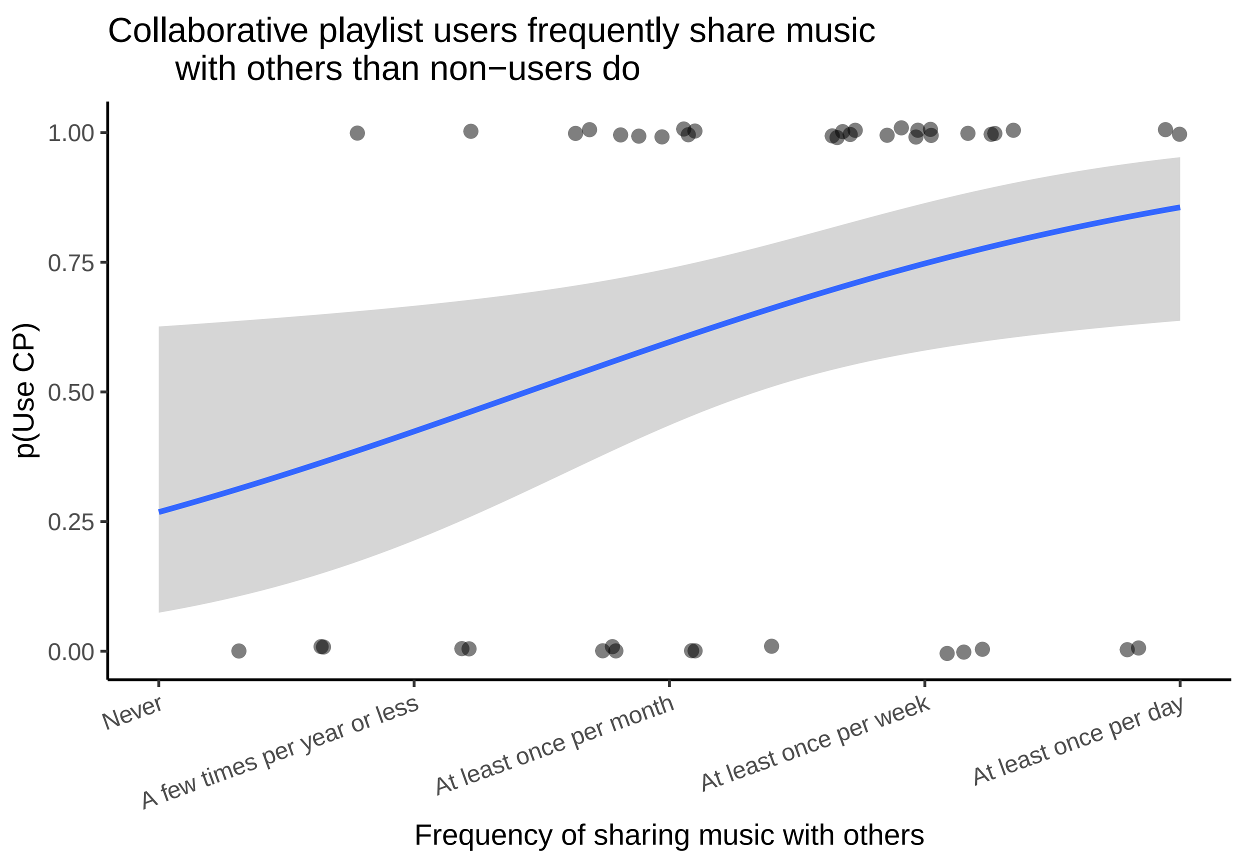
```
df.musicFreq_2 <- df.raw %>%
  select(Q2.3, grep("Q7.6", names(df.raw))) %>%
  select(-Q7.6_7) %>%
  lapply(FUN = function(x) recode(x,
    "'Never' = 0;
    'A few times per year or less' = 1;
    'At least once per month' = 2;
    'At least once per week' = 3;
    'At least once per day' = 4;
    'Yes' = 1;
    'No' = 0;")) %>%
  lapply(as.character) %>%
  lapply(as.numeric)

ggplot(data = data.frame(df.musicFreq_2),
       mapping = aes(y = (Q2.3),
                     x = as.factor(Q7.6_5))) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_y_continuous(limits = c(0, 1),labels = c("Non-CP Users",
                                   "CP Users"), breaks = c(0, 1)) +
  scale_x_discrete(labels = c("Never", "A few times per year or less", "At least once per month",
                              "At least once per week", "At least once per day")) +
  theme_minimal(base_size = 9) +
  xlab("Frequency of Sharing Music With Others") +
  ggtitle("Frequency of Sharing Music With Others") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_text(angle = 20, hjust = 1))
```



Looking at the plot without the first two frequencies of "Never" and "A few times per year or less" (as we have 1 or 2 data points to represent these answer choices), we see there's a slight trend. CP users seem to more frequently share music.

```
glm(formula = Q2.3 ~ 1 + Q7.6_5,
    family = "binomial",
    data = df.musicFreq_2) %>% summary()

ggplot(data = data.frame(df.musicFreq_2),
       mapping = aes(x = Q7.6_5,
                     y = Q2.3)) +
  geom_smooth(method = "glm",
              method.args = list(family = "binomial")) +
  geom_jitter(height = 0.01, size = 2, alpha = 0.5) +
  theme(axis.text.x = element_text(angle = 20, hjust = 1)) +
  scale_x_continuous(limits = c(0, 4),
                     breaks = c(0, 1, 2, 3, 4),
                     labels = c("Never", "A few times per year or less",
                                "At least once per month",
                                "At least once per week", "At least once per day")) +
  labs(title = "Collaborative playlist users frequently share music
       with others than non-users do",
       x = "Frequency of sharing music with others",
       y = "p(Use CP)")
```

```
Call:
glm(formula = Q2.3 ~ 1 + Q7.6_5, family = "binomial", data = df.musicFreq_2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9687  -1.0502   0.7627   0.7627   1.6221

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.0032     0.7750  -1.294   0.1955
Q7.6_5        0.6964     0.3100   2.246   0.0247 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 63.262  on 48  degrees of freedom
Residual deviance: 57.487  on 47  degrees of freedom
AIC: 61.487

Number of Fisher Scoring iterations: 4
```

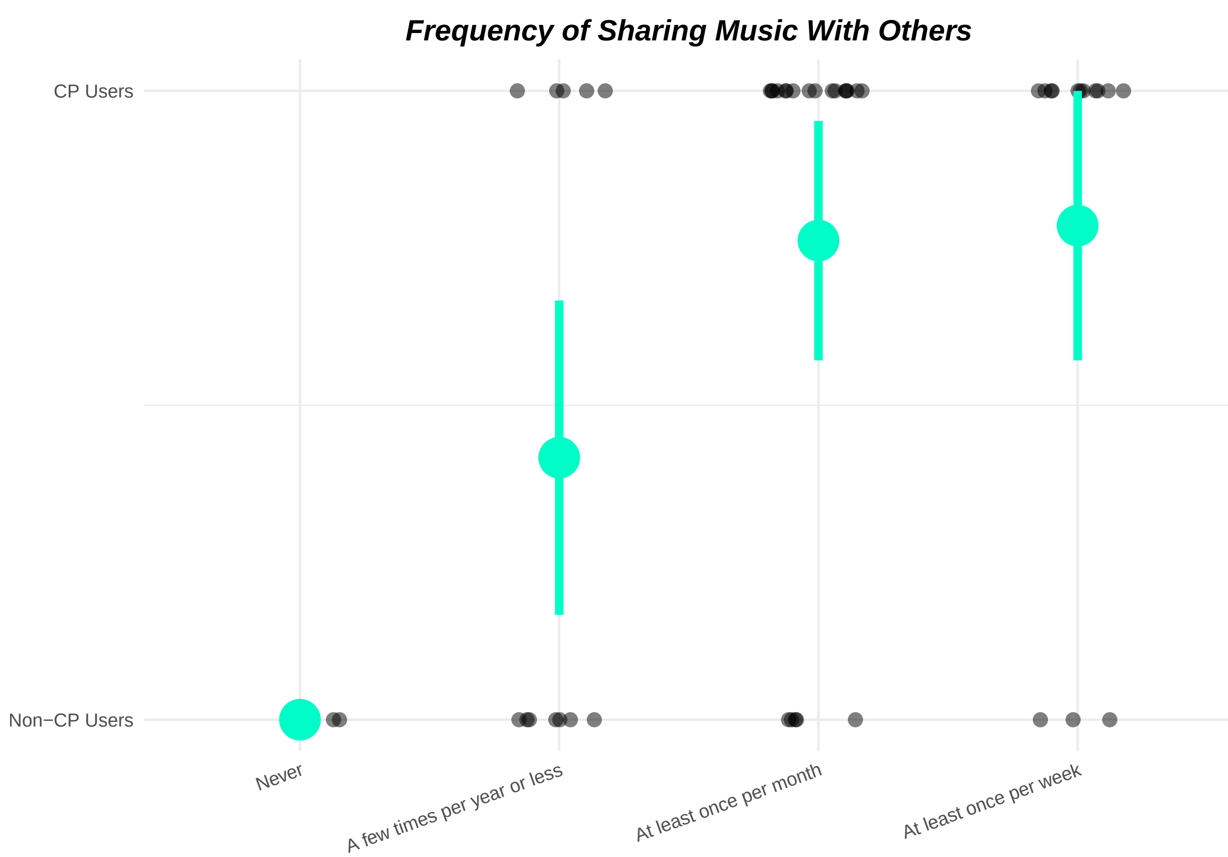


Results show that users of CP more frequently share music with others compared to non-CP users ($p < 0.05$), whether it is through the digital platform of collaborative playlists or through other platforms for sharing music links (e.g., Facebook, Twitter).

3.3.1.3 H1c Music Discovery: Users of CP more frequently engage in discovering new music compared to non-CP users.

We hypothesized that another reason why users might participate in CP was because of their interest in discovering new music.

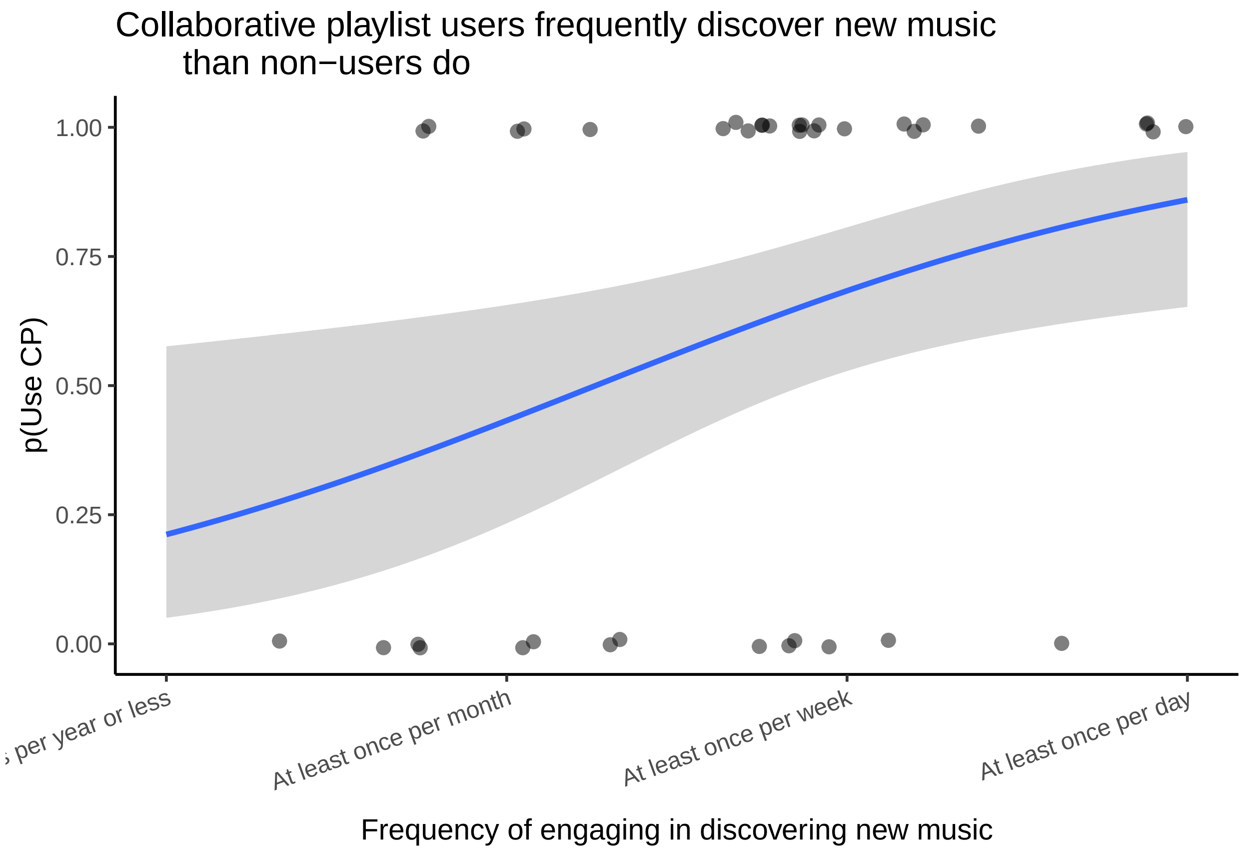
```
ggplot(data = data.frame(df.musicFreq_2),
       mapping = aes(y = (Q2.3),
                     x = as.factor(Q7.6_2))) +
  geom_jitter(height = 0, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_y_continuous(limits = c(0, 1),labels = c("Non-CP Users",
                                  "CP Users"), breaks = c(0, 1)) +
  scale_x_discrete(labels = c("Never", "A few times per year or less",
                              "At least once per month",
                                "At least once per week", "At least once per day")) +
  theme_minimal(base_size = 9) +
  xlab("Frequency of Sharing Music With Others") +
  ggtitle("Frequency of Sharing Music With Others") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_text(angle = 20, hjust = 1))
```



From plotting the data, we see there is a trend associated between use of CP and frequency of music discovery. CP users appear to more frequently engage in music discovery.

```r
glm(formula = Q2.3 ~ 1 + Q7.6_2,
    family = "binomial",
    data = df.musicFreq_2) %>% summary()

ggplot(data = data.frame(df.musicFreq_2),
       mapping = aes(x = Q7.6_2,
                     y = Q2.3)) +
  geom_smooth(method = "glm",
              method.args = list(family = "binomial")) +
  geom_jitter(height = 0.01, size = 2, alpha = 0.5) +
  theme(axis.text.x = element_text(angle = 20)) +
  scale_x_continuous(limits = c(1, 4),
                     breaks = c(1, 2, 3, 4),
                     labels = c("A few times per year or less",
                                "At least once per month",
                                "At least once per week", "At least once per day")) +
  scale_color_manual(values="#00fdc8") +
  labs(title = "Collaborative playlist users frequently discover new music
       than non-users do",
       x = "Frequency of engaging in discovering new music",
       y = "p(Use CP)") +
  theme(axis.text.x = element_text(angle = 20, hjust = 1))
```

```
Call:
glm(formula = Q2.3 ~ 1 + Q7.6_2, family = "binomial", data = df.musicFreq_2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9817  -1.0642   0.5500   0.8724   1.2950

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.3569     1.2279  -1.919   0.0549 .
Q7.6_2        1.0423     0.4233   2.462   0.0138 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 63.262  on 48  degrees of freedom
Residual deviance: 56.054  on 47  degrees of freedom
AIC: 60.054

Number of Fisher Scoring iterations: 3
```



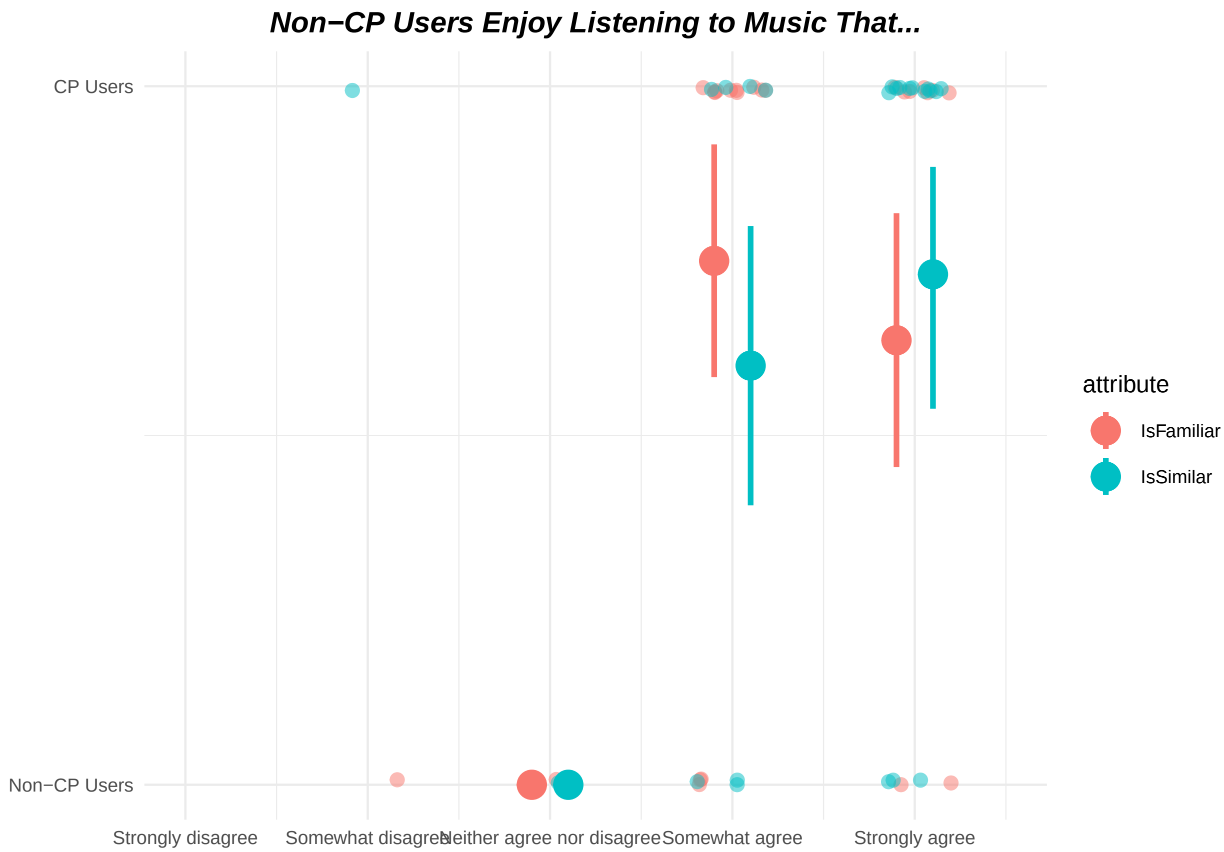
A logistic regression shows that frequency of music discovery is significantly associated with use of CP ($p < 0.05$). CP users appear to more frequently engage in discovering new music compared to non-CP users.

3.3.1.4 H1d Music Taste Diversity: non-CP users prefer to listen to familiar music.

We hypothesized that non-CP users might not engage in CP because they prefer listening to familiar music.

```
# How often do you engage in the following activities... [collaborative playlist]?
musicTaste_Q7_10 = df.raw_all %>%
  mutate(id = row_number()) %>%
  select(c('id','Q2.3','Q7.10_1','Q7.10_2','Q7.10_3','Q7.10_9','Q7.10_8',
           'Q7.10_10','Q7.10_11', 'Q7.10_13')) %>%
  setnames("Q2.3", "UseCP") %>%
  setnames("Q7.10_1", "SuggestedByKnown") %>%
  setnames("Q7.10_2", "SuggestedByStrangers") %>%
  setnames("Q7.10_3", "IsRecommendedAI") %>%
  setnames("Q7.10_8", "EvokesMemories") %>%
  setnames("Q7.10_9", "IsFamiliar") %>%
  setnames("Q7.10_10", "IsNew") %>%
  setnames("Q7.10_11", "IsSimilar") %>%
  setnames("Q7.10_13", "FeelsCollaborative")

musicTaste_Q7_10[,2:10] <- lapply( musicTaste_Q7_10[,2:10],
                                   FUN = function(x) recode(x,
                                   "'Yes' = 1;
```

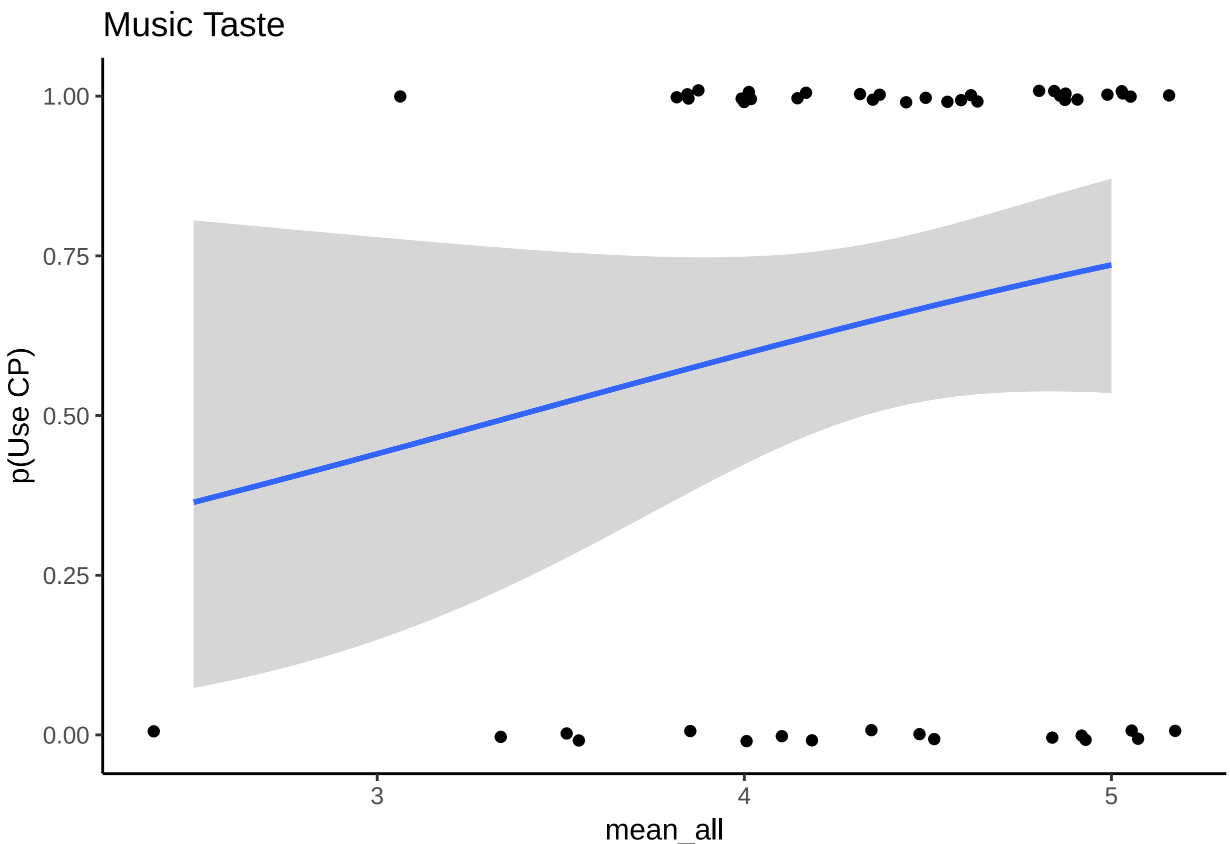
```
                                    'No' = 0;
                                    'Strongly agree' = 5;
                                    'Somewhat agree' = 4;
                                    'Neither agree nor disagree' = 3;
                                    'Somewhat disagree' = 2;
                                    'Strongly disagree' = 1;") )
musicTaste_Q7_10[,2:10] <- lapply( musicTaste_Q7_10[,2:10],
                                   FUN = function(x) as.numeric(as.character(x)))

musicTaste_Q7_10 = musicTaste_Q7_10 %>%
  rowwise() %>%
  mutate(mean_all = mean(c(IsFamiliar, IsSimilar)))

ggplot(data = data.frame(musicTaste_Q7_10 %>% gather(attribute, response, c(6,9))),
       mapping = aes(y = UseCP,
                     x = response,
                     color = attribute)) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", size = 1,
               position = position_dodge(width = 0.4)) +
  scale_y_continuous(limits = c(0, 1),labels = c("Non-CP Users",
                                  "CP Users"), breaks = c(0, 1)) +
  scale_x_continuous(labels = c("Strongly disagree", "Somewhat disagree",
                              "Neither agree nor disagree", "Somewhat agree",
                              "Strongly agree"),
                   breaks = c(1,2,3,4,5),
                   limits = c(1,5.5)) +
  theme_minimal(base_size = 9) +
  xlab("Non-CP Users Enjoy Listening to Music That...") +
  ggtitle("Non-CP Users Enjoy Listening to Music That...") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.x = element_blank(),
        axis.title.y = element_blank())
```




```
fit.glm = glm(formula = UseCP ~ 1 + mean_all,
              family = "binomial",
              data = musicTaste_Q7_10)
fit.glm %>% summary()

ggplot(data = musicTaste_Q7_10,
       mapping = aes(x = mean_all,
                     y = UseCP)) +
  geom_smooth(method = "glm",
              method.args = list(family = "binomial")) +
  geom_jitter(height = 0.01) +
  ggtitle("Music Taste") +
  labs(y = "p(Use CP)")
```

```
Call:
glm(formula = UseCP ~ 1 + mean_all, family = "binomial", data = musicTaste_Q7_10)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6317  -1.3477   0.7833   0.8950   1.2812

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.1383     2.2767  -0.939    0.348
mean_all      0.6325     0.5176   1.222    0.222
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 63.262  on 48  degrees of freedom
Residual deviance: 61.725  on 47  degrees of freedom
AIC: 65.725

Number of Fisher Scoring iterations: 4
```



The model shows that there is no statistically significant difference in preference for familiar or similar music between CP and non-CP users ($p > 0.2$).

3.3.2 Q2 Do patterns of user behavior change depending on a collaborator's level of engagement?

For our second part of the survey analysis, we chose to focus on users of CP. We were interested in understanding the relationship between user engagement and sense of ownership. We address "engagement" with two factors: perceived role (perception) and frequency of contribution (behavior).

3.3.2.1 H2a Playlist ownership can be measured by a user's perceived role and frequency of contribution

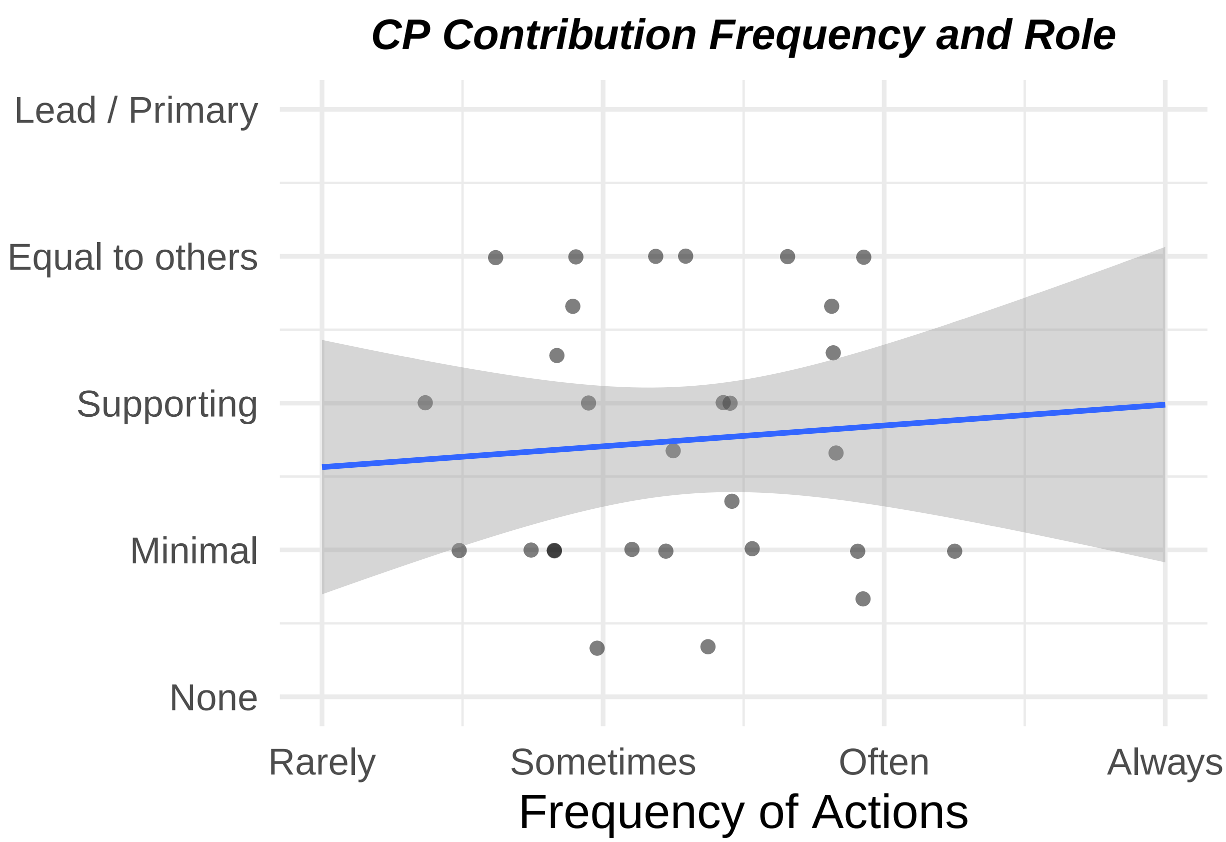
```
# Q5.20 Ownership
df.ownershiP_Q5_20 = df.raw_yes %>%  select(starts_with('Q5.20')) %>%
        mutate(id = row_number()) %>%
```

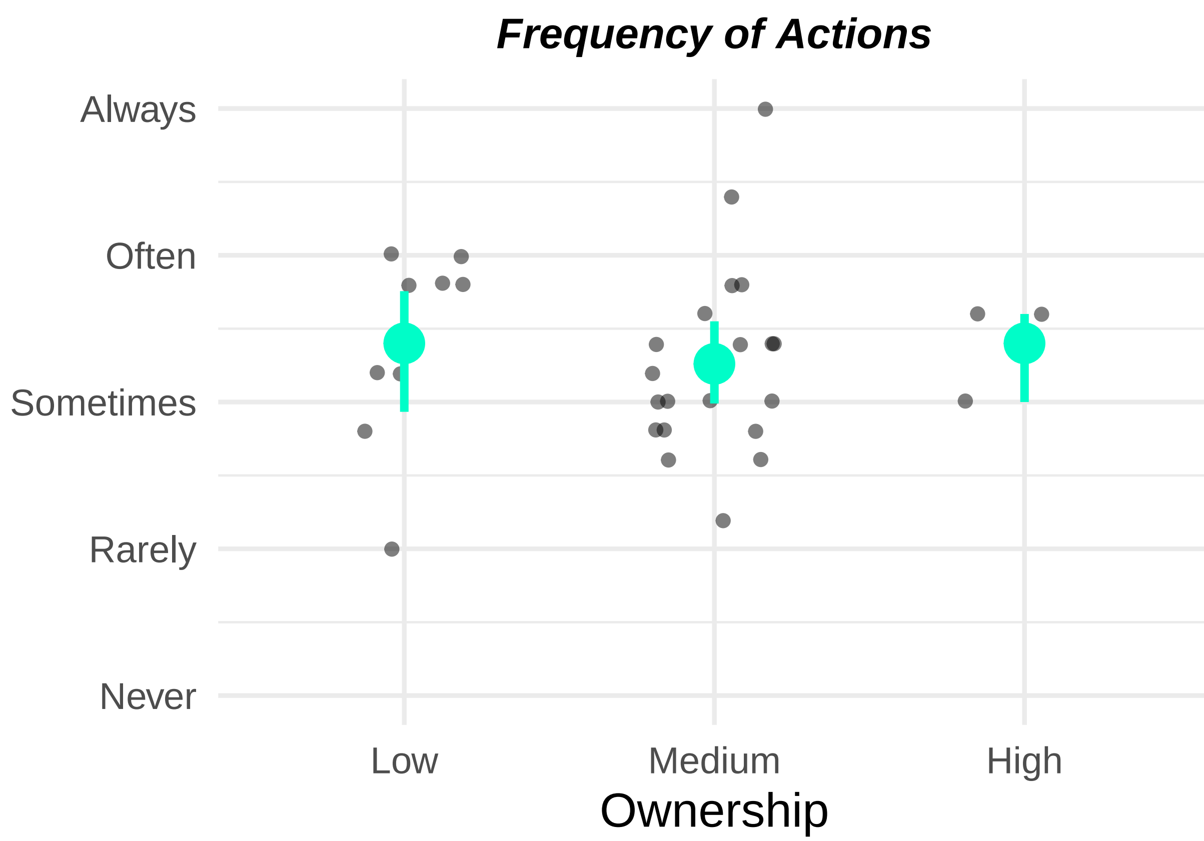
```
        select(id, 'Q5.20') %>%
        setnames("Q5.20", "ownership")
df.ownershiP_Q5_20$ownership<-recode(df.ownershiP_Q5_20$ownership,
                      "'The playlist belongs to me.' = 1;
                       'The playlist belongs to me and my collaborator(s).' = 0;
                       'The playlist belongs to my collaborator(s).' = -1;
                       else = 0")
df.ownershiP_Q5_20$ownership <- as.numeric(as.character(df.ownershiP_Q5_20$ownership))


# Q5.16 Frequency
df.frequency_Q5_16 = df.raw_yes %>%  select(starts_with('Q5.16_')) %>%
  mutate(id = row_number()) %>%
  setnames("Q5.16_1", "check") %>%
  setnames("Q5.16_2", "listen_playAlone") %>%
  setnames("Q5.16_3", "listen_playOthers") %>%
  setnames("Q5.16_4", "contribute") %>%
  setnames("Q5.16_5", "share") %>%
  select(-'Q5.16_6_TEXT') %>%
  select(-'Q5.16_6')

df.frequency_Q5_16[,1:5] <- lapply( df.frequency_Q5_16[,1:5],
                                   FUN = function(x) recode(x,
                                   "'Always' = 5;
                                    'Often' = 4;
                                    'Sometimes' = 3;
                                    'Rarely' = 2;
                                    'Never' = 1;") )
df.frequency_Q5_16[,1:5] <- lapply( df.frequency_Q5_16[,1:5],
                                   FUN = function(x) as.numeric(as.character(x)))
# add grand mean for frequency
df.frequency_Q5_16 = df.frequency_Q5_16 %>%
  mutate('mean_16' = rowMeans(df.frequency_Q5_16[,1:5]))


df.role_Q5_18 = df.raw_yes %>%  select(starts_with('Q5.18_')) %>%
  mutate(id = row_number()) %>%
  setnames("Q5.18_1", "add") %>%
  setnames("Q5.18_2", "delete") %>%
  setnames("Q5.18_3", "reorder") %>%
  select(-starts_with('Q5.18_4'))

# Q5.18 Role
df.role_Q5_18[,1:3] <- lapply( df.role_Q5_18[,1:3],
                                   FUN = function(x) recode(x,
                                   "'Lead / Primary' = 5;
                                    'Equal to others' = 4;
                                    'Supporting' = 3;
                                    'Minimal' = 2;
                                    'None' = 1;") )
df.role_Q5_18[,1:3] <- lapply( df.role_Q5_18[,1:3],
                                   FUN = function(x) as.numeric(as.character(x)))
# add grand mean for role
```

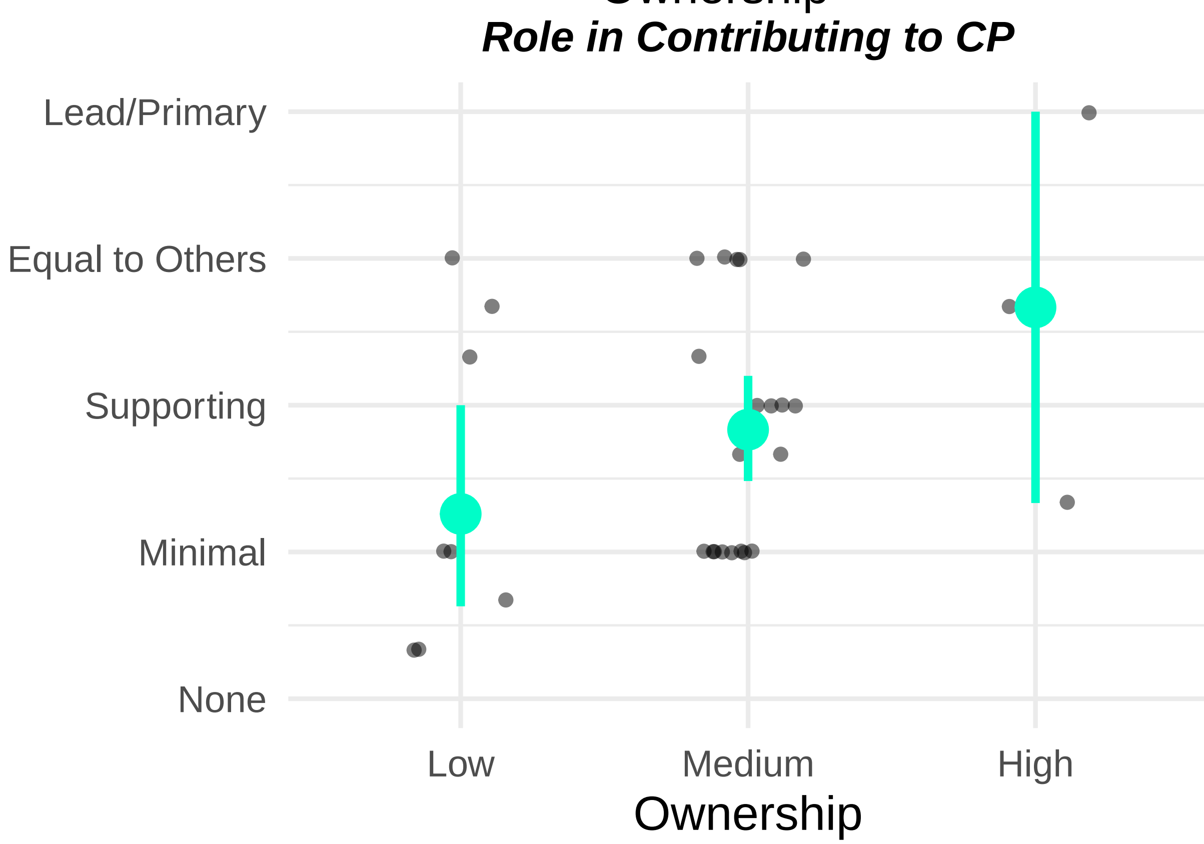
```r
df.role_Q5_18 = df.role_Q5_18 %>%
  mutate('mean_18'= rowMeans(df.role_Q5_18[,1:3]))

df.Q2 = full_join(df.frequency_Q5_16, df.ownershiP_Q5_20, by = "id")
df.Q2 = full_join(df.Q2, df.role_Q5_18, by = "id")
```

```r
# scatter plot
ggplot(data = df.Q2,
       mapping = aes(x = mean_16,
                     y = mean_18)) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  scale_y_continuous(limits = c(1, 5),
                     breaks = c(1, 2, 3, 4, 5),
                     labels = c("None", "Minimal", "Supporting",
                                "Equal to others", "Lead / Primary")) +
  scale_x_continuous(limits = c(1, 4),
                     breaks = c(1, 2, 3, 4),
                     labels = c("Rarely", "Sometimes",
                                "Often", "Always")) +
  geom_smooth(method = "lm") +
  theme_minimal(base_size = 18) +
  xlab("Frequency of Actions") +
  ylab("Role in Contributing to CP") +
  ggtitle("CP Contribution Frequency and Role") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5, size = 16),
        axis.title.y = element_blank())

# frequency
ggplot(data = df.Q2,
       mapping = aes(x = factor(ownership),
                     y = as.numeric(mean_16))) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_x_discrete(labels = c("High", "Medium", "Low"),
                          breaks = c(1,0,-1)) +
  scale_y_continuous(limits = c(0, 4),
                     breaks = c(0, 1, 2, 3, 4),
                     labels = c("Never", "Rarely", "Sometimes",
                                "Often", "Always")) +
  theme_minimal(base_size = 18) +
  xlab("Ownership") +
  ggtitle("Frequency of Actions") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5, size = 16),
        axis.title.y = element_blank())

# role
ggplot(data = df.Q2,
       mapping = aes(x = factor(ownership),
                     y = as.numeric(mean_18))) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  stat_summary(fun.data = "mean_cl_boot", colour = "#00fdc8", size = 1.5) +
  scale_x_discrete(labels = c("High", "Medium", "Low"),
                   breaks = c(1, 0, -1)) +
```

```
scale_y_continuous(limits = c(1, 5),
                   breaks = c(1, 2, 3, 4, 5),
                   labels = c("None", "Minimal", "Supporting",
                              "Equal to Others", "Lead/Primary")) +
theme_minimal(base_size = 18) +
xlab("Ownership") +
ggtitle("Role in Contributing to CP") +
theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5, size = 16),
      axis.title.y = element_blank())
```




Frequency of Actions
Always
Often
Sometimes
Rarely
Never
Low
Medium
High
Ownership


Role in Contributing to CP
Lead/Primary
Equal to Others
Supporting
Minimal
None
Low
Medium
High
Ownership

```r
lm.compact = lm(ownership ~ 1 + mean_18,
                data = df.Q2)
summary(lm.compact) %>% print()

lm.augmented = lm(ownership ~ 1 + mean_16*mean_18,
                data = df.Q2)
summary(lm.augmented) %>% print()
anova(lm.augmented)

# store the model information
df.compact_summary = tidy(lm.compact)
df.augmented_summary = tidy(lm.augmented)

# create a data frame that contains the residuals
df.compact_model = augment(lm.compact) %>%
  clean_names()
df.augmented_model = augment(lm.augmented) %>%
  clean_names()

# Calculate SSE for both models
sse_compact = df.compact_model %>%
  summarize(SSE = sum(resid^2))
sse_augmented = df.augmented_model %>%
  summarize(SSE = sum(resid^2))

pc = 1 # compact model parameter is just the mean_16 so 1 parameter
pa = 2 # augmented model includes mean_16 + mean_18
n = count(df.Q2) %>% as.numeric() # number of observations

# Proportional reduction of error
pre = as.numeric(1 - (sse_augmented/sse_compact)) %>%  print()

# F-statistic
f = (pre/(pa-pc))/((1-pre)/(n-pa))

# p-value
p_value = 1 - pf(f, df1 = pa-pc, df2 = n-pa)

# check if it's worth it
anova(lm.compact, lm.augmented) %>%  print()

#### P-value for role
anova(lm.compact)
```

```
Call:
lm(formula = ownership ~ 1 + mean_18, data = df.Q2)

Residuals:
    Min      1Q  Median      3Q     Max
-1.1004 -0.4286  0.1299  0.3603  1.2835

Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.82098    0.29043  -2.827  0.00829 **
mean_18      0.23036    0.09941   2.317  0.02750 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5545 on 30 degrees of freedom
Multiple R-squared:  0.1518,	Adjusted R-squared:  0.1235
F-statistic: 5.369 on 1 and 30 DF,  p-value: 0.0275


Call:
lm(formula = ownership ~ 1 + mean_16 * mean_18, data = df.Q2)

Residuals:
     Min       1Q   Median       3Q      Max
-0.81750 -0.42915  0.05669  0.33246  1.28662

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.1995     0.9934  -2.214   0.0351 *
mean_16           0.6168     0.4291   1.438   0.1616
mean_18           0.9125     0.4021   2.270   0.0311 *
mean_16:mean_18  -0.3000     0.1726  -1.738   0.0931 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5432 on 28 degrees of freedom
Multiple R-squared:  0.2403,	Adjusted R-squared:  0.1589
F-statistic: 2.952 on 3 and 28 DF,  p-value: 0.04972

Analysis of Variance Table

Response: ownership
                Df Sum Sq Mean Sq F value  Pr(>F)
mean_16          1 0.0230 0.02298  0.0779 0.78225
mean_18          1 1.6985 1.69850  5.7564 0.02333 *
mean_16:mean_18  1 0.8917 0.89170  3.0220 0.09313 .
Residuals       28 8.2618 0.29506
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
[1] 0.1043233
Analysis of Variance Table

Model 1: ownership ~ 1 + mean_18
Model 2: ownership ~ 1 + mean_16 * mean_18
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     30 9.2241
2     28 8.2618  2   0.96229 1.6306 0.2139
Analysis of Variance Table

Response: ownership
          Df Sum Sq Mean Sq F value Pr(>F)
mean_18    1 1.6509 1.65089  5.3693 0.0275 *
```

```
Residuals 30 9.2241 0.30747
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results show that including the person's perceived role leads to a 10.4323296 percent reduction in the model error. Using an alpha level of 0.05, we can conclude that a user's perceived role significantly affects a user's sense of playlist ownership compared to a person's frequency of contributions to the playlist, p = 0.03848.

```r
df.Q2lmer <- df.Q2 %>%
  select(id, mean_16, mean_18) %>%
  gather(question, response, 2:3)
lmem.Q2 <- lmer(response ~ question + (1|id), data = df.Q2lmer)
lmem.Q2 %>% summary()

AIC(lm.compact, lm.augmented, lmem.Q2)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: response ~ question + (1 | id)
   Data: df.Q2lmer

REML criterion at convergence: 160.2

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.90700 -0.80815 -0.03732  0.58696  2.57976

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 0.05484  0.2342
 Residual             0.64042  0.8003
Number of obs: 64, groups:  id, 32

Fixed effects:
                Estimate Std. Error      df t value Pr(>|t|)
(Intercept)       2.3125     0.1474 61.6167  15.689   <2e-16 ***
questionmean_18   0.4375     0.2001 31.0000   2.187   0.0364 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr)
questnmn_18 -0.679
             df       AIC
lm.compact    3  57.00677
lm.augmented  5  57.48115
lmem.Q2       4 168.15127
```

We also compared the model to a linear mixed effects model (LMEM) with random intercepts for participants. Results show that the compact linear model is still a better model with an AIC of 57 compared to and AIC of 168 for the LMEM model.

3.3.2.2 H2b Individual actions and ownership

```
lm.individualRole = lm(ownership ~ 1 + add + delete + reorder,
                   data = df.Q2)
summary(lm.individualRole) %>% print ()
anova(lm.individualRole)
```

```

Call:
lm(formula = ownership ~ 1 + add + delete + reorder, data = df.Q2)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.9675 -0.1162  0.1008  0.1698  0.8788 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.54802    0.38424  -4.029 0.000389 ***
add          0.33310    0.10405   3.201 0.003394 ** 
delete       0.09004    0.07592   1.186 0.245603    
reorder     -0.04427    0.07226  -0.613 0.545016    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5018 on 28 degrees of freedom
Multiple R-squared:  0.3518,	Adjusted R-squared:  0.2823 
F-statistic: 5.064 on 3 and 28 DF,  p-value: 0.006295

Analysis of Variance Table

Response: ownership
          Df Sum Sq Mean Sq F value    Pr(>F)    
add        1 3.4712  3.4712 13.7867 0.0009019 ***
delete     1 0.2596  0.2596  1.0311 0.3185890    
reorder    1 0.0945  0.0945  0.3754 0.5450156    
Residuals 28 7.0497  0.2518                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

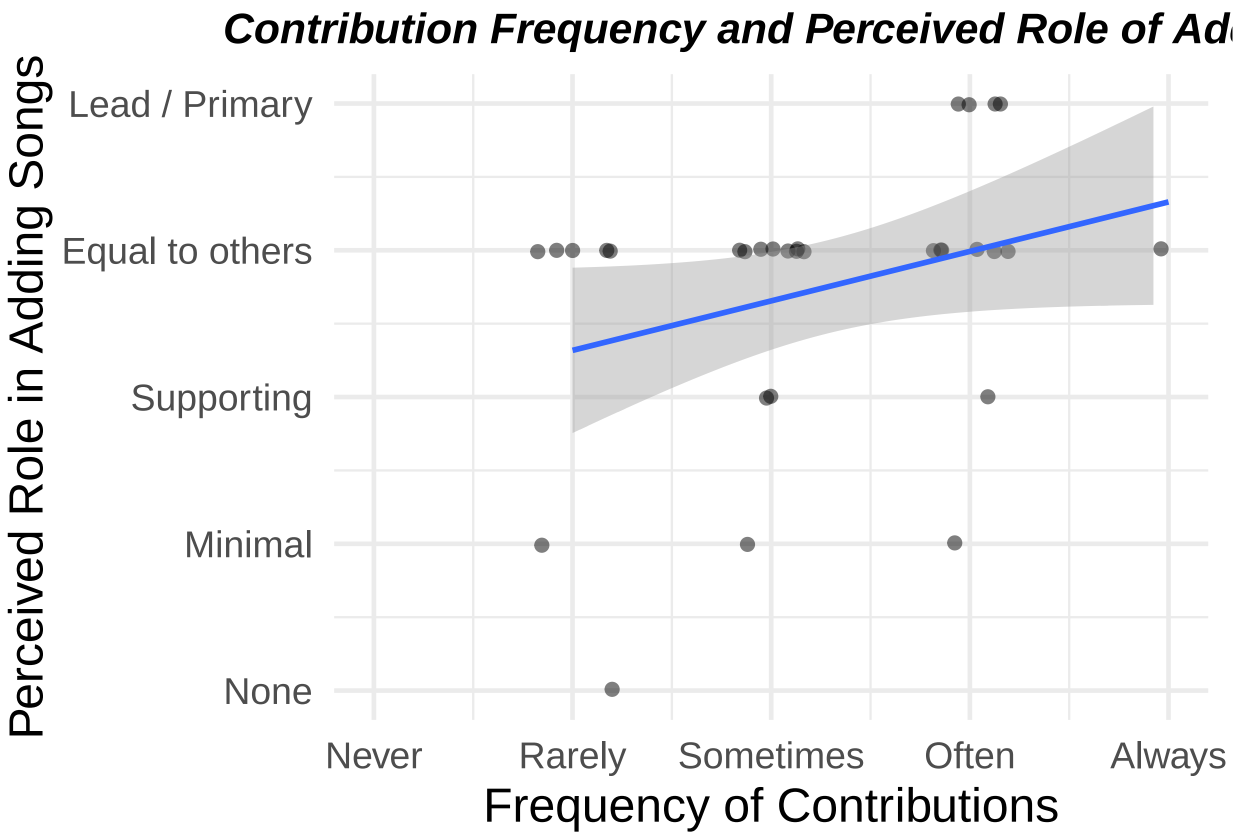
Results show that one's role in adding music is indicative of ownership that they feel towards the collaborative playlist (mine/our's vs their's) ($p < 0.05$).

The scatter plot below of contribution frequency of adding, deleting, and reordering songs in the CP with adding songs does not show any linear relationship between the two.

```
ggplot(data = df.Q2,
       mapping = aes(x = contribute,
                     y = add)) +
  geom_jitter(height = 0.01, width = 0.2, size = 2, alpha = 0.5) +
  scale_y_continuous(limits = c(1, 5),
                     breaks = c(1, 2, 3, 4, 5),
                     labels = c("None", "Minimal", "Supporting",
                                "Equal to others", "Lead / Primary")) +
```

```r
  scale_x_continuous(limits = c(0, 4),
                     breaks = c(0, 1, 2, 3, 4),
                     labels = c("Never", "Rarely", "Sometimes",
                                "Often", "Always")) +
  geom_smooth(method = "lm") +
  theme_minimal(base_size = 18) +
  xlab("Frequency of Contributions") +
  ylab("Perceived Role in Adding Songs") +
  ggtitle("Contribution Frequency and Perceived Role of Adding") +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5, size = 16))

lm(contribute ~ add, df.Q2) %>% summary()
```

```
Call:
lm(formula = contribute ~ add, data = df.Q2)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.36058 -0.36058 -0.04327  0.63942  1.63942 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)   1.0913     0.6455   1.691   0.1013  
add           0.3173     0.1674   1.896   0.0677 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8534 on 30 degrees of freedom
Multiple R-squared:  0.107,	Adjusted R-squared:  0.07722 
F-statistic: 3.594 on 1 and 30 DF,  p-value: 0.06766
```

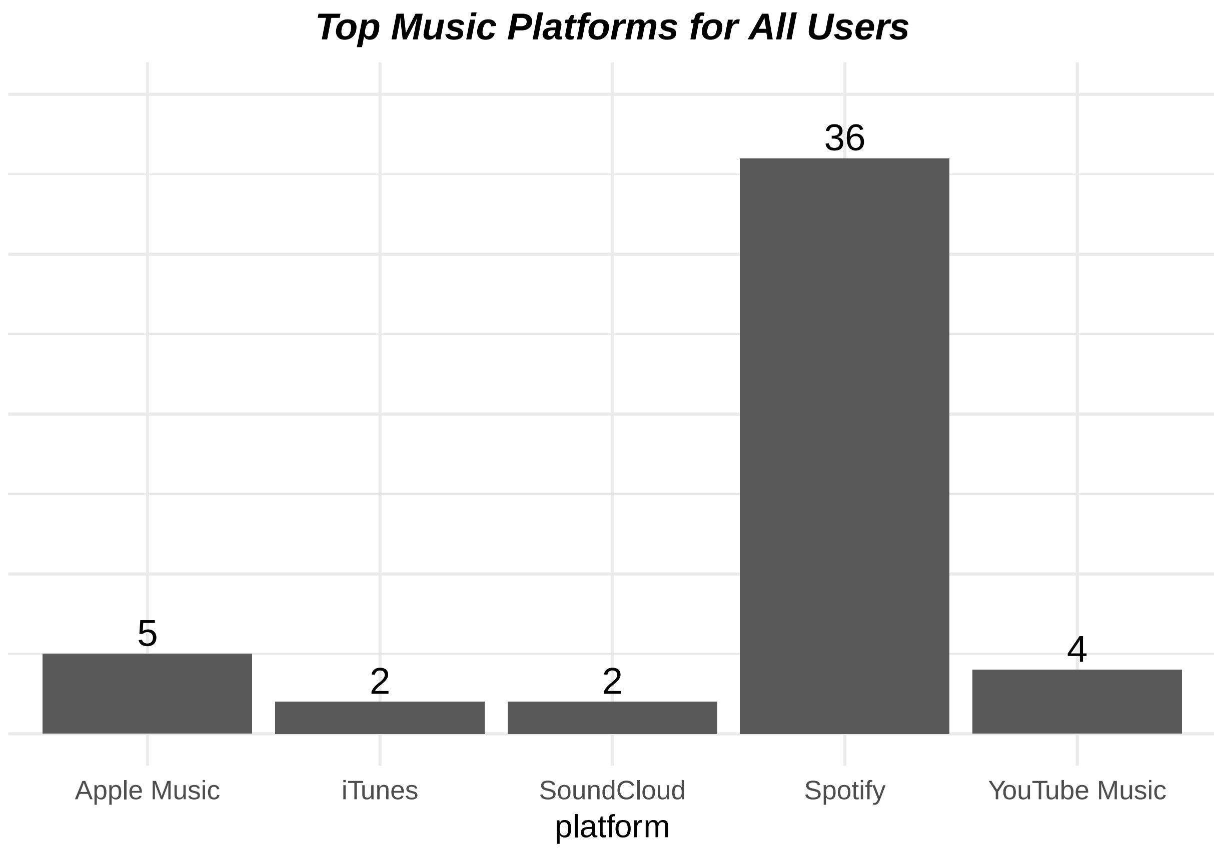


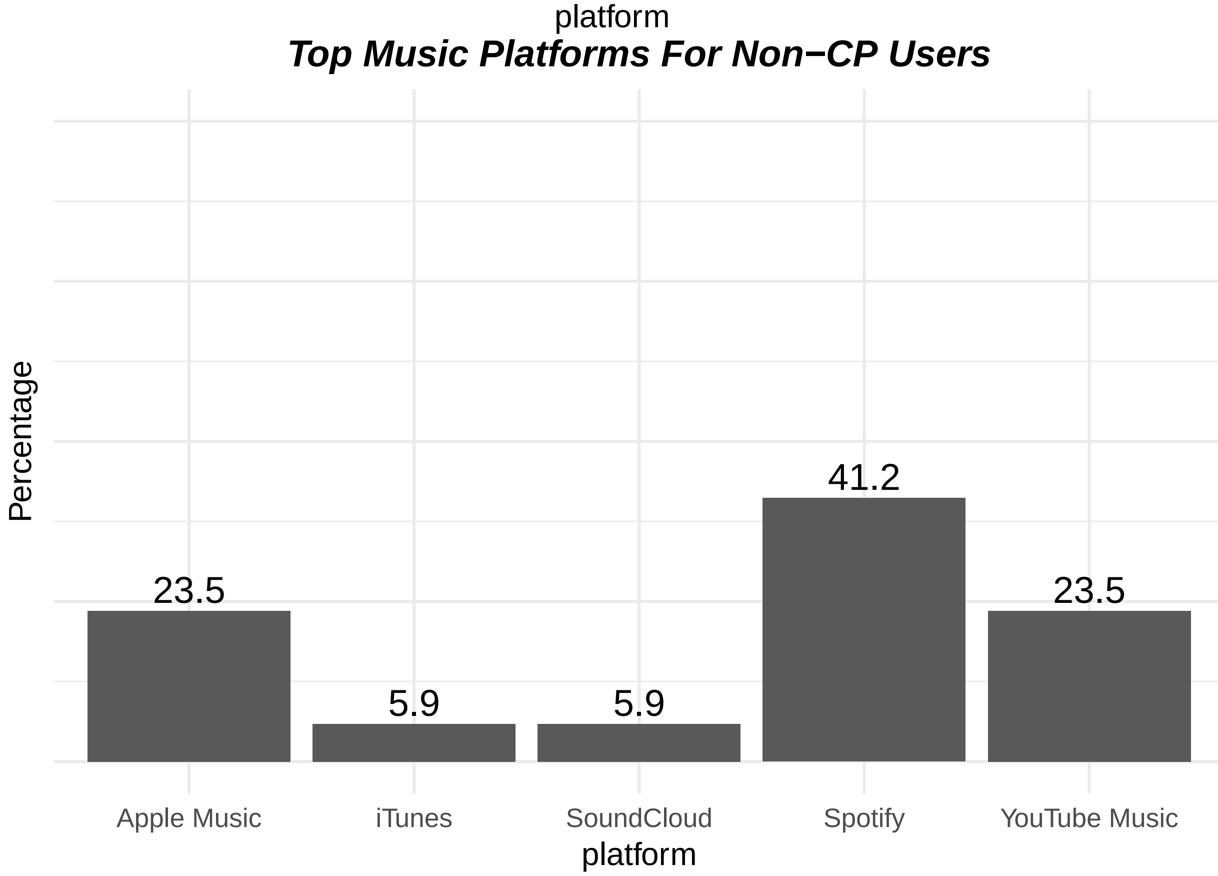
3.4 Exploratory analysis

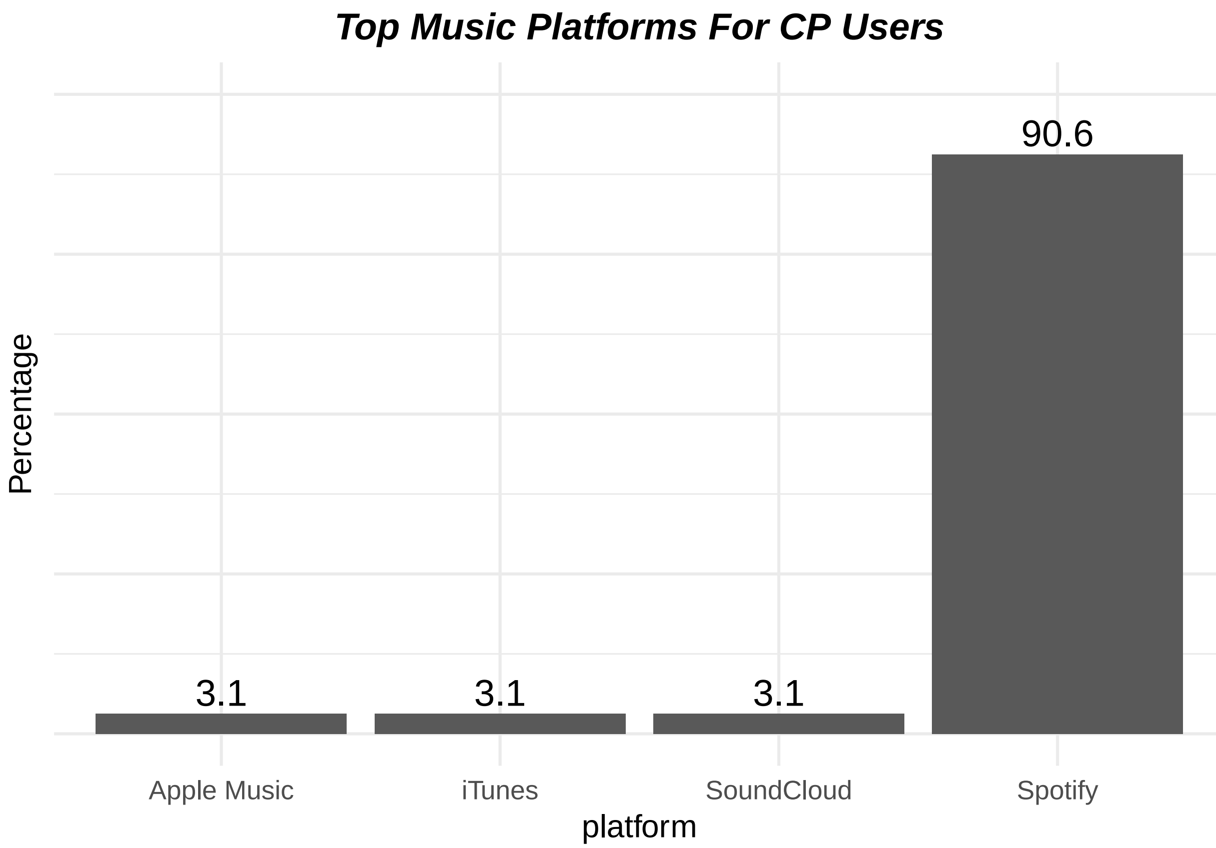
3.4.1 Choice of Music Platform

```
# get top 8 platforms
df.musicplatform = df.raw_all %>%
  mutate(id = row_number()) %>%
  select(c('id','Q2.3', starts_with('Q7.7_1'))) %>%
  setnames("Q2.3", "UseCP") %>%
  setnames("Q7.7_1", "platform")

total = nrow(df.musicplatform)

df.musicplatform %>%
  # gather(key = "activity", value = "platform", 3:9) %>%
  group_by(platform) %>%
  filter(platform != "N/A" && platform != "") %>%
  summarize(tally = n(), percent = tally/total) %>%
  top_n(5, wt = tally) %>%
  arrange(desc(tally)) %>%
  ggplot(data = .,
       aes( x = platform,
            y = tally)) +
  geom_bar(stat="identity") +
  geom_text(aes(label=tally), vjust=-0.3, color="black", size=5.0) +
```

```r
  ggtitle("Top Music Platforms for All Users") +
  theme_minimal(base_size = 12) +
  ylim(0,40) +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.y = element_blank(),
        axis.text.y = element_blank())

df.musicplatform %>%
  filter(UseCP == "No") %>%
  group_by(platform) %>%
  filter(platform != "N/A" && platform != "") %>%
  summarize(tally = n(), percent = 100 * (tally/nrow(.)) ) %>%
  top_n(5, wt = percent) %>%
  arrange(desc(percent)) %>%
  ggplot(data= . ,
         aes( x = platform,
              y = percent)) +
    geom_bar(stat="identity") +
    ylim(0,100) +
    ylab("Percentage") +
    geom_text(aes(label=round(percent, digits = 1)), vjust=-0.3, color="black", size=5.0) +
    ggtitle("Top Music Platforms For Non-CP Users") +
    theme_minimal(base_size = 12) +
    theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
          # axis.text=element_text(size=14),
          axis.text.y = element_blank())


df.musicplatform %>%
  filter(UseCP == "Yes") %>%
  group_by(platform) %>%
  filter(platform != "N/A" && platform != "") %>%
  summarize(tally = n(), percent =  100 * (tally/nrow(.)) ) %>%
  top_n(5, wt = percent) %>%
  arrange(desc(percent)) %>%
  ggplot(data= . ,
         aes( x = platform,
              y = percent)) +
    geom_bar(stat="identity") +
    ylim(0,100) +
    ylab("Percentage") +
    geom_text(aes(label=round(percent, digits = 1)), vjust=-0.3, color="black", size=5.0) +
    ggtitle("Top Music Platforms For CP Users") +
    theme_minimal(base_size = 12) +
    theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
          axis.text.y = element_blank())
```


Top Music Platforms for All Users
36
5
2
2
4
Apple Music
iTunes
SoundCloud
Spotify
YouTube Music
platform


Top Music Platforms For Non−CP Users
Percentage
41.2
23.5
23.5
5.9
5.9
Apple Music
iTunes
SoundCloud
Spotify
YouTube Music
platform

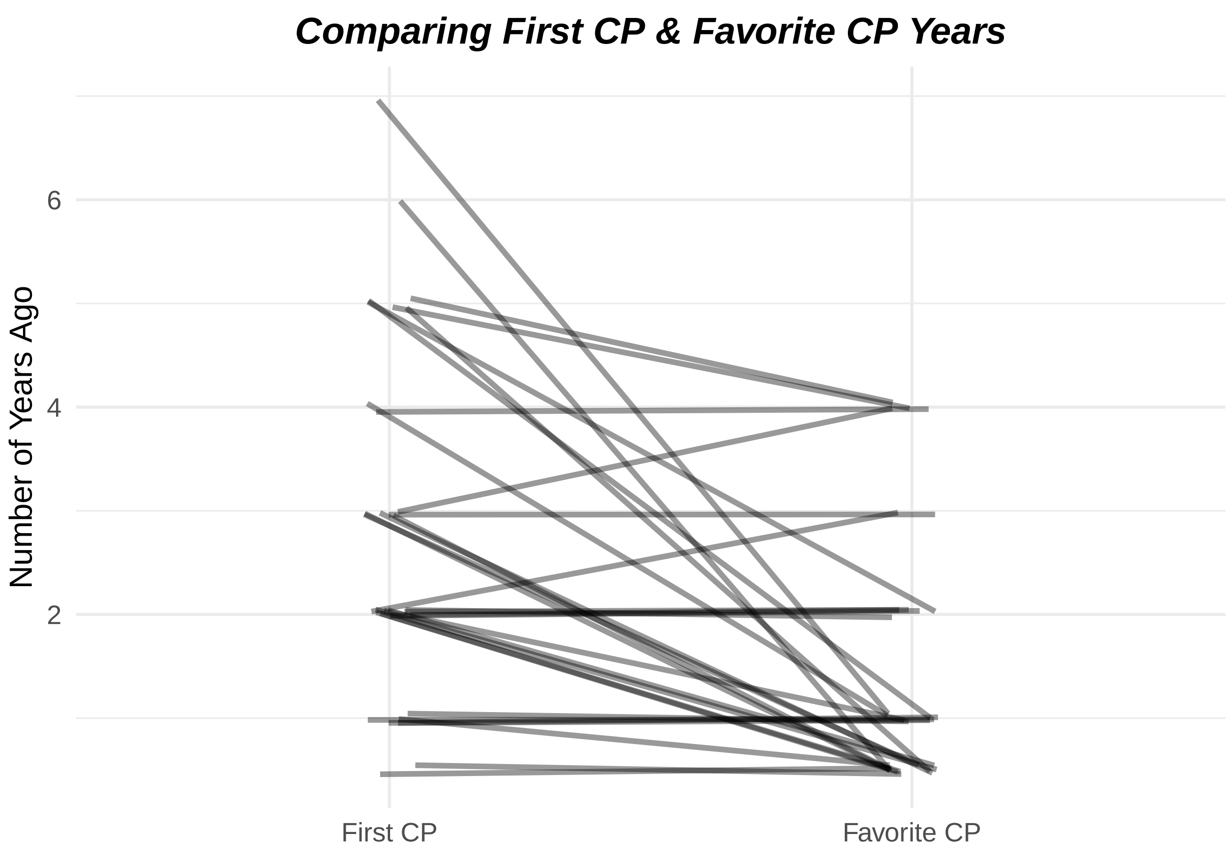


Spotify is the most widely used platform for engaging with music. More than 90% of CP users use Spotify as their main music platform. The results are a bit more diverse for non-CP users. Only 40% of non-CP users use Spotify.

3.4.2 First CP vs Fave CP

```
df.useCP <- df.raw %>%
  filter(Q2.3 == "Yes") %>%
  select(-grep("Q3.", names(df.raw)))

df.useCP[df.useCP==""] <- NA

df.initiateCP <- df.useCP %>%
  select(ResponseId, Q4.3, Q5.3) %>%
  group_by(ResponseId) %>%
  gather(firstORfave, when, 2:3)

df.initiateCP$when[df.initiateCP$when == "Less than 1"] = "0.5"
df.initiateCP$when <- as.numeric(df.initiateCP$when)
ggplot(df.initiateCP, mapping = aes(x = factor(firstORfave), y = when, group = ResponseId)) +
    geom_line(alpha = 0.4,
              size = 1,
              position = position_jitter(w = 0.05, h = 0.05)) +
  ylab("Number of Years Ago") +
  ggtitle("Comparing First CP & Favorite CP Years") +
```

```
  scale_x_discrete(labels = c("First CP", "Favorite CP")) +
    theme_minimal(base_size = 12) +
    theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
          axis.title.x = element_blank())
```

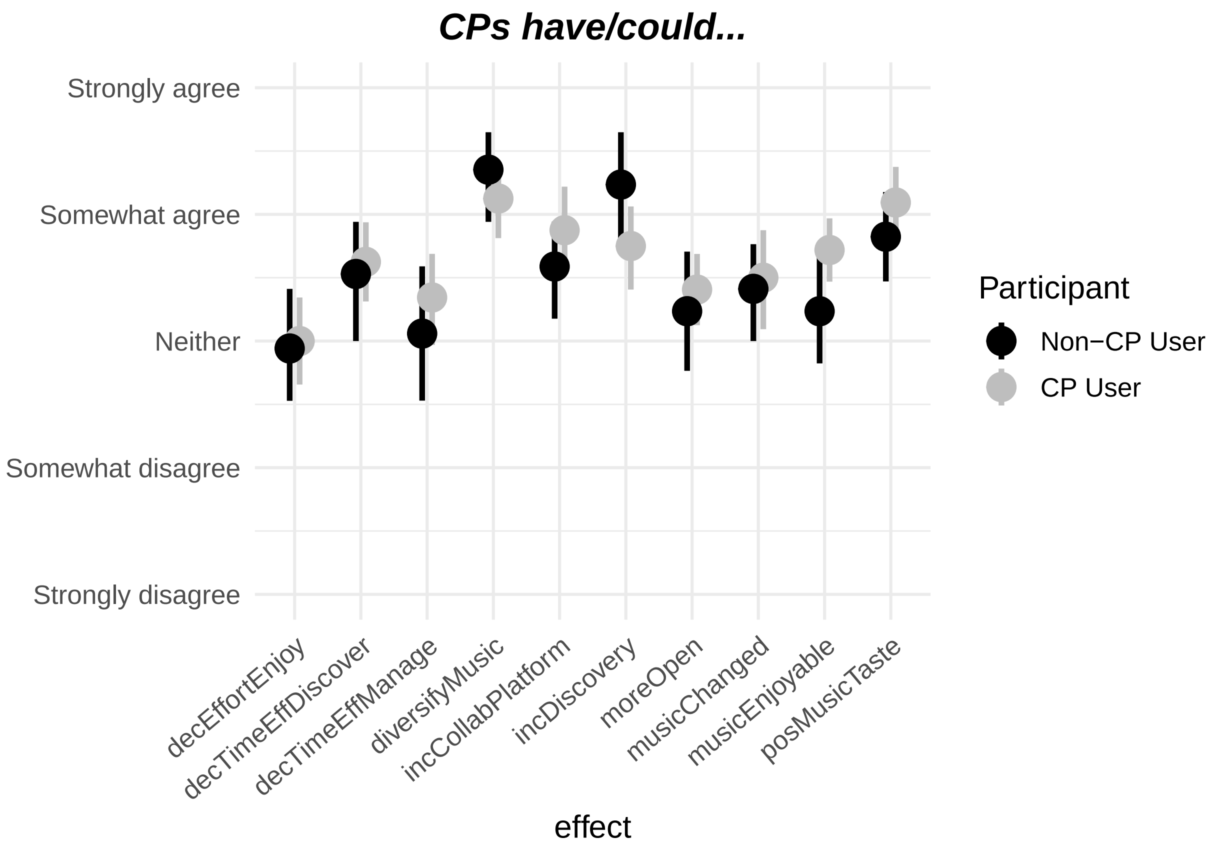


There is a downward trend in recency vs. favorite CP. This indicates user's favorite CP tends to me the most recent CP.

3.4.3 CPs have/could...

```
df.general <- df.raw %>%
  select(grep("Q3.8", names(df.raw)), grep("Q4.5", names(df.raw)))

df.gen1 <- df.general[,1:10] # don't use CP
df.gen1[df.gen1==""] <- NA
df.gen1 <- df.gen1 %>%
  na.omit()

df.gen2 <- df.general[,11:20] # use CP
df.gen2[df.gen2==""] <- NA
df.gen2 <- df.gen2 %>%
  na.omit()

n_gen1 <- nrow(df.gen1)
```

```r
n_gen2 <- nrow(df.gen2)

GenWords <- c("diversifyMusic","incDiscovery","decTimeEffDiscover","decTimeEffManage",
              "decEffortEnjoy","musicEnjoyable","musicChanged","posMusicTaste",
              "incCollabPlatform","moreOpen")

names(df.gen1) <- GenWords
names(df.gen2) <- GenWords

df.gen_both <- rbind(df.gen1, df.gen2)

df.gen_both[1:n_gen1,"experience"] <- replicate(n_gen1, 0) # don't use CP
df.gen_both[(n_gen1+1):(n_gen1+n_gen2),"experience"] <- replicate(n_gen2, 1) # use CP
df.gen_both <- gather(df.gen_both, "effect", "rating", 1:10)

# mapping word ratings into numerical ratings
df.gen_both$num_rating[df.gen_both$rating == "Strongly agree"] <- 5
df.gen_both$num_rating[df.gen_both$rating == "Somewhat agree"] <- 4
df.gen_both$num_rating[df.gen_both$rating == "Neither agree nor disagree"] <- 3
df.gen_both$num_rating[df.gen_both$rating == "Somewhat disagree"] <- 2
df.gen_both$num_rating[df.gen_both$rating == "Strongly disagree"] <- 1

## means w/ errors
ggplot(df.gen_both, mapping = aes(x = effect, y = as.numeric(num_rating),
                                  # group = as.factor(experience),
                                  color = as.factor(experience))) +
  theme(axis.text.x = element_text(angle = 40, hjust = 1, size = 10)) +
  stat_summary(fun.data = "mean_cl_boot", size = 1,
               position = position_dodge(width = 0.3)) +
  guides(color = guide_legend(title = "Participant")) +
  scale_color_manual(values = c("orange","blue"), labels = c("0","1")) +
  scale_y_continuous(labels = c("Strongly agree", "Somewhat agree", "Neither",
                                "Somewhat disagree", "Strongly disagree"),
                     breaks = c(5, 4, 3, 2, 1),
                     limits = c(1,5)) +
  scale_color_manual(labels = c("Non-CP User", "CP User"), values = c("black", "gray")) +
  ggtitle("CPs have/could...") +
  theme_minimal(base_size = 12) +
  theme(plot.title = element_text(color="black", face="bold.italic", hjust=0.5),
        axis.title.y = element_blank(),
        axis.text.x = element_text(angle = 40, hjust = 1, size = 10))
```



In comparing how CP usage has changed users' music choices, habits, and perceptions with changes that non-CP users expect if they were to use CP, we see some differences between actual vs expected values. The most notable difference is in incDiscovery (i.e., increase in music discovery), with the mean value for what CP users actually report is lower than that of what non-CP users expect.

4 Discussion

Our first analysis was around understanding differences in music habits and values between CP and non-CP users. We were interested in four aspects: (1) music as a means of social interaction, (2) music sharing, (3) music discovery, and (4) music taste. Our analysis confirmed our hypothesis that users of CP value more music co-enjoyment as a means for social interaction (H1a). CP users value more the importance of sharing and listening to music with others. Their value on social interaction might be a strong factor for them to engage in CPs, since CPs are a form of fostering social connectedness.

Results also confirmed our hypothesis that users of CP more frequently engage in actions around music. CP users more frequently engage in sharing (H1b) and discovering new music (H1c) compared to non-CP users. Engaging in CPs might motivate users to be more active in music discovery since they are more likely be exposed to new music that is added by collaborators. They might also be more likely to share music, as a result of being contributors to a CP. Another factor that might have influenced this result is that perhaps non-CP users simply don't share or discover music as frequently compared to CP users because they are happier listening to their own music. However, we were not able to test for this with our current analysis.

We had also hypothesized that CP and non-CP users might have different taste in music (H1d). In particular our hypothesis was that non-CP users preferred familiar music and thus might be a reason why they do not engage in CPs. However, our analysis did not confirm the hypothesis. More research is needed to understand what prevents non-CP users from engaging in CPs. Perhaps there are other factors not considered in this survey that would be better explain characterize non-CP users, such as music genres or reasons for experiencing music (e.g., fulfill function, for entertainment).

Our hypothesis on playlist ownership as measured by a user's perceived role and frequency of contribution (H2a) was half correct. While perceived role was a significant predictor of playlist ownership, the frequency of contributions was neither a meaningful predictor for ownership nor correlated with their perceived role. This reflects how user behavior may not necessarily be correlated to user perception. For the user's perceived role of contribution, however, it is in the same domain of "perception" as sense of ownership. Therefore, we can explain for why there is a strong correlation between the two. However, we cannot claim causality–that greater sense of ownership is caused by greater perceived role of one's contributions to the CP, or vice versa. Therefore, further investigation is needed to parse out which causes the other, or whether there is no one cause but rather are factors that positively reinforce each other.

Looking further into the individual activities of adding, deleting, and reordering songs, we see that the user's perceived role in adding of songs is most indicative of their sense of ownership (H2b), despite the frequency of these contributions not being significantly correlated with that of the perceived roles. We believe that this may be because the making of a CP is an additive process and therefore, the more one engages in adding songs to the CP the higher the engagement towards and involvement in the CP, which may then lead to higher sense of ownership.

In our exploratory analysis, we found that Spotify is the most widely used platform for engaging with music (E1). More than 90% of CP users use Spotify as their main music platform. This might be another motivating factor for participating in CPs since Spotify is one of the music platforms with more developed functionality for sharing and contributing to playlists.

We also were curious to see whether the CP users' favorite playlists were their first (E2), and therefore compared their answers for "How many years ago was your first interaction with a CP" with that of "How many years ago was [their favorite CP] initiated." While the upward directionality of some of the results does not make sense, because it implies that the users' favorite playlist is older than when they first started interacting with a CP, the other results are interesting. We observe more downward trends rather than a horizontal one, and this indicates to us that CP users' favorite playlists tend to be more recent ones than that which are their first CP.

Our final exploratory analysis looks at how CP usage has changed users' music choices, habits, and perceptions (E3) and we find that while we have found in H1c that music discovery is more frequent for CP users, what CP users actually report in terms of increasing music discovery is not as high as what non-CP users would expect were they to use CPs. We will need to uncover more of the data that we have collected through the online survey to better understand this particular point.

5 Conclusion

Overall, the hypotheses that we were aiming to answer were addressed using our online survey, of whether there are differences in music habits and values between CP and non-CP users and of user's engagement in a CP related to their sense of ownership. We have found areas that need more unpacking of the results that we will continue on.

We have additionally found some other findings that were interesting to see, such as seeing the use of CP being more widespread than what we expected and how there are differences between the actual vs predicted effects of CP usage.

References

Axelsson, Fredrik, and Mattias Östergren. 2002. "SoundPryer: Joint Music Listening on the Road." In *Adjunct Proceedings*, 39.

Bardhi, Fleura, and Giana M Eckhardt. 2012. "Access-Based Consumption: The Case of Car Sharing." *Journal of Consumer Research* 39 (4): 881–98.

Brown, Barry, Abigail J Sellen, and Erik Geelhoed. 2001. "Music Sharing as a Computer Supported Collaborative Application." In *ECSCW 2001*, 179–98. Springer.

"Collaborate on Playlists with Spotify's Collaboration Feature." n.d. *Dummies*. https://www.dummies.com/social-media/spotify/collaborate-on-playlists-with-spotifys-collaboration-feature/.

Danielson, Scott. 2011. *Turntable.fm: A Creative Twist on Social Music*. web.appstorm.net. http://web.appstorm.net/reviews/media-reviews/turntable-fm-a-creative-twist-on-social-music/.

Frank, Jay. 2009. *Futurehit. Dna: How the Digital Revolution Is Changing Top 10 Songs*. BookBaby.

Jung, Malte F, Nikolas Martelaro, and Pamela J Hinds. 2015. "Using Robots to Moderate Team Conflict: The Case of Repairing Violations." In *Proceedings of the Tenth Annual Acm/Ieee International Conference on Human-Robot Interaction*, 229–36. ACM.

Kim, Taemie, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. "Meeting Mediator: Enhancing Group Collaborationusing Sociometric Feedback." In *Proceedings of the 2008 Acm Conference on Computer Supported Cooperative Work*, 457–66. ACM.

Koelsch, Stefan. 2013. "From Social Contact to Social Cohesion—the 7 Cs." *Music and Medicine* 5 (4): 204–9.

Lee, Jin Ha, Hyerim Cho, and Yea-Seul Kim. 2016. "Users' Music Information Needs and Behaviors: Design Implications for Music Information Retrieval Systems." *Journal of the Association for Information Science and Technology* 67 (6): 1301–30.

Lenz, Eva, Sarah Diefenbach, Marc Hassenzahl, and Sébastien Lienhard. 2012. "Mo. Shared Music, Shared Moment." In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, 736–41. ACM.

Liu, KuanTing, and Roger Andersson Reimer. 2008. "Social Playlist: Enabling Touch Points and Enriching Ongoing Relationships Through Collaborative Mobile Music Listening." In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, 403–6. ACM.

Park, So Yeon, and Blair Kaneshiro. 2017. "An Analysis of User Behavior in Co-Curation of Music Through Collaborative Playlists."

Valentine, Melissa A, Daniela Retelny, Alexandra To, Negar Rahmati, Tulsee Doshi, and Michael S Bernstein. 2017. "Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds as Organizations." In *Proceedings of the 2017 Chi Conference on Human Factors in Computing Systems*, 3523–37. ACM.