Coaching and organisational strucutures in eSports

Load data

```
## Directory and file settings
dir <- "data"
filename <- "160913_Coaching-and-organisational-structures-in-eSports.csv"

## Read data
dat <- read.csv2(file.path(dir, filename),
                 header = T,
                 sep = ",",
                 stringsAsFactors = F)
```

Data structure:

```
str(dat)
```

```
## 'data.frame':    65 obs. of  19 variables:
##  $ Zeitstempel
##  $ What.s.your.age.
##  $ How.much.time.do.you.play.games.per.day.
##  $ How.much.time.do.you.watch.competitive.games.per.day.
##  $ What.s.your.role.in.eSport.
##  $ Do.you.prefer.playing.in.teams.or.solo.
##  $ Have.you.or.your.team.received.coaching.
##  $ Which.of.these.games.do.you.play.competitively.
##  $ Which.of.these.games.is.your.most.important.competitive.eSports.game.
##  $ Regarding.your.choice..which.is.the.strongest.region.for.the.eSport.game.
##  $ What.do.you.think.might.be.the.reason.
##  $ How.do.you.rate.the.overall.importance.of.structural.organisation..management..coaching..analysis
##  $ How.do.you.rate.the.importance.of.managers.in.eSports.
##  $ What.s.your.opinion.on.managers.
##  $ How.do.you.rate.the.importance.of.coaches.in.eSports.
##  $ What.s.your.opinion.on.coaches.
##  $ How.do.you.rate.the.importance.of.analysts.in.eSports.
##  $ What.s.your.opinion.on.analysts.
##  $ How.do.you.rate.the.importance.of.management.and.coaching.for.further.development.of.eSport.popul
```

Recode variable names

```
## Save old variable names
colnames_old <- colnames(dat)

## Prepare new variable names
colnames_new <- c(
  "time",
  "age",
  "time_play",
  "time_spect",
  "role",
  "team_solo",
  "rec_coach",
  "games_playcompete",
  "games_important",
  "region",
  "region_reason",
  "importance_overall",
  "importance_manager",
  "importance_manager_txt",
  "importance_coach",
  "importance_coach_txt",
  "importance_analyst",
  "importance_analyst_txt",
  "future"
)

## Recode variable names
colnames(dat) <- colnames_new

## For testing purpose
#print(data.frame("Old varnames" = colnames_old, "New varnames" = colnames_new))
```

Data cleaning

Values ranges

- Age (optional):
  - "under 12"
  - "12 - 15"
  - "16 - 18"
  - "19 - 21"
  - "21 - 24" **WHY 21 AGAIN????**
  - "25 - 27"
  - "28 - 30"
  - "31 - 35"
- Role:
  - Player
  - Manager
  - Analyst
  - Coach
  - Scout
  - Caster
  - Sonstiges:
- Preference:
  - "Solo"
  - "Team"
- Importance: 1-10 (not important - essential)
- Future: 1- 10 (not important - essential)

- Time (play, watch):
  - “under 1 hour”
  - 1-3 hours"
  - “3-5 hours”
  - “5-7 hours”
  - “8-10 hours”
- Received coaching:
  - “currently getting coached”
  - “got coached in the last 3 month”
  - “got coached in the last 6 month”
  - “got coached in the last year”
  - “got coached but way earlier”
  - “didn’t receive any coaching”
- Games (play ~ multiple):
  - “None”
  - “League of Legends”
  - “Counterstrike GO”
  - “Dota 2”
  - “Smite”
  - “Hearthstone”
  - “World of Warcraft”
  - “Overwatch”
  - “Sonstiges”
- Region:
  - “Europe”
  - “North America”
  - “South America”
  - “Korea”
  - “China”
  - “Asia (without China and Korea)”
  - “Australia”
  - “Afrika”
  - “Sonstiges”
- Reason(multiple):
  - “mentality of the players”
  - “time investment”
  - “structural organisation”
  - “coaching”
  - “popularity of eSports in the region”
  - “salaries and investments”
  - “Sonstiges”

Remove trolls

Identified trolls (to be removed: * Timestamp: 29.08.2016 20:19:10

```r
## Remove troll
case2rm <- which(dat$time == "29.08.2016 20:19:10")
dat <- dat[-case2rm, ]
```

Question: What's your role in eSports?

Given answers in free text field “Sonstiges”

Recode scheme:

- “Teamorganizer” to “Manager”"
- “Viewer” to “Player” or remove case
  - **Final decision: Remove case**
- “Casual Player” to “Player”"
- “Many of the above”
  - **STILL UNCLEAR**

```r
## Identify misc cases
casefinder <-
  !dat$role %in% c("Player", "Manager", "Analyst", "Coach", "Scout", "Caster")
print(dat$role[casefinder])
```

```
## [1] "Many of the above, including Org Owner"
## [2] "Casual Player"
## [3] "None"
## [4] "Viewer"
## [5] "None. "
## [6] "/"
## [7] "Teamorganizer"
```

```r
## Recode into new variable role2
dat$role2 <- dat$role
dat$role2[which(dat$role2 == "Teamorganizer")] <- "Manager"
dat$role2[which(dat$role2 == "Casual Player")] <- "Player"

## Remove cases
case2rm <- which(dat$role == "Viewer")
dat <- dat[-case2rm, ]

case2rm <- which(dat$role == "Many of the above, including Org Owner")
dat <- dat[-case2rm, ]

case2rm <- which(grepl("None", dat$role))
dat <- dat[-case2rm, ]
case2rm <- which(dat$role == "/")
dat <- dat[-case2rm, ]
```

Questions regarding games (play/spectate):

Recode into new variables-per-game

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
## Games played competetively
dat <-
  dat %>%
  ## Workaround: Re-initialise as data.frame
  data.frame() %>%
  mutate(games_playcompete_none = ifelse(grepl("None", games_playcompete), 1, 0),
         games_playcompete_lol = ifelse(grepl("League of Legends", games_playcompete), 1, 0),
         games_playcompete_csgo = ifelse(grepl("Counterstrike GO", games_playcompete), 1, 0),
         games_playcompete_dota2 = ifelse(grepl("Dota 2", games_playcompete), 1, 0),
         games_playcompete_smite = ifelse(grepl("Smite", games_playcompete), 1, 0),
         games_playcompete_heartstone = ifelse(grepl("Hearthstone", games_playcompete), 1, 0),
         games_playcompete_wow = ifelse(grepl("World of Warcraft", games_playcompete), 1, 0),
         games_playcompete_overwatch = ifelse(grepl("Overwatch", games_playcompete), 1, 0),
         games_playcompete_skill = ifelse(grepl("S.K.I.L.L.", games_playcompete), 1, 0),
         games_playcompete_startcraft2 = ifelse(grepl("Starcraft 2", games_playcompete), 1, 0),
         games_playcompete_bf34 = ifelse(grepl("Battlefield 3/4", games_playcompete), 1, 0),
         games_playcompete_codbo3 = ifelse(grepl("Call of Duty: Black Ops 3", games_playcompete), 1, 0)
         games_playcompete_codmw2 = ifelse(grepl("call of duty modern warfare 2", games_playcompete), 1
         games_playcompete_rss = ifelse(grepl("Rainbow Six Siege", games_playcompete), 1, 0)
         )

## Region reason
dat <-
  dat %>%
  ## Workaround: Re-initialise as data.frame
  data.frame() %>%
  mutate(region_reason_mentality = ifelse(grepl("mentality of the players", region_reason), 1, 0),
         region_reason_time = ifelse(grepl("time investment", region_reason), 1, 0),
         region_reason_structuralorg = ifelse(grepl("structural organisation", region_reason), 1, 0),
         region_reason_coaching = ifelse(grepl("coaching", region_reason), 1, 0),
         region_reason_popularity = ifelse(grepl("popularity of eSports in the region", region_reason),
         region_reason_salaries = ifelse(grepl("salaries and investments", region_reason), 1, 0)
         )

## Age
dat$age[dat$age == ""] <- NA

## Coaching
dat$rec_coached10 <- as.character(dat$rec_coach)
dat$rec_coached10[which(grepl("didn't", dat$rec_coached10))] <- 0
dat$rec_coached10[which(dat$rec_coached10 != "no coaching")] <- 1

dat$rec_coach_currently10 <- as.character(dat$rec_coach)
dat$rec_coach_currently10[which(grepl("didn't", dat$rec_coach_currently10))] <- 0
dat$rec_coach_currently10[which(grepl("currently", dat$rec_coach_currently10))] <- 1

dat$rec_coach2 <- as.character(dat$rec_coach)
dat$rec_coach2[which(grepl("didn't", dat$rec_coach2))] <- "never been coached"
dat$rec_coach2[which(grepl("currently", dat$rec_coach2))] <- "currently coached"
dat$rec_coach2[which(grepl("got", dat$rec_coach2))] <- "been coached"
```

Analysis

Research questions of interest

- importance_coach vs. rec_coach
- importance_X vs. time_play
- importance_X vs. time_spect
- importance_X vs. role
- importance_X vs. team_solo
- future vs. time_play
- future vs. time_spect
- future vs. role
- future vs. team_solo

How much time do you play games per day?

```
dat$time_play <- factor(dat$time_play)
print(levels(dat$time_play))
```

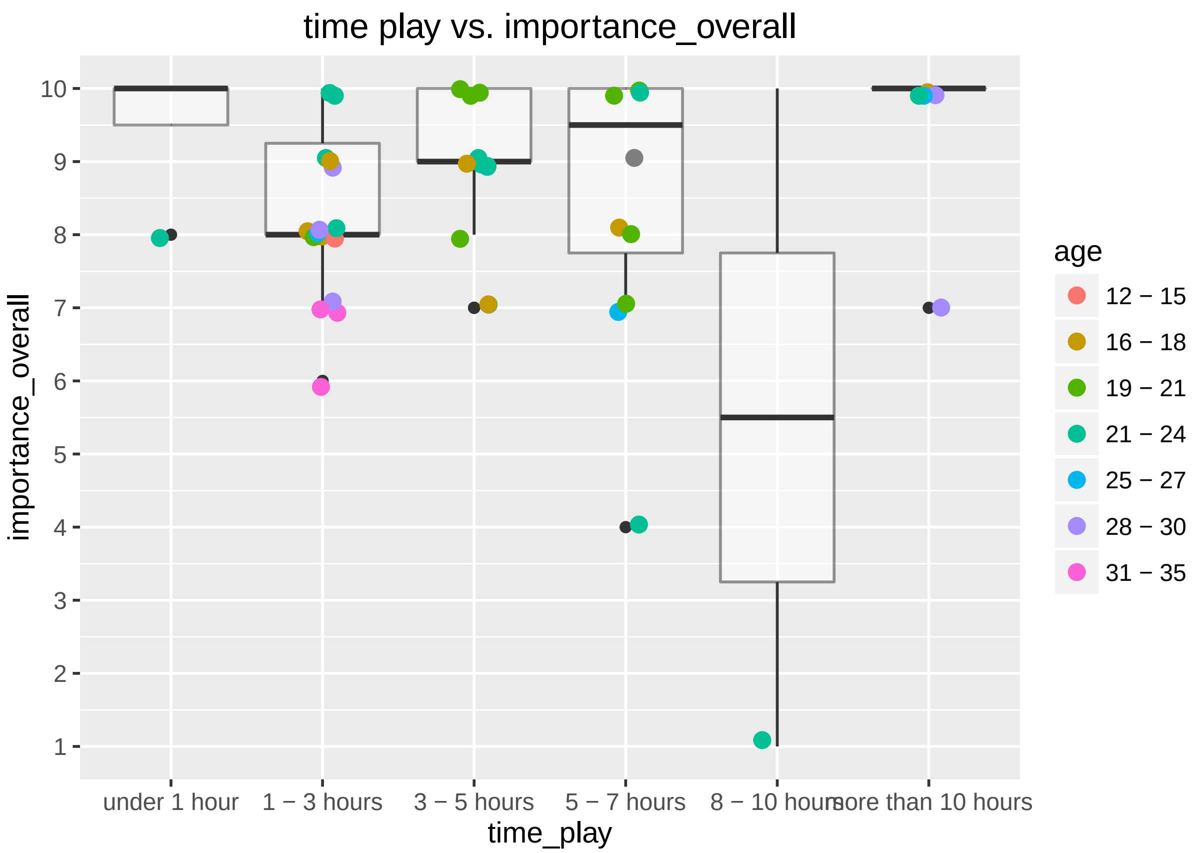
```
## [1] "1 - 3 hours"        "3 - 5 hours"        "5 - 7 hours"
## [4] "8 - 10 hours"       "more than 10 hours" "under 1 hour"
```

```
dat$time_play <- factor(dat$time_play, levels = levels(dat$time_play)[c(6, 1:5)])

dat$time_spect <- factor(dat$time_spect)
print(levels(dat$time_spect))
```

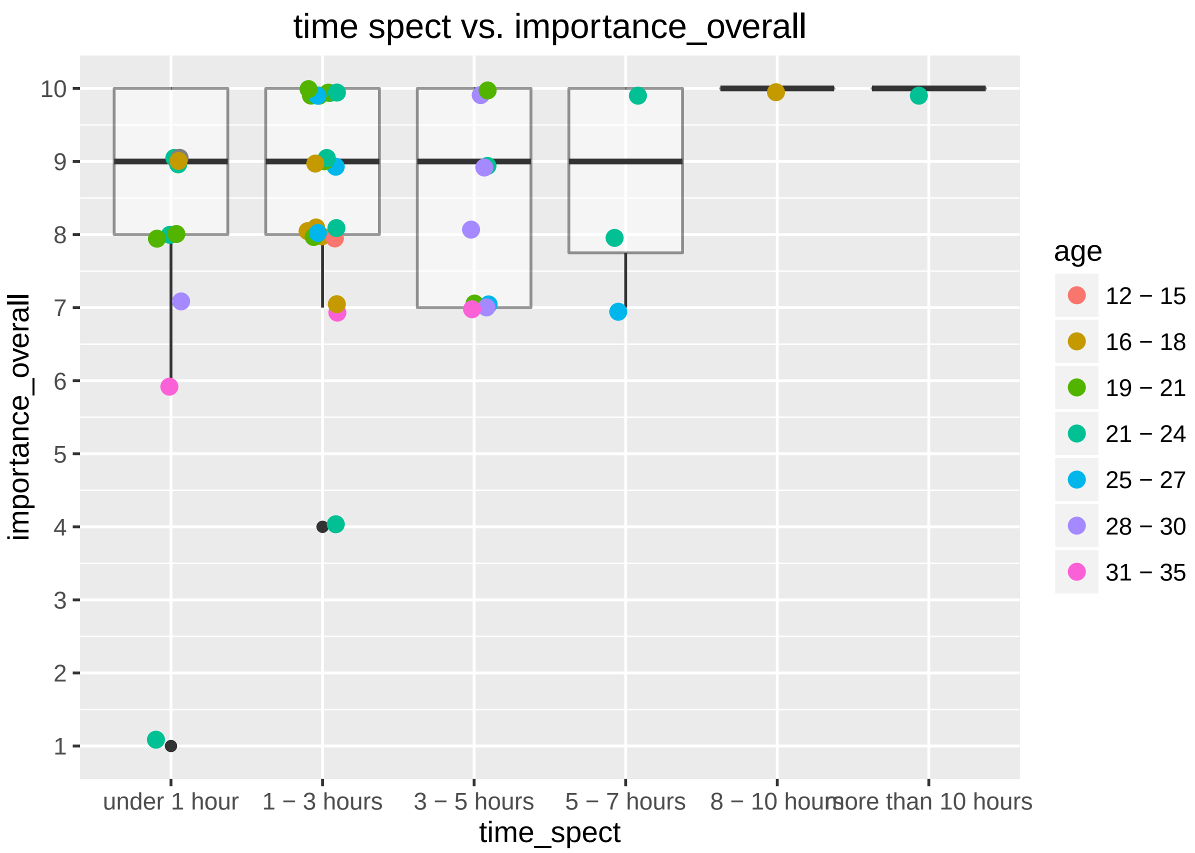
```
## [1] "1 - 3 hours"        "3 - 5 hours"        "5 - 7 hours"
## [4] "8 - 10 hours"       "more than 10 hours" "under 1 hour"
```

```
dat$time_spect <- factor(dat$time_spect, levels = levels(dat$time_spect)[c(6, 1:5)])

library(ggplot2)

var4y_list <- c(
  "importance_overall",
  "importance_manager",
  "importance_coach",
  "importance_analyst"
)

for(var4y in var4y_list) {

  set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes_string(x = "time_play",
                            y = var4y),
                 alpha = 0.5) +
    geom_point(data = dat,
               aes_string(x = "time_play",
                          y = var4y,
                          colour = "age"),
               position = position_jitter(w = 0.25, h = 0.25),
               size = 2.5) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    #theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("time play vs.", var4y))
  plot(plotdat)

  set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes_string(x = "time_spect",
                            y = var4y),
                 alpha = 0.5) +
    geom_point(data = dat,
             aes_string(x = "time_spect",
                        y = var4y,
                        colour = "age"),
             position = position_jitter(w = 0.25, h = 0.25),
               size = 2.5) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    #theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("time spect vs.", var4y))
  plot(plotdat)
}
```

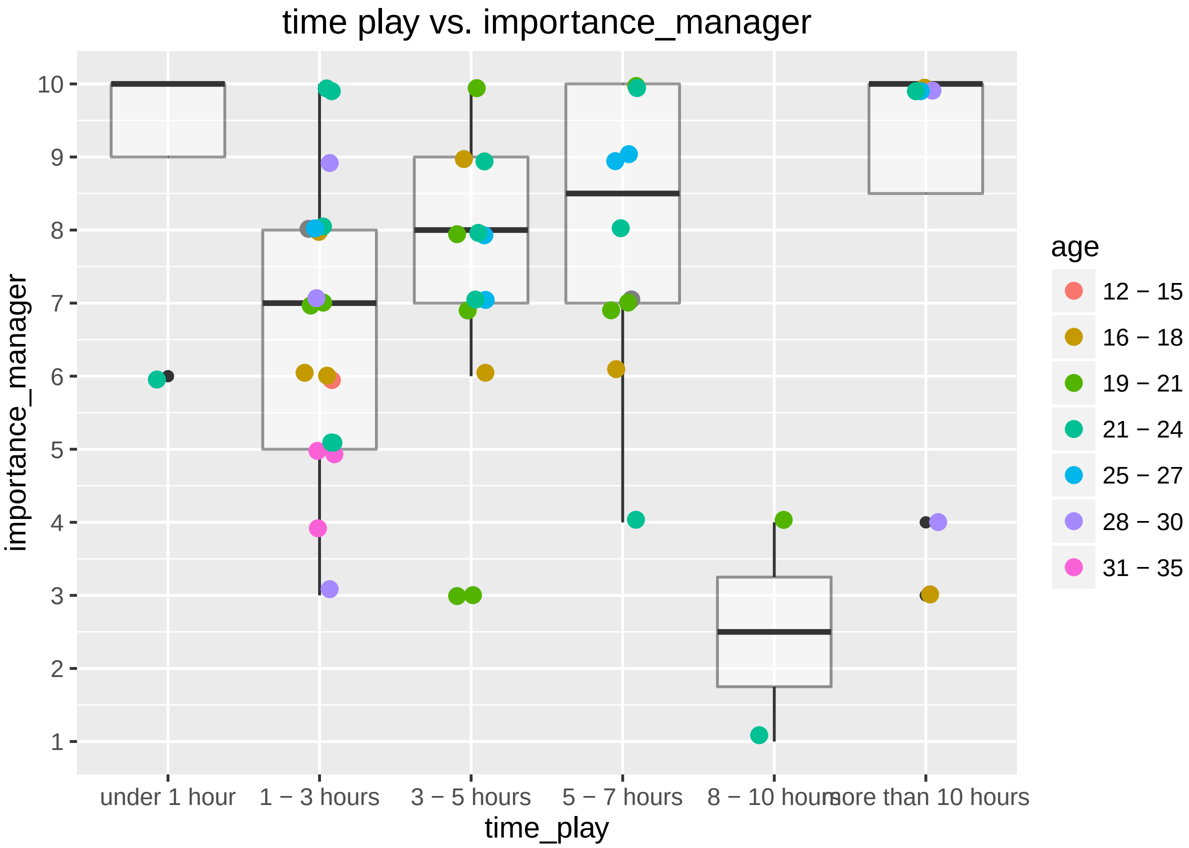
```
## Warning: Removed 14 rows containing missing values (geom_point).
```



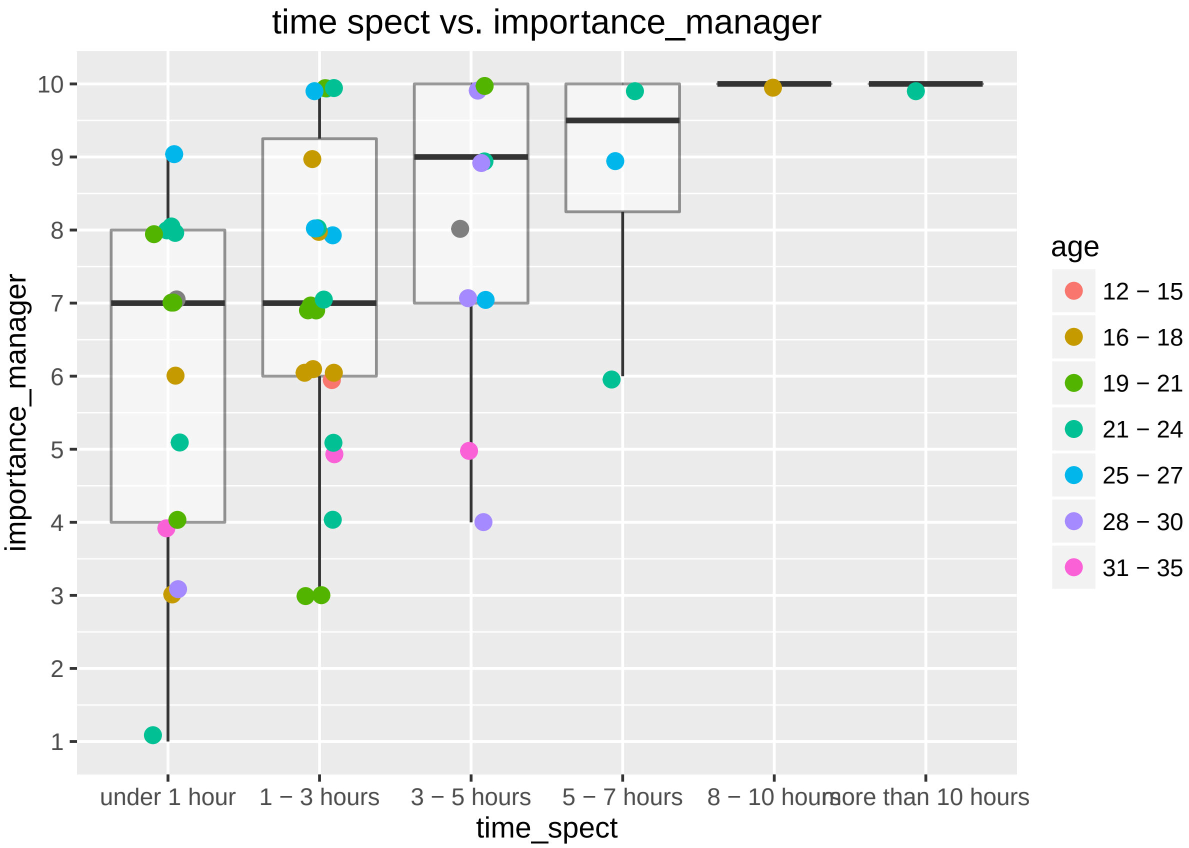
```
## Warning: Removed 14 rows containing missing values (geom_point).
```



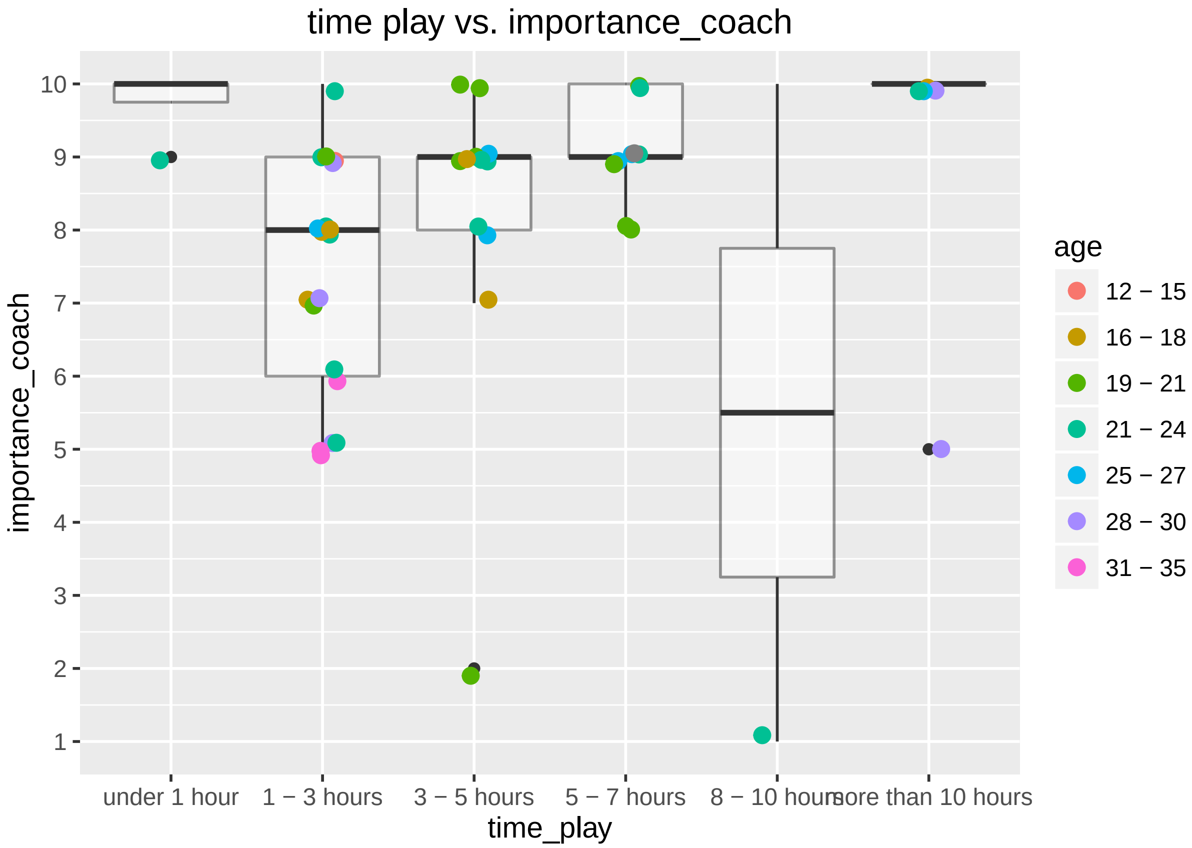
```
## Warning: Removed 8 rows containing missing values (geom_point).
```



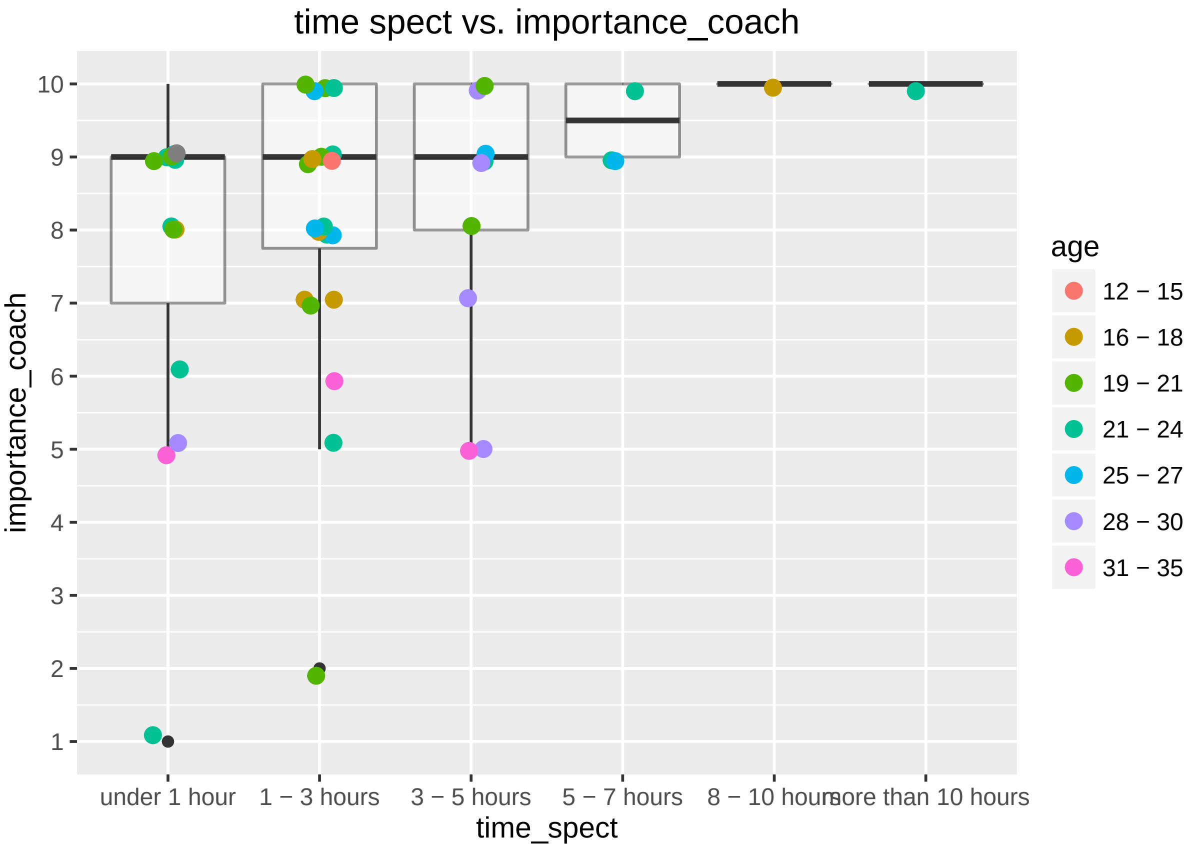
```
## Warning: Removed 8 rows containing missing values (geom_point).
```



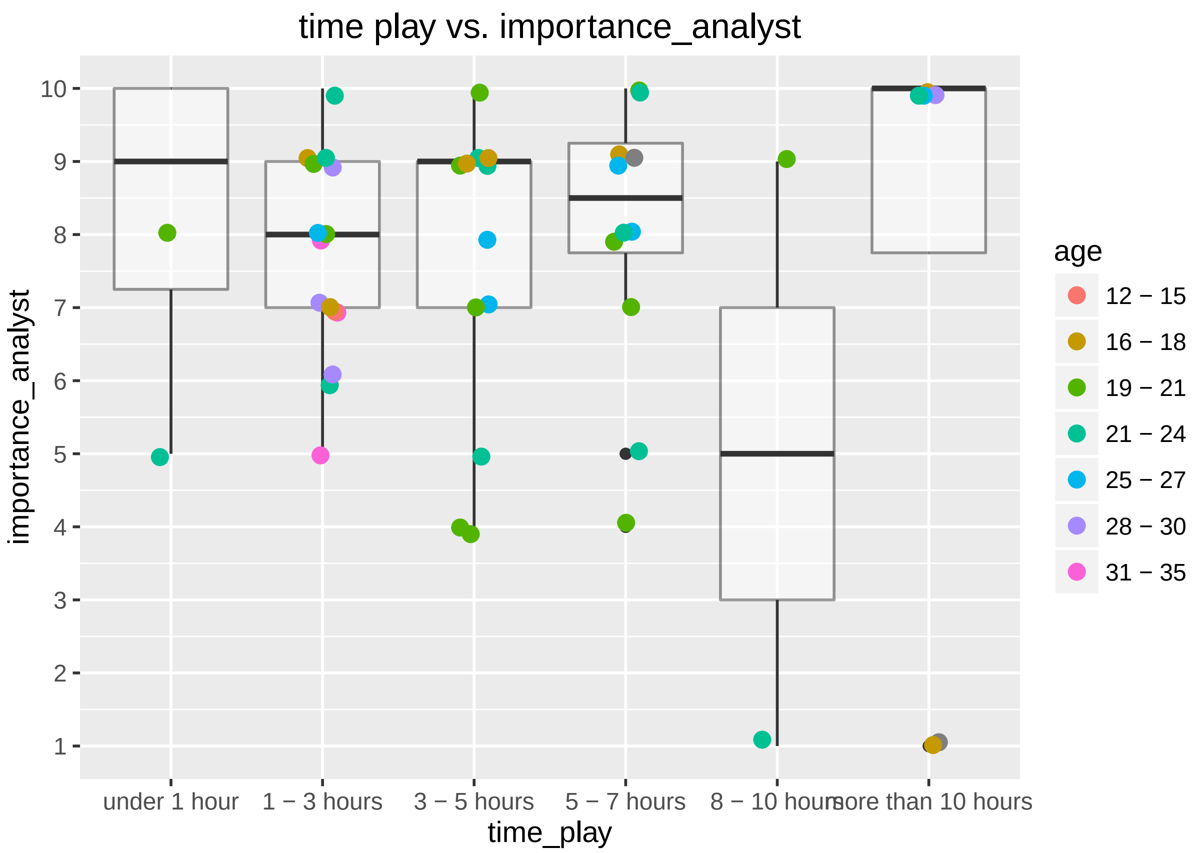
```
## Warning: Removed 12 rows containing missing values (geom_point).
```



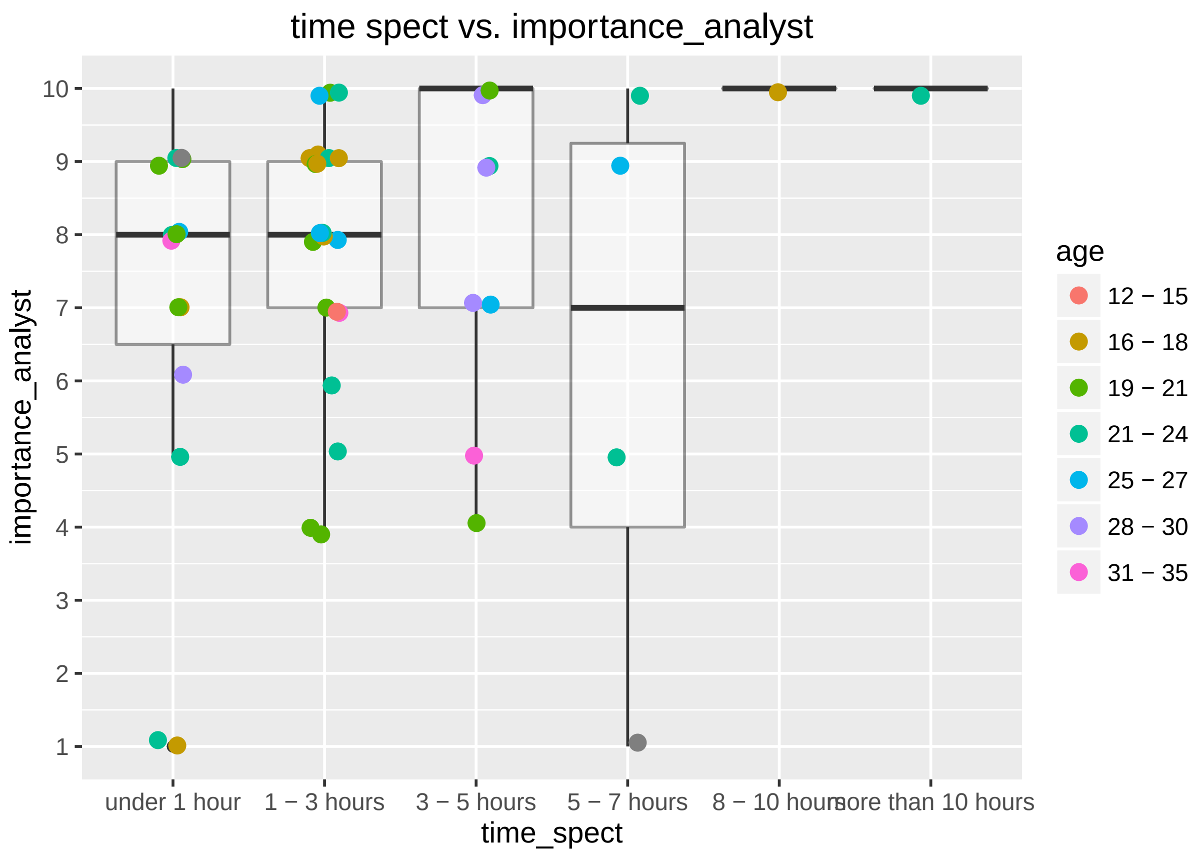
```
## Warning: Removed 12 rows containing missing values (geom_point).
```



```
## Warning: Removed 9 rows containing missing values (geom_point).
```



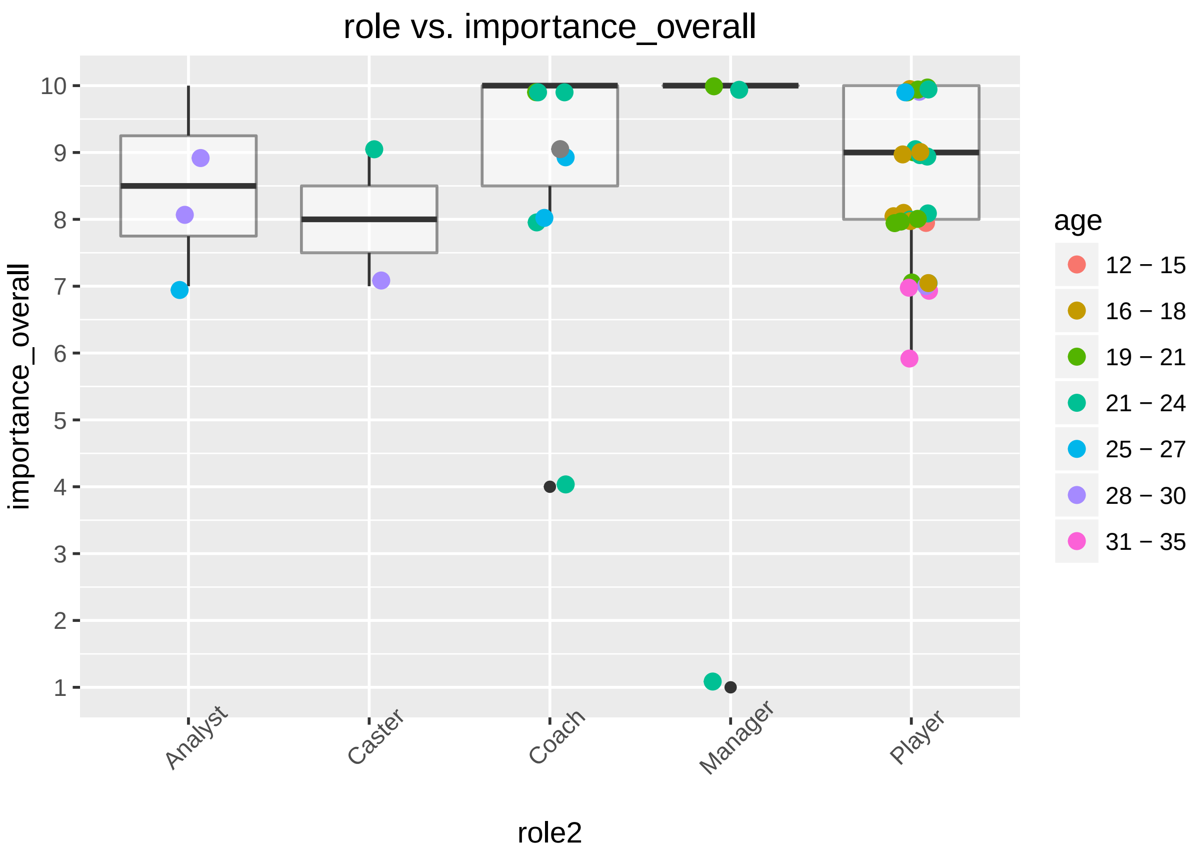
```
## Warning: Removed 9 rows containing missing values (geom_point).
```



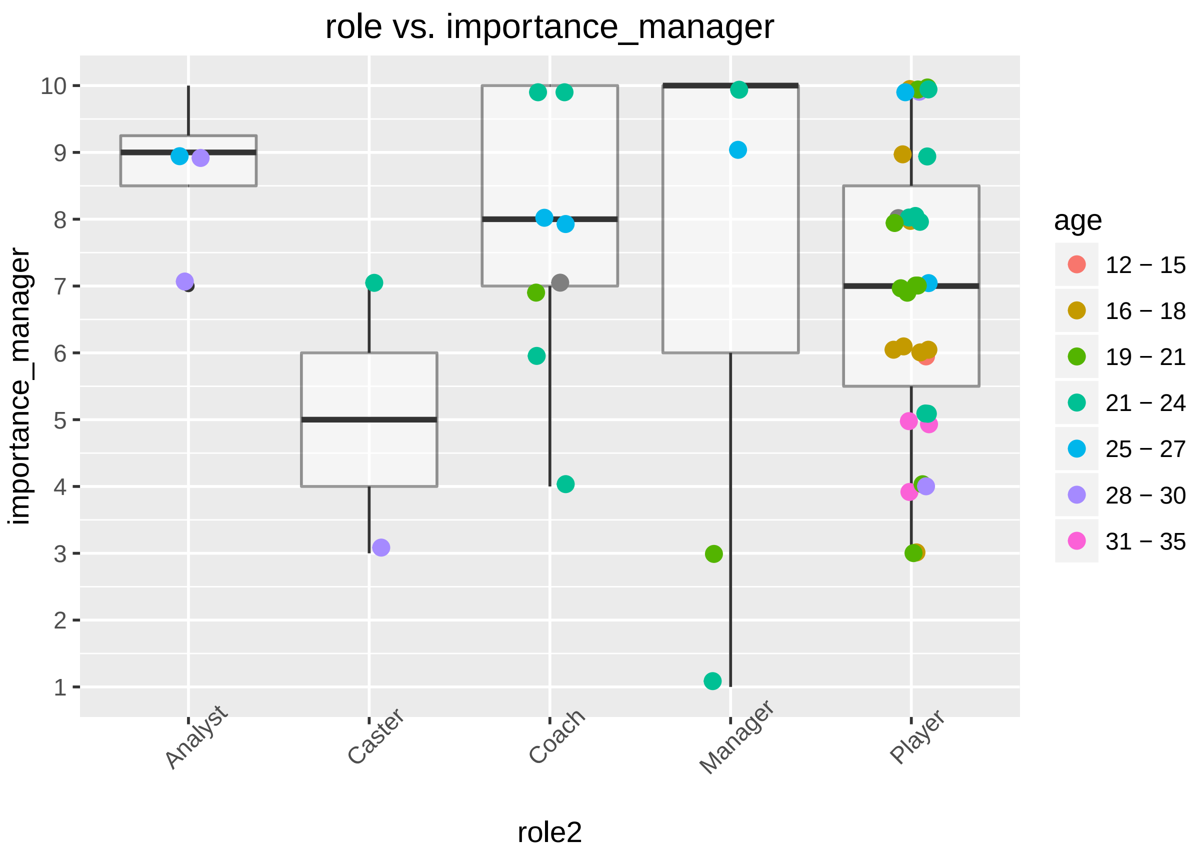
Role

```
for(var4y in var4y_list) {
  set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes_string(x = "role2",
                            y = var4y),
                 alpha = 0.5) +
    geom_point(data = dat,
               aes_string(x = "role2",
                          y = var4y,
                          colour = "age"),
               position = position_jitter(w = 0.25, h = 0.25),
               size = 2.5) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("role vs.", var4y))
  plot(plotdat)
}
```

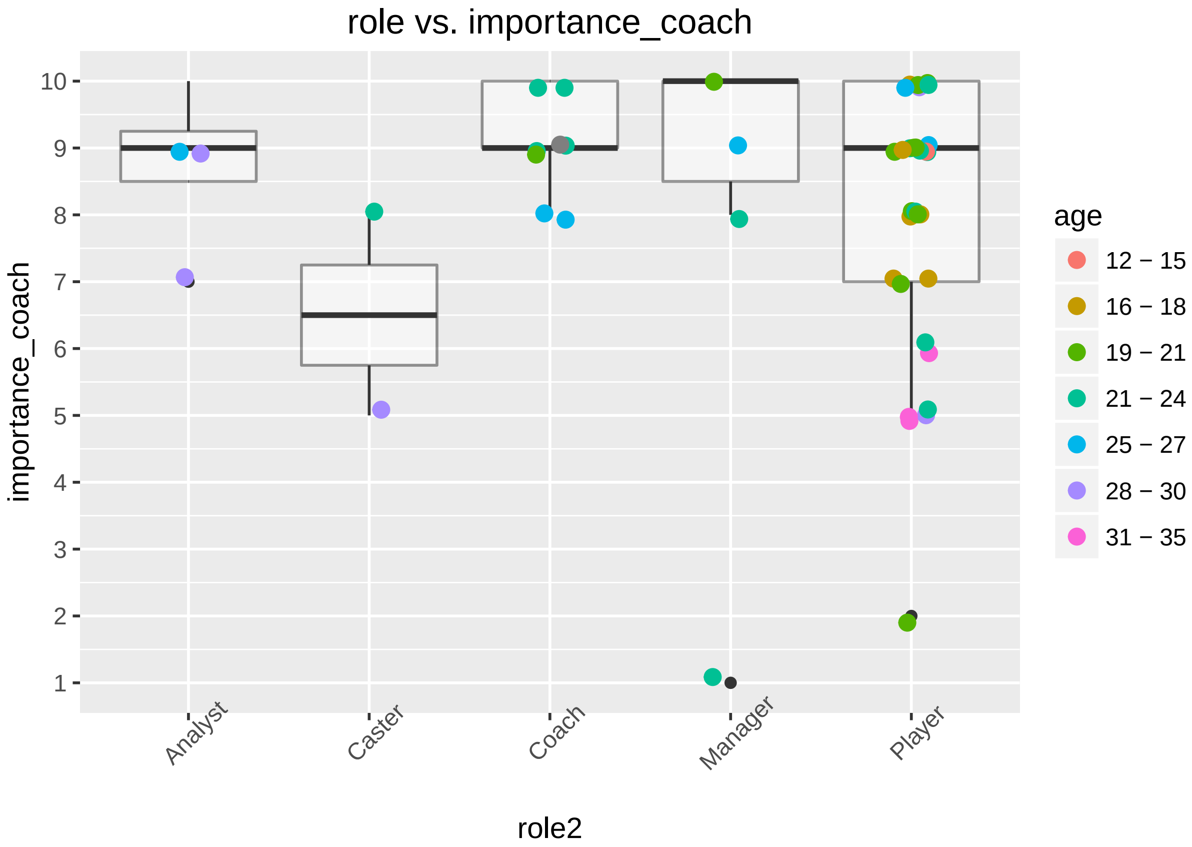
```
## Warning: Removed 14 rows containing missing values (geom_point).
```



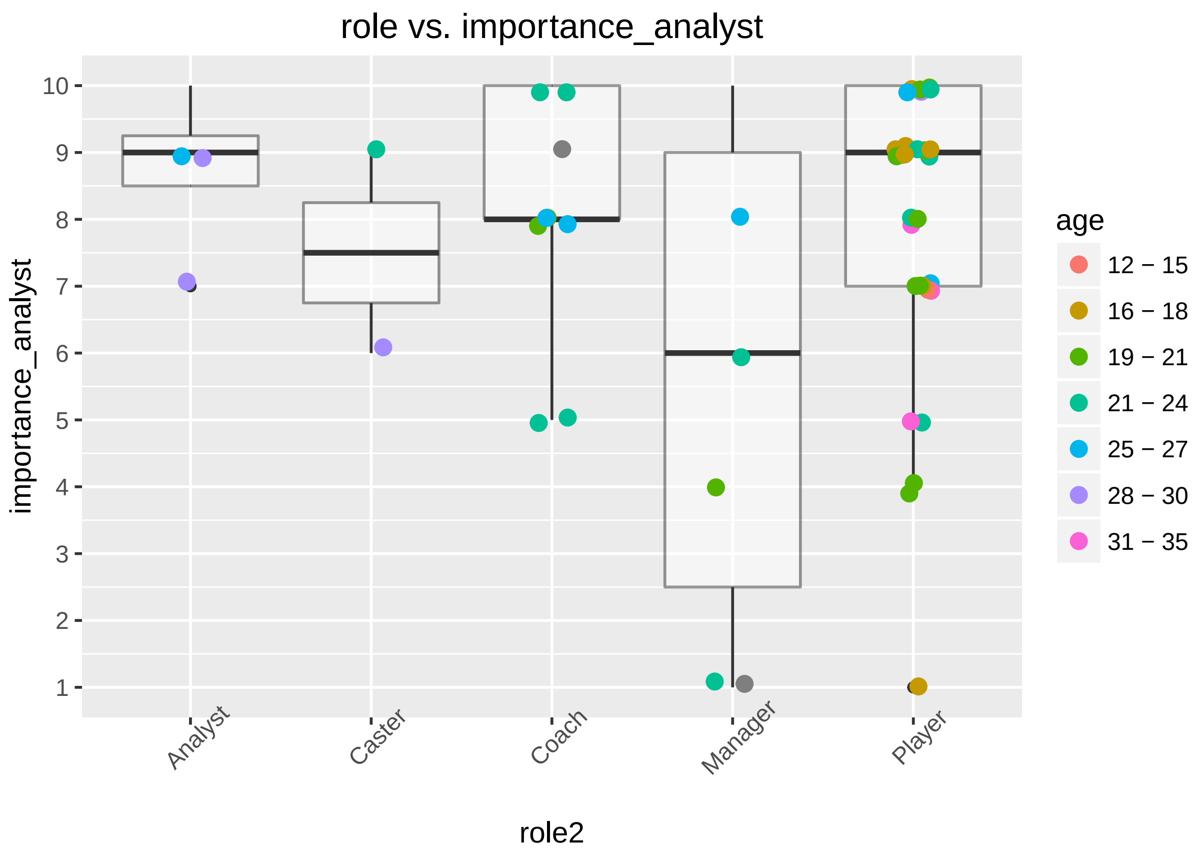
```
## Warning: Removed 8 rows containing missing values (geom_point).
```



```
## Warning: Removed 12 rows containing missing values (geom_point).
```

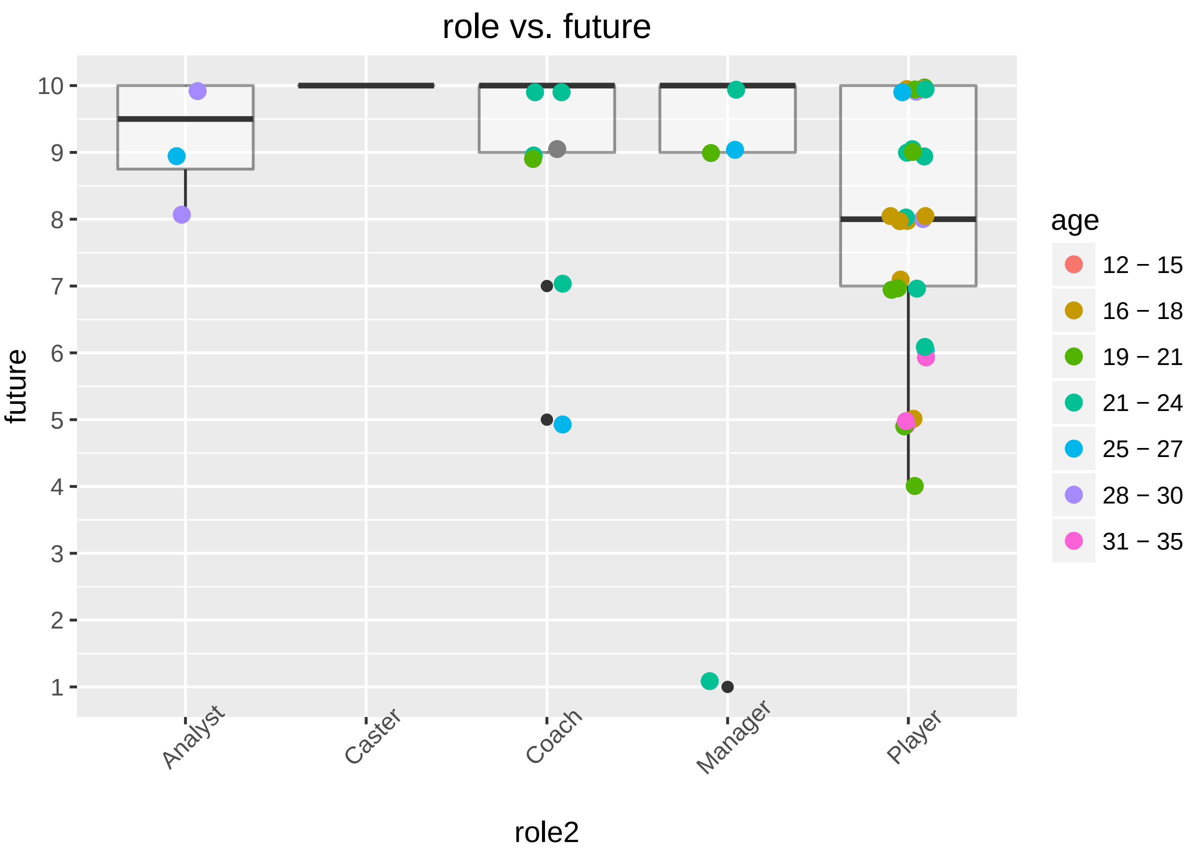


```
## Warning: Removed 9 rows containing missing values (geom_point).
```



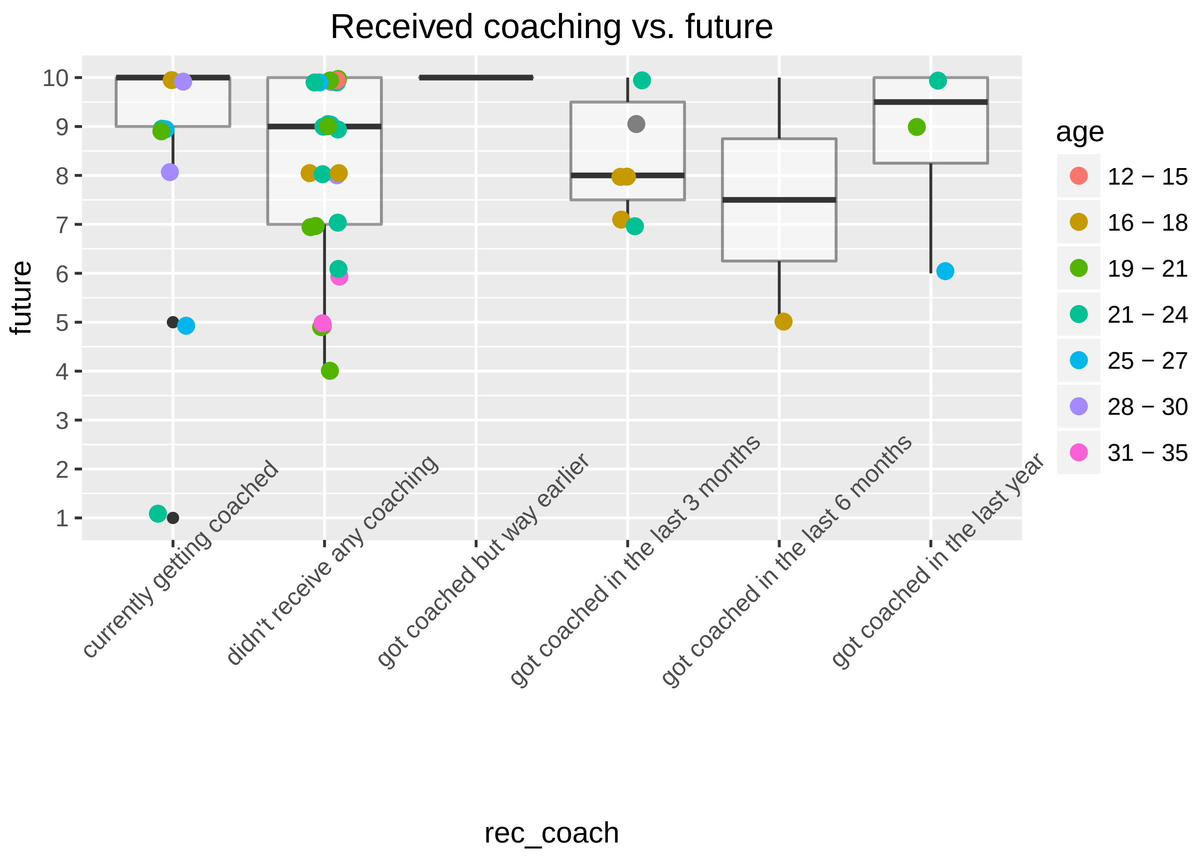
```
set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes(x = role2,
                     y = future),
                 alpha = 0.5) +
    geom_point(data = dat,
               aes(x = role2,
                   y = future,
                   colour = age),
               position = position_jitter(w = 0.25, h = 0.25),
               size = 2.5) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("role vs. future"))
  plot(plotdat)
```

```
## Warning: Removed 16 rows containing missing values (geom_point).
```



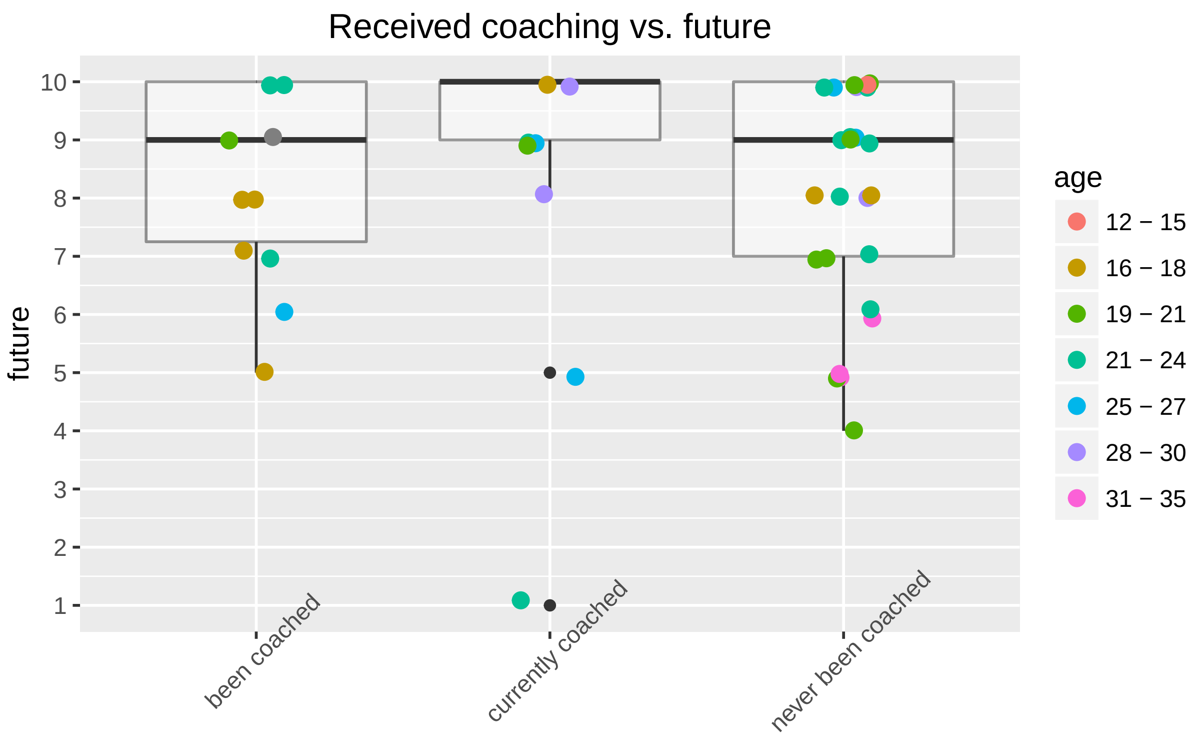
```r
set.seed(42)
plotdat <-
  ggplot() +
  geom_boxplot(data = dat,
               aes(x = rec_coach,
                   y = future),
               alpha = 0.5) +
  geom_point(data = dat,
             aes(x = rec_coach,
                 y = future,
                 colour = age),
             position = position_jitter(w = 0.25, h = 0.25),
             size = 2.5) +
  scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
  theme(axis.text.x = element_text(angle = 45)) +
  ggtitle(paste("Received coaching vs. future"))
plot(plotdat)
```

```
## Warning: Removed 16 rows containing missing values (geom_point).
```



```r
set.seed(42)
plotdat <-
  ggplot() +
  geom_boxplot(data = dat,
               aes(x = rec_coach2,
                   y = future),
               alpha = 0.5) +
  geom_point(data = dat,
             aes(x = rec_coach2,
                 y = future,
                 colour = age),
             position = position_jitter(w = 0.25, h = 0.25),
             size = 2.5) +
  scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
  theme(axis.text.x = element_text(angle = 45)) +
  ggtitle(paste("Received coaching vs. future"))
plot(plotdat)
```

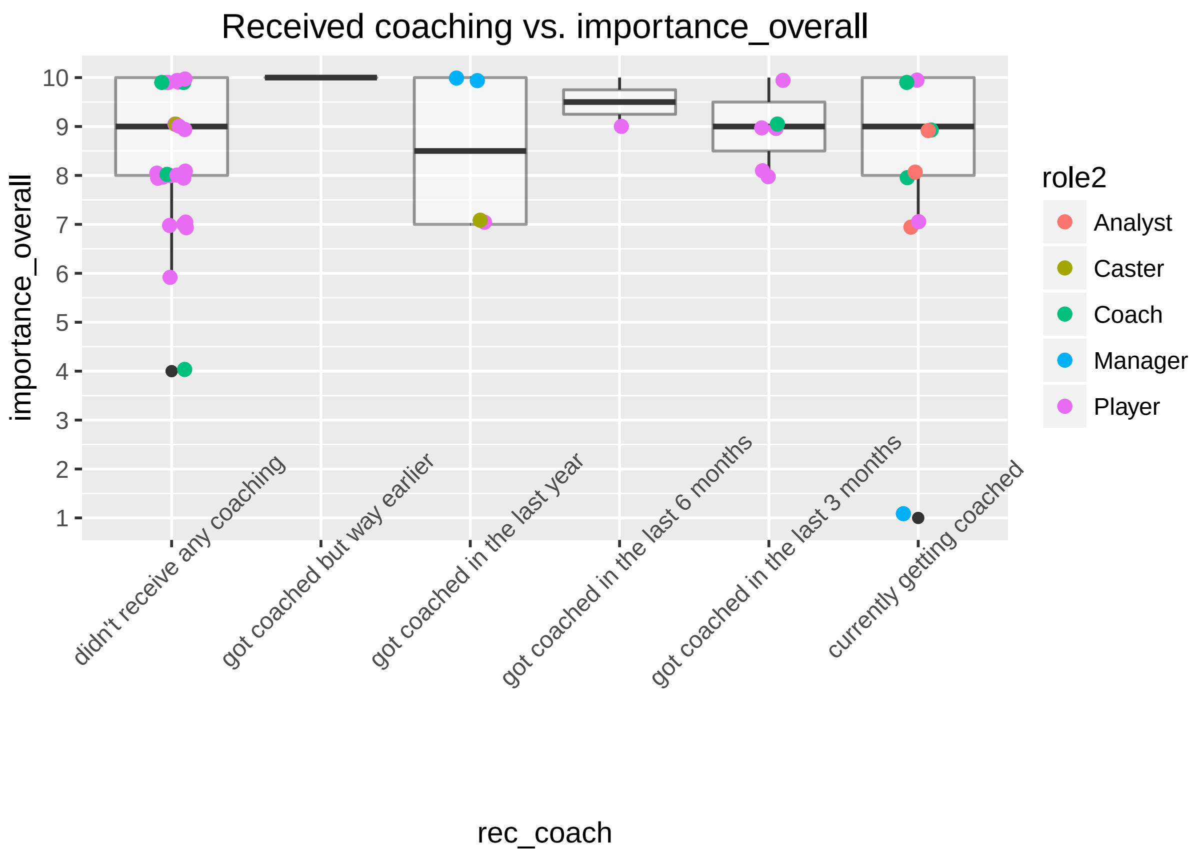
```
## Warning: Removed 16 rows containing missing values (geom_point).
```



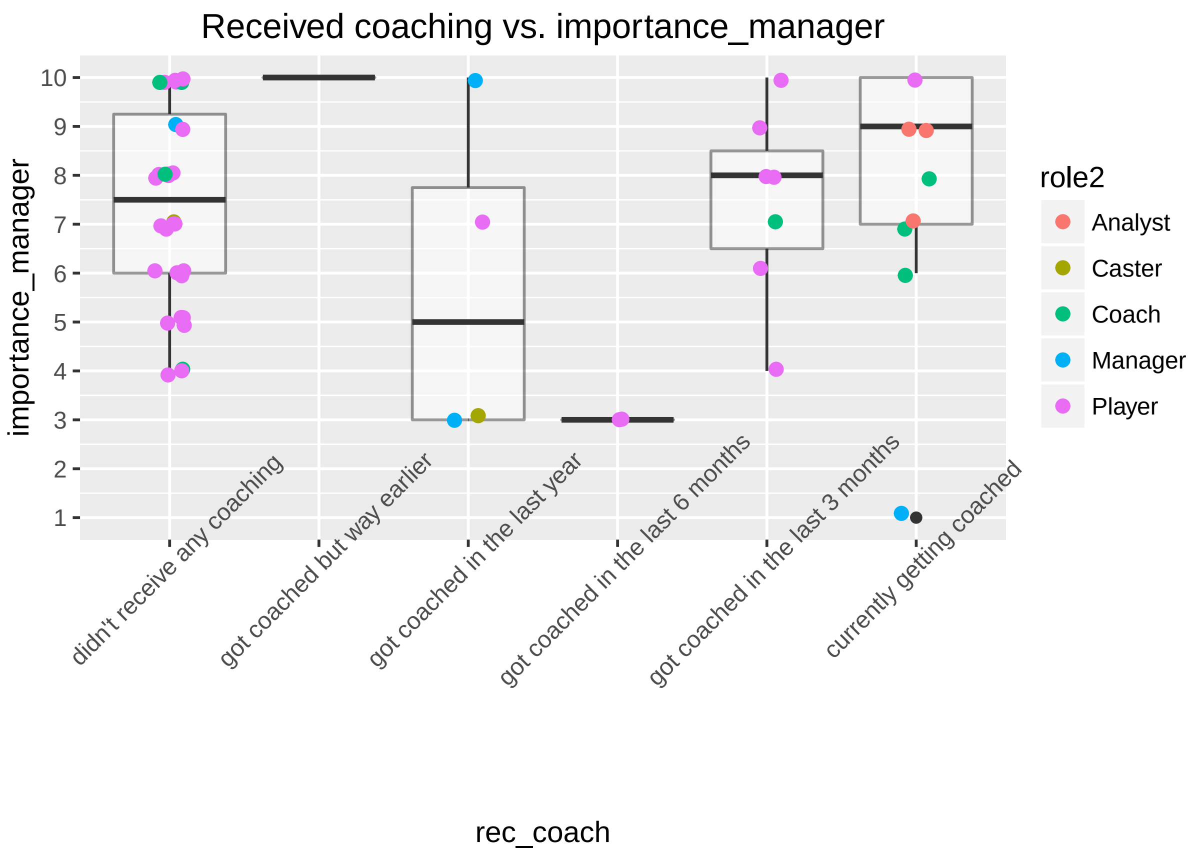
rec_coach2

```
dat$rec_coach <- factor(dat$rec_coach)
dat$rec_coach <- factor(dat$rec_coach, levels = levels(dat$rec_coach)[c(2, 3, 6, 5, 4, 1)])

for(var4y in var4y_list) {
  set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes_string(x = "rec_coach",
                            y = var4y),
                 alpha = 0.5) +
    geom_point(data = dat,
               aes_string(x = "rec_coach",
                          y = var4y,
                          colour = "role2"),
               position = position_jitter(w = 0.25, h = 0.25),
               size = 2) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("Received coaching vs.", var4y))
  plot(plotdat)
}
```

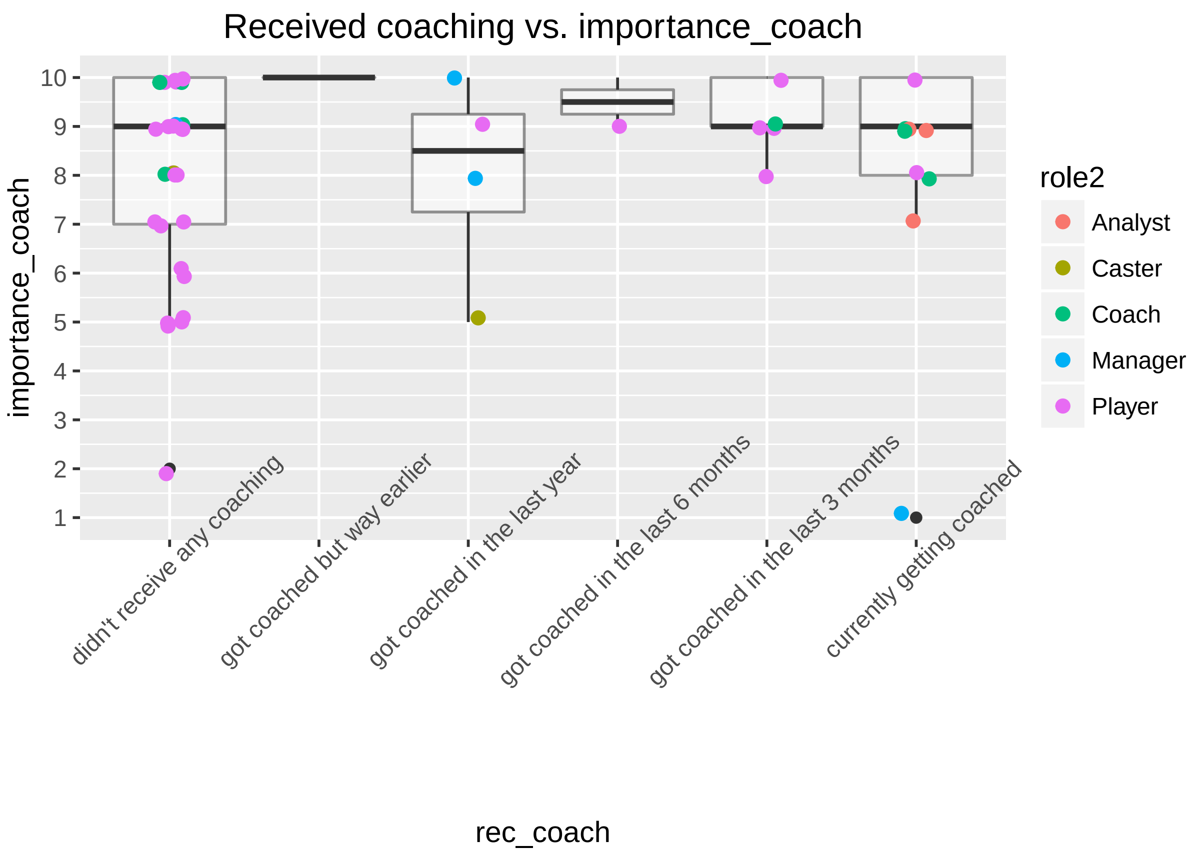
```
## Warning: Removed 14 rows containing missing values (geom_point).
```



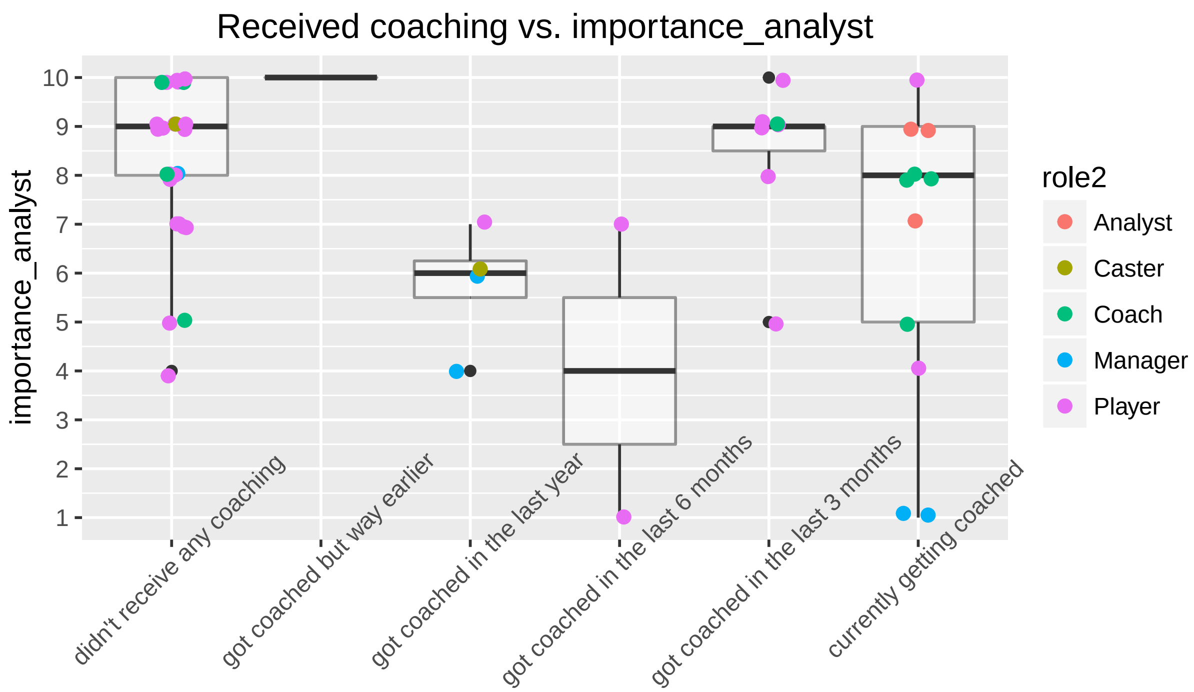
```
## Warning: Removed 8 rows containing missing values (geom_point).
```



```
## Warning: Removed 12 rows containing missing values (geom_point).
```



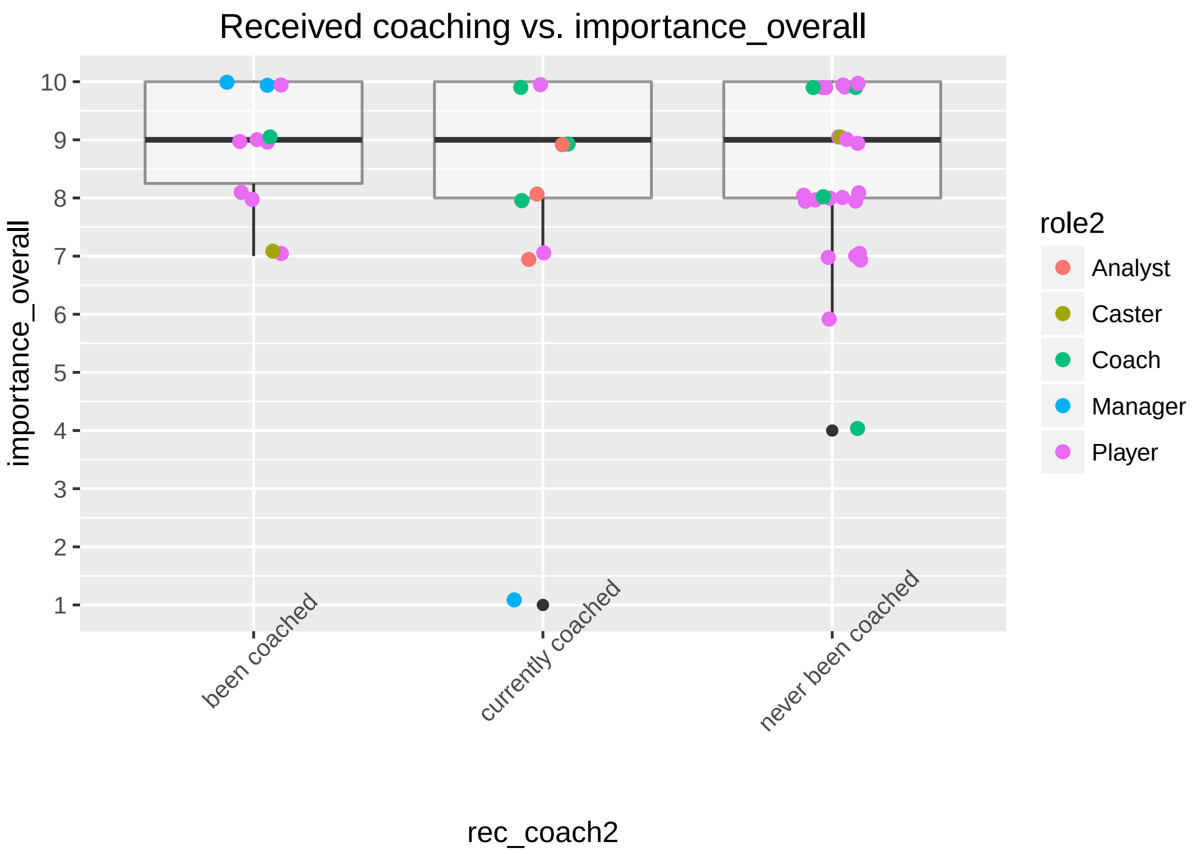
```
## Warning: Removed 9 rows containing missing values (geom_point).
```



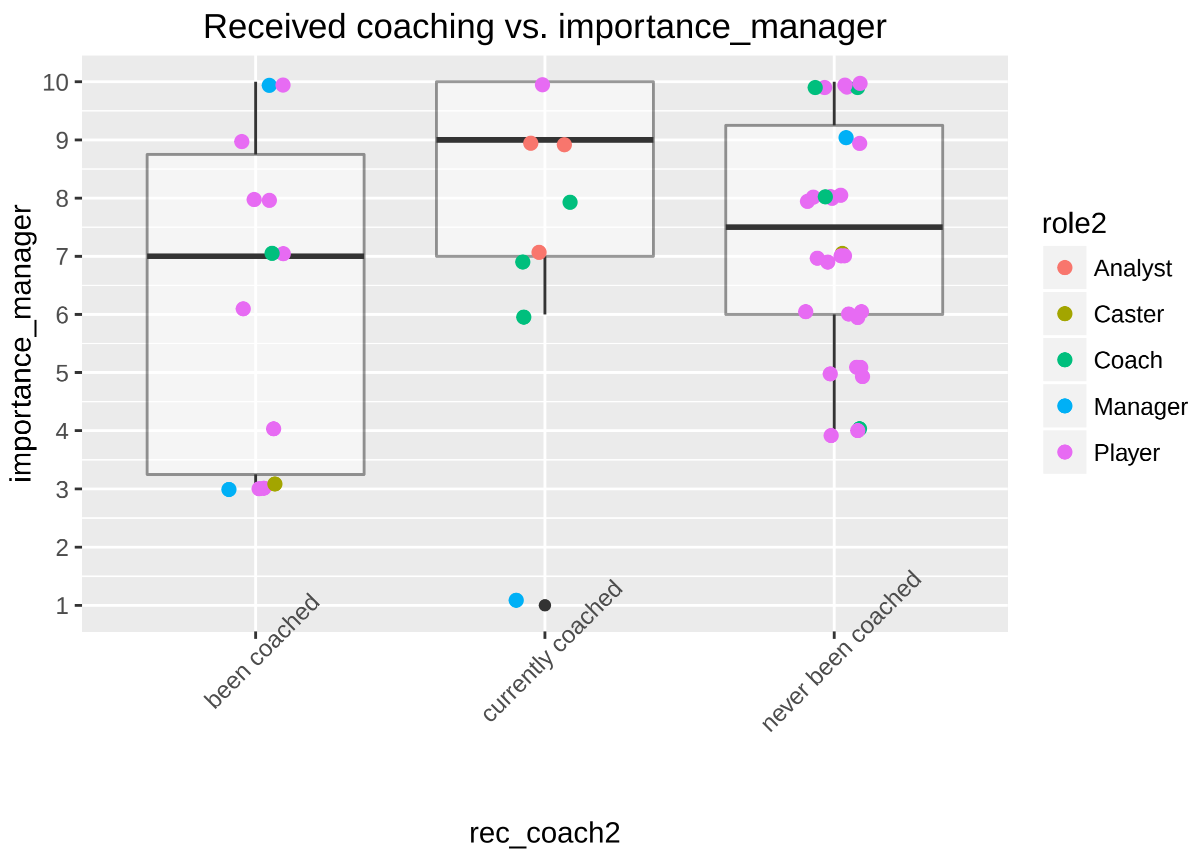
rec_coach

```
for(var4y in var4y_list) {
  set.seed(42)
  plotdat <-
    ggplot() +
    geom_boxplot(data = dat,
                 aes_string(x = "rec_coach2",
                            y = var4y),
                 alpha = 0.5) +
    geom_point(data = dat,
               aes_string(x = "rec_coach2",
                          y = var4y,
                          colour = "role2"),
               position = position_jitter(w = 0.25, h = 0.25),
               size = 2) +
    scale_y_continuous(limits = c(1, 10), breaks = seq(1, 10, 1)) +
    theme(axis.text.x = element_text(angle = 45)) +
    ggtitle(paste("Received coaching vs.", var4y))
  plot(plotdat)
}
```

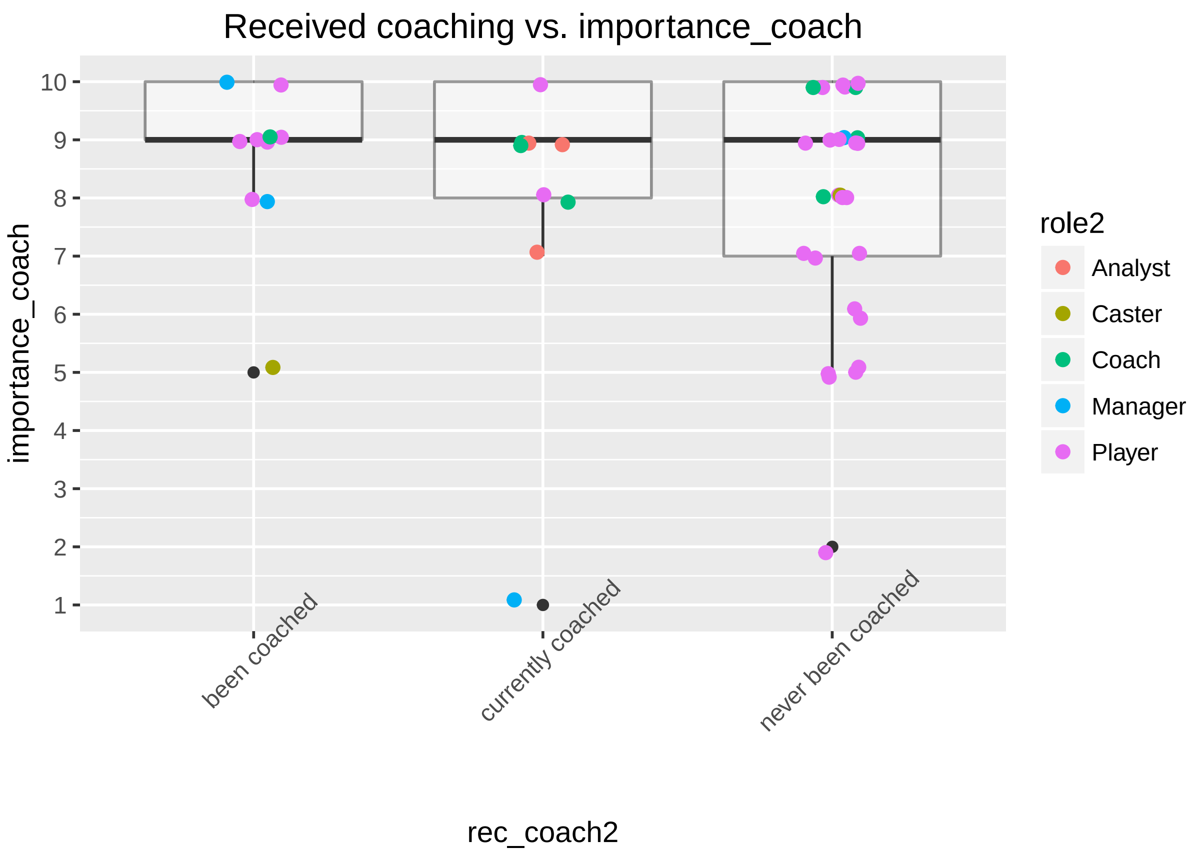
```
## Warning: Removed 14 rows containing missing values (geom_point).
```



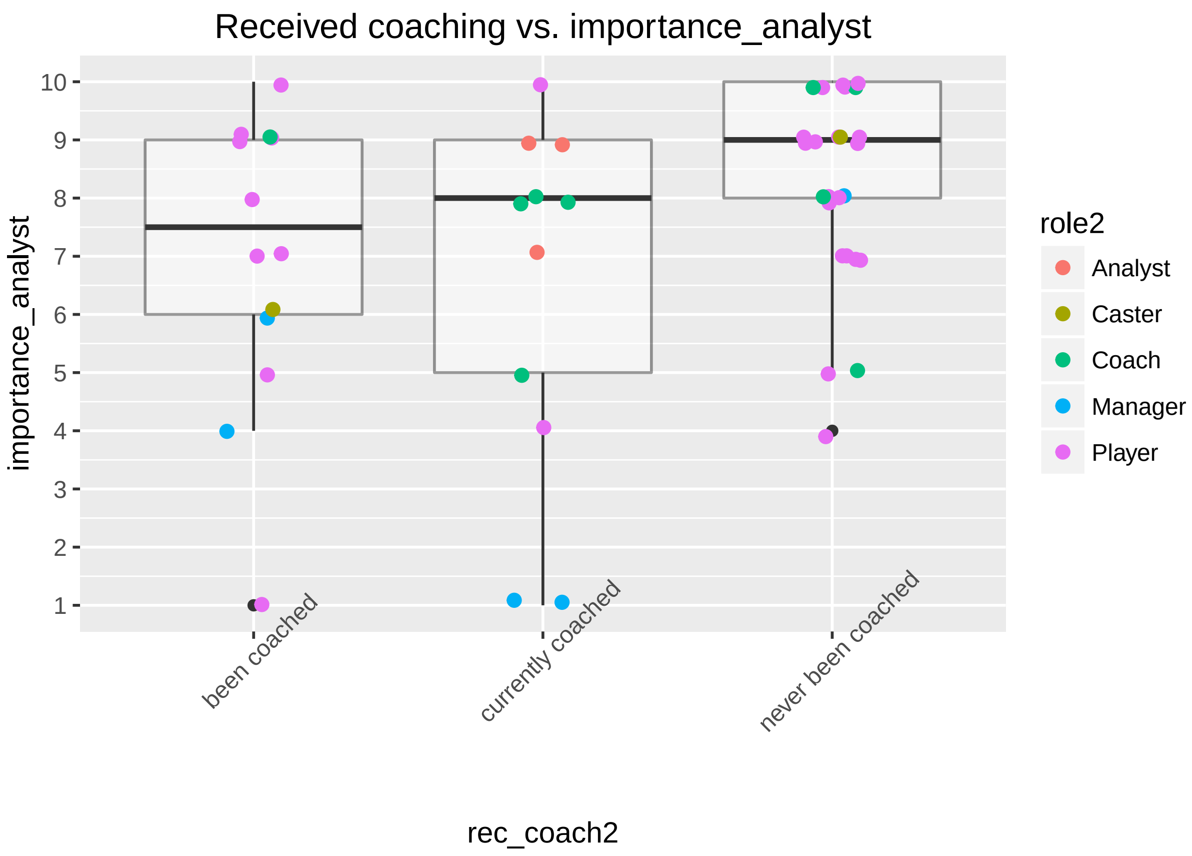
```
## Warning: Removed 8 rows containing missing values (geom_point).
```



```
## Warning: Removed 12 rows containing missing values (geom_point).
```



```
## Warning: Removed 9 rows containing missing values (geom_point).
```


Received coaching vs. importance_analyst
importance_analyst
10
9
8
7
6
5
4
3
2
1
been coached
currently coached
never been coached
rec_coach2
role2
Analyst
Caster
Coach
Manager
Player