Christmas soda test 2025 Rmarkdown report

2025-10-03

The Christmas Soda test webpage/dashboard

This Christmas Soda Test also has its own webpage/dashboard! Click here to have a look.

The webpage is interactive and lets you:

- View the data analysis of the Christmas Soda Test (Page 1).
- Discover Christmas sodas you might also like based on your favorite Christmas soda. This recommendation is based on the result from this test (Page 2).
- Examine the raw data from the test (Page 3)

Info about the Christmas Soda test

At this page you'll find the data analysis of the Christmas Soda test. This is an analysis of a Christmas Soda test in 2023. This Christmas Soda test was arranged among friends, and made for fun. In the test, 27 different Christmas sodas were scored by 16 participants. Each Christmas soda was rated from 1 to 6 with 6 being the best. The Christmas sodas had different colors that included red, brown, green and orange. Both females and males participated in the test, and the participants were seated at two different tables. The organizer of the Christmas Soda test did not sit at a table, and therefore this person's seating is labeled with NA.

Import data and cleand and organize the data

```
# Import data and clean and organize the data ----------------------------------
brus_data_pca = read.csv("data/brus_data_pca.csv")
brus_data_pca = brus_data_pca %>% rename("Person" = "X")
brus_data_pca = brus_data_pca %>% rename("7_fjell" = "X7_fjell")
brus_data_pca = brus_data_pca %>% rename("table_seating" = "bordplassering")
brus_data_pca$table_seating = as.factor(brus_data_pca$table_seating)

brus_data_pca_longer = brus_data_pca %>% pivot_longer(cols = "Coop_Raud":"Tante_hedwigs",
                                                      names_to = "soda",
                                                      values_to = "score")

brus_data_soda_rows = read.csv("data/data_brustype.csv")
brus_data_soda_rows = brus_data_soda_rows %>% rename("Soda" = "X")
brus_data_soda_rows = brus_data_soda_rows %>%
  mutate(across(c(Person1:Person16), as.numeric))
rownames(brus_data_soda_rows) = brus_data_soda_rows$Soda
brus_data_soda_rows$brus_color[brus_data_soda_rows$brus_color == "brown"] = "#895129"
```

```r
#The brown color was to similar to red.

glimpse(brus_data_pca_longer)
```

```
## Rows: 432
## Columns: 7
## $ Person        <chr> "Person1", "Person1", "Person1", "Person1", "Person1", "~
## $ table_seating <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ gender        <chr> "male", "male", "male", "male", "male", "male", "male", ~
## $ PC1           <dbl> 0.4621666, 0.4621666, 0.4621666, 0.4621666, 0.4621666, 0~
## $ PC2           <dbl> -0.3237505, -0.3237505, -0.3237505, -0.3237505, -0.32375~
## $ soda          <chr> "Coop_Raud", "7_fjell", "Rudolf_og_nisseMOR", "Rampeniss~
## $ score         <int> 3, 3, 4, 2, 3, 2, 2, 4, 1, 4, 2, 2, 1, 5, 1, 3, 5, 3, 3,~
```

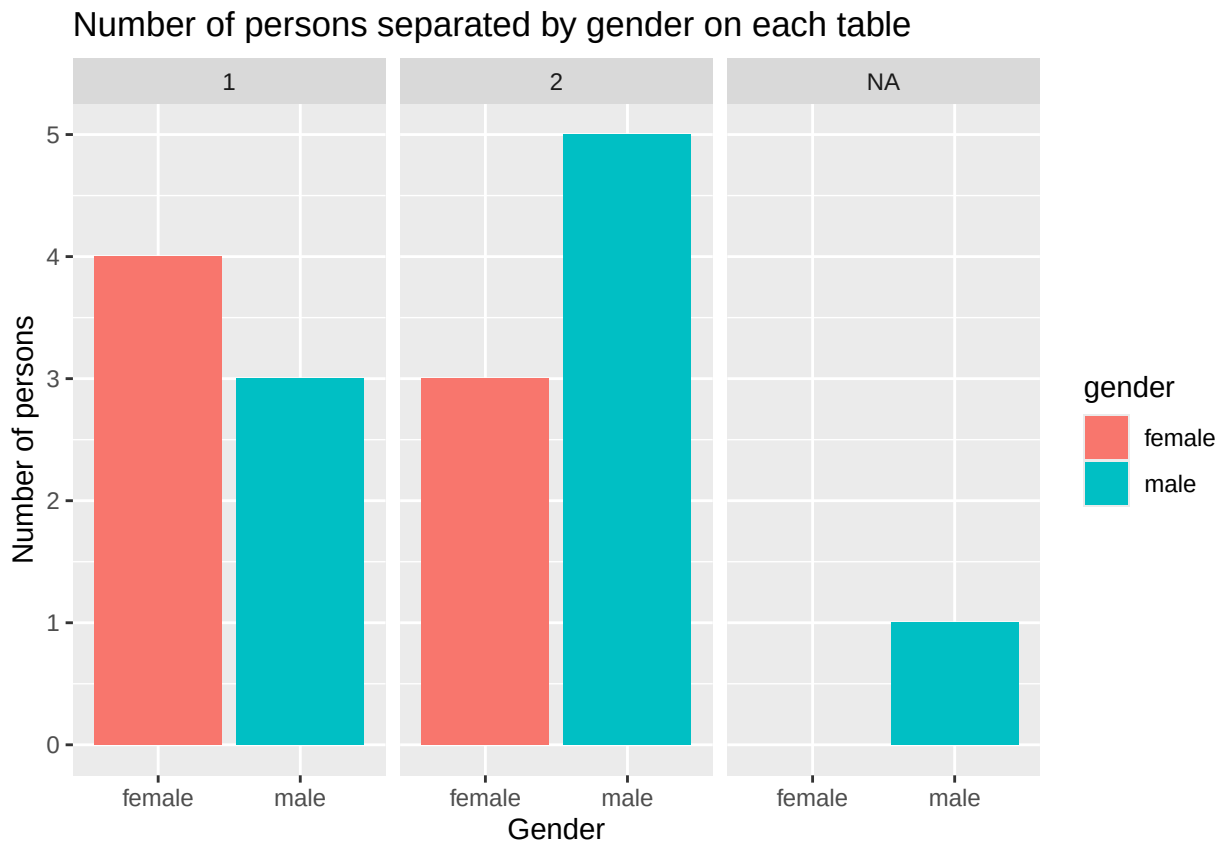
```r
glimpse(brus_data_soda_rows)
```

```
## Rows: 27
## Columns: 19
## $ Soda       <chr> "Coop_Raud", "7_fjell", "Rudolf_og_nisseMOR", "Rampenissen"~
## $ Person1    <dbl> 3, 3, 4, 2, 3, 2, 2, 4, 1, 4, 2, 2, 1, 5, 1, 3, 5, 3, 3, 3,~
## $ Person2    <dbl> 5, 3, 5, 4, 6, 4, 2, 5, 2, 4, 3, 3, 1, 5, 2, 3, 4, 3, 3, 5,~
## $ Person3    <dbl> 3, 2, 4, 2, 5, 3, 2, 4, 1, 5, 1, 2, 1, 4, 1, 3, 3, 3, 5, 4,~
## $ Person4    <dbl> 6, 2, 4, 3, 4, 2, 4, 6, 1, 4, 3, 2, 4, 3, 1, 6, 5, 5, 3, 5,~
## $ Person5    <dbl> 5, 3, 4, 2, 5, 2, 2, 4, 3, 4, 1, 1, 2, 2, 1, 3, 4, 3, 3, 5,~
## $ Person6    <dbl> 1, 2, 2, 3, 1, 2, 1, 2, 4, 3, 1, 1, 2, 3, 2, 1, 3, 3, 1, 3,~
## $ Person7    <dbl> 2, 3, 4, 3, 5, 1, 2, 5, 3, 4, 1, 1, 3, 3, 1, 3, 4, 2, 6, 4,~
## $ Person8    <dbl> 2, 2, 3, 4, 6, 2, 2, 4, 1, 5, 2, 5, 3, 5, 1, 2, 4, 3, 3, 4,~
## $ Person9    <dbl> 3, 2, 3, 2, 5, 2, 3, 5, 1, 4, 2, 3, 1, 5, 1, 3, 5, 2, 4, 3,~
## $ Person10   <dbl> 5, 3, 5, 2, 5, 2, 6, 6, 1, 4, 1, 5, 2, 6, 1, 4, 6, 2, 2, 4,~
## $ Person11   <dbl> 3, 2, 5, 4, 6, 2, 2, 5, 1, 4, 4, 2, 3, 6, 2, 3, 6, 2, 3, 2,~
## $ Person12   <dbl> 3, 2, 6, 2, 5, 3, 5, 5, 1, 4, 3, 5, 3, 4, 1, 3, 4, 3, 5, 4,~
## $ Person13   <dbl> 3, 2, 4, 3, 6, 2, 6, 5, 1, 4, 1, 6, 2, 5, 1, 5, 4, 3, 5, 3,~
## $ Person14   <dbl> 4, 3, 4, 4, 4, 2, 5, 5, 1, 4, 2, 5, 3, 5, 1, 3, 4, 3, 4, 3,~
## $ Person15   <dbl> 2, 4, 4, 3, 6, 2, 2, 5, 1, 4, 2, 1, 3, 4, 4, 2, 6, 5, 4, 5,~
## $ Person16   <dbl> 4, 2, 5, 2, 6, 2, 2, 6, 1, 3, 2, 2, 3, 4, 1, 3, 6, 3, 1, 4,~
## $ brus_color <chr> "red", "red", "red", "green", "red", "red", "red", "red", "~
## $ mean_score <dbl> 3.3750, 2.5000, 4.1250, 2.8125, 4.8750, 2.1875, 3.0000, 4.7~
```

Overview of the dataset

```r
# Summary of gender on each table
 gender_table = brus_data_pca %>%
   group_by(gender,
            table_seating) %>%
   tally()

 gender_table_plot = gender_table %>% ggplot(aes(x = gender, y = n, fill = gender)) +
   geom_bar(stat = "identity") +
   facet_wrap(~ table_seating) +
   xlab("Gender") +
```

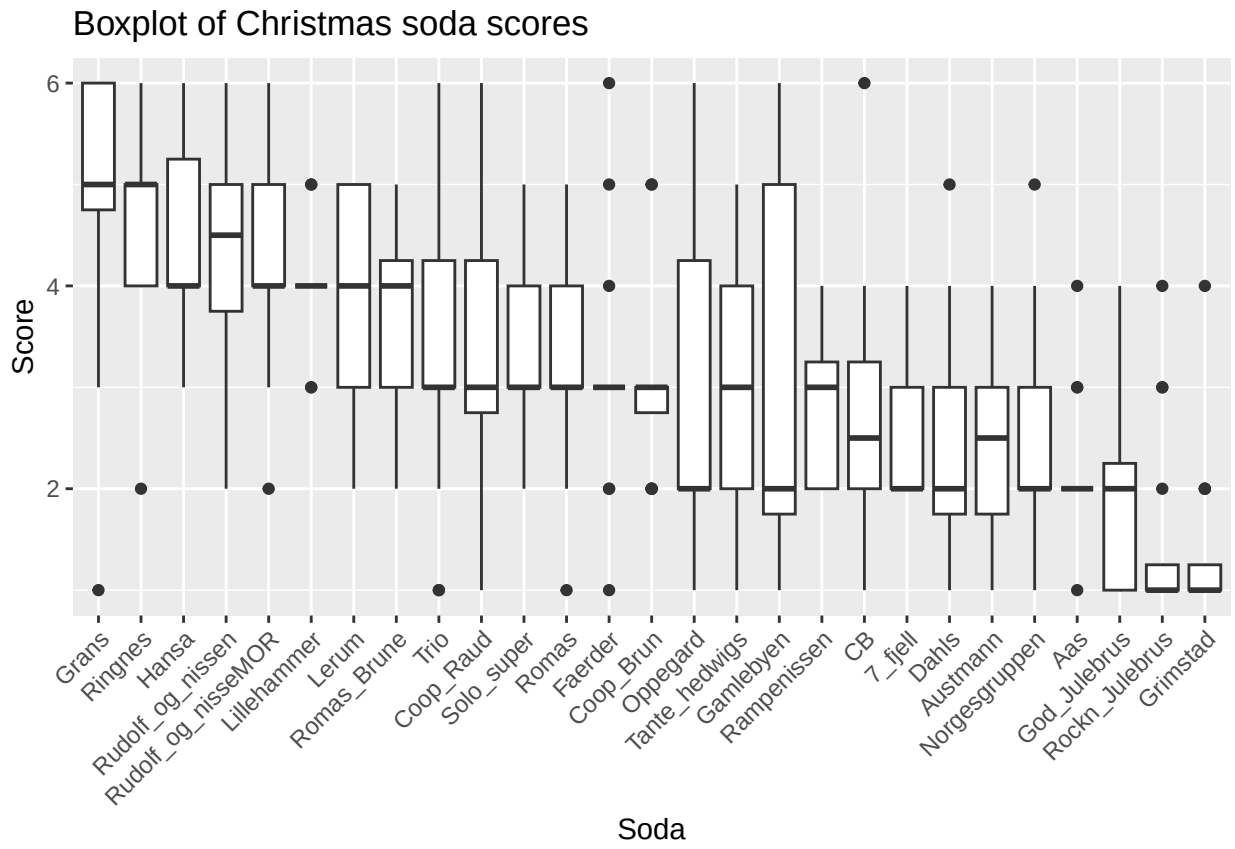
```
    ylab("Number of persons") +
    ggtitle("Number of persons separated by gender on each table")

gender_table_plot
```



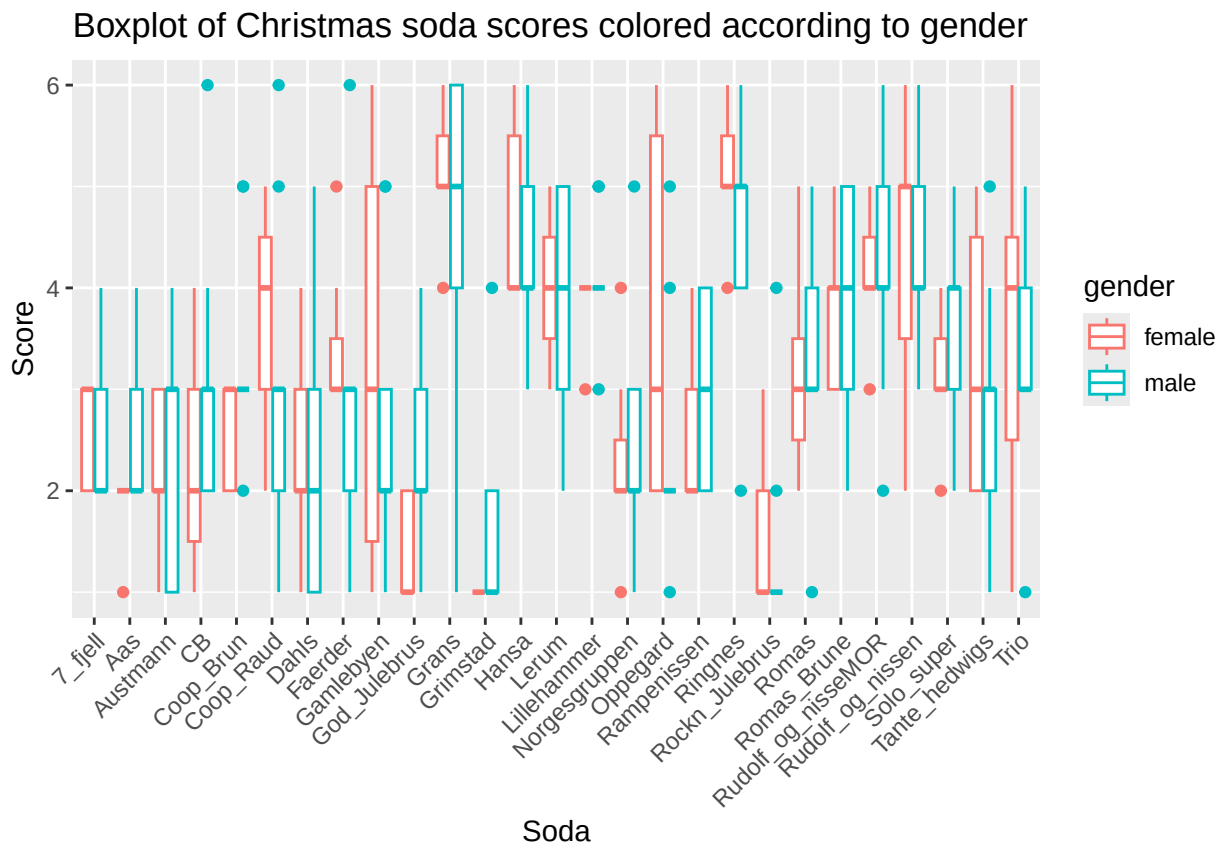
On table 1 there were 4 females and 3 males, while on table 2 there were 3 females and 5 males. One male did not sit with a table, and was in charge of the Christmas soda test. Therefore the table for this person was assigned with NA.

Overview of the christmas soda scores

```
boxplot_score_all_plot = brus_data_pca_longer %>%
 ggplot(aes(x = reorder(soda, score, decreasing = T), y = score)) +
   geom_boxplot() +
   xlab("Soda") +
   ylab("Score") +
   ggtitle("Boxplot of Christmas soda scores") +
   theme(axis.text.x = element_text(angle = 45, hjust = 1))

boxplot_score_all_plot
```

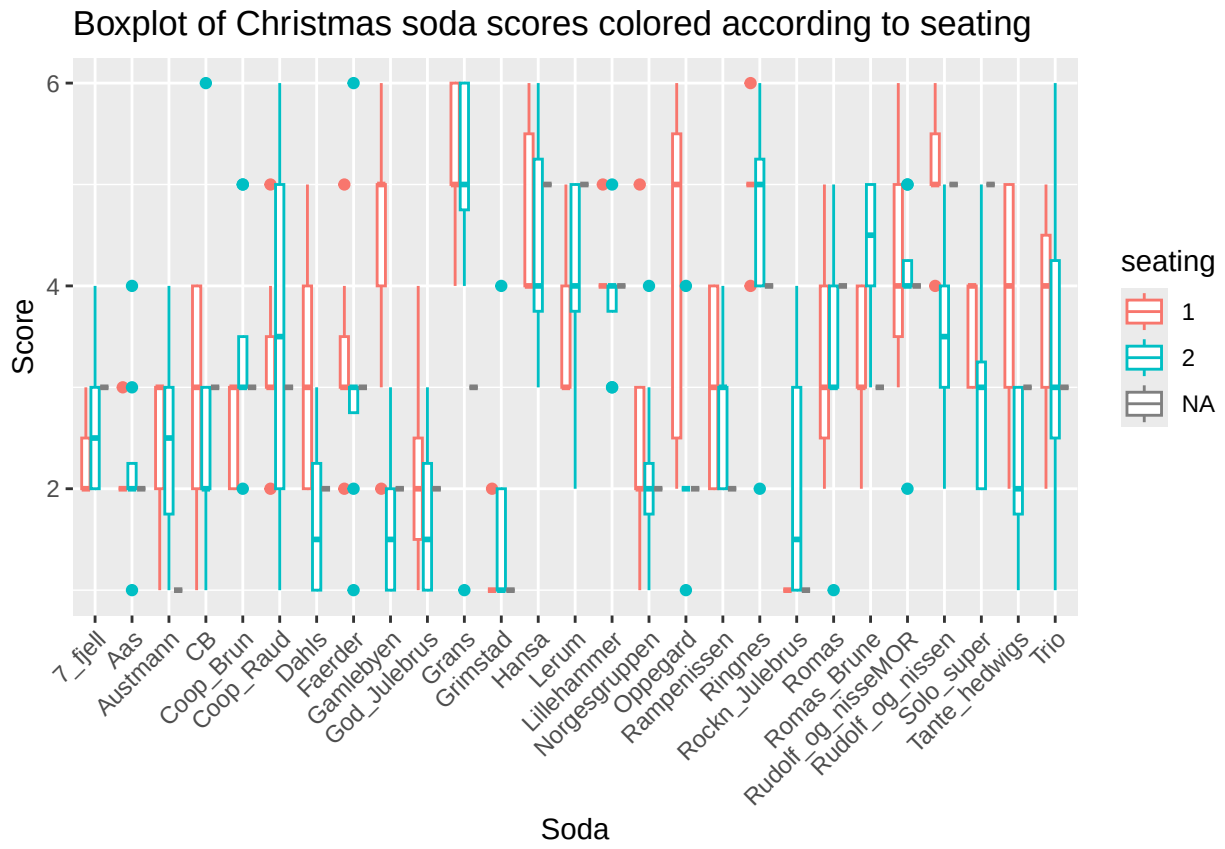



```
# Plot Christmas Soda score colored after gender

    ggplot(brus_data_pca_longer, aes(x = soda, y = score, colour = gender)) +
      geom_boxplot() +
      xlab("Soda") +
      ylab("Score") +
      ggtitle("Boxplot of Christmas soda scores colored according to gender") +
      theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



The winner of this year's Christmas soda test was Grans!

```
# Plot Christmas Soda score colored after seating

    ggplot(brus_data_pca_longer, aes(x = soda, y = score, colour = table_seating)) +
      geom_boxplot() +
      xlab("Soda") +
      ylab("Score") +
      ggtitle("Boxplot of Christmas soda scores colored according to seating") +
      theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
      labs(color = "seating" )
```

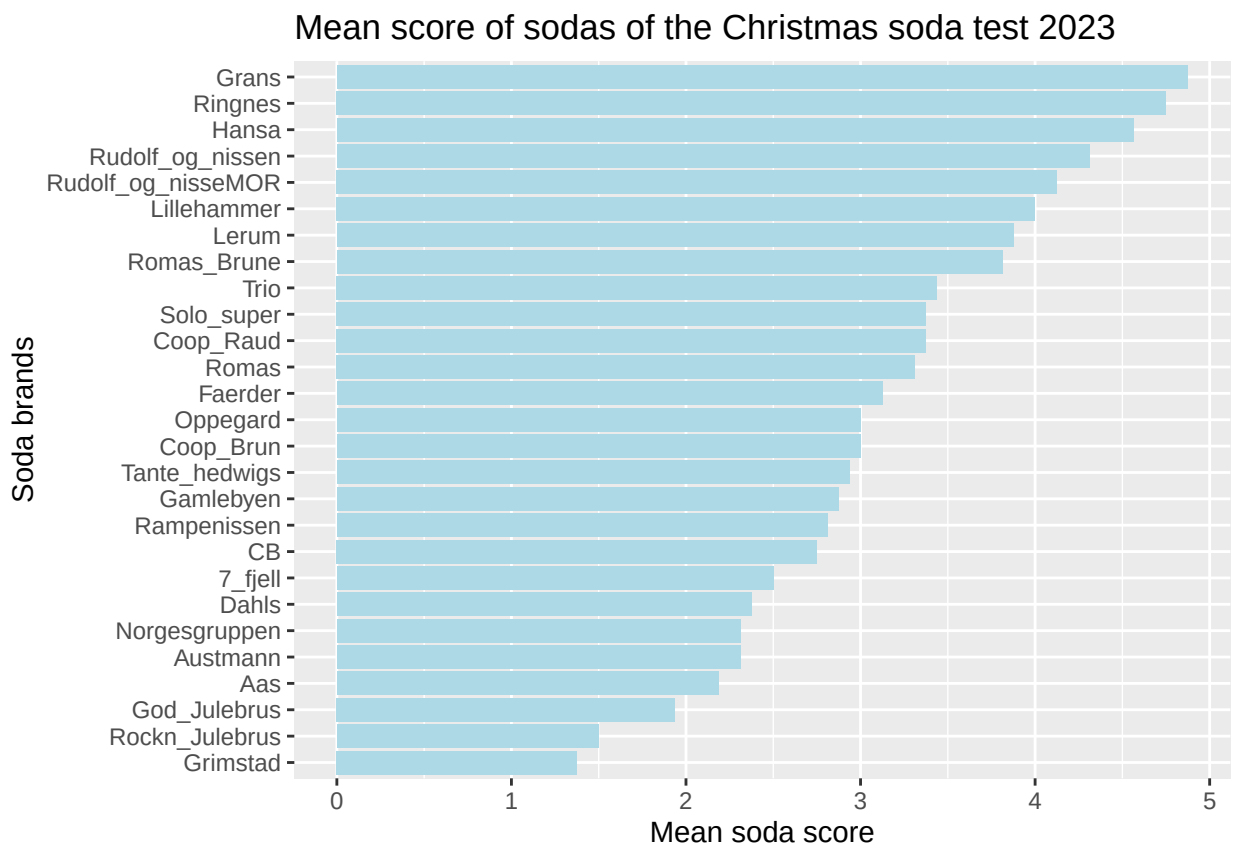


When separating based on table seating, we can spot that there are some differences in which sodas the participants at the two different tables prefered. This is evident for several Christmas sodas such as Dahls, Rudolf and Nissen, Solo Super and Trio that table 1 gave a higher score than table 2. The difference was especially evident for Gamlebyen, Oppegard and Tante Hedwigs. While table 2 gave a higher score for Christmas sodas such as 7 fjell, Coop Raud, Lerum and Romas Brune.

The mean and median christmas sodas scores in order

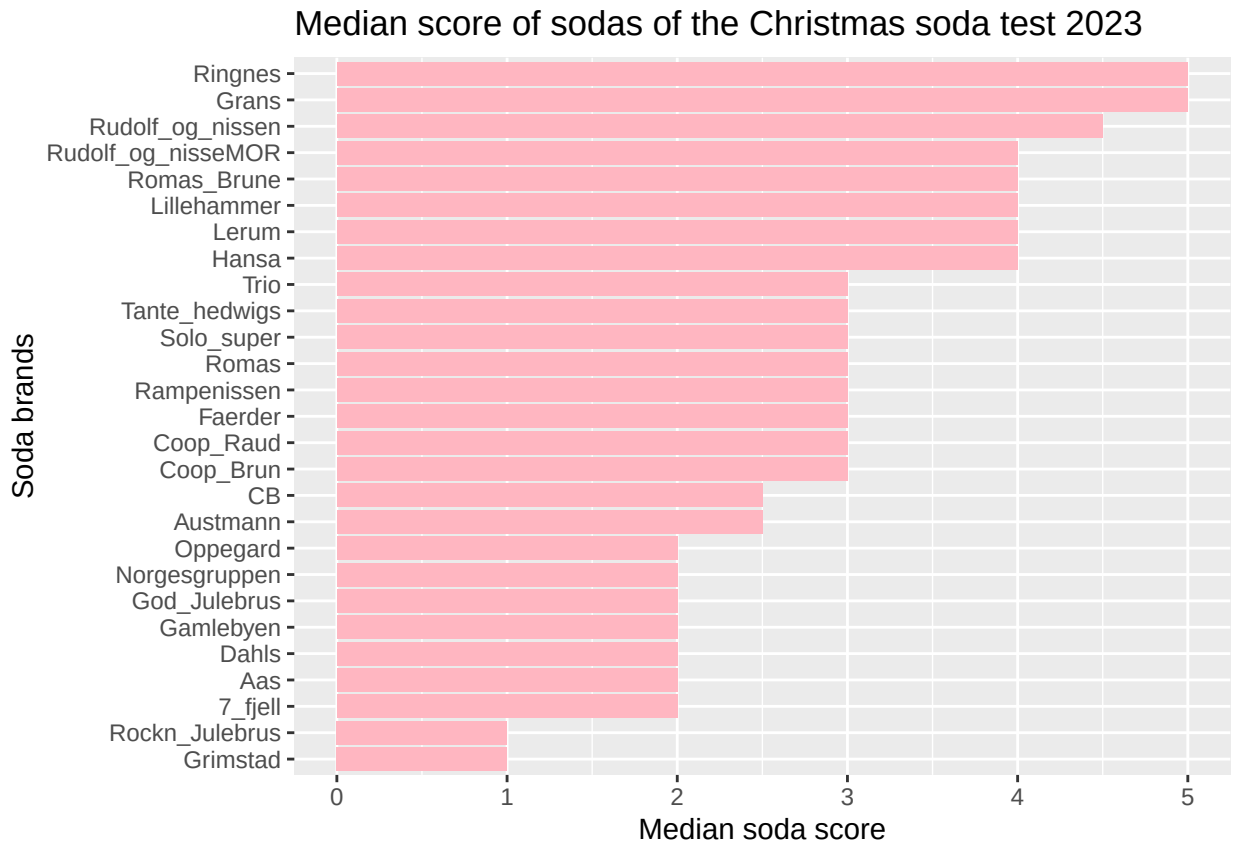
Mean Christmas Soda scores.

```
## Mean soda scores
 mean_soda_scores = brus_data_pca_longer %>%
    group_by(soda) %>%
    summarise(mean_score = mean(score),
              median_score = median(score))

 mean_soda_scores_plot = mean_soda_scores %>%
   ggplot(aes(x = reorder(soda, mean_score, decreasing = F),
                                                    y = mean_score)) +
   geom_bar(stat = "identity", fill = "lightblue") +
   labs(y = "Mean soda score", x = "Soda brands",
        title = "Mean score of sodas of the Christmas soda test 2023") +
   coord_flip()

 mean_soda_scores_plot
```



Median Christmas Soda scores

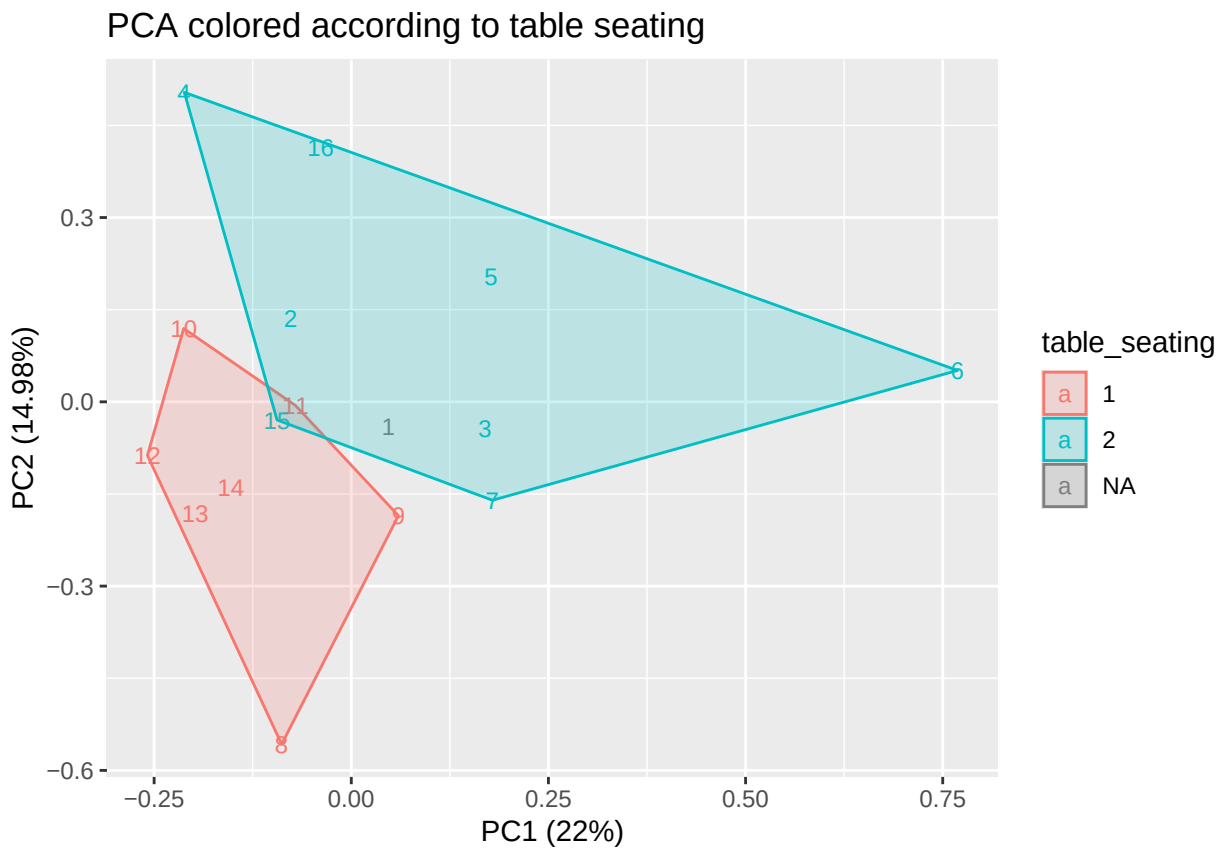
```
## Median soda scores

median_soda_scores_plot = mean_soda_scores %>%
 ggplot(aes(x = reorder(soda, median_score, decreasing = F),
                                                    y = median_score)) +
  geom_bar(stat = "identity", fill = "lightpink") +
  labs(y = "Median soda score", x = "Soda brands",
       title = "Median score of sodas of the Christmas soda test 2023") +
  coord_flip()

median_soda_scores_plot
```



PCA of christmas soda data with persons in focus

PCA colored after table seating

```
# PCA of persons in the test based on their Christmas soda scores colored after
# seating.

   pc = prcomp(brus_data_pca[,-c(1,29:32)],
               scale. = T) # only select columns that includes soda types.

   brus_data_pc = cbind(brus_data_pca, pc$x[,1:2])

   pc_autoplot_table_seating = autoplot(pc,
                                        data = brus_data_pca,
                                        colour = "table_seating",
                                        shape = F,
                                        label.size = 3,
                                        frame = T) +
     ggtitle("PCA colored according to table seating")

   pc_autoplot_table_seating
```



In this PCA plot each number represents a participant. E.g. 1 represents person 1.

This PCA plot describes an interesting finding. Based on this plot it seems like that the table seating played an important role in what score the different sodas obtained. In this plot we can see that the individuals from the same tables tended to score the Christmas soda in a more similar way. Suggesting that the participants have had an impact on each others scoring

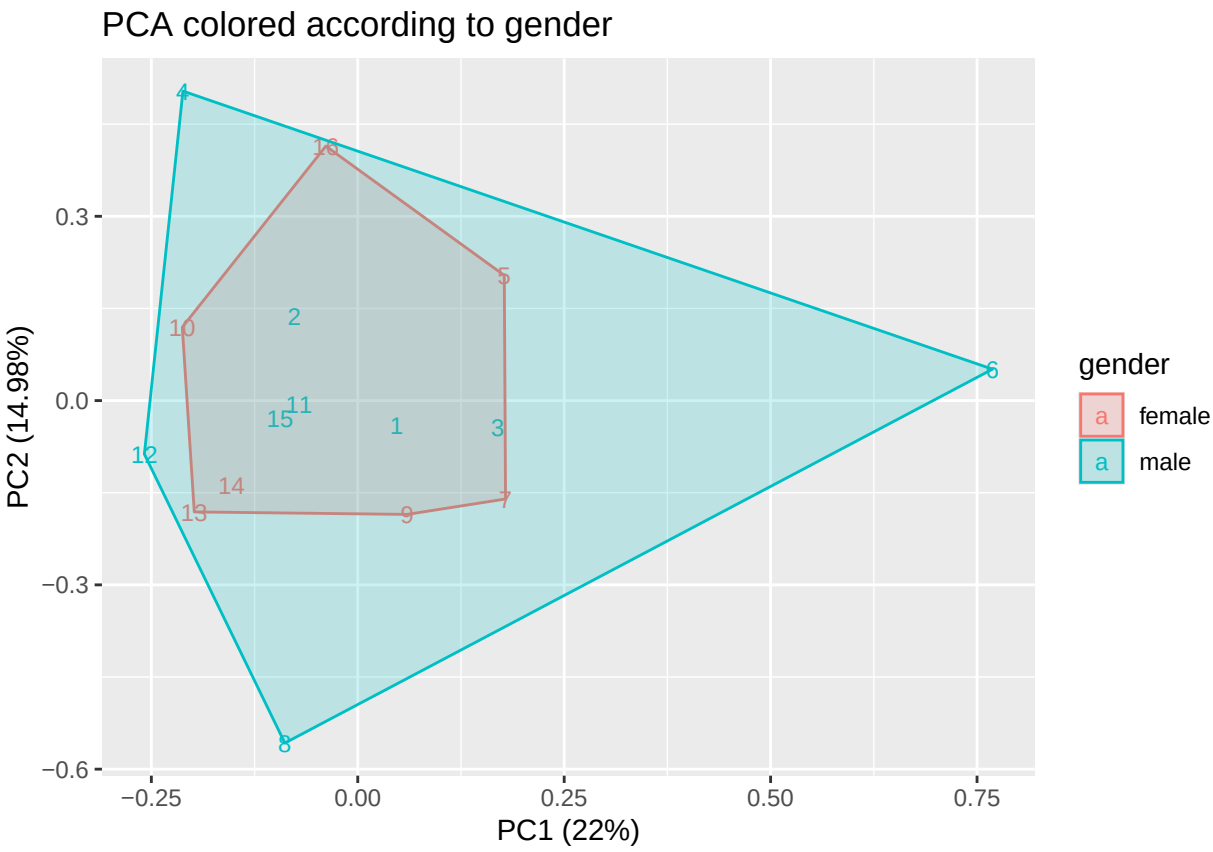
An interesting observations is that person 6 had a distinct scoring pattern compared to the rest of the participants, and are located to the very left in the PCA plot.

Person 1 that did not sit with a table, had similar scoring pattern as the participants seated at table 2.

PCA colored after gender

```
 # PCA of persons in the test based on their Christmas soda scores colored after
# gender.

  pc = prcomp(brus_data_pca[,-c(1,29:32)],
              scale. = T) # only select columns that includes soda types.

  brus_data_pc = cbind(brus_data_pca, pc$x[,1:2])

  pc_autoplot_gender = autoplot(pc,
                                        data = brus_data_pca,
                                        colour = "gender",
                                        shape = F,
```

```
                                        label.size = 3,
                                        frame = T) +
    ggtitle("PCA colored according to gender")

  pc_autoplot_gender
```



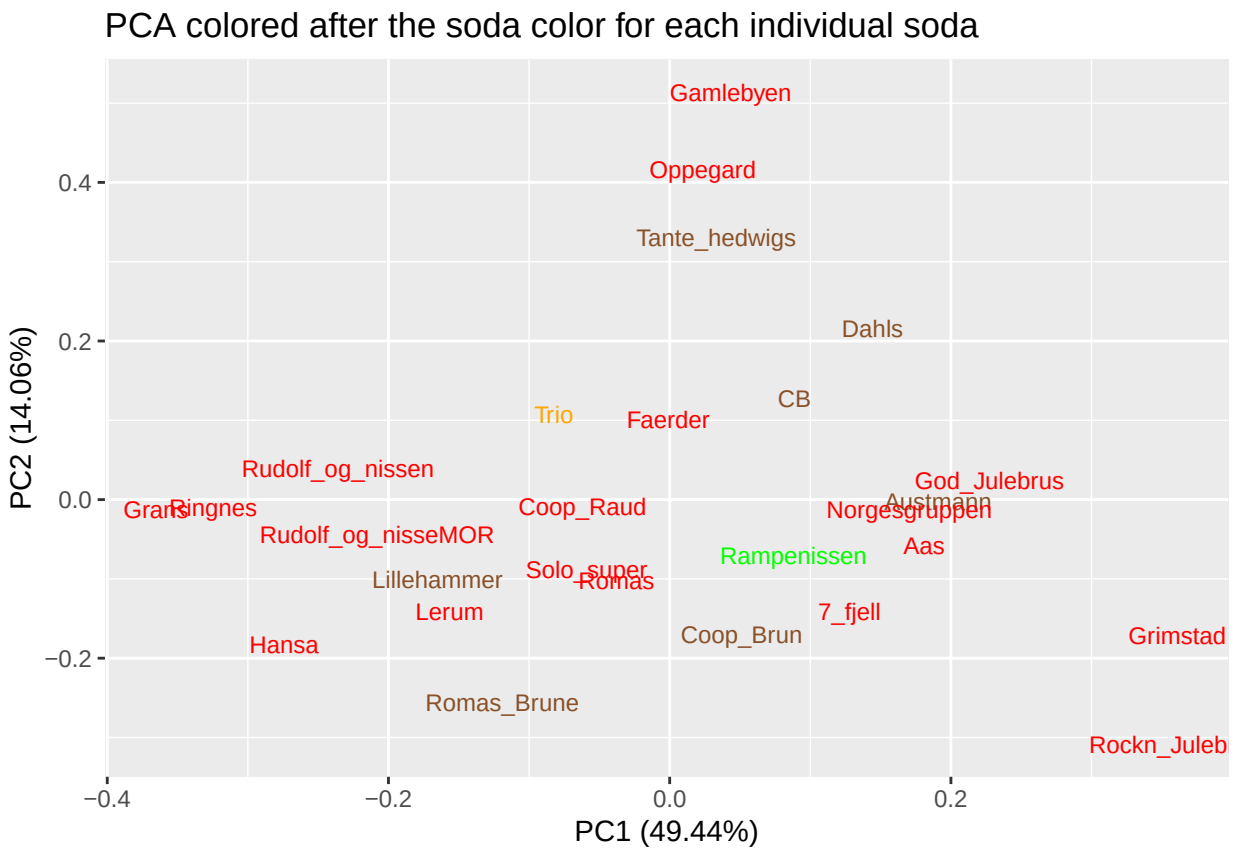
Likewise as the PCA plot above, in this PCA plot each number represents a participant. E.g. 1 represents person 1.

However, there is no clear difference in how the different genders scored the sodas. This PCA plot suggest that females had a more similar scoring than the males.

PCA of christmas soda data with sodas in focus

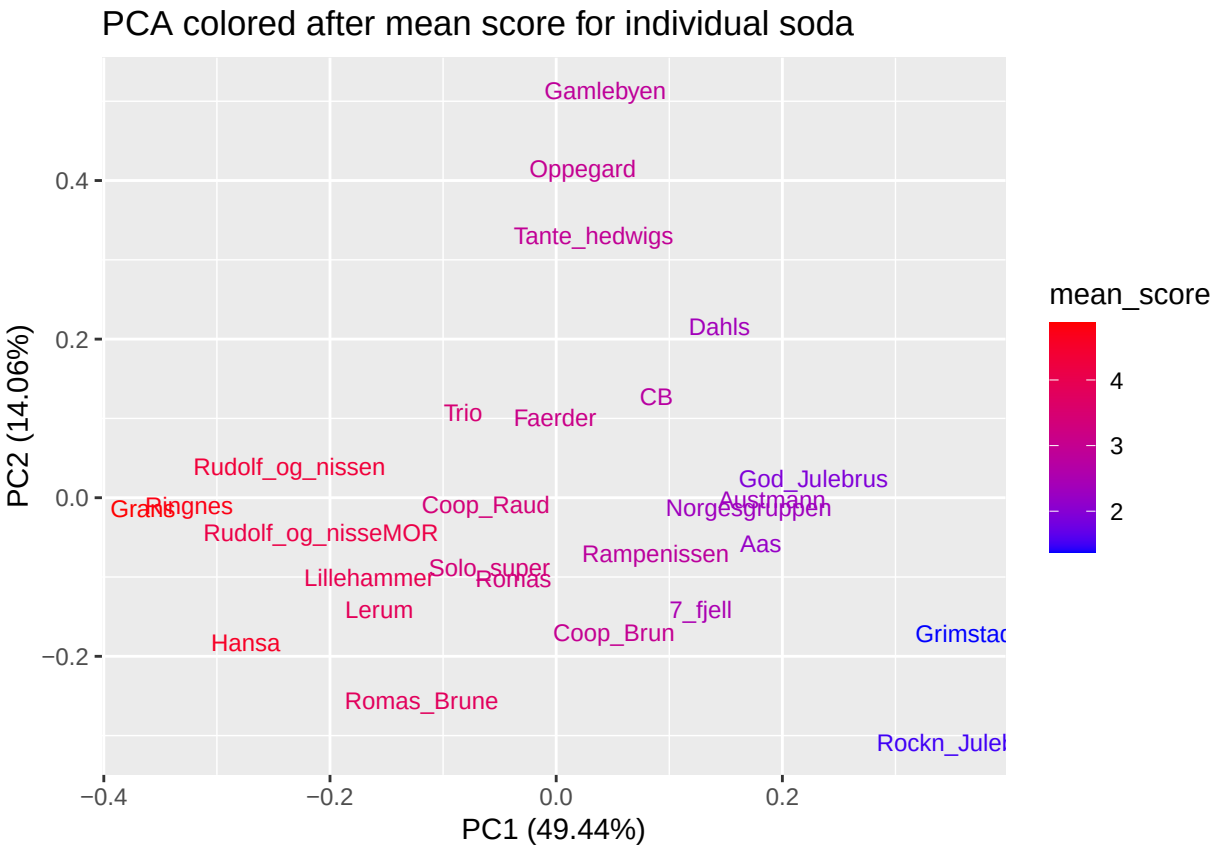
```
# PCA soda in focus and soda color

  pc = prcomp( brus_data_soda_rows[, c(-1,-18, -19)],
               scale. = T) # only select columns that includes soda types.

  # Plot PCA and color after brus color
   pc_autoplot_brus_color = autoplot(pc,
                                     data = brus_data_soda_rows,
                                     shape = F,
                                     colour = brus_data_soda_rows$brus_color,
```

```
                                    label.size = 3,
                                    frame = F) +
    ggtitle("PCA colored after the soda color for each individual soda");
  pc_autoplot_brus_color
```



The color of the different Christmas sodas were either red, brown, green or orange. This PCA plot where the Christmas sodas are colored according to which color the soda has does not display any particular difference in scoring based on color. Suggesting that soda color does not play an important role in scoring. This is interesting to know, since the participants could see the color of the soda during the test.

```
# Plot PCA and color after score
  pc = prcomp( brus_data_soda_rows[, c(-1,-18, -19)],
               scale. = T) # only select columns that includes soda types.
pc_autoplot_mean_score = autoplot(pc,
                                  data = brus_data_soda_rows,
                                  shape = F,
                                  colour = "mean_score",
                                  label.size = 3,
                                  frame = F)+
  scale_color_gradient(low = "blue", high = "red") +
  ggtitle("PCA colored after mean score for individual soda");
pc_autoplot_mean_score
```



This PCA plot where the Christmas sodas are colored according to their mean score, display the localization and clustering of sodas with similar scoring. For instance are Grans and Ringes with the best mean scores clustered together, likewise are the sodas with the lowest mean scores Rock'n Julebrus and Grimstad. This plot is to give you a feeling of where the different sodas are within the PCA according to mean score.

Heatmap of Christmas soda ratings

The heatmap displays dendograms with hierarchical clustering. For the Christmas sodas (clustering of rows), we observe two main branches/clusters. These sodas consist of the sodas with the highest rating (like Grans, Ringnes and Hansa), while the other cluster consist of sodas with a lower score. This cluster is again divided into two branches/cluster, where the smallest cluster consist of the two sodas with the lowest score, which were Grimstad and Rock'n Julebrus.

```
#Heatmap

## Heatmaply is dependent on HTML, and this is not supported by PDF.
if (knitr::is_latex_output()) {
  heatmap_ratings = heatmaply(brus_data_soda_rows %>%
              select(-"Soda", -"mean_score") %>%
              rename("soda_color" = "brus_color"),
            main = "Christmas soda ratings",
            ylab = "Christmas soda brands",
            xlab = "Participant",
            key.title = "Score",
            width = 1000,
```

```
            height = 500)

htmlwidgets::saveWidget(heatmap_ratings, "heatmaply_plot.html",
                        selfcontained = TRUE)

webshot::webshot("heatmaply_plot.html", "heatmaply_plot.png",
                 vwidth = 800, vheight = 400)

knitr::include_graphics("heatmaply_plot.png")
} else {
   heatmap_ratings = heatmaply(brus_data_soda_rows %>%
             select(-"Soda", -"mean_score") %>%
             rename("soda_color" = "brus_color"),
           main = "Christmas soda ratings",
           ylab = "Christmas soda brands",
           xlab = "Participant",
           key.title = "Score",
           width = 1000,
           height = 500)
   heatmap_ratings
}
```