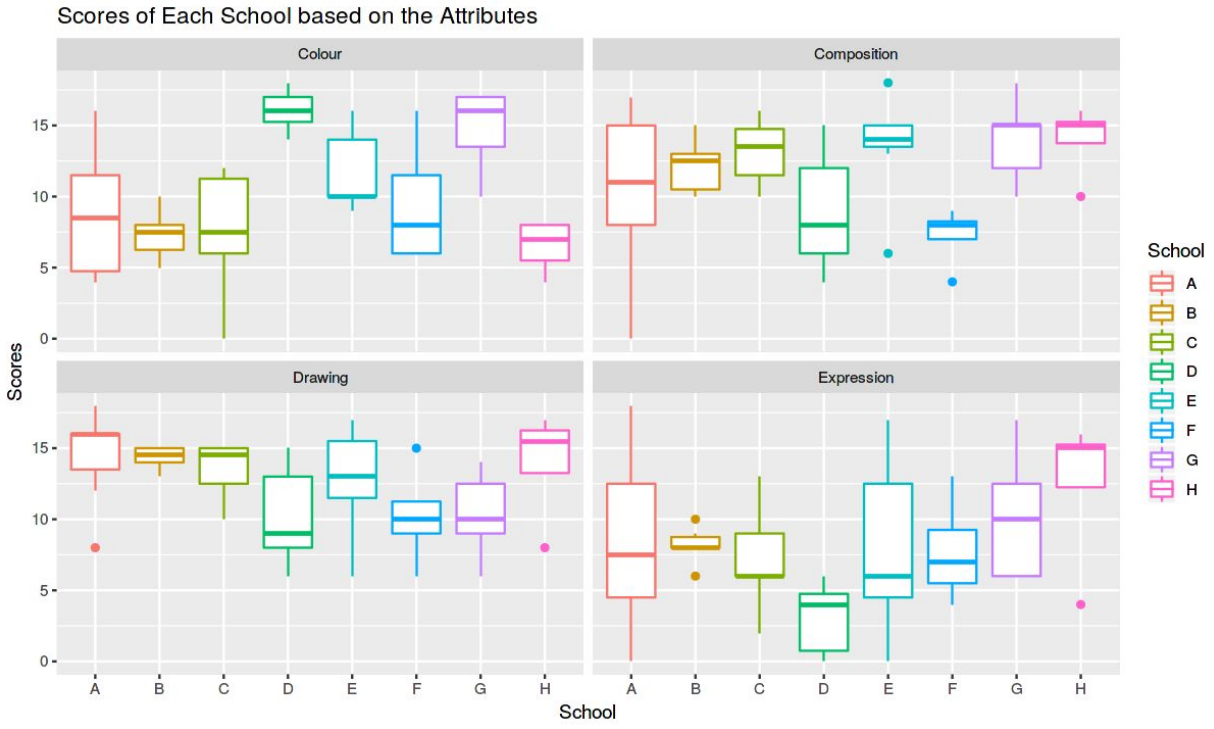
Plot 1

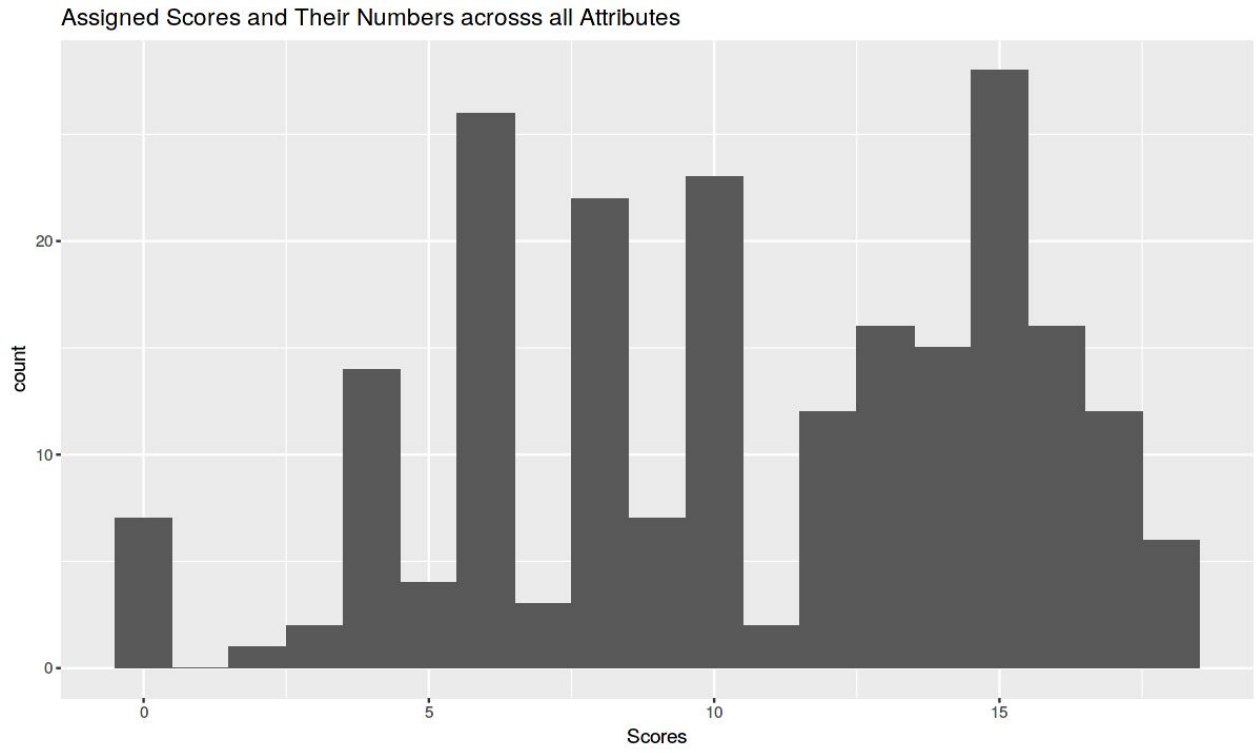


Boxplots are a measure of how well the data within a data set is distributed. This breaks down the dataset into three quartiles. This graph represents the minimum, average, median, quartile number one and quartile number three in the data set.

This boxplot shows the score distribution of each school on the basis of attributes. School D scored highest score in color. There is a Zero (0) score for school C. For color, there is no outlier in the plot. Median color score is highest in School D & G and lowest in school H. In Composition, school G has the highest score and school A has the lowest score. And median composition score is highest in School H and lowest in school F. There are few outliers in composition plot.

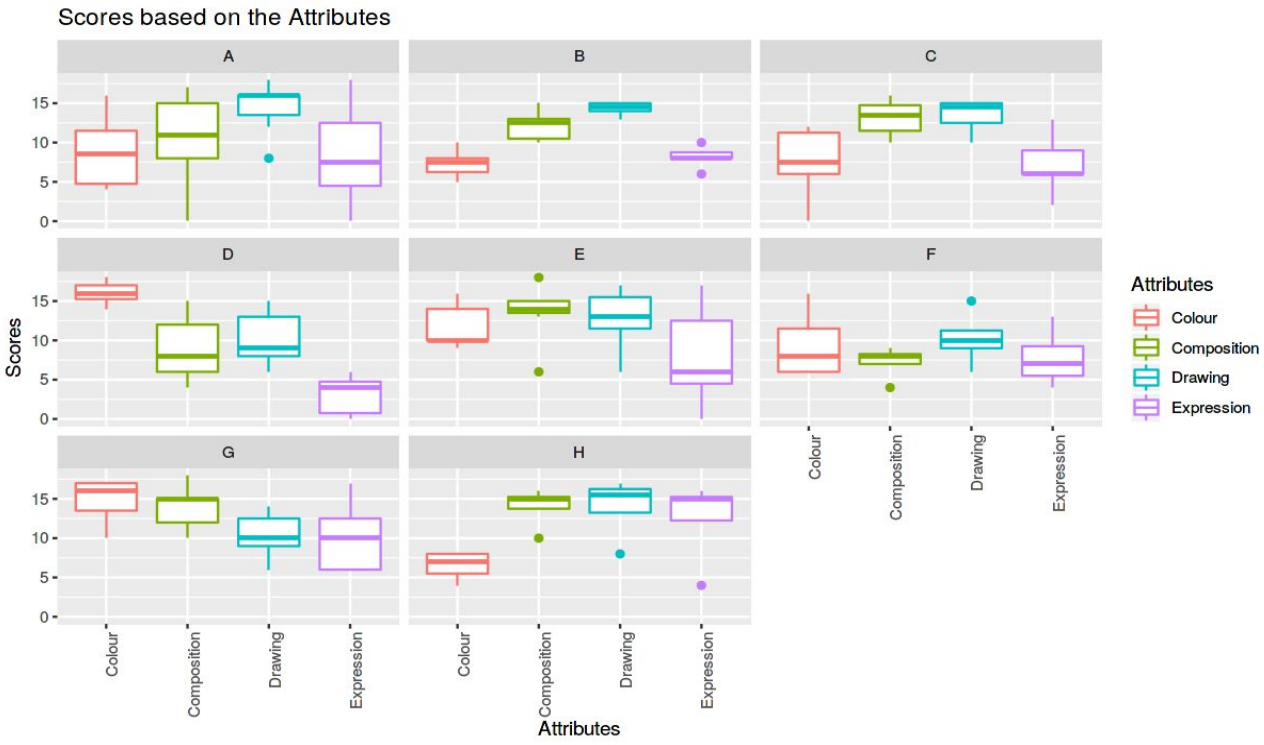
In drawing plot, we can see that school A has the highest score and School D, E & F have the lowest scores. School B has the highest median Drawing score and school D has the lowest median Drawing score. There are few outliers in drawing plot. In Expression plot, we can see that school A has the highest score and School A, D & E have the lowest scores. School H has the highest median Expression score and school D has the lowest median Expression score. There are few outliers in Expression plot.

Plot 2



This histogram shows the number count for each of the score assigned for all attributes. We can see that one (1) was not assigned. Two (2) was assigned only once. Both three (3) and eleven (11) were assigned twice. On the other hand, fifteen (15) was assigned most of the time. Six (6) was assigned for second highest time. Twelve (12) & seventeen (17) and thirteen (13) & sixteen (16), numbers of these two pairs were assigned in equal number of times.

Plot 3



This boxplot shows the score distribution of each attribute on the basis of schools. In school A, both Drawing and Expression were given the highest scores. On the other hand, Composition and Expression were given the lowest scores. Drawing held the highest median score and Expression held the lowest median score. There is an outlier in the graph of school A. In school B, both Drawing and Composition were given the highest scores. On the other hand, Color was given lowest scores. Drawing held the highest median score and Color held the lowest median score. There are outliers in the graph of school B.

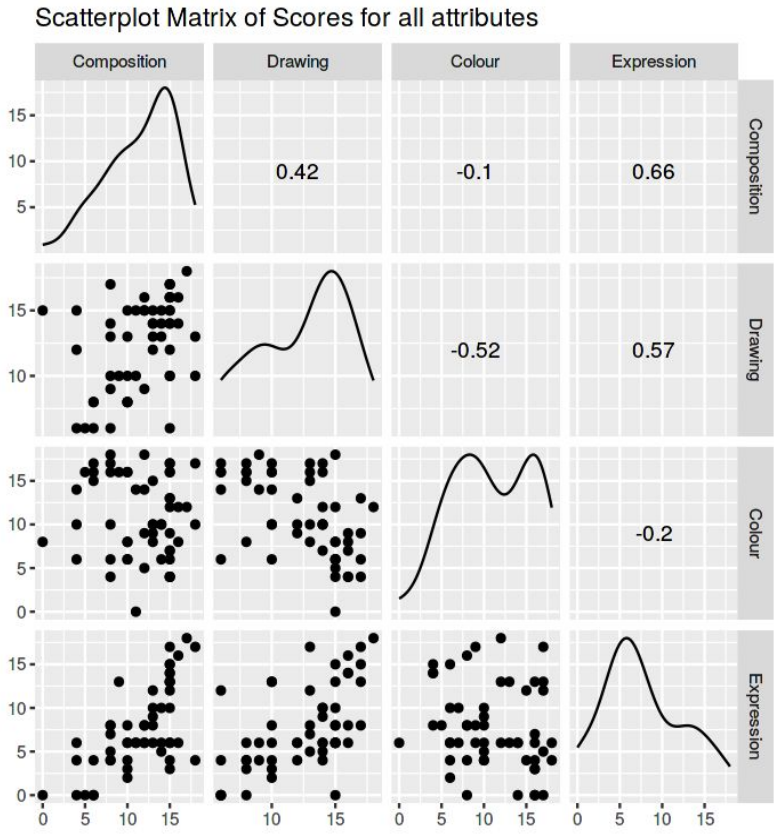
In school C, both Composition was given highest scores. On the other hand, Color was given lowest scores. Drawing held the highest median score and Expression held the lowest median score. There is no outlier in the graph of school C. In school D, Color was given the highest scores. On the other hand, Expression was given lowest scores. Color held the highest median score and Expression held the lowest median score. There is no outlier in the graph of school D.

In school E, Composition was given highest scores. On the other hand, Expression was given lowest scores. Composition held the highest median score and Expression held the lowest median score. There are outliers in the graph of school E. In school F, Color was given highest scores. On the other hand, Composition and Expression were given lowest scores. Drawing held the highest median score and Expression held the lowest median score. There are outliers in the graph of school F.

In school G, both Composition was given highest scores. On the other hand, Expression was given lowest scores. Color held the highest median score and Drawing held the lowest median score. There are outliers in the graph of school G. In school H, Drawing was given highest scores. On the other hand, Color and Expression were given lowest scores. Drawing held the highest median score and Color held the lowest median score. There are outliers in the graph of school H.

We can see that in most of the cases Drawing had the highest median score and Expression had the lowest median score.

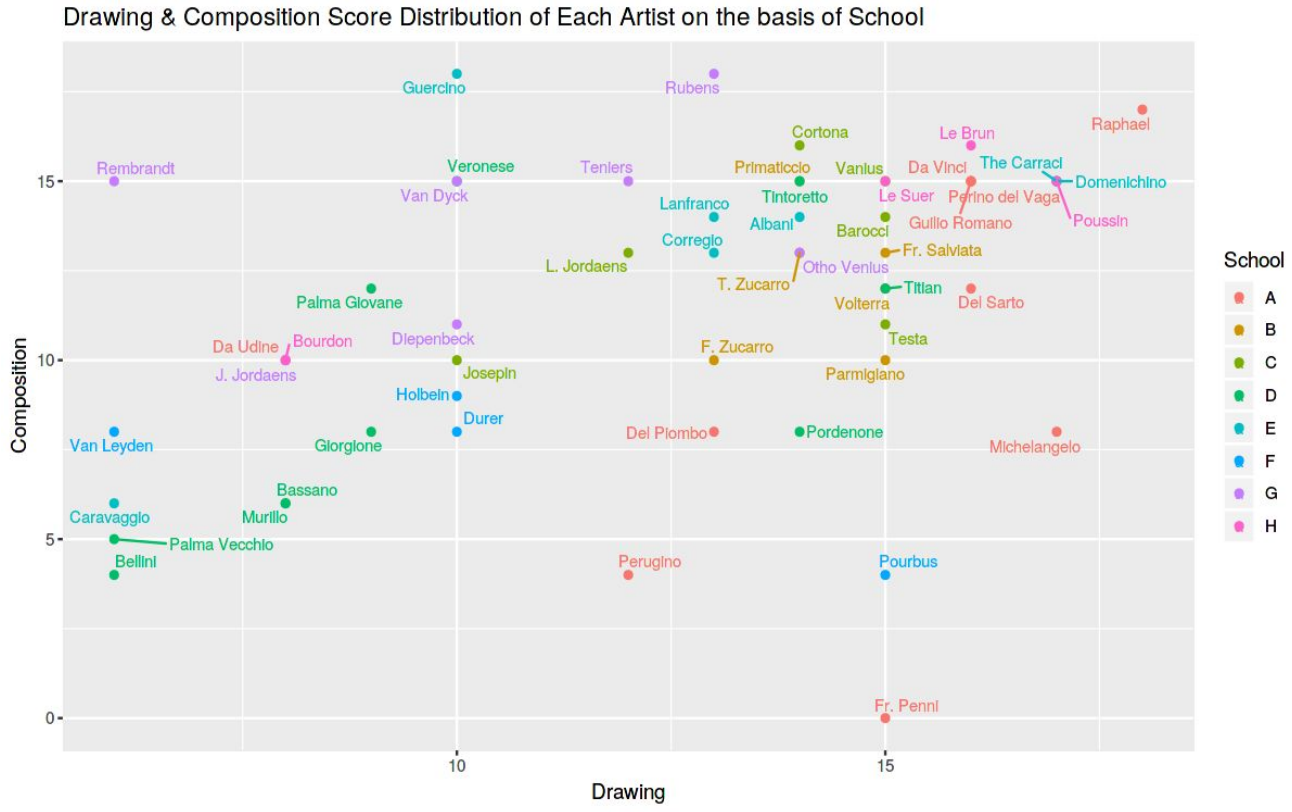
Plot 4



In a Scatterplot Matrix, Each scatterplot shows the relationship between a pair of variables. It shows the correlation between the variables. We can see the highest positive correlation between Expression and Composition. That means if a painter maintains composition then the picture becomes more expressive and a painter is getting around same scores both in Composition and Expression. On the other hand, the highest negative correlation is between Drawing and Color. That means the

score difference between Drawing and Color is higher. If a painter is getting higher scores in Drawing, he is likely to get lower scores in Color. There are also negative correlation between Color & Expression and Color & Composition. So, we can say that painters are failing to choose appropriate Color in their paintings, that's why they are getting lower scores in color.

Plot 5



A scatter plot incorporates values of two quantitative variables in a data set and shows them within a Cartesian diagram as geometric points. This scatterplot shows Drawing & Composition Score Distribution of Each Artist on the basis of School. X-axis represents Drawing scores and Y-axis represents Composition scores. The values are colored on the basis of school and labeled on the basis of Artist name. I've also showed artist name without overlapping with points by using ggrepel library.

Appendix

```
painters.df <- read.csv(file = "painters.csv",
```

```
                header = TRUE,
                sep = ",")
paintersLong.df <- gather(data=painters.df, Attributes,Scores, -Artist, -School)
paintersLong.df
```

Plot 1:

```
bp1 <- ggplot(data=paintersLong.df,
          aes(x=School,
              y= Scores,
              colour=School))
bp1 + geom_boxplot()+
  ggtitle("Scores of Each School based on the Attributes")+
  facet_wrap(~Attributes)
```

Plot 2:

```
hist<- ggplot(data = paintersLong.df,
          aes(x=Scores))+
  geom_histogram(binwidth = 1)+ggtitle("Assigned Scores and Their Numbers
acrosss all Attributes")
hist
```

Plot 3:

```
bp2 <- ggplot(data=paintersLong.df,
          aes(x=Attributes,
              y= Scores,
```

```
            colour=Attributes))
bp2 + geom_boxplot()+
  ggtitle("Scores based on the Attributes")+
  facet_wrap(~School)+
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```

Plot 4:

```
spm <- ggscatmat(data=painters.df,columns=2:5)+ggtitle("Scatterplot Matrix of Scores for all attributes")
spm
```

Plot 5:

```
sp <- ggplot(data=painters.df,
          aes(x=Drawing,
              y=Composition,
              colour=School,
              label=Artist))+
geom_point()+
xlab("Drawing")+ylab("Composition")+
ggtitle("Scores for All the Attributes")+
geom_text_repel(aes(label=Artist),size=3)
sp
```