

Portfolio 4

Yijin Guan

Introduction

This dataset is a collection of artworks of the 50 most influential artists of all time. We can learn the year, nationality, genre, link of Wikipedia and the number of their paintings. I want to learn more about the genres of the top 50 artists, so I create two visualizations to compare the genres with the artists and their nationalities.

```
artist<-read_csv("https://uwmadison.box.com/shared/static/wznh67zldnsm1wjrlzhl1glwtjimv5n.csv")
head(artist)
```

```
## # A tibble: 6 x 8
##   id name                years  genre nationality bio  wikipedia paintings
##   <dbl> <chr>                 <chr> <chr> <chr> <chr> <chr> <dbl>
## 1     0 Amedeo Modigliani 1884 - 1~ Expr~ Italian  "Ame~ http://e~      193
## 2     1 Vasiliy Kandinskiy 1866 - 1~ Expr~ Russian  "Was~ http://e~       88
## 3     2 Diego Rivera      1886 - 1~ Soci~ Mexican  "Die~ http://e~       70
## 4     3 Claude Monet      1840 - 1~ Impr~ French   "Osc~ http://e~       73
## 5     4 Rene Magritte    1898 - 1~ Surr~ Belgian  "Ren~ http://e~      194
## 6     5 Salvador Dali    1904 - 1~ Surr~ Spanish  "Sal~ http://e~      139
```

Visualization 1

Since there may be multiple genres for one artist, I first split the genre with comma and add new rows for the detached genre. Then I use *tbl_graph* class stores node and edge attributes in a single data structure and visualize the result. Most of the 50 artists belonged to the Impressionism and Post-Impressionism. Symbolism, Surrealism, Northern Renaissance, Baroque and High Renaissance are also very popular genres, while other genres correspond to only one or two artists. From an artist's point of view, he or she belongs to at most two genres.

```
a1<-artist %>%
  select(genre,name)
y <- data.frame(genre= character(),Name= character(), stringsAsFactors=FALSE)

for (i in 1:dim(a1)[1]) {
  x <- str_split(a1[i, 1], ",") %>% unlist()
  y <- rbind(y, data.frame(genre = x,Name = rep(a1$name[i], length(x)) ))
}
```

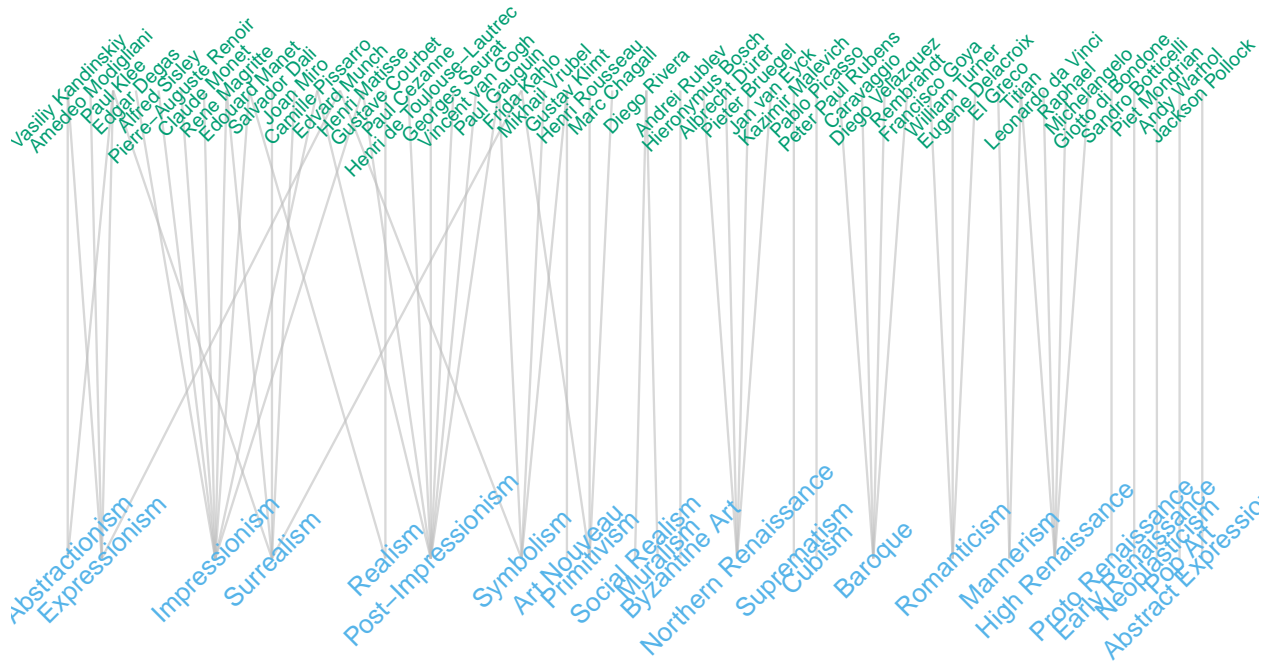
```
G_art<-tbl_graph(edges = y)%>%
  activate(nodes) %>%
  mutate(type = name %in% unique(y$genre))
```

```

ggraph(G_art,'auto') +
geom_edge_link(colour = "grey", width = 0.5, alpha = 0.6) +
geom_node_text(aes(label = name, size = type, col = type),angle=45) +
scale_size_discrete(range = c(3.2, 4.2)) +
scale_y_reverse(expand = c(0, 0.3, 0, 0.2)) +
scale_color_manual(values = c("#019e73", "#56b4e9"))+
theme_void() +
theme(legend.position = "none")

```

```
## Using `sugiyama` as default layout
```



Visualization 2

To learn more about the genres of paintings, I tried to study the relationship between nationality and genre. Since some of the artists belonged to multiple genres or have multiple nationalities, I had to deal with the data to make sure that each row of data contains only one genre and one country. In the node-link diagram, nodes that are close to one another in the shortest-path-sense also appear close to one another on the page. The genres in the United States (Pop Art and Abstract Expressionism) are relatively independent compared to those in European countries. What surprised me was that there are many great painters in Mexico, and their genres are closely linked to Europe instead of America.

```

pop<-load.image("D:/Information/msds_wisc/STAT479/portfolio/4/Andy_Warhol.jpg")
impressionism<-load.image("D:/Information/msds_wisc/STAT479/portfolio/4/Edouard_Manet.jpg")
proto<-load.image("D:/Information/msds_wisc/STAT479/portfolio/4/Giotto.jpg")
cubism<-load.image("D:/Information/msds_wisc/STAT479/portfolio/4/Pablo.jpg")
byza<-load.image("D:/Information/msds_wisc/STAT479/portfolio/4/Andrei.jpg")

```

```

a2<-artist %>%
  select(nationality,genre)
y2 <- data.frame(Nation= character(),genre= character(), stringsAsFactors=FALSE)

for (i in 1:dim(a2)[1]) {
  x <- str_split(a2[i, 2], ",") %>% unlist()
  y2 <- rbind(y2, data.frame(Nation = rep(a2$nationality[i], length(x)), genre = x))
}

y3 <- data.frame(Nation= character(),genre= character(), stringsAsFactors=FALSE)
for (i in 1:dim(y2)[1]) {
  x <- str_split(y2[i, 1], ",") %>% unlist()
  y3 <- rbind(y3, data.frame(genre = rep(y2$genre[i], length(x)), Nation = x))
}

G2<-tbl_graph(edges = y3)%>%
  activate(nodes) %>%
  mutate(type = name %in% unique(y3$Nation))
ggraph(G2,'kk') +
  geom_edge_link(colour = "grey", width = 0.5, alpha = 0.6) +
  geom_node_text(aes(label = name, size = type, col = type)) +
  scale_size_discrete(range = c(4, 5)) +
  scale_color_manual(values = c("#019e73", "#e7a208"))+
  annotation_raster(pop,0.3,1,5,6.8)+
  annotation_raster(impressionism,-4,-3,-0.5,1.5)+
  annotation_raster(proto,1.3,2.5,2.3,4.5)+
  annotation_raster(cubism,-0.2,0.4,-3.1,-1.3)+
  annotation_raster(byza,-1.8,-0.8,3.3,5)+
  theme_void() +
  theme(legend.position = "none")

```

