Data Visualization: Assignment 2

Student 1: Sofya Aksenyuk, 150284

Student 2: Uladzimir Ivashka, 150281
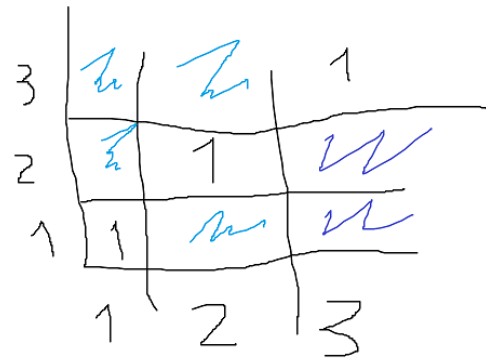
Three preliminary visualizations

Sketch 1

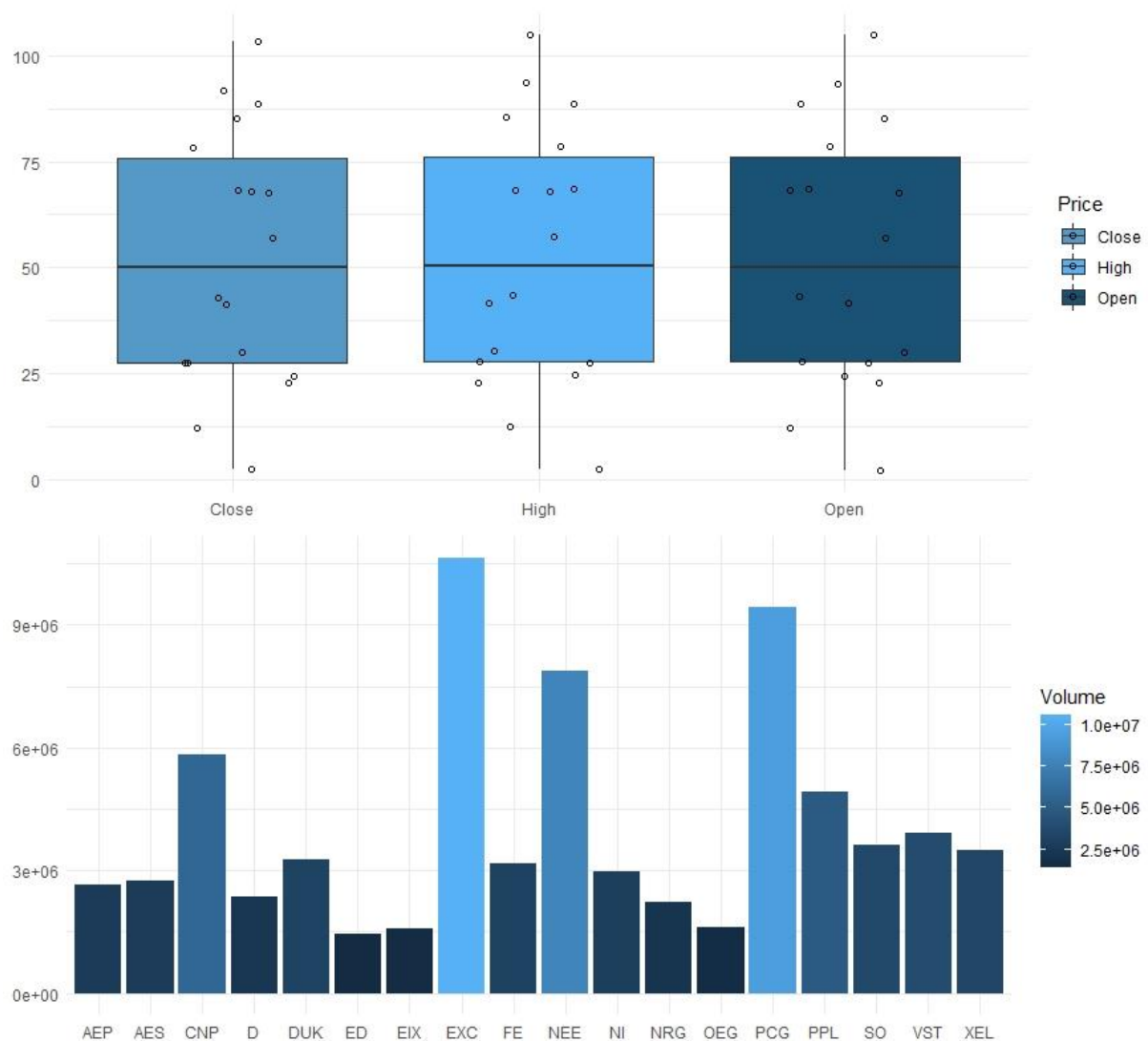# Implementation 1

Implementation code 1

```
library(ggpubr)
library(ggplot2)

ut <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/utilities.csv")

data <- ut[, c(3,7,8,9)]

one <- data %>% pivot_longer(-Name) %>%
  ggplot(aes(x=name,y=value,fill=name)) +
  geom_boxplot()  + theme_minimal() + theme(axis.title = element_blank()) +
  geom_jitter(shape=1, position=position_jitter(0.2)) +
  scale_fill_manual(values=c("#5499C7", "#56b1f7", "#1A5276"), name = "Price")

two <- ggplot(ut, aes(x=Symbol, y=Volume, fill=Volume)) +
  geom_bar(stat="identity", position="identity") + theme_minimal() +
  theme(axis.title = element_blank())

ggarrange(one, two, nrow=2)
```
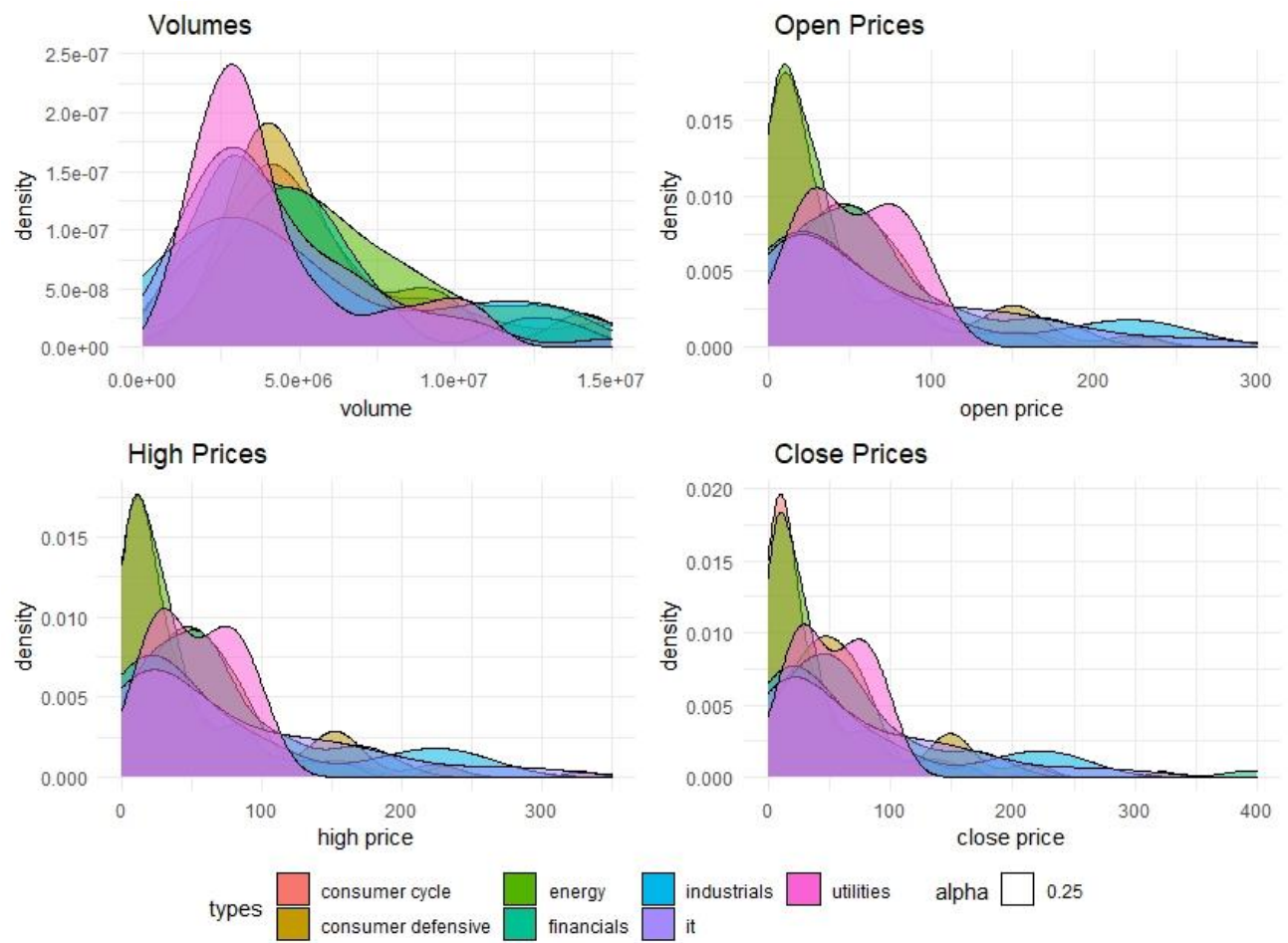
Implementation 2

Implementation code 2

```r
ind <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/industrials.csv")
cd <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/consumerdefensive.csv")
ut <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/utilities.csv")
fin <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/financials.csv")
en <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/energy.csv")
cc <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/consumercycle.csv")
it <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Sectors/it.csv")

types = c(c(rep("industrials", length(ind[,1]))), c(rep("consumer defensive", length(cd[,1]))),
          c(rep("utilities", length(ut[,1]))), c(rep("financials", length(fin[,1]))),
          c(rep("energy", length(en[,1]))), c(rep("consumer cycle", length(cc[,1]))),
          c(rep("it", length(it[,1]))))

volumes = c(ind[,4], cd[,4], ut[,4], fin[,4], en[,4], cc[,4], it[,4])
opens = c(ind[,5], cd[,7], ut[,7], fin[,5], en[,7], cc[,7], it[,5])
highs = c(ind[,6], cd[,8], ut[,8], fin[,6], en[,8], cc[,8], it[,6])
closes = c(ind[,7], cd[,9], ut[,9], fin[,7], en[,9], cc[,9], it[,7])

all <- data.frame(types, volumes, opens, highs, closes)
vol <- ggplot(all, aes(volumes, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 15000000) +
  theme_minimal() +
  labs(title ="    Volumes") +
  xlab("volume")

op <- ggplot(all, aes(opens, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 300) +
  theme_minimal() +
  labs(title = "    Open Prices") +
  xlab("open price")

hi <- ggplot(all, aes(highs, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 350) +
  theme_minimal() +
  labs(title = "    High Prices") +
  xlab("high price")

cl <- ggplot(all, aes(closes, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 400) +
  theme_minimal() +
  labs(title = "    Close Prices") +
  xlab("close price")


ggarrange(vol, op, hi, cl, nrow=2, ncol=2, common.legend = TRUE, legend = "bottom")
```
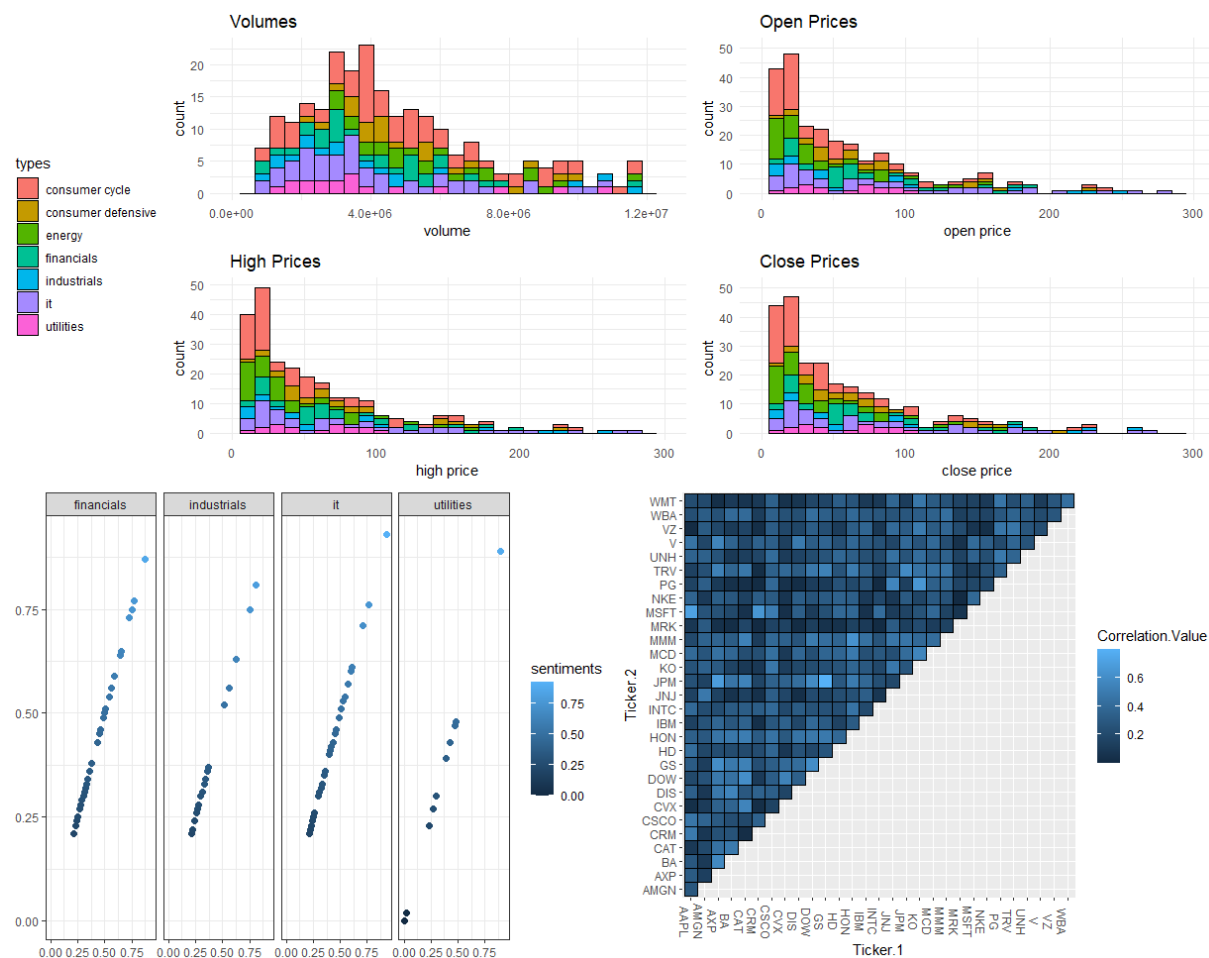
Implementation 3

Implementation code 3

```r
cor <- read.csv("C:/Users/1625203/Desktop/DV/Dataset/Dataset/Correlations.csv")

sentiments = c(fin[,8], ind[,8], it[,8], ut[,11])

sectors = c(c(rep("financials", length(fin[,8]))),
            c(rep("industrials", length(ind[,8]))),
            c(rep("it", length(it[,8]))),
            c(rep("utilities", length(ut[,8])))))

sent_data <- data_frame(sentiments, sectors)

sent_plot <- ggplot(sent_data, aes(sentiments, sentiments, color=sentiments)) +
  geom_point(size=2) +
  facet_grid(. ~ sectors) +
  theme_bw() +
  theme(axis.title = element_blank())


vol2 <- ggplot(all, aes(volumes, fill=types)) +
  geom_histogram(color="black") +
  xlim(0, 12500000) +
  theme_minimal() +
  labs(title ="    Volumes") +
  xlab("volume")

op2 <- ggplot(all, aes(opens, fill=types)) +
  geom_histogram(color="black") +
  xlim(0, 300) +
  theme_minimal() +
  labs(title = "    Open Prices") +
  xlab("open price")

hi2 <- ggplot(all, aes(highs, fill=types)) +
  geom_histogram(color="black") +
  xlim(0, 300) +
  theme_minimal() +
  labs(title = "    High Prices") +
  xlab("high price")

cl2 <- ggplot(all, aes(closes, fill=types)) +
  geom_histogram(color="black") +
  xlim(0, 300) +
  theme_minimal() +
  labs(title = "    Close Prices") +
  xlab("close price")

cor_plot <- ggplot(cor, aes(Ticker.1, Ticker.2, fill=Correlation.Value)) +
  geom_tile(color="black") +
  theme(axis.text.x = element_text(angle = 270))

temp <- ggarrange(vol2, op2, hi2, cl2, nrow=2, ncol=2, common.legend=TRUE, legend = "left")
temp2 <- ggarrange(sent_plot, cor_plot, nrow=1, ncol=2)
ggarrange(temp, temp2, nrow=2, ncol=1)
```
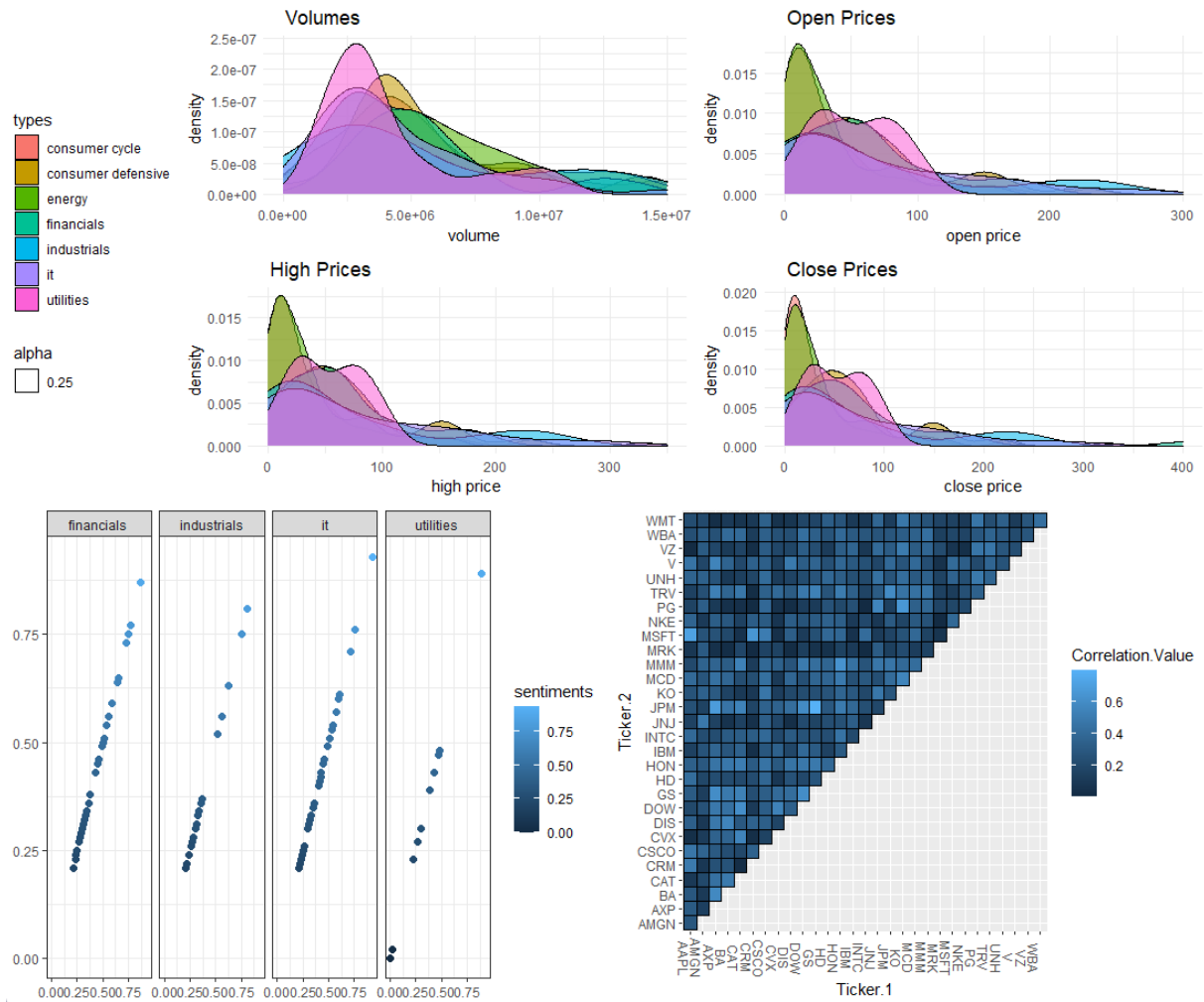
Selected final visualization

Implementation

Implementation code

```r
types = c(c(rep("industrials", length(ind[,1]))), c(rep("consumer defensive", length(cd[,1]))),
          c(rep("utilities", length(ut[,1]))), c(rep("financials", length(fin[,1]))),
          c(rep("energy", length(en[,1]))), c(rep("consumer cycle", length(cc[,1]))), c(rep("it", length(it[,1]))))

volumes = c(ind[,4], cd[,4], ut[,4], fin[,4], en[,4], cc[,4], it[,4])
opens = c(ind[,5], cd[,7], ut[,7], fin[,5], en[,7], cc[,7], it[,5])
highs = c(ind[,6], cd[,8], ut[,8], fin[,6], en[,8], cc[,8], it[,6])
closes = c(ind[,7], cd[,9], ut[,9], fin[,7], en[,9], cc[,9], it[,7])

all <- data.frame(types, volumes, opens, highs, closes)
vol <- ggplot(all, aes(volumes, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 15000000) +
  theme_minimal() +
  labs(title ="   Volumes") +
  xlab("volume")

op <- ggplot(all, aes(opens, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 300) +
  theme_minimal() +
  labs(title = "   Open Prices") +
  xlab("open price")

hi <- ggplot(all, aes(highs, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 350) +
  theme_minimal() +
  labs(title = "   High Prices") +
  xlab("high price")

cl <- ggplot(all, aes(closes, fill=types, alpha=0.25)) +
  geom_density() +
  xlim(0, 400) +
  theme_minimal() +
  labs(title = "   Close Prices") +
  xlab("close price")

sentiments = c(fin[,8], ind[,8], it[,8], ut[,11])
sectors = c(c(rep("financials", length(fin[,8]))), c(rep("industrials", length(ind[,8]))),
            c(rep("it", length(it[,8]))), c(rep("utilities", length(ut[,8]))))

sent_data <- data_frame(sentiments, sectors)
sent_plot <- ggplot(sent_data, aes(sentiments, sentiments, color=sentiments)) +
  geom_point(size=2) +
  facet_grid(. ~ sectors) +
  theme_bw() +
  theme(axis.title = element_blank())

cor_plot <- ggplot(cor, aes(Ticker.1, Ticker.2, fill=Correlation.value)) +
  geom_tile(color="black") +
  theme(axis.text.x = element_text(angle = 270))

temp <- ggarrange(vol, op, hi, cl, nrow=2, ncol=2, common.legend=TRUE, legend = "left")
temp2 <- ggarrange(sent_plot, cor_plot, nrow=1, ncol=2)
ggarrange(temp, temp2, nrow=2, ncol=1)
```