R 기본분석 1

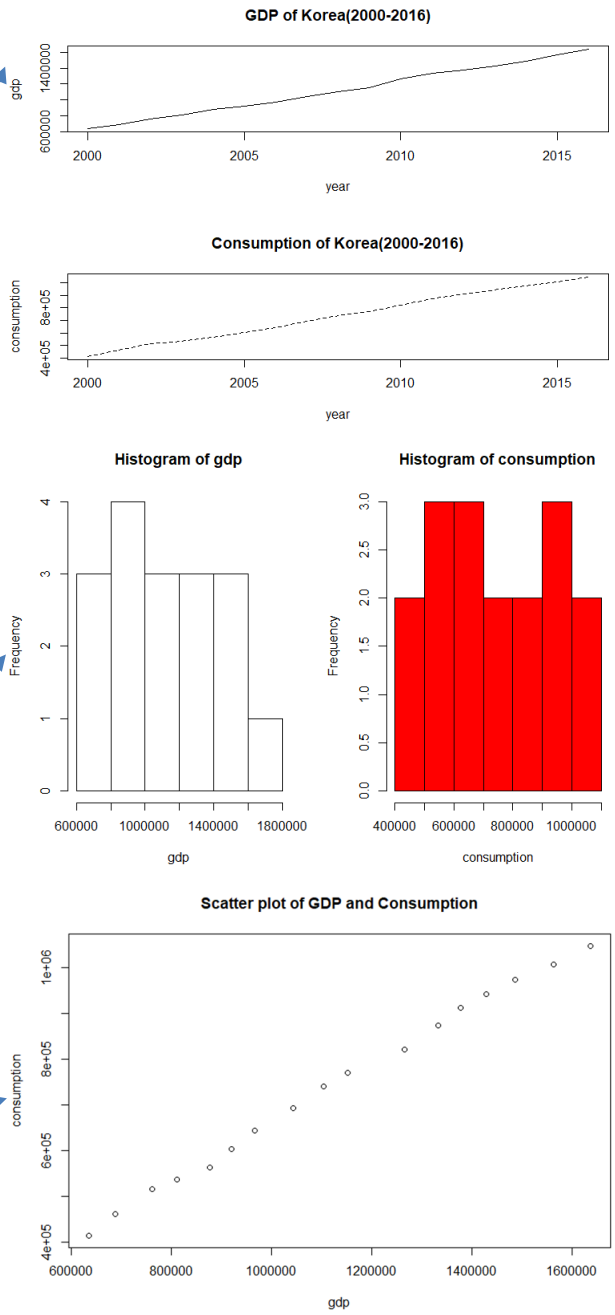
I. 함수를 이용한 그림그리기

II. ggplot2를 이용한 그림그리기

1. 선 그래프/히스토그램/산포도

4-3-1.R

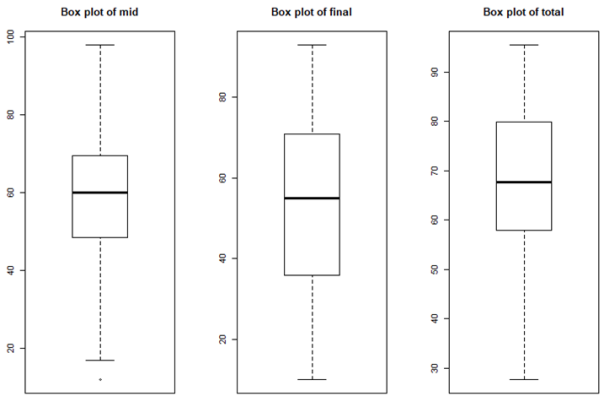
```
library(openxlsx)
sample1<-read.xlsx("http://kanggc.iptime.org/
book/data/sample1-n.xlsx")

year<-sample1$year
gdp<-sample1$gdp
consumption<-sample1$consumption

gdp
consumption

par(mfrow=c(2,1))
plot(year, gdp, type="l", main="GDP of Korea(
2000-2016)")
plot(year, consumption, type="l", lty=2,main="
Consumption of Korea(2000-2016)")

par(mfrow=c(1,2))
hist(gdp)
hist(consumption, breaks=8, col="red")

par(mfrow=c(1,1))
plot(gdp, consumption, main="Scatter plot of
GDP and Consumption")
```

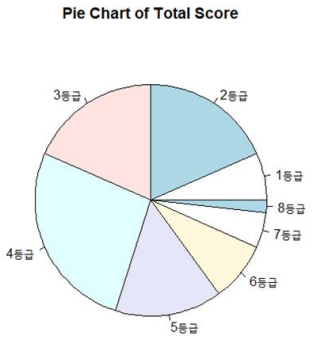


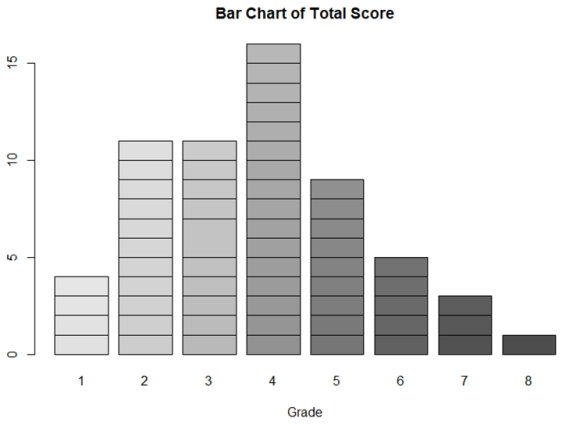
2. 상자그래프/원그래프/막대그래프

4-3-2.R

```
library(openxlsx)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")
mid<-sample1$mid
final<-sample1$final
total<-sample1$total
grade<-sample1$grade
summary(sample1)

par(mfrow=c(1,3))
boxplot(mid, main="Box plot of mid")
boxplot(final, main="Box plot of final")
boxplot(total, main="Box plot of total")

par(mfrow=c(1,1))
table(grade)
slices<-c(4,11,11,16,9,5,3,1)
lbls<-c("1등급","2등급","3등급","4등급","5등급","6등급","7등급","8등급")
pie(slices, labels=lbls, main="Pie Chart of Total Score")

counts<-table(total, grade)
barplot(counts, main="Bar Chart of Total Score", xlab="Grade")
```







3. 잎-줄기 그래프

4-3-3.R

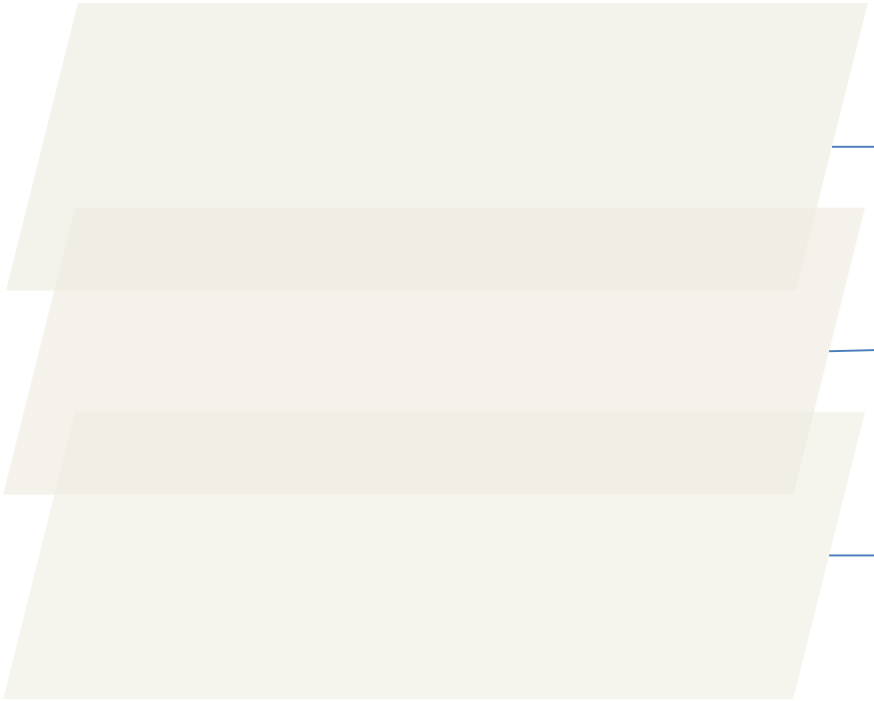
```
library(openxlsx)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")
mid<-sample1$mid
final<-sample1$final
total<-sample1$total
grade<-sample1$grade
total
stem(total)
stem(total, scale=0.5)
stem(total, scale=2)
```

```
The decimal point is 1 digit(s) to the right of the

2 | 8
3 | 01
4 | 017789
5 | 1135788
6 | 00112333355778888
7 | 012445666789
8 | 11123444457
9 | 0235
```

```
The decimal point is 1 digit(s) to the right of the

2 | 801
4 | 0177891135788
6 | 00112333355778888012445666789
8 | 111234444570235
```

```
The decimal point is 1 digit(s) to the right of the

2 | 8
3 | 01
3 |
4 | 01
4 | 7789
5 | 113
5 | 5788
6 | 001123333
6 | 55778888
7 | 01244
7 | 5666789
8 | 111234444
8 | 57
9 | 023
9 | 5
```

0. ggplot2 레이어 구조



3단계 : 설정추가(축, 범위, 색, 표시)

2단계 : 그래프 추가(점, 막대, 선)

1단계 : 배경 설정(축)

- ggplot2는 데이터를 시각화하는 패키지로 보통 3단계로 구성되어 있음
  - 1단계 : 배경 설정으로 데이터 축을 설정
  - 2단계 : 그래프 추가(점, 막대, 선 등)
  - 3단계 : 세부 설정 추가(축 범위, 색, 표식 등)

- ggplot2의 함수 구조의 예를 들면 다음과 같음

  ggplot(data=data1, aes(x=var1, y=var2))+geom_point()+xlim(3,6)

  (1단계) (2단계) (3단계)

  - 1단계에서 data는 사용할 데이터, aes의 괄호 안은 x축 변수, y축 변수
  - 2단계는 그래프의 종류를 나타내는데 주로 사용하는 종류는 다음과 같음

    geom_point() : 산포도

    geom_smooth() : 평활그래프

    geom_bar() : 막대그래프(빈도 막대그래프로 X축만 설정)

    geom_col() : 막대그래프(집단간 차이를 나타냄)

    geom_boxplot() : 상자그래프

    geom_histogram() : 히스토그램

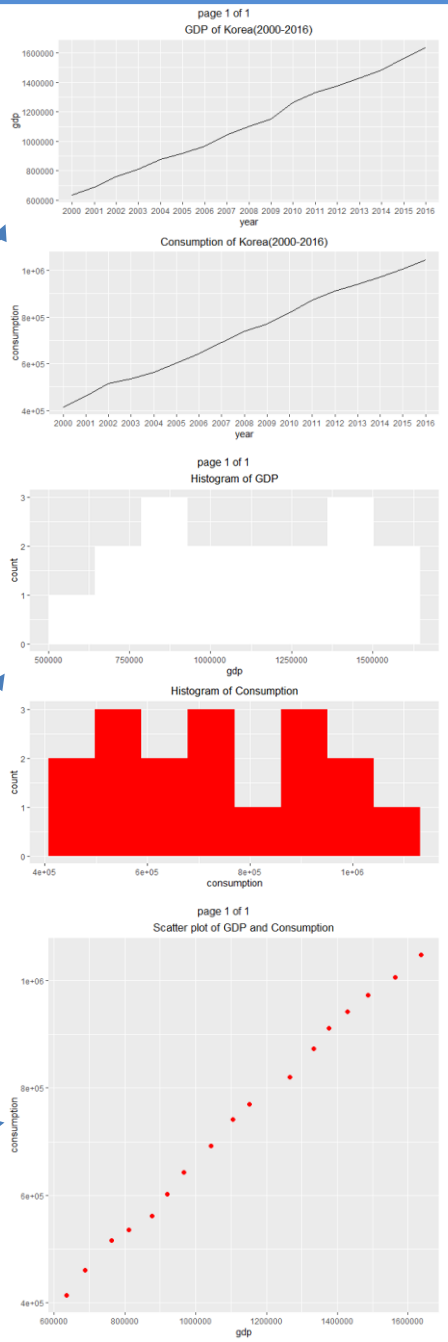
    geom_line() : 선그래프

  - 3단계에서 xlim의 괄호 안은 x축에 그릴 데이터의 범위를 나타냄

1. 선 그래프/히스토그램/산포도

4-3-4.R

```
library(openxlsx)
library(ggplot2)
library(gridExtra)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/sample1-n.xlsx")
year<-sample1$year
gdp<-sample1$gdp
consumption<-sample1$consumption

plot1<-ggplot(data=sample1, aes(x=year, y=gdp, group=1)) + geom_line() + ggtitle("GDP o
f Korea(2000-2016)")+ theme(plot.title = element_text(hjust = 0.5))
plot2<-ggplot(data=sample1, aes(x=year, y=consumption, group=1)) + geom_line() + ggtitl
e("Consumption of Korea(2000-2016)") + theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)

plot3<-ggplot(data=sample1, aes(x=gdp)) + geom_histogram(fill="white",bins=8) + ggtitle("
Histogram of GDP")+ theme(plot.title = element_text(hjust = 0.5))
plot4<-ggplot(data=sample1, aes(x=consumption)) + geom_histogram(fill="red",bins=8) +
ggtitle("Histogram of Consumption") + theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot3, plot4), nrow=2, ncol=1)

plot5<-ggplot(data=sample1, aes(x=gdp, y=consumption)) + geom_point(colour="red", siz
e=2) + ggtitle("Scatter plot of GDP and Consumption")+ theme(plot.title = element_text(hjus
t = 0.5))
marrangeGrob(grobs=list(plot5), nrow=1, ncol=1)
```



2. 상자그래프/원그래프/막대그래프-1

4-3-5.R

```
library(openxlsx)
library(ggplot2)
library(gridExtra)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")
mid<-sample1$mid
final<-sample1$final
total<-sample1$total
grade<-sample1$grade
plot1<-ggplot(data=sample1, aes(x=1, y=mid)) + geom_boxplot() + ggtitle("Boxplot of Mid")+ theme
(plot.title = element_text(hjust = 0.5))
plot2<-ggplot(data=sample1, aes(x=1, y=final)) + geom_boxplot() + ggtitle("Boxplot of Final")+ the
me(plot.title = element_text(hjust = 0.5))
plot3<-ggplot(data=sample1, aes(x=1, y=total)) + geom_boxplot() + ggtitle("Boxplot of Total")+ the
me(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2, plot3), nrow=1, ncol=3)

t.grade<-data.frame(table(sample1$grade))
plot4<-ggplot(t.grade, aes(x="", y=Freq, fill=Var1)) + geom_bar(width=1, stat="identity") +  coord_p
olar(theta="y") + ggtitle("Pie Chart of Total Score")+ theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot4), nrow=1, ncol=1)

plot5<-ggplot(data=sample1, aes(x=grade)) + geom_bar() + ggtitle("Bar Chart of Grade")+ theme(pl
ot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot5), nrow=1, ncol=1)
```

3. 막대그래프-2

4-3-6.R

```
library(openxlsx)
library(dplyr)
library(ggplot2)
library(gridExtra)

df<-read.xlsx("http://kanggc.iptime.org/book/data/subtotal-e.xlsx")
df

dept_name_1 <- df %>%
  group_by(dept, class) %>%
  summarise(mean_total = mean(total))
dept_name_1

plot1<-ggplot(data=dept_name_1, aes(x=dept, y=mean_total, fill=class
)) +
  geom_col(position="dodge2") + ggtitle("Bar Chart of Total Mean by d
ept & class")+ theme(plot.title = element_text(hjust = 0.5))
plot2<-ggplot(data=dept_name_1, aes(x=class, y=mean_total, fill=dept
)) +
  geom_col(position="dodge2") + ggtitle("Bar Chart of Total Mean by cl
ass & dept")+ theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)
```