



中國人民大學

RENMIN UNIVERSITY OF CHINA

R 语言大作业

基于 finviz.com 网页功能开发

组别: 第七组

组长: 赵海榕 2019100440

组员: 米月 2019102477

一、前期准备

(一) 下载并运行所需的 packages

```
install.packages("plotly")
install.packages("portfolio")
install.packages("dplyr")
install.packages("RColorBrewer")
install.packages("quantmod")
install.packages("ggplot2")
install.packages("readxl")
library(quantmod)
library(RColorBrewer)
library(plotly)
library(portfolio)
library(dplyr)
library(ggplot2)
library(readxl)
```

(二) 导入数据

本文主要选取沪深 300 股票数据展开研究。

```
setwd("C:/Users/thinkpad/Desktop")
#导入沪深 300 数据
HS300<- read.csv("C:/Users/thinkpad/Desktop/R/HS300.csv")
#order 表示股票代码， name 表示股票名称， sector 表示所处行业， mar_cap 表示股票市值， close 表示股票收盘价， pri_lim 表示价格涨跌幅
HS300$order<-as.character(HS300$order)
HS300$mar_cap<-as.numeric(HS300$mar_cap)#设置数据类型
HS300 <- HS300[-302:-301,] #删除空行
head(HS300)
```

	order <chr>	name <fctr>	sector <fctr>	mar_cap <dbl>	close <fctr>	pri_lim <dbl>
1	600547.SH	山东黄金	有色金属	300	32.2000	-2.6602
2	601899.SH	紫金矿业	有色金属	273	3.3100	-1.4881
3	603993.SH	洛阳钼业	有色金属	269	3.4400	-1.4327
4	601600.SH	中国铝业	有色金属	253	3.4600	-1.7045
5	600362.SH	江西铜业	有色金属	223	14.1600	-1.6667
6	600111.SH	北方稀土	有色金属	177	10.1600	-1.4549

6 rows

二、功能实现

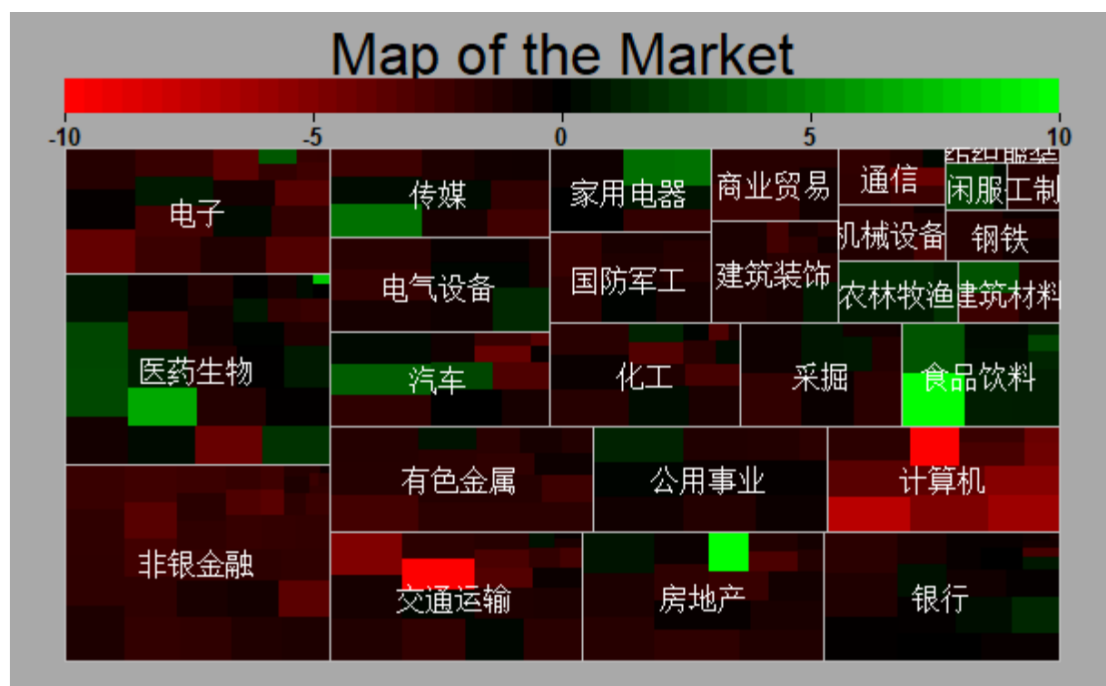
功能一：板块层级图

#使用 portfolio 包中的 map.market 函数绘制板块层级图

#1. 绘制沪深 300 的板块层级图

```
map.market(id      = HS300$order,
           area     = HS300$mar_cap,
           group    = HS300$sector,
           color    = HS300$pri_lim,
           lab      = c("group"=TRUE, "id"=FALSE),
           main     = "Map of the Market")
```

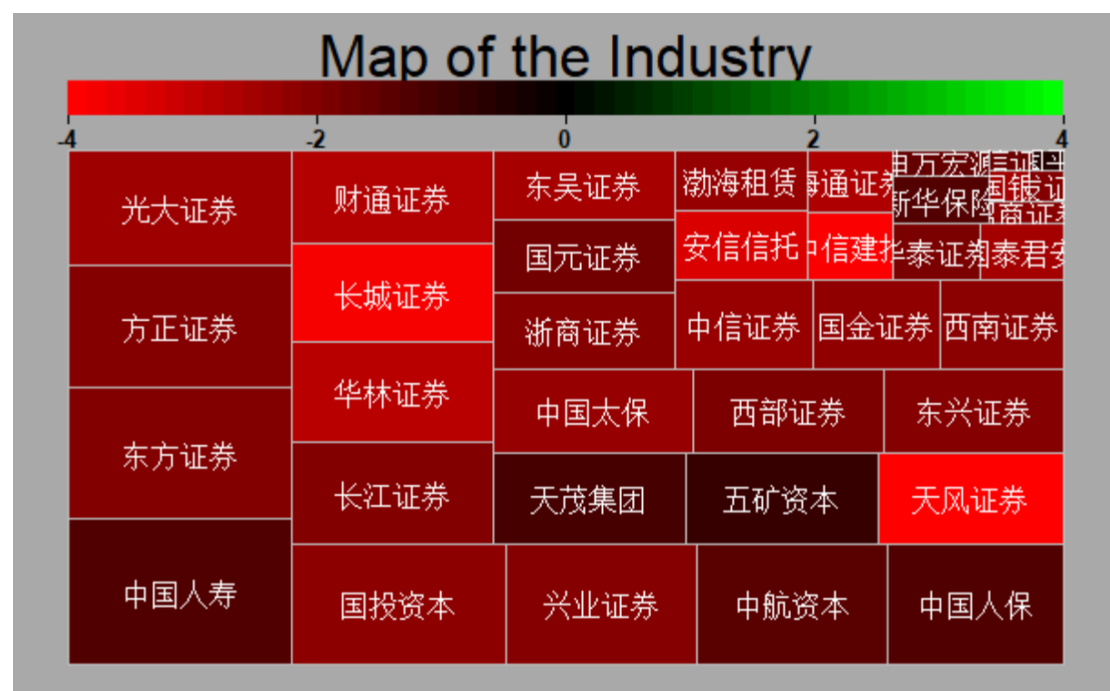
#分组变量是沪深 300 股票所处行业，面积代表股票市值，颜色代表股票涨跌幅



#2. 绘制沪深 300 中非银金融行业的板块层级图

```
finids<-HS300[HS300$sector=='非银金融',]
finids$order<-as.character(finids$order)
finids$mar_cap<-as.numeric(finids$mar_cap)
map.market(id      = finids$order,
            area    = finids$mar_cap,
            color   = finids$pri_lim,
            group   = finids$name,
            lab     = c("group"=TRUE, "id"=FALSE),
            main    = "Map of the Industry")
```

#分组变量是非银金融行业股票名称，面积代表股票市值，颜色代表股票涨跌幅



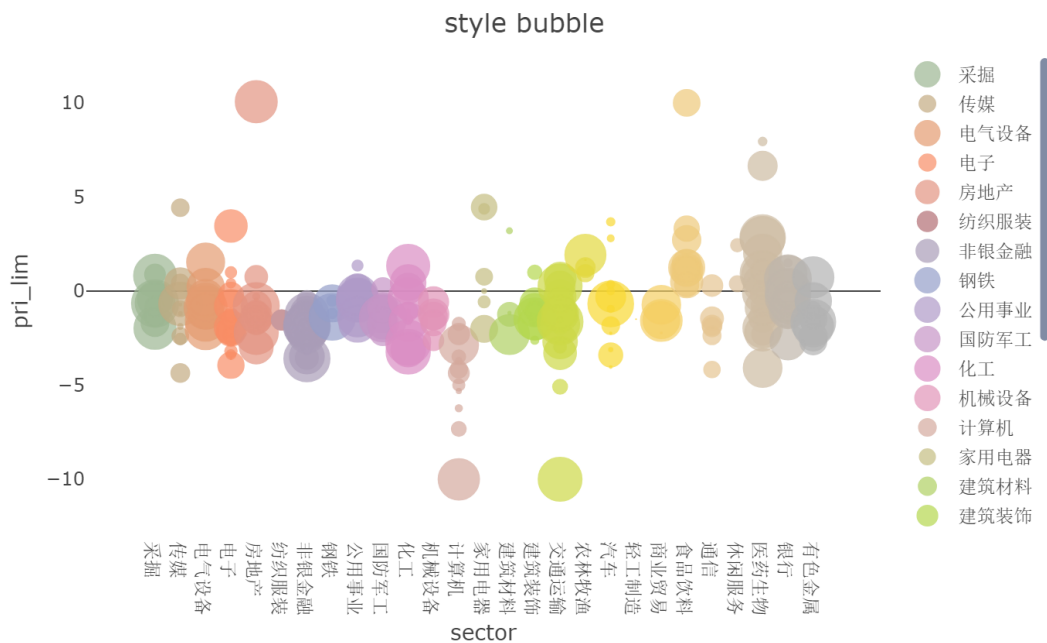
功能二：气泡图

```
p<-plot_ly(HS300,
            x=~sector,
            y=~pri_lim,
            text=~name,
            type = 'scatter',
            marker = list(size = ~mar_cap/10),
            mode = 'markers',
            color = ~sector
          ) %>%
```

```

layout(title = 'style bubble',
       xaxis = list(showgrid = FALSE),
       yaxis = list(showgrid = FALSE)
)
p

```



功能三：K 线图

#从雅虎财经读取谷歌的数据

```
getSymbols("GOOG",src="yahoo",from="2019-01-01",to='2019-10-27')
```

#1. 使用 chartSeries()函数的默认形式绘制 K 线图

```
chartSeries(GOOG)
```



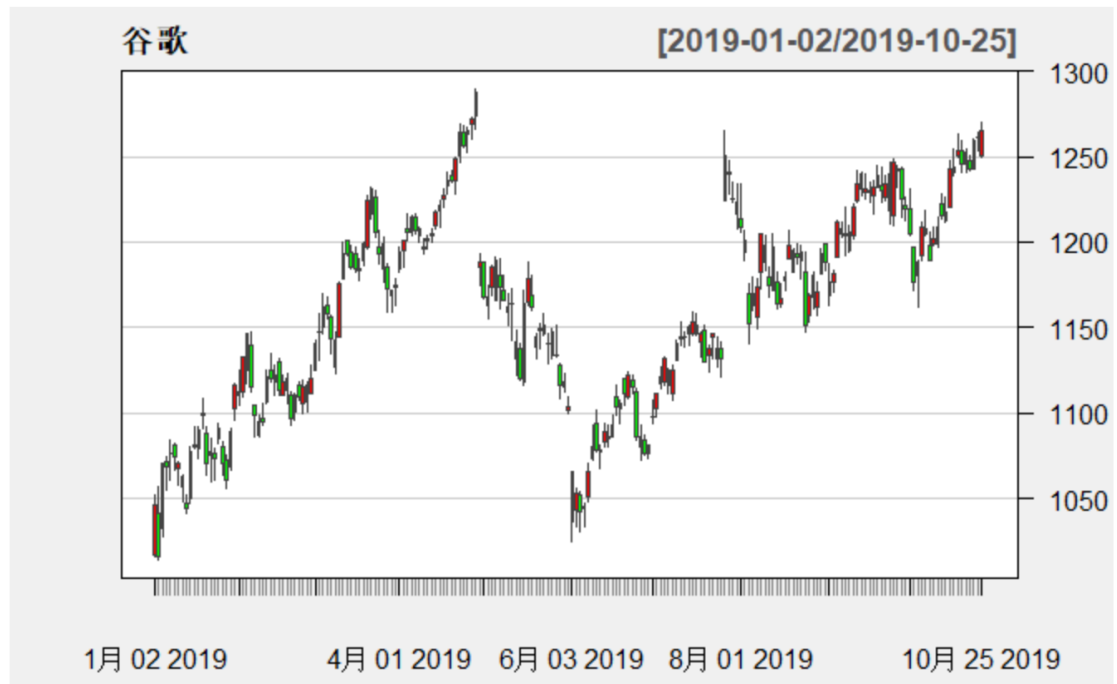
#2. 设定参数，白色背景，蜡烛图为红色和绿色

```
chartSeries(GOOG,theme = 'white',
            name = '谷歌',
            up.col = 'red',
            dn.col = 'green')
```



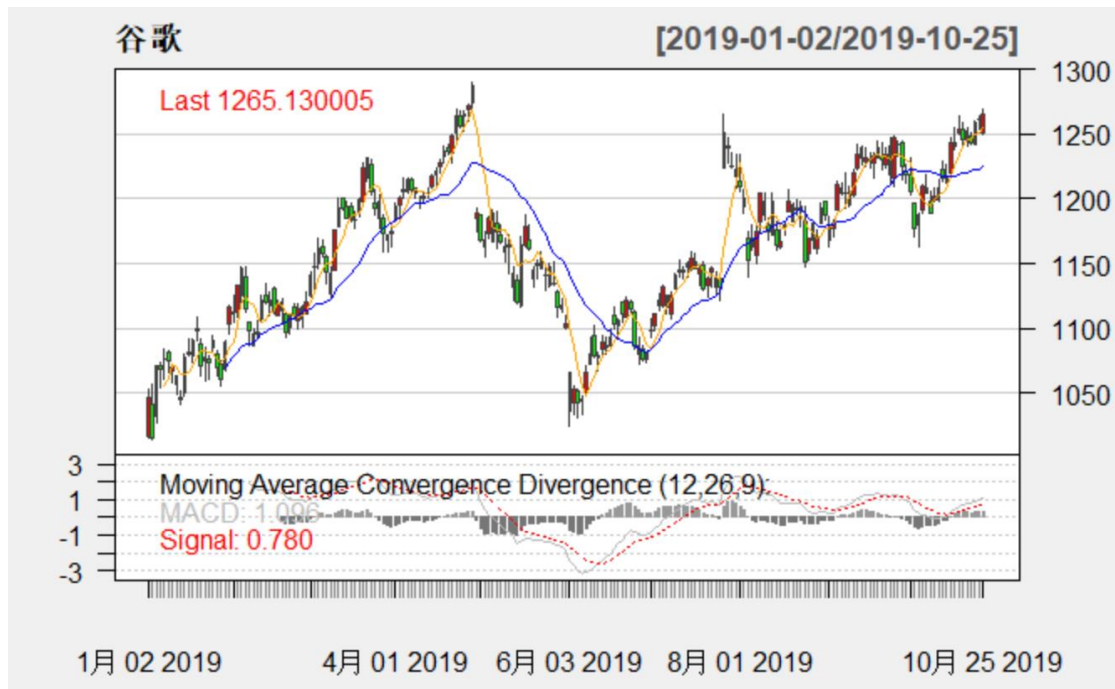
#3. 只绘制 K 线图，不绘制成交量

```
chartSeries(GOOG,theme = 'white',  
            name = '谷歌',  
            TA=NULL,  
            up.col = 'red',  
            dn.col = 'green')
```



#4. 加入 5 日、10 日均线

```
chartSeries(GOOG,theme = 'white',  
            name = '谷歌',  
            up.col = 'red',  
            dn.col = 'green',  
            TA =c(addMACD(),  
                  addSMA(n=5,col="orange"),  
                  addSMA(n=20,col="blue")))
```



功能四：1 DAY PERFORMANCE 条形图

#绘制沪深 300 成分股中按行业分类的 1 DAY PERFORMANCE 条形图

#读取 2019-11-01 当天沪深 300 的成分股数据

```
hs300<- read_excel("C:/Users/thinkpad/Desktop/R/hs_300.xlsx")
```

筛选包括公司代码、名称、行业、收盘价格、业绩、市值的数据

```
data1<-select(hs300,code,names,sector,price,change,mc)
```

计算各行业的业绩平均值

```
sector_per_avg=tapply(data1$change, data1$sector,mean)
```

取绝对值

```
sector_per_avg_ab=abs(sector_per_avg)
```

设置一个字符向量代表行业

```
a<-c('A','B','C','D','E','F','G','H','I','J','K','L','M','N','O','P','Q','R')
```

合并行业和行业业绩均值两列

```
sum_sector<-cbind.data.frame(a,sector_per_avg)
```

```
sum_sector_ab<-cbind.data.frame(a,sector_per_avg_ab)
```


画出分行业 1 DAY PERFORMANCE 条形图

```
ggplot(data=sum_sector_ab, aes(x=a,y=sector_per_avg_ab,fill=sector_per_avg))+  
  geom_bar(stat="identity",position="identity")+  
  coord_flip()+# 把 x 轴和 y 轴互换  
  
  geom_text(aes(label = round(sector_per_avg,2), vjust = 0.5, hjust = -0.2))+# 添  
加条形图标签
```

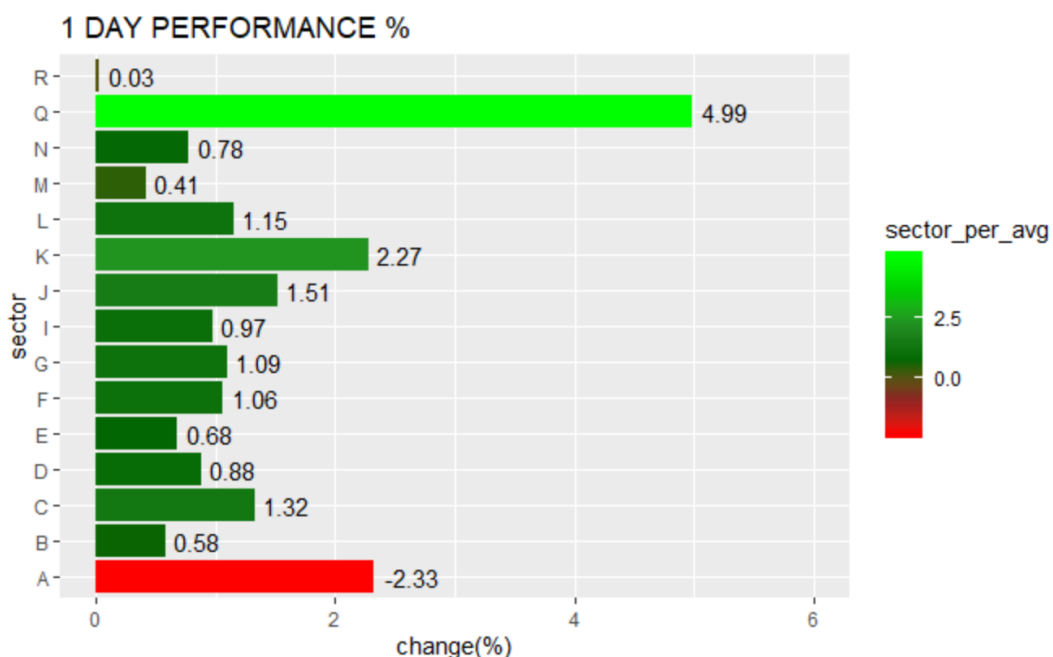
```
scale_y_continuous(limits=c(0,6))+# 设置 y 轴坐标范围为 0-6
```

```
scale_fill_gradientn(colours=c("red","brown4","darkgreen","forestgreen","green3"  
,"green"))+
```

```
# 设置条形图的颜色
```

```
labs(x="sector",y="change(%)")+# 添加横纵坐标标题
```

```
ggtitle('1 DAY PERFORMANCE %')# 添加图表标题
```



功能五：股票涨跌数量堆积条形图

绘制沪深 300 成分股中股价上涨、下跌、持平的股票数量的堆积条形图

```
# 读取 2019-11-05 数据
dat <- read_excel("C:/Users/thinkpad/Desktop/R/HS300comp.xlsx")
# 修改列名称
colnames(dat) <-
  c("Stkcd", "Stknm", "SMA50", "Industry", "Return", "Close", "Mktvl")
# 新增两列，分别表示当日涨跌和与均线的关系
dat <- dat %>%
  mutate(BullBear = ifelse(Return > 0, "Advancing",
                           ifelse(Return == 0, "Zero", "Declining")),
         UpDown50 = ifelse(Close >= SMA50, "Above", "Below"))
```

```
# 提取当日涨跌相关的数据
```

```
dat.home.1 <- dat %>%
  group_by(BullBear) %>%
  summarise(Stocks = n()) %>%
  mutate(Stocks_ratio = Stocks / 300,
         HS300 = 1,
         ys = c(0.1, 0.9, 0.5),
         yl = paste(round(Stocks_ratio * 100, 1), "%", sep = ""))
dat.home.1$yl[3] <- ""
```

```
# 设定当日涨跌变量的因子水平
```

```
dat.home.1$BullBear <-
  factor(dat.home.1$BullBear,
        levels = c("Declining", "Zero", "Advancing"),
        ordered = T)
```

```
# 显示数据
```

```
dat.home.1
```

	BullBear <ord>	Stocks <int>	Stocks_ratio <dbl>	HS300 <dbl>	ys <dbl>	yl <chr>
	Advancing	211	0.70333333	1	0.1	70.3%
	Declining	75	0.25000000	1	0.9	25%
	Zero	14	0.04666667	1	0.5	

```
3 rows
```

```
# 画出当日涨跌股票数量分布的堆积条形图
```

```
ggplot(dat.home.1, aes(x = HS300, y = Stocks_ratio, fill = BullBear)) +
  geom_bar(stat = "identity", width = 0.5) + # width 设置宽度
```

```

scale_fill_manual(values = c("#de7e7e", "grey", "#83ca83")) + # 设置颜色

annotate("text", x = 1.5, y = 0.15, # 上方文字显示
        label = dat.home.1$BullBear[1],
        size = 6) +
annotate("text", x = 1.5, y = 0.05,
        label = dat.home.1$Stocks[1],
        size = 6,
        colour = "#00B060") +
annotate("text", x = 1.5, y = 0.85,
        label = dat.home.1$BullBear[2],
        size = 6) +
annotate("text", x = 1.5, y = 0.95,
        label = dat.home.1$Stocks[2],
        size = 6,
        colour = "#FF4500") +
geom_text(y = dat.home.1$ys, # 百分比数值显示
        label = dat.home.1$yl,
        size = 9,
        colour = "white") +
xlim(0, 2) +
coord_flip() + # 坐标轴翻转

theme(panel.background = element_blank(), # 设置主题
      axis.title = element_blank(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.position = "none")

```



功能六：股票业绩分布条形图

绘制沪深 300 成分股中业绩高于和低于 50 日均线水平的股票分布条形图

```
# 提取与 50 日均线相关的数据
```

```
dat.home.2 <- dat %>%  
  group_by(UpDown50) %>%  
  summarise(Stocks = n()) %>%  
  mutate(Stocks_ratio = Stocks / 300,  
         HS300 = 1,  
         ys = c(0.1, 0.9),  
         yl = paste(round(Stocks_ratio * 100, 1), "%", sep = ""))
```

```
dat.home.2$UpDown50 <-  
  factor(dat.home.2$UpDown50,  
        levels = c("Below", "Above"),  
        ordered = T)
```

```
# 显示数据
```

```
dat.home.2
```

UpDown50 <ord>	Stocks <int>	Stocks_ratio <dbl>	HS300 <dbl>	ys <dbl>	yl <chr>
Above	145	0.4833333	1	0.1	48.3%
Below	155	0.5166667	1	0.9	51.7%

2 rows

```
# 画出超过、低于 50 日均线的股票分布条形图
```

```
ggplot(dat.home.2, aes(x = HS300, y = Stocks_ratio, fill = UpDown50)) +  
  geom_bar(stat = "identity", width = 0.5) +  
  scale_fill_manual(values = c("#de7e7e", "#83ca83")) + # 设置颜色
```

```
annotate("text", x = 1.5, y = 0.15, # 上方文字显示
```

```
  label = dat.home.2$UpDown50[1],  
  size = 6) +
```

```
annotate("text", x = 1.5, y = 0.05,  
  label = dat.home.2$Stocks[1],  
  size = 6,  
  colour = "#00B060") +
```

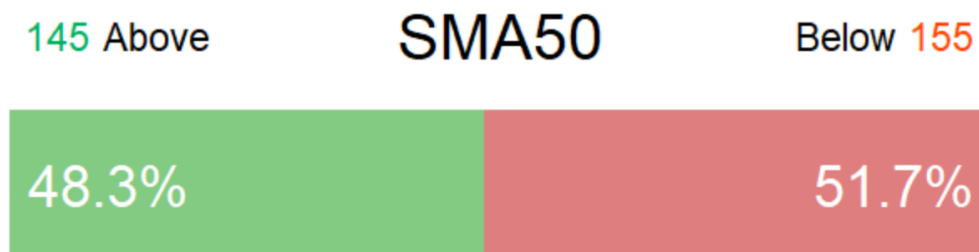
```
annotate("text", x = 1.5, y = 0.85,  
  label = dat.home.2$UpDown50[2],  
  size = 6) +
```

```
annotate("text", x = 1.5, y = 0.95,  
  label = dat.home.2$Stocks[2],  
  size = 6,  
  colour = "#FF4500") +
```

```

geom_text(y = dat.home.2$ys,
          label = dat.home.2$yl,
          size = 9,
          colour = "white") +
annotate("text", x = 1.5, y = 0.5, # 百分比数字显示
          label = "SMA50", size = 10) +
xlim(0, 2) +
coord_flip() +
theme(panel.background = element_blank(), # 设置主题
      axis.title = element_blank(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.position = "none")

```



功能七：1 DAY PERFORMANCE 谱图

#绘制沪深 300 成分股中显示行业及公司的 1 DAY PERFORMANCE

计算行业股票总市值

```

ind.mktvl <- dat %>%
  group_by(Industry) %>%
  summarise(sum.mktvl = sum(Mktvl))

```

计算绘图所需数据

```

dat.2 <- dat %>%
  left_join(ind.mktvl, by = "Industry") %>%
  mutate(Mktvl_ratio = Mktvl / sum.mktvl, # 计算市值在行业内占比

```

```

        Label = ifelse(Mktvl_ratio > 0.09, Stknm, "")) %>% # 市值在业内占比

```

达到 0.09 显示名称

```

arrange(Industry, desc(Return)) %>%
  plyr::ddply("Industry", transform, Label_y = cumsum(Mktvl_ratio) -
0.5*Mktvl_ratio)

```

画出分行业及行业内部分公司的 1 DAY PERFORMANCE SPECTRUM

```

ggplot(dat.2, aes(x = Industry, y = Mktvl_ratio, fill = Return)) +
  geom_bar(stat = "identity", width = 0.75) +
  scale_fill_gradient2(low = "#00ff00", mid = "#000200", high = "#ff0000") + #

```

设置颜色

```

geom_text(aes(y = Label_y, label = Label), size = 2.9, colour = "white") + # 显示

```

股票名称

```

scale_y_continuous(expand = c(0, 0)) +
coord_flip() +

```

```

labs(title = "1 DAY PERFORMANCE SPECTRUM") + # 设置标题文本

```

```

theme(panel.background = element_blank(), # 设置主题

```

```

axis.title = element_blank(),
axis.ticks = element_blank(),
axis.text.x = element_blank(),
legend.position = "none",
plot.title = element_text(hjust = -0.4))

```

