

数据分析、展现与R语言 第9周

2013.03.23

【声明】 本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

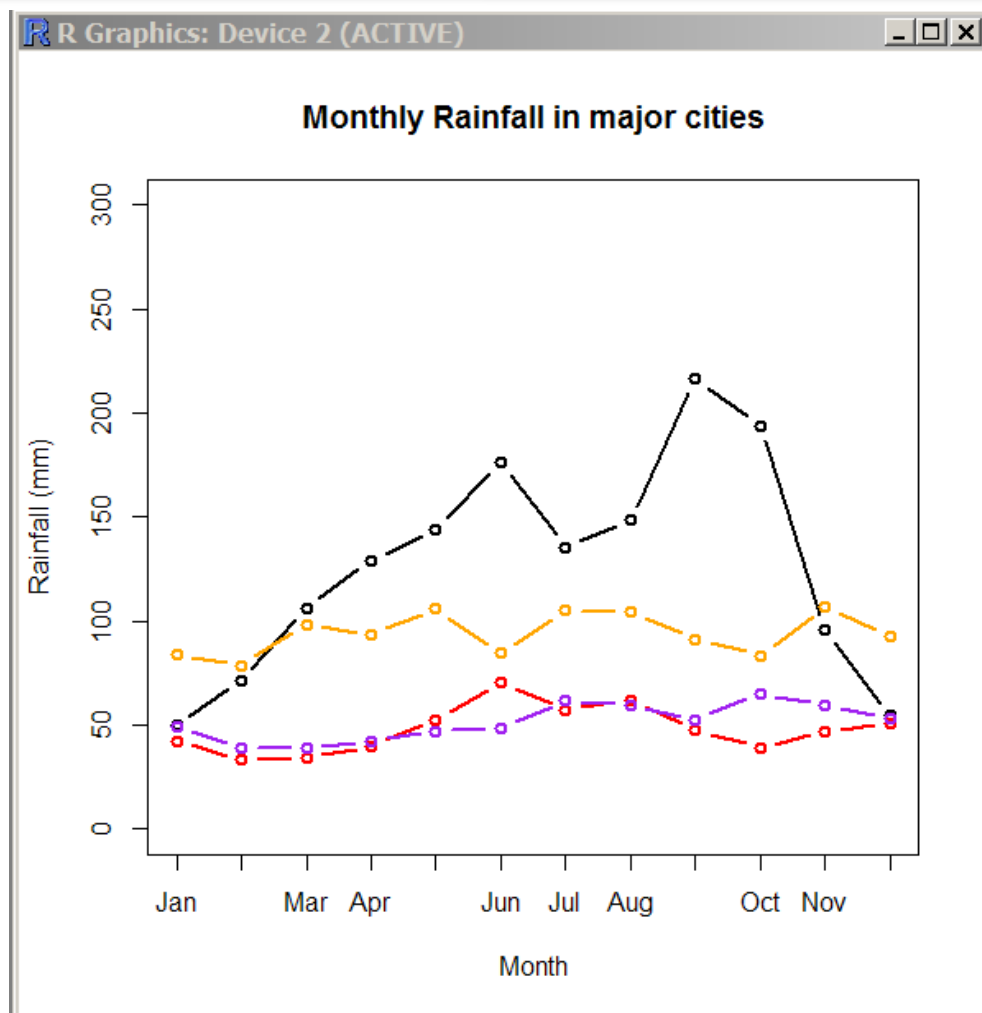
课程详情访问炼数成金培训网站

<http://edu.dataguru.cn>

■ 使用cityrain.csv数据

```
rain<-read.csv("cityrain.csv")
plot(rain$Tokyo,type="b",lwd=2,
xaxt="n",ylim=c(0,300),col="black",
xlab="Month",ylab="Rainfall (mm)",
main="Monthly Rainfall in major cities")
axis(1,at=1:length(rain$Month),labels=rain$Month)
lines(rain$Berlin,col="red",type="b",lwd=2)
lines(rain$NewYork,col="orange",type="b",lwd=2)
lines(rain$London,col="purple",type="b",lwd=2)
legend("topright",legend=c("Tokyo","Berlin","New York","London"),
lty=1,lwd=2,pch=21,col=c("black","red","orange","purple"),
ncol=2,bty="n",cex=0.8,
text.col=c("black","red","orange","purple"),
inset=0.01)
```

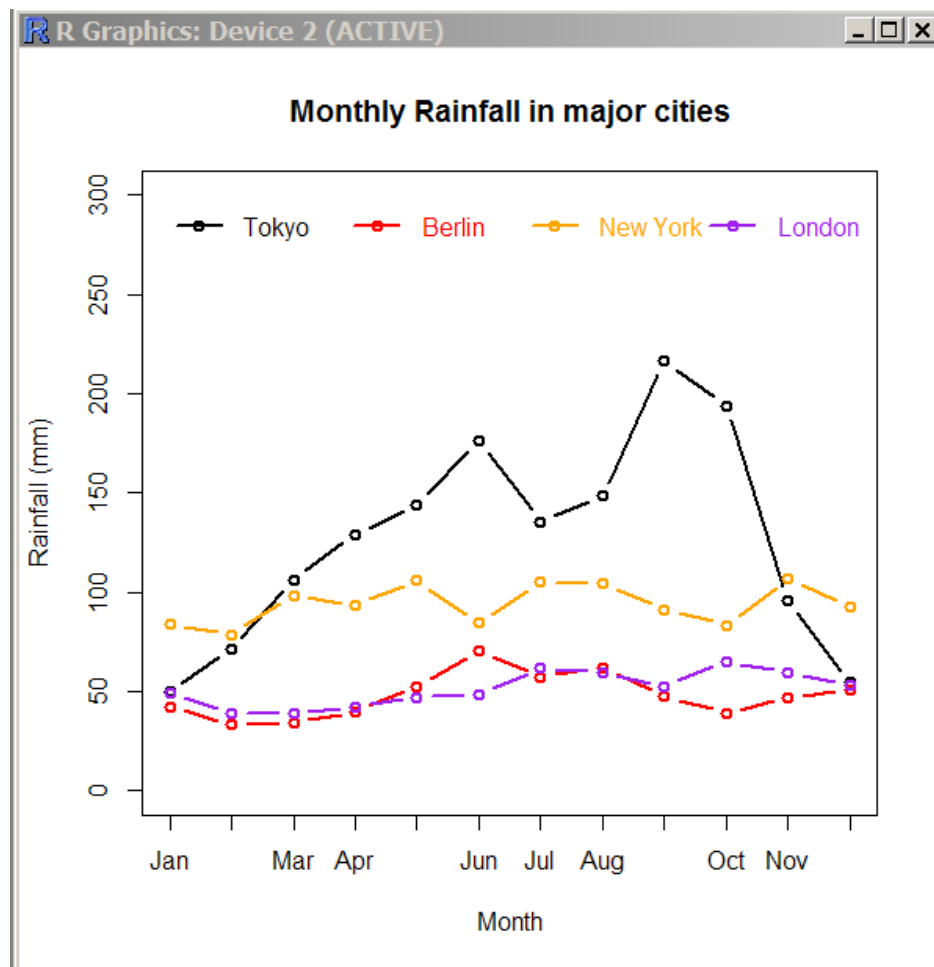
2013.03.23



2013.03.23

给线图增加legend (图例)

```
legend(1,300,legend=c("Tokyo","Berlin","New York","London"),  
      lty=1,lwd=2,pch=21,col=c("black","red","orange","purple"),  
      horiz=TRUE,bty="n",bg="yellow",  
      cex=1,  
      text.col=c("black","red","orange","purple"))
```



2013.03.23

用线图描画时间序列—GDP的例子

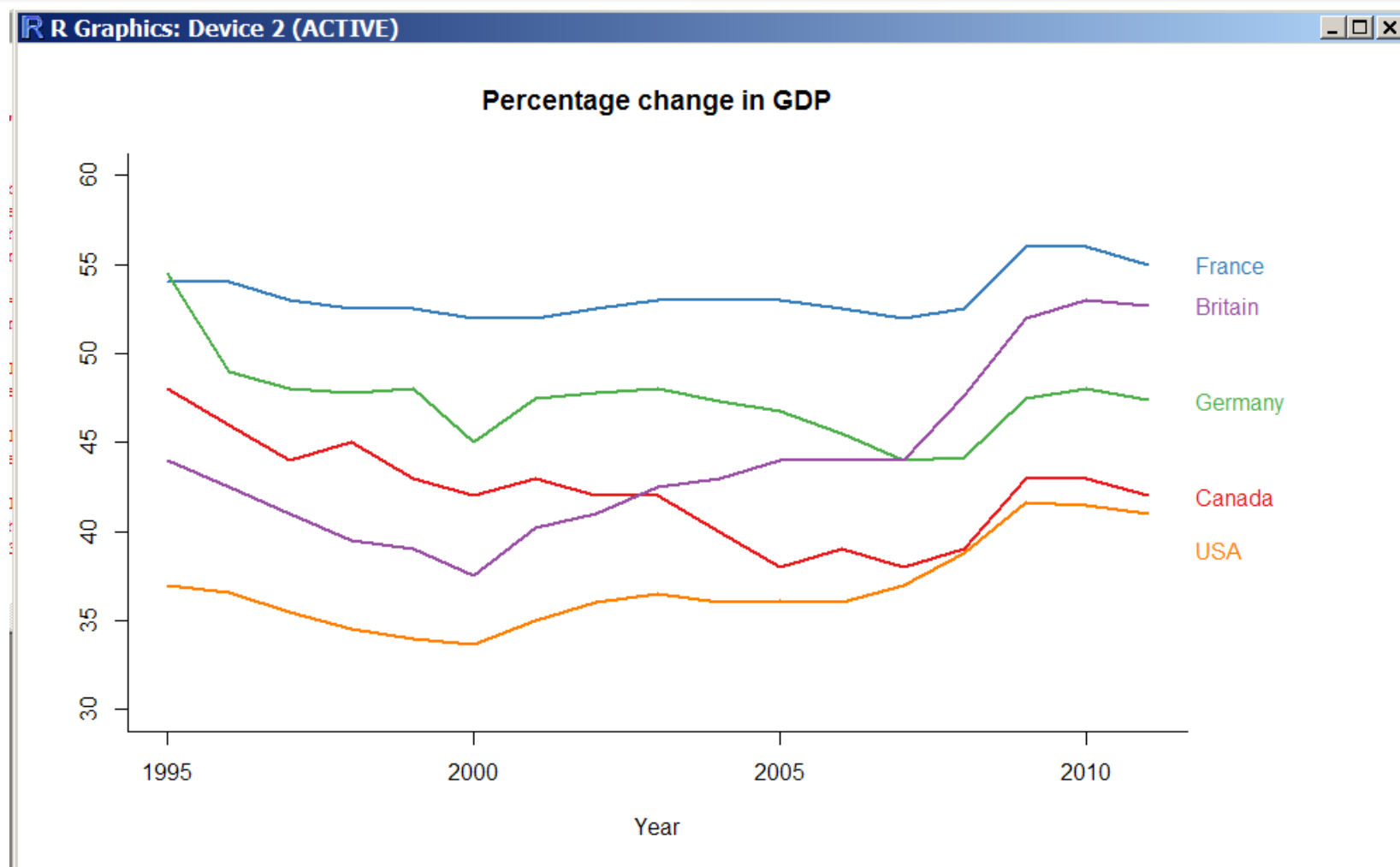
- 注意，工作目录转向第四章数据

```
gdp<-read.table("gdp_long.txt",header=T)
library(RColorBrewer)
pal<-brewer.pal(5,"Set1")
par(mar=par()$mar+c(0,0,0,2),bty="l")
plot(Canada~Year,data=gdp,type="l",lwd=2,lty=1,ylim=c(30,60),
col=pal[1],main="Percentage change in GDP",ylab="")
mtext(side=4,at=gdp$Canada[length(gdp$Canada)],text="Canada",
col=pal[1],line=0.3,las=2)
```

代码（续）

```
lines(gdp$France~gdp$Year,col=pal[2],lwd=2)
mtext(side=4,at=gdp$France[length(gdp$France)],text="France",
col=pal[2],line=0.3,las=2)
lines(gdp$Germany~gdp$Year,col=pal[3],lwd=2)
mtext(side=4,at=gdp$Germany[length(gdp$Germany)],text="Germany",
col=pal[3],line=0.3,las=2)
lines(gdp$Britain~gdp$Year,col=pal[4],lwd=2)
mtext(side=4,at=gdp$Britain[length(gdp$Britain)],text="Britain",
col=pal[4],line=0.3,las=2)
lines(gdp$USA~gdp$Year,col=pal[5],lwd=2)
mtext(side=4,at=gdp$USA[length(gdp$USA)]-2,
text="USA",col=pal[5],line=0.3,las=2)
```

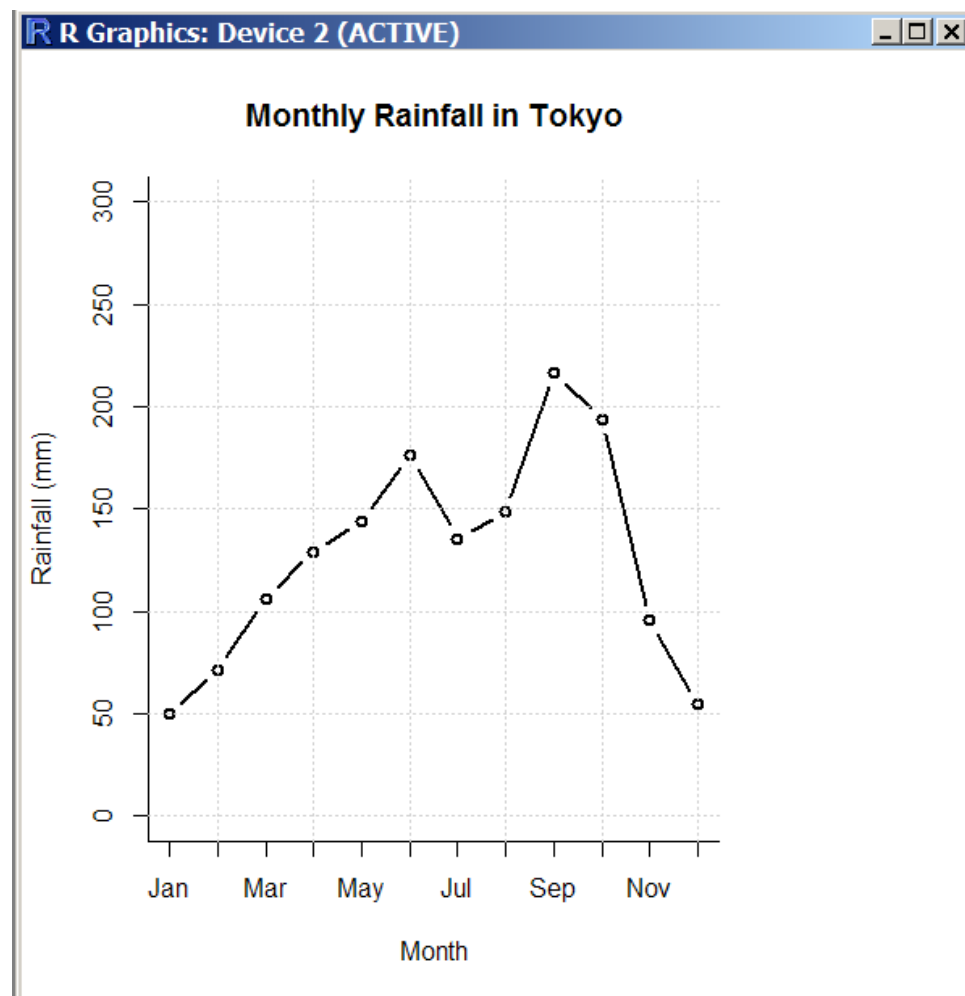
GDP变化图



2013.03.23

画底纹格子

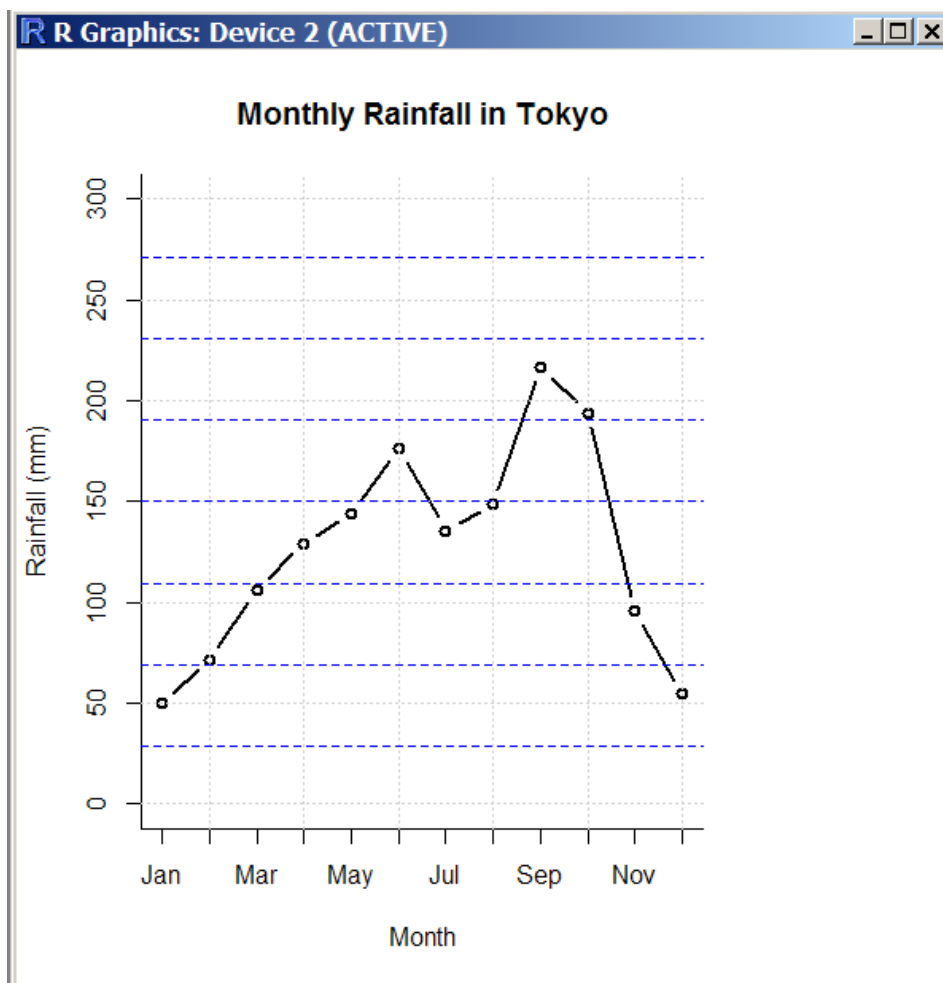
```
rain<-read.csv("cityrain.csv")  
plot(rain$Tokyo,type="b",lwd=2,  
xaxt="n",ylim=c(0,300),col="black",  
xlab="Month",ylab="Rainfall  
  (mm)",  
main="Monthly Rainfall in Tokyo")  
axis(1,at=1:length(rain$Month),label  
  ls=rain$Month)  
grid()
```



2013.03.23

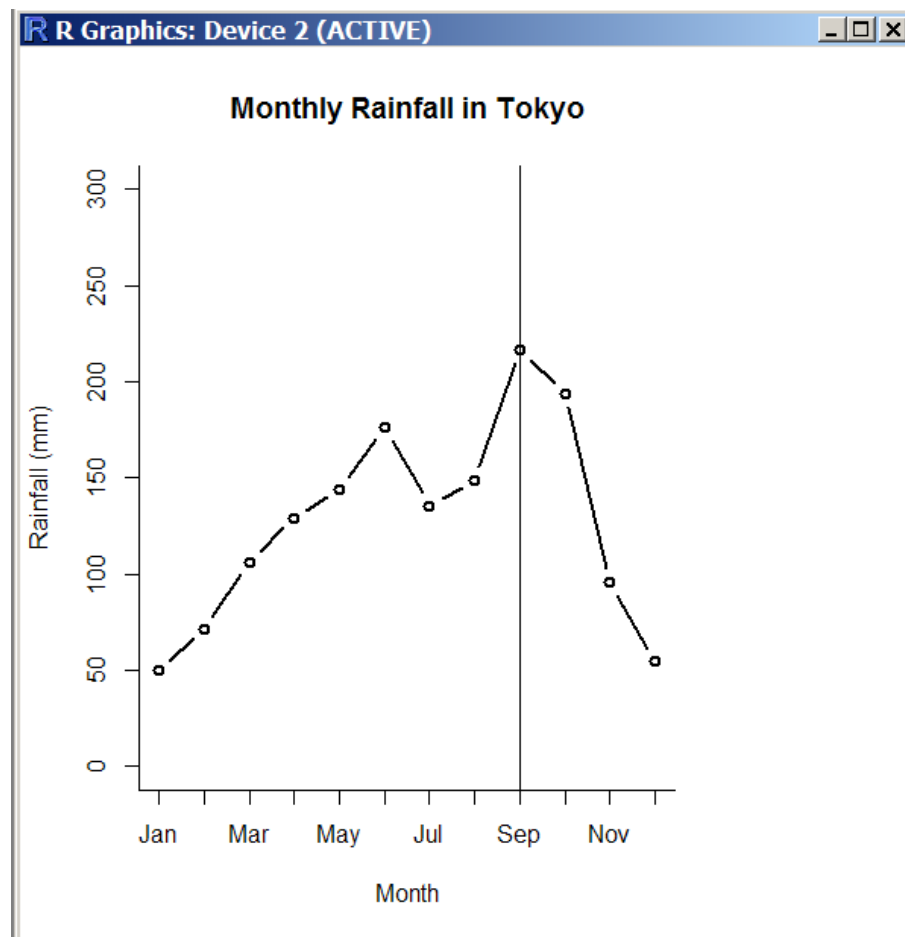
水平线底纹

```
grid(nx=NA, ny=8,  
lwd=1,lty=2,col="blue")
```



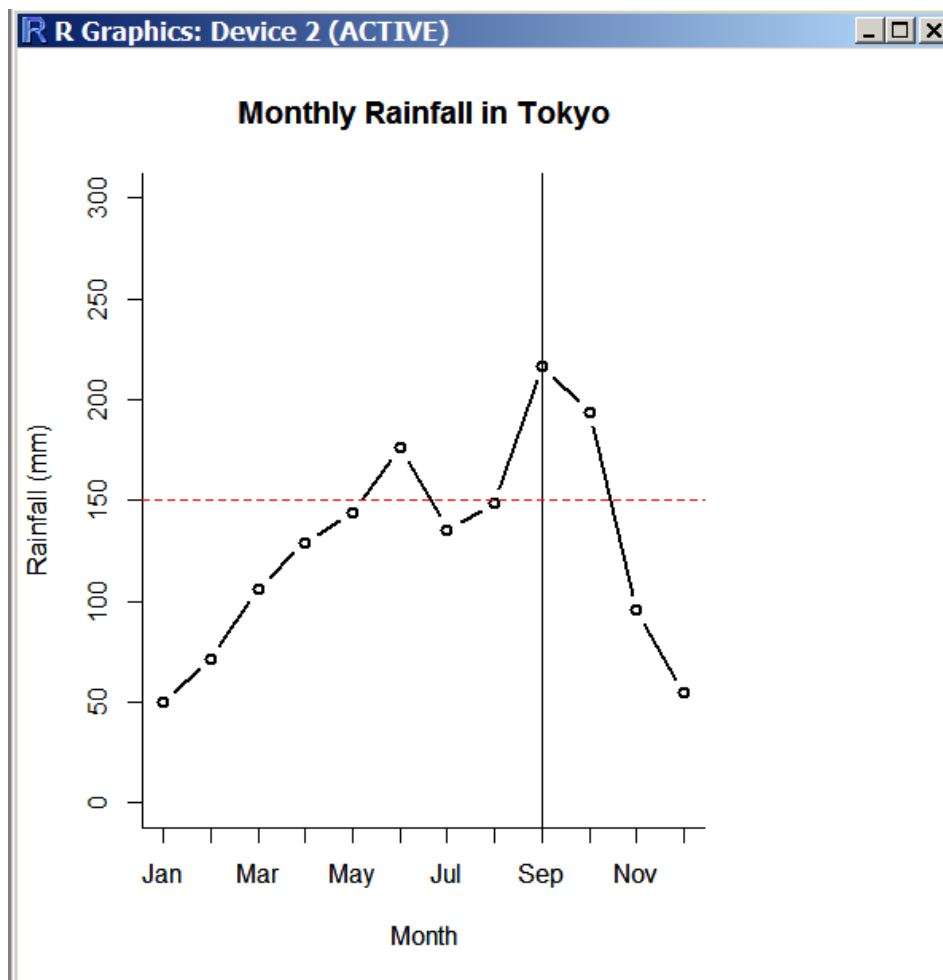
2013.03.23

```
rain <- read.csv("cityrain.csv")  
plot(rain$Tokyo,type="b",lwd=2,  
xaxt="n",ylim=c(0,300),col="black",  
xlab="Month",ylab="Rainfall (mm)",  
main="Monthly Rainfall in Tokyo")  
axis(1,at=1:length(rain$Month),labels  
=rain$Month)  
abline(v=9)
```



垂直线和水平线

```
abline(h=150,col="red",lty=2)
```



2013.03.23

Sparkline

```
rain <- read.csv("cityrain.csv")
par(mfrow=c(4,1),mar=c(5,7,4,2),omi=c(0.2,2,0.2,2))
for(i in 2:5)
{
plot(rain[,i],ann=FALSE,axes=FALSE,type="l",
col="gray",lwd=2)
mtext(side=2,at=mean(rain[,i]),names(rain[i]),
las=2,col="black")
mtext(side=4,at=mean(rain[,i]),mean(rain[i]),
las=2,col="black")
points(which.min(rain[,i]),min(rain[,i]),pch=19,col="blue")
points(which.max(rain[,i]),max(rain[,i]),pch=19,col="red")
}
```

Tokyo 126.8

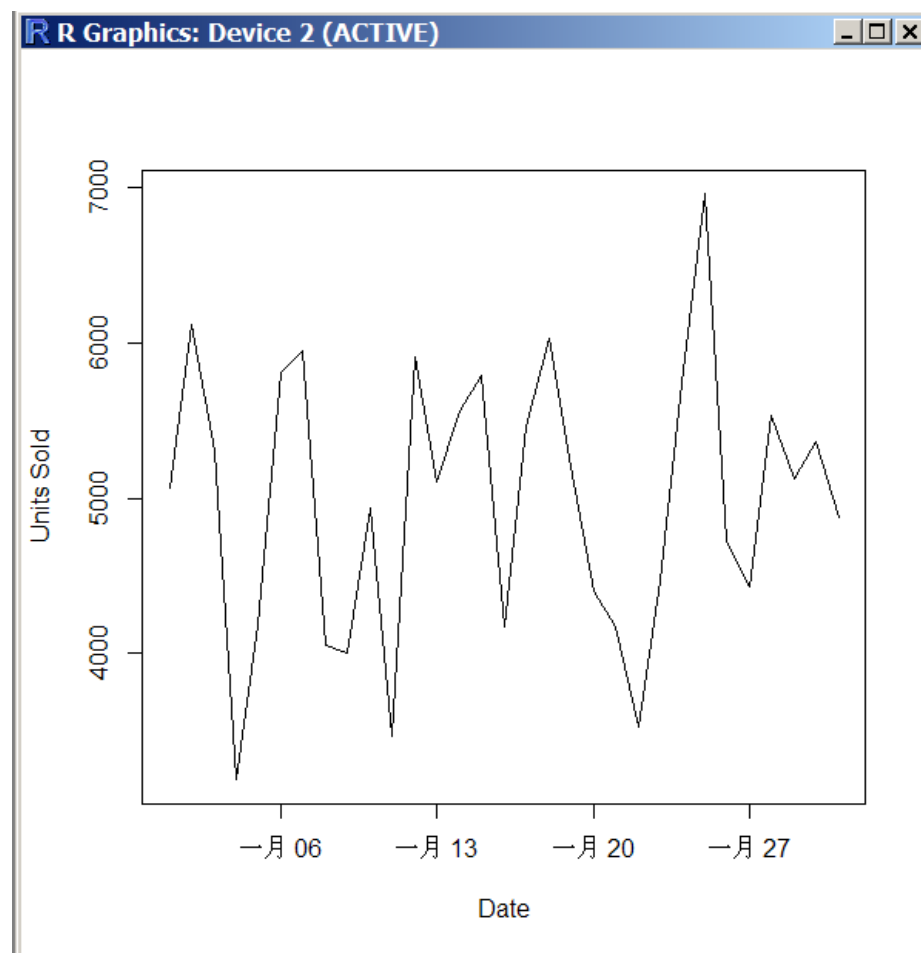
NewYork 93.975

London 51.35

Berlin 48.075

时间序列图

```
sales<-read.csv("dailysales.csv")  
plot(sales$units~as.Date(sales$d  
    ate,"%d/%m/%y"),type="l",  
xlab="Date",ylab="Units Sold")
```

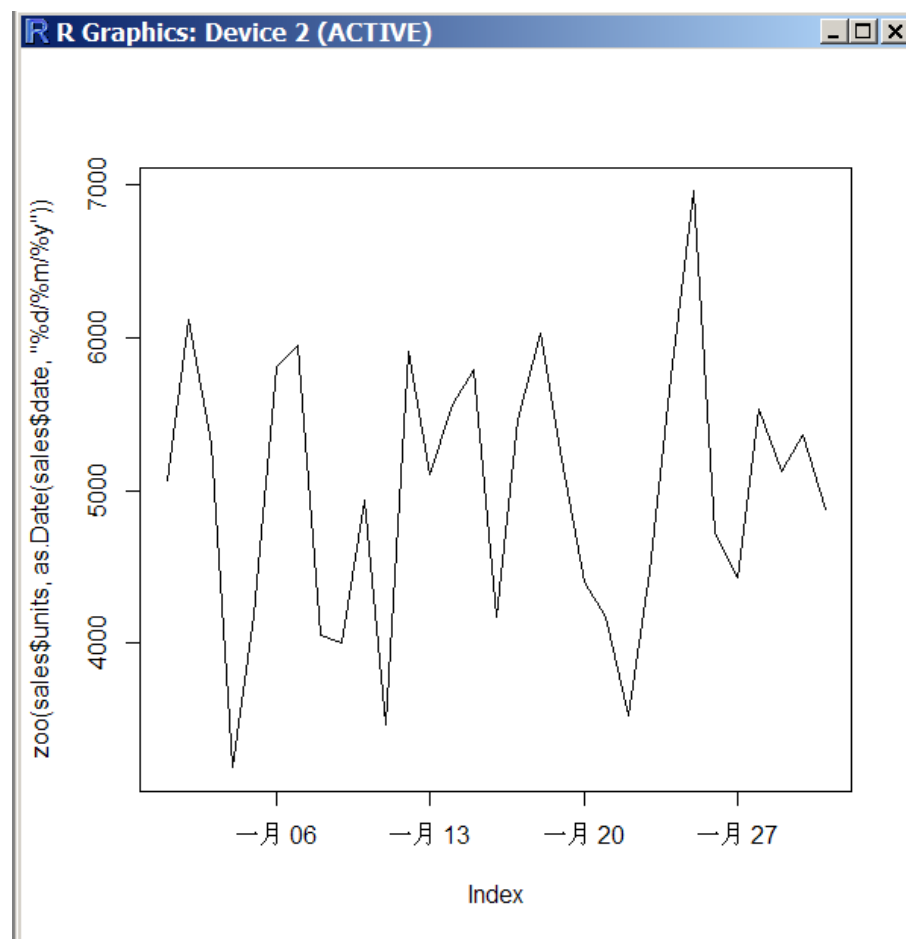


2013.03.23

处理时间序列数据的包zoo

```
library(zoo)
```

```
plot(zoo(sales$units,as.Date(sales  
$date,"%d/%m/%y")))
```

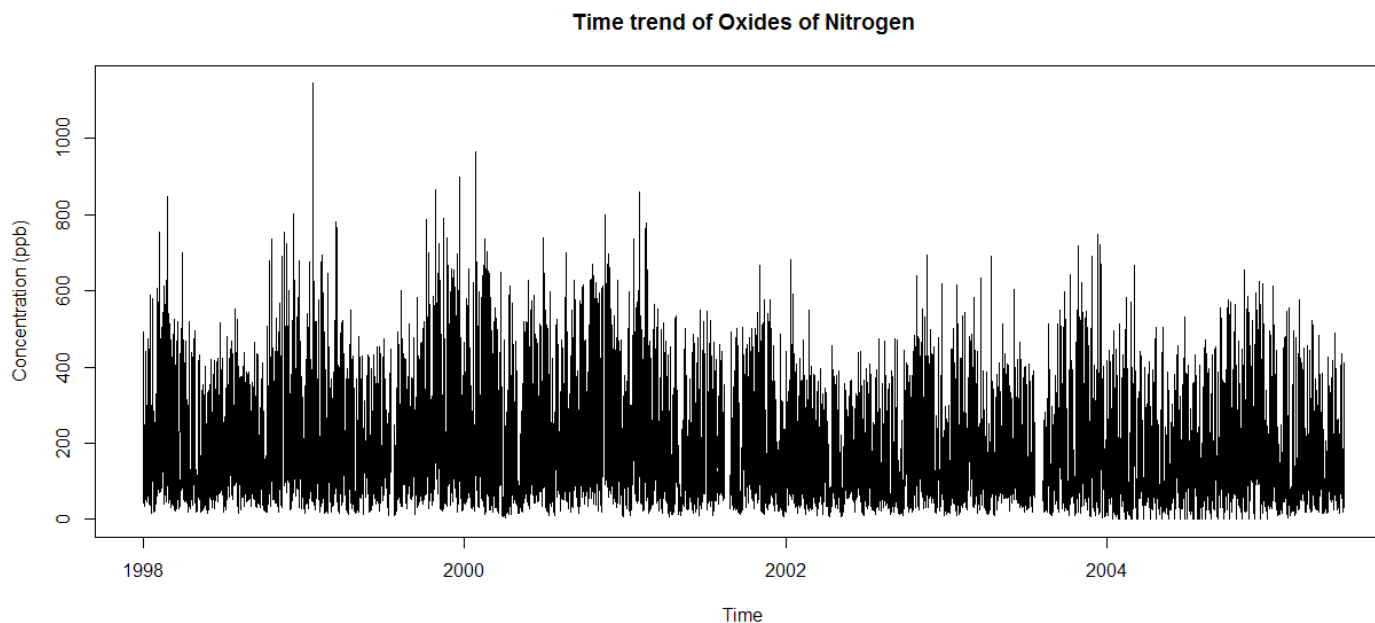


2013.03.23

数据较多的例子

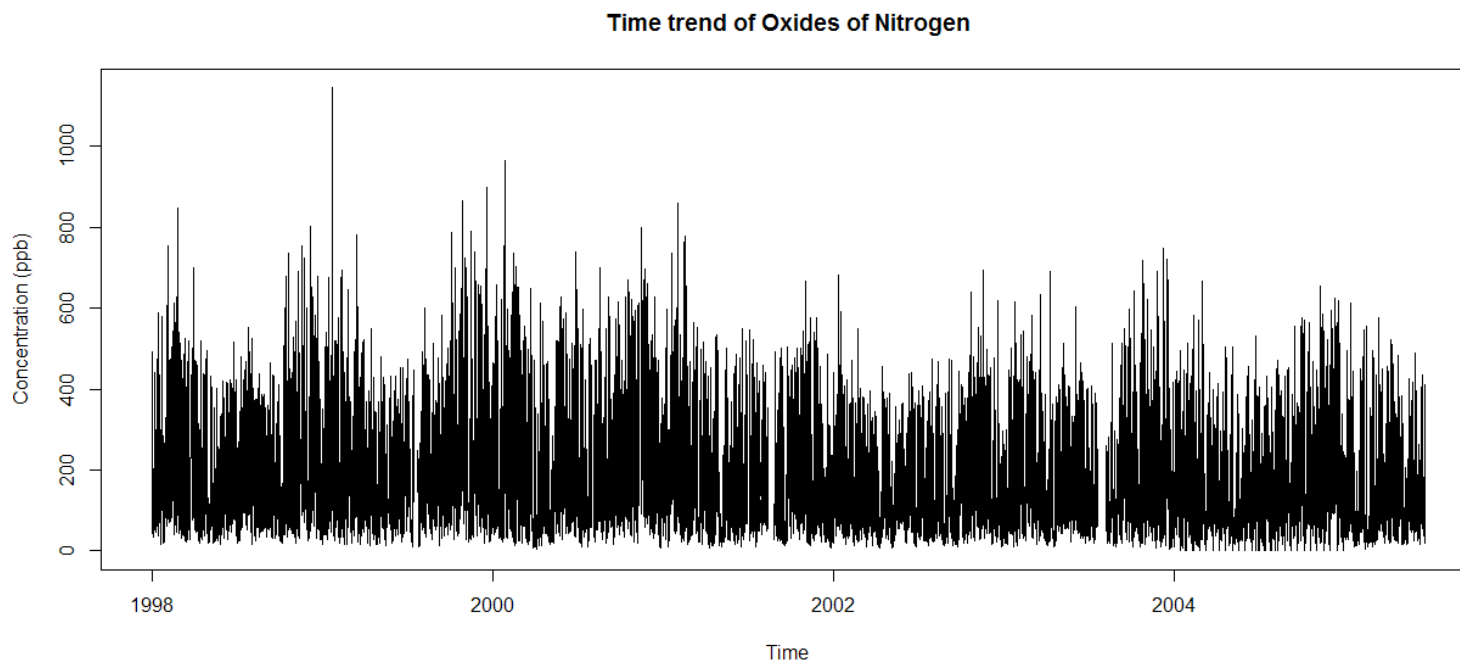
```
air<-read.csv("openair.csv")  
plot(air$nox~as.Date(air$date,"%d/%m/%Y %H:%M"),type="l",  
xlab="Time", ylab="Concentration (ppb)",  
main="Time trend of Oxides of Nitrogen")
```

需要第四章数据



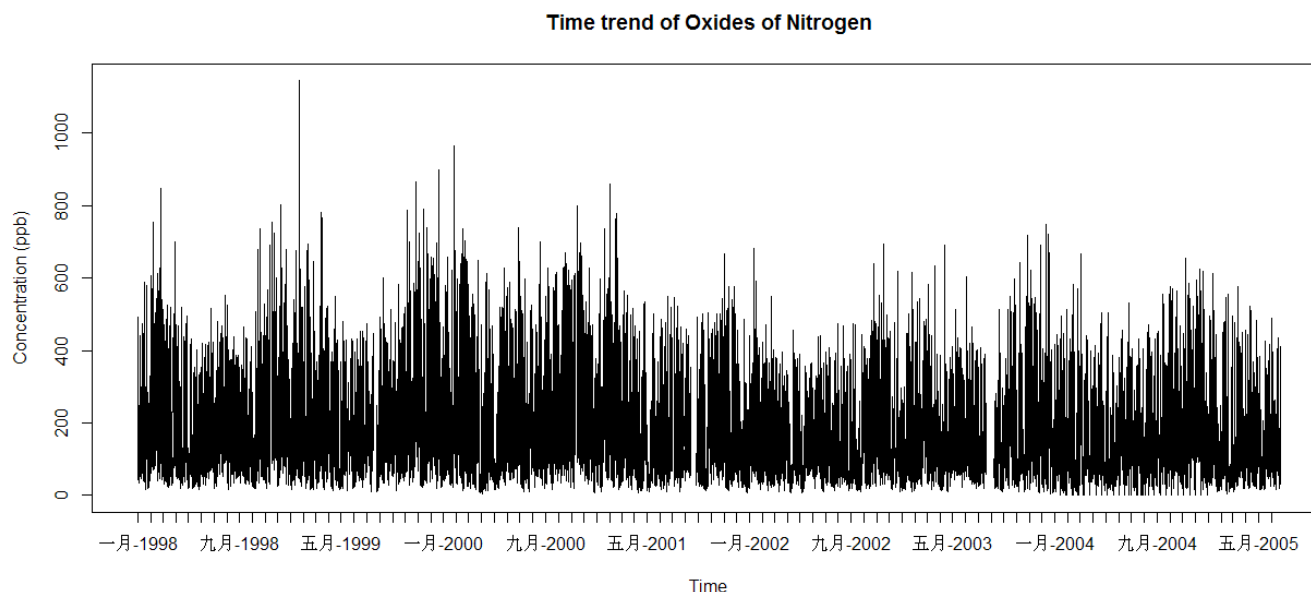
2013.03.23


```
plot(zoo(air$nox,as.Date(air$date,"%d/%m/%Y %H:%M")),  
xlab="Time", ylab="Concentration (ppb)",  
main="Time trend of Oxides of Nitrogen")
```



2013.03.23

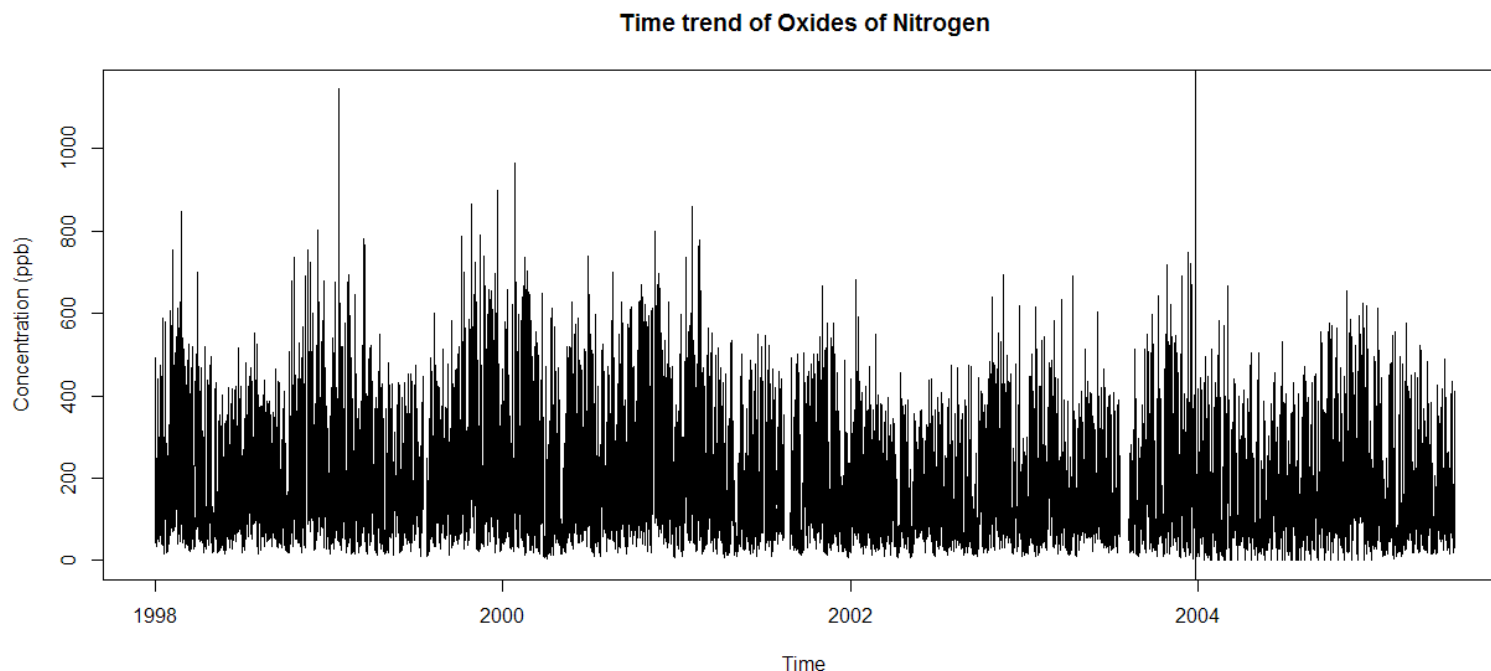
```
plot(air$nox~as.Date(air$date,"%d/%m/%Y %H:%M"),type="l",  
xaxt="n",  
xlab="Time", ylab="Concentration (ppb)",  
main="Time trend of Oxides of Nitrogen")  
xlabels<-strptime(air$date, format = "%d/%m/%Y %H:%M")  
axis.Date(1, at=xlabels[xlabels$mday==1], format="%b-%Y")
```



2013.03.23

标记特定的时间事件

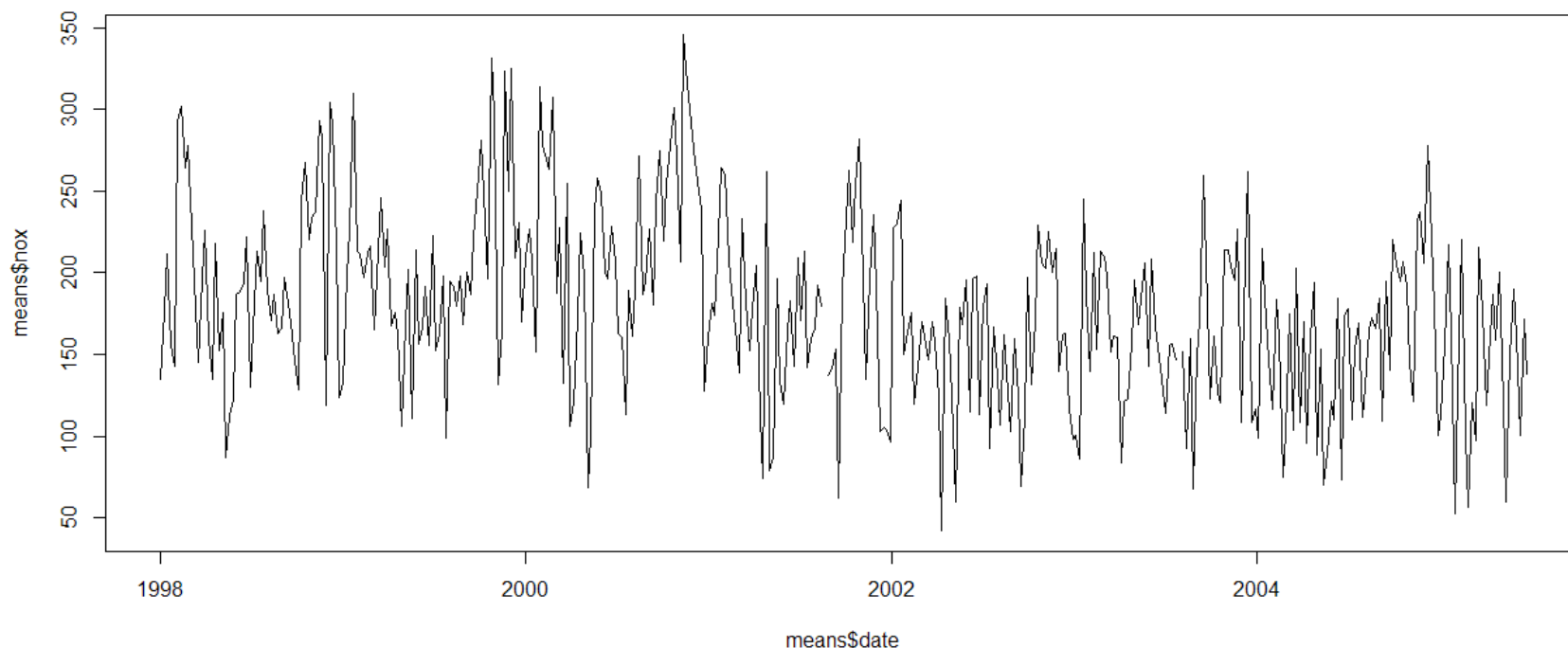
```
plot(air$nox~as.Date(air$date,"%d/%m/%Y %H:%M"),type="l",  
xlab="Time", ylab="Concentration (ppb)",  
main="Time trend of Oxides of Nitrogen")  
abline(v=as.Date("25/12/2003", "%d/%m/%Y"))
```



2013.03.23

求出均值后画时间序列

```
air$date = as.POSIXct(strptime(air$date, format = "%d/%m/%Y %H:%M ", "GMT"))  
means <- aggregate(air["nox"], format(air["date"], "%Y-%U"), mean, na.rm = TRUE)  
means$date <- seq(air$date[1], air$date[nrow(air)], length = nrow(means))  
plot(means$date, means$nox, type = "l")
```



2013.03.23

```
install.packages("quantmod")
```

```
install.packages("tseries")
```

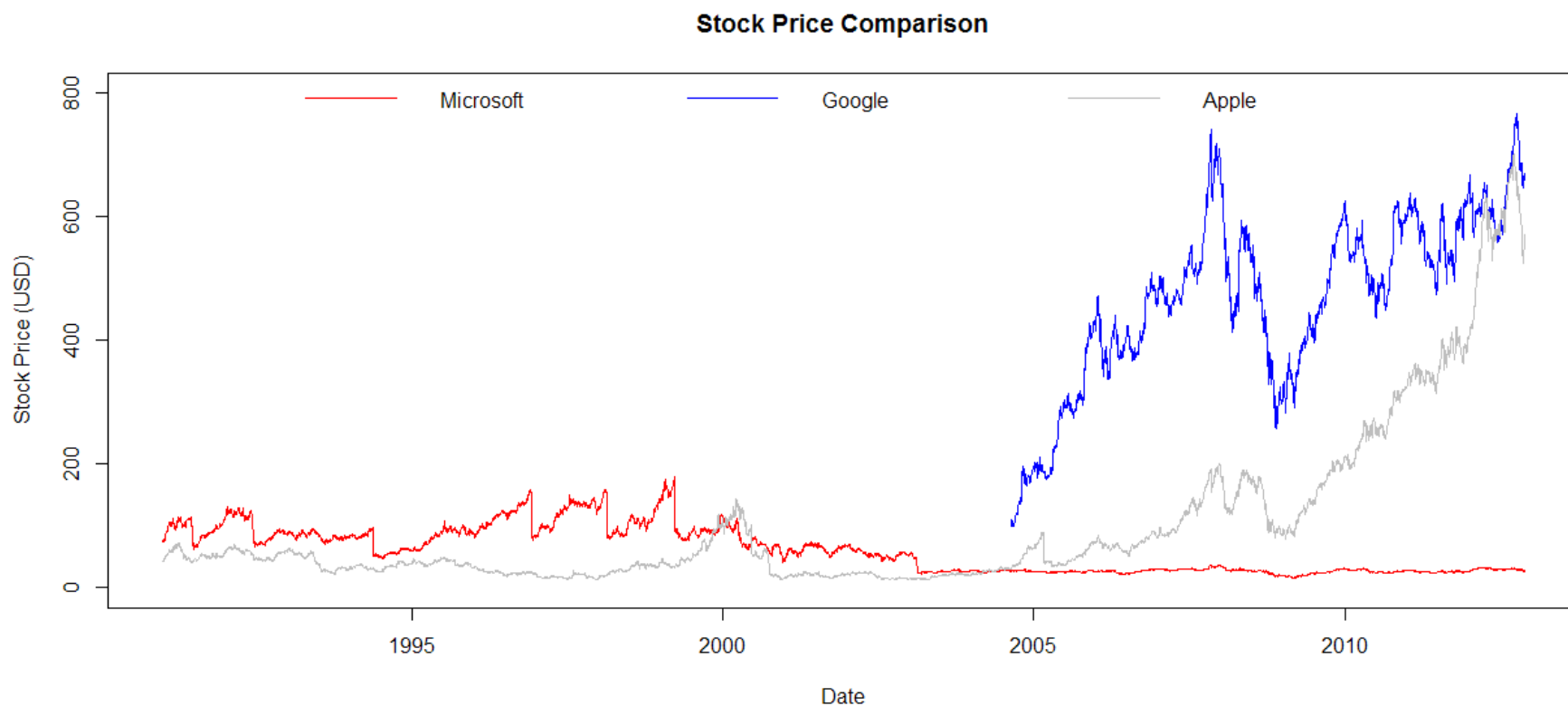
```
library(quantmod)
```

```
library(tseries)
```

抓取股票数据并画出趋势图

```
aapl<-get.hist.quote(instrument = "aapl", quote = c("Cl", "Vol"))
goog <- get.hist.quote(instrument = "goog", quote = c("Cl", "Vol"))
msft <- get.hist.quote(instrument = "msft", quote = c("Cl", "Vol"))
plot(msft$Close,main = "Stock Price Comparison",
ylim=c(0,800), col="red", type="l", lwd=0.5,
pch=19,cex=0.6, xlab="Date" ,ylab="Stock Price (USD)")
lines(goog$Close,col="blue",lwd=0.5)
lines(aapl$Close,col="gray",lwd=0.5)
legend("top",horiz=T,legend=c("Microsoft","Google","Apple"),
col=c("red","blue","gray"),lty=1,bty="n")
```

2013.03.23

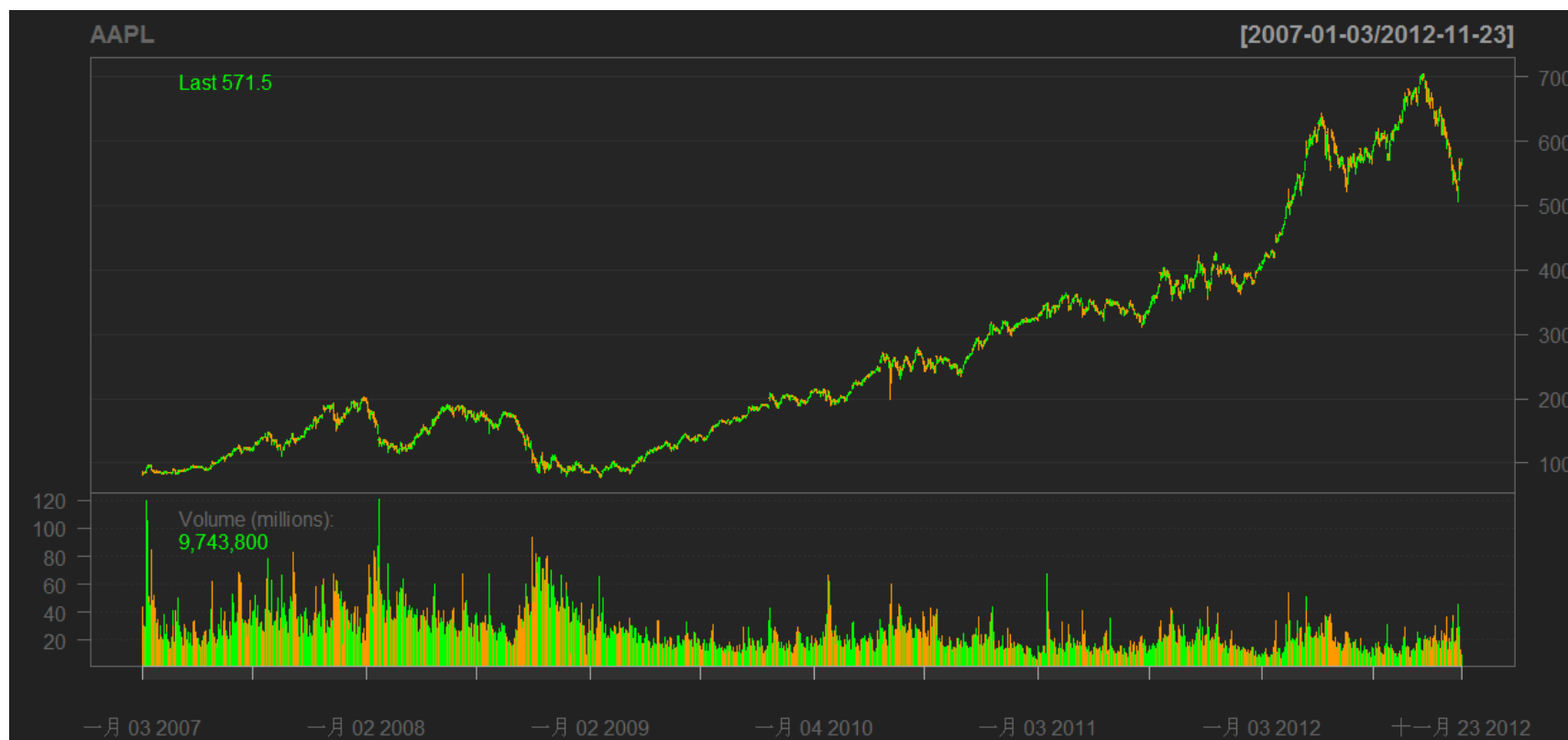


2013.03.23

使用quantmod 包的功能画图

```
getSymbols("AAPL",src="yahoo")
```

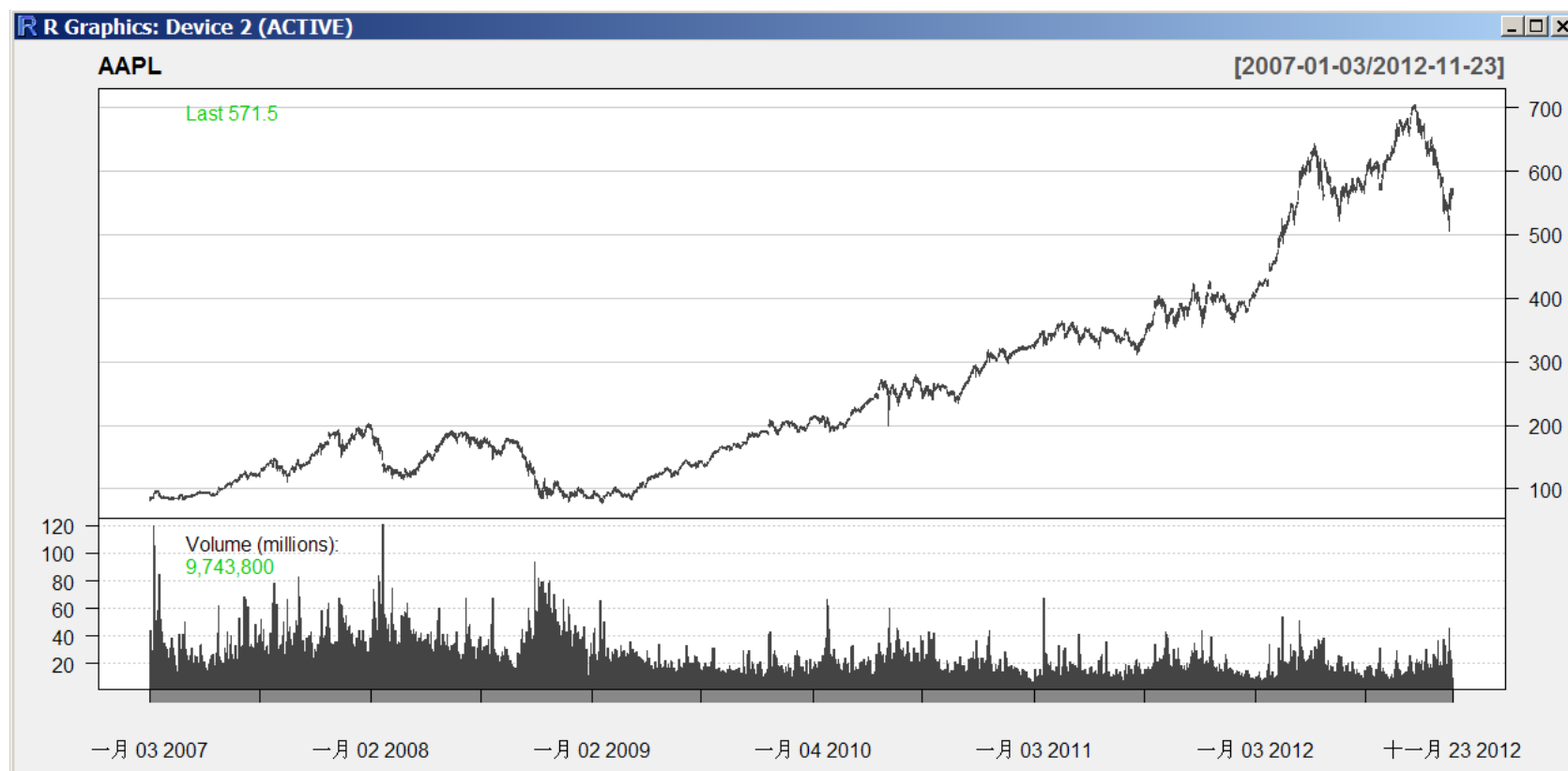
```
barChart(AAPL)
```



2013.03.23

蜡烛图

```
candleChart(AAPL,theme="white")
```



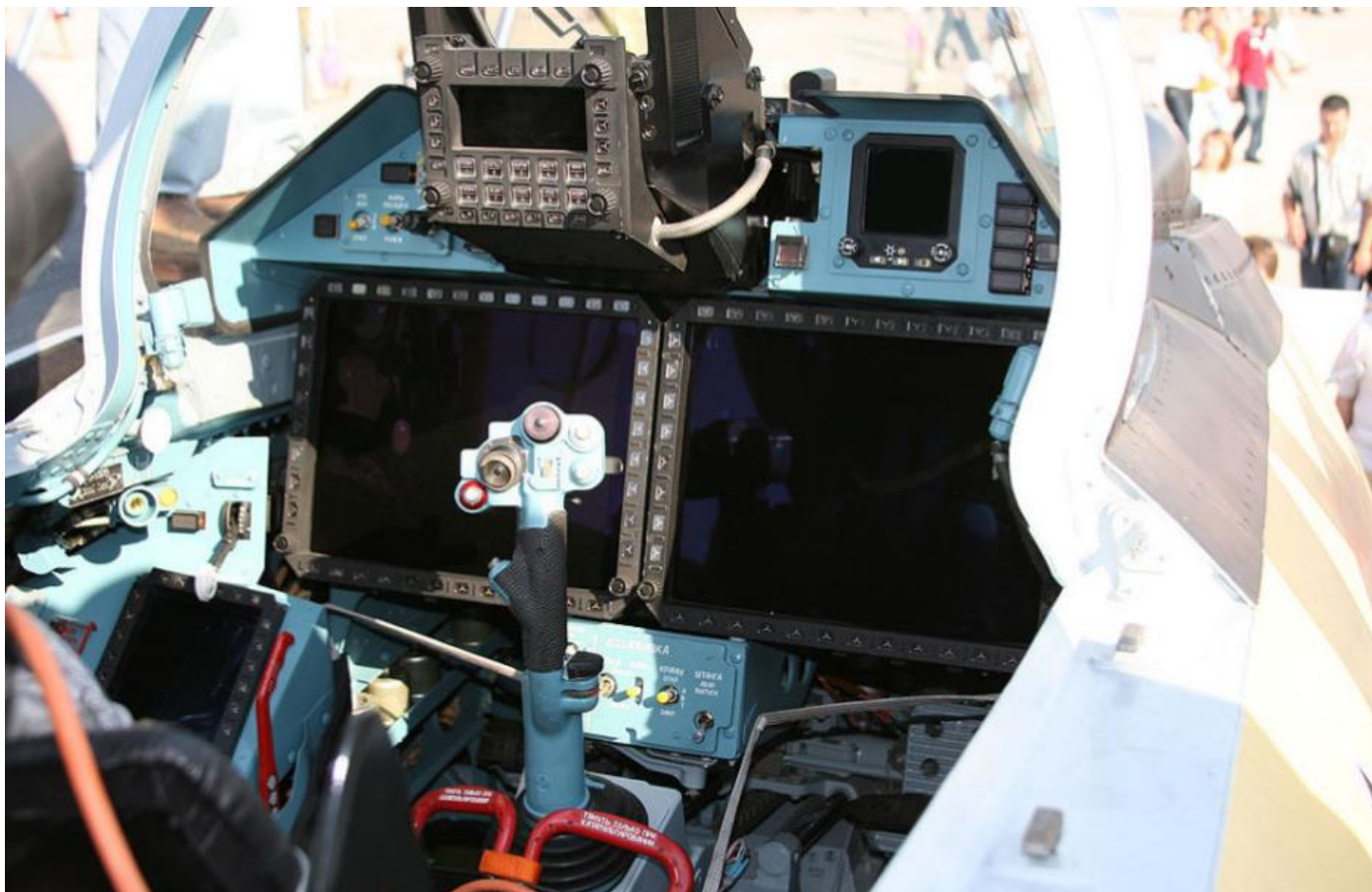
2013.03.23

案例：Su35



2013.03.23

■ MFI-35显示器



2013.03.23

Su35的液晶面板



2013.03.23

老式座舱



U9娱乐

www.gaibar.com

2013.03.23

其它俄罗斯战机座舱



2013.03.23

其它俄罗斯战机座舱

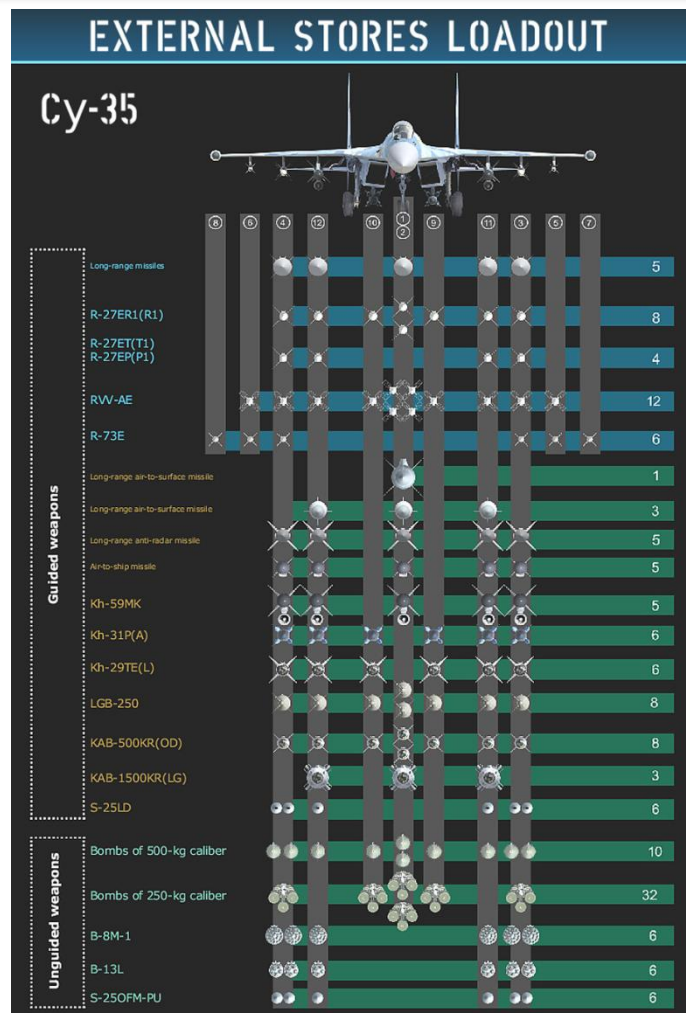


2013.03.23



2013.03.23

武器挂载信息图



2013.03.23

- Dataguru（炼数成金）是专业数据分析网站，提供教育，媒体，内容，社区，出版，数据分析业务等服务。我们的课程采用新兴的互联网教育形式，独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围，重竞争压力的特点，同时又发挥互联网的威力打破时空限制，把天南地北志同道合的朋友组织在一起交流学习，使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本，直线下降至百元范围，造福大众。我们的目标是：低成本传播高价值知识，构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情，请看我们的培训网站 <http://edu.dataguru.cn>



Thanks

FAQ时间