



R数据可视化—ggplot2包 第5周

2013.2.8

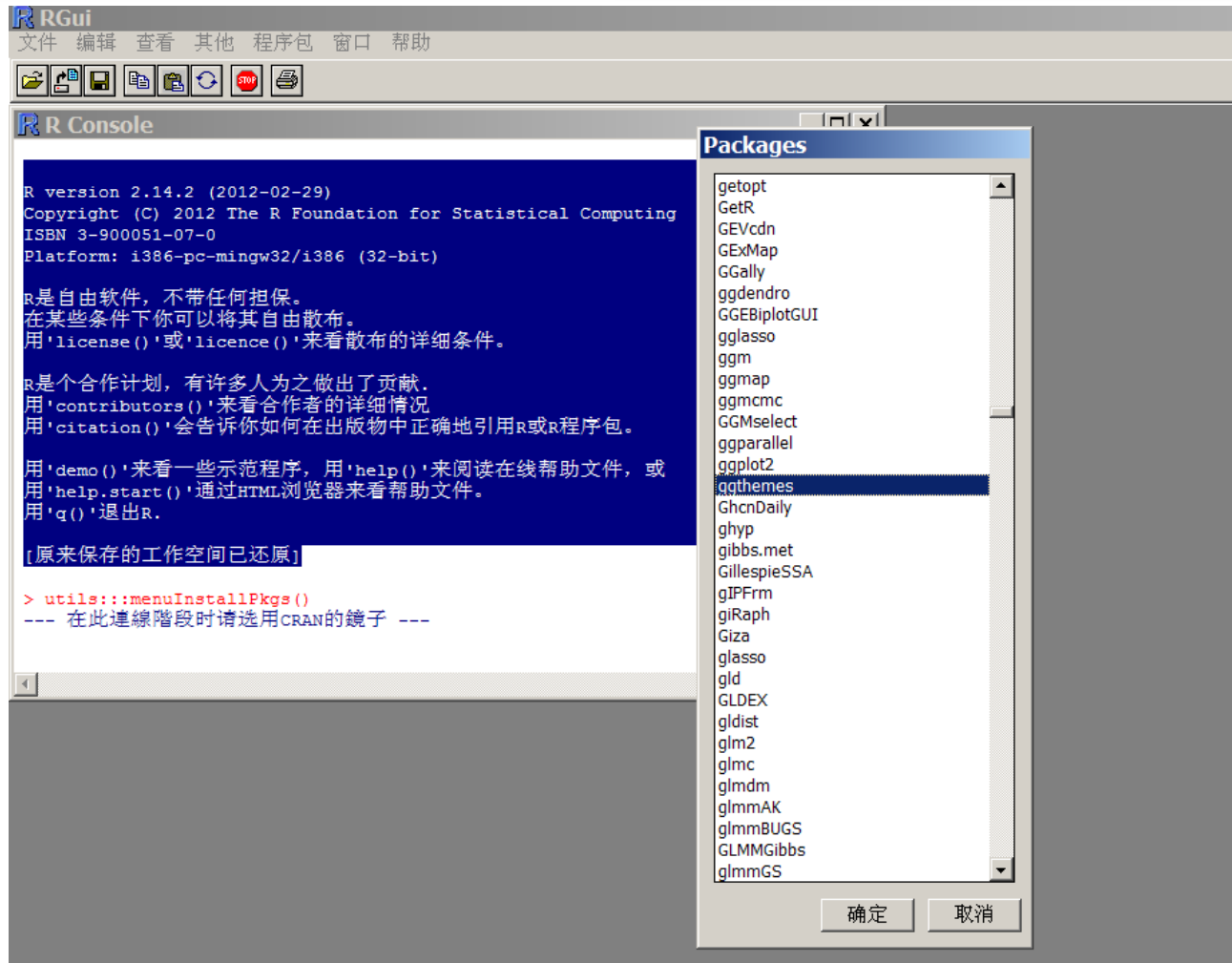
法律声明

【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>

ggthemes包

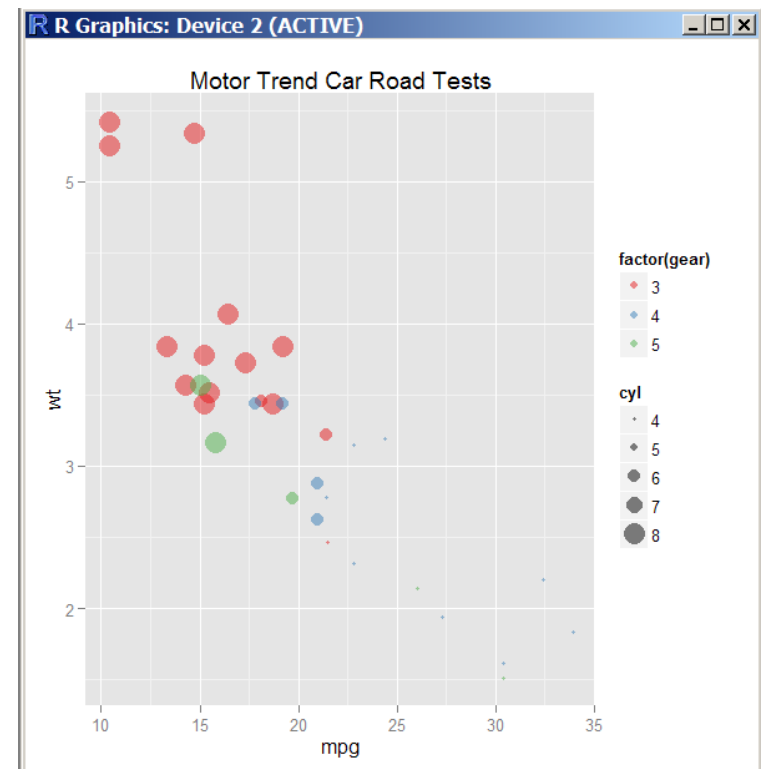


2013.2.8

例子

■ 传统ggplot2风格

```
ggplot(mtcars,aes(x=mpg,y=wt,size=cyl,colour=factor(gear)))+  
geom_point(alpha=.5)+ #透明度  
scale_size_area()+ #区域和数值成比例  
scale_colour_brewer(palette="Set1")+  
ggtitle("Motor Trend Car Road Tests")
```

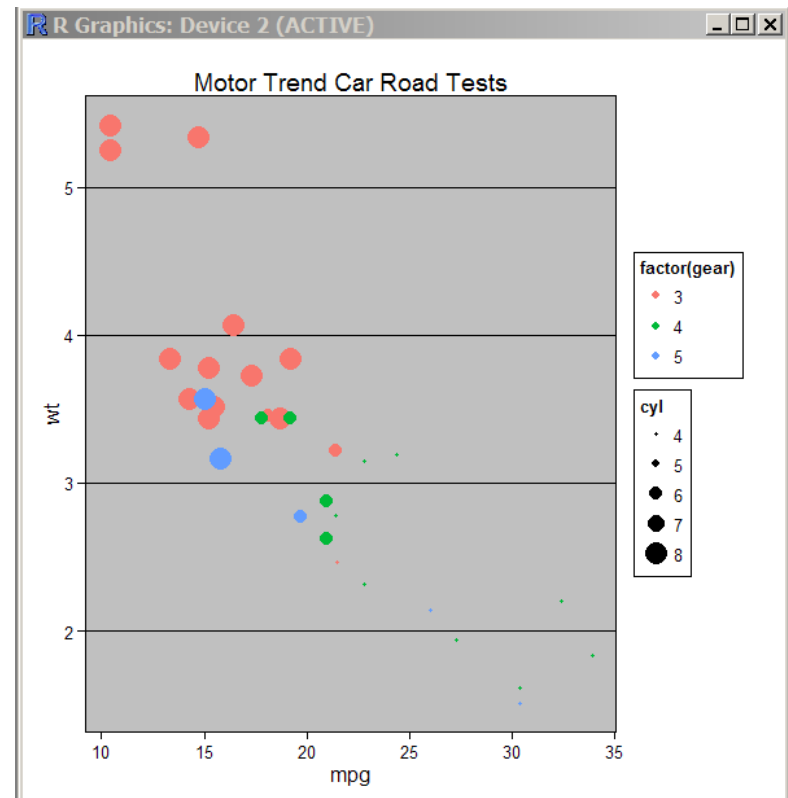


2013.2.8

例子

■ Excel风格

```
ggplot(mtcars,aes(x=mpg,y=wt,size=cyl,colour=factor(gear)))+  
geom_point()+  
theme_excel() +  
scale_fill_excel()+  
ggtitle("Motor Trend Car Road Tests")
```

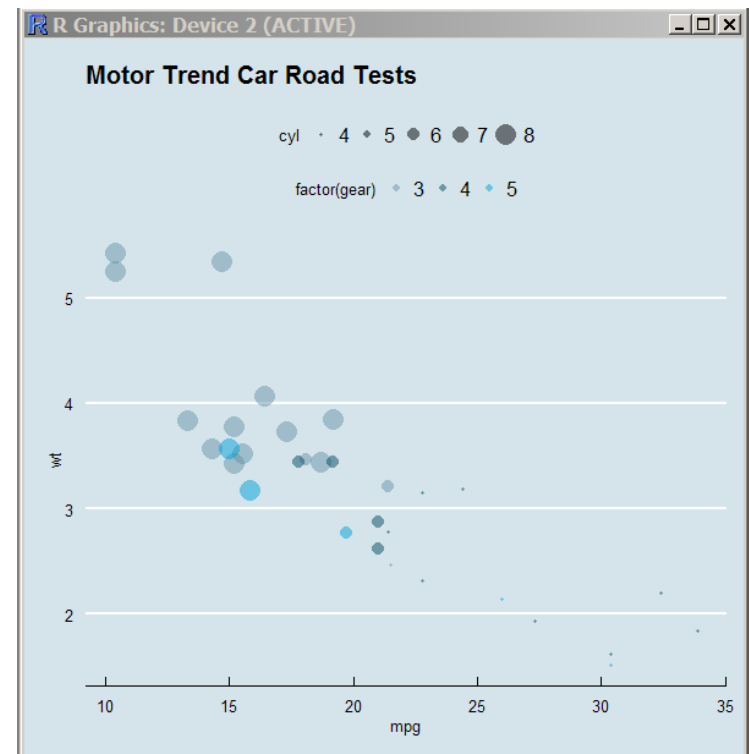


2013.2.8

例子

■ Economist风格

```
ggplot(mtcars,aes(x=mpg,y=wt,size=cyl,colour=factor(gear)))+  
  geom_point(alpha=.5)+  
  theme_economist() +  
  scale_colour_economist()+  
  ggtitle("Motor Trend Car Road Tests")
```

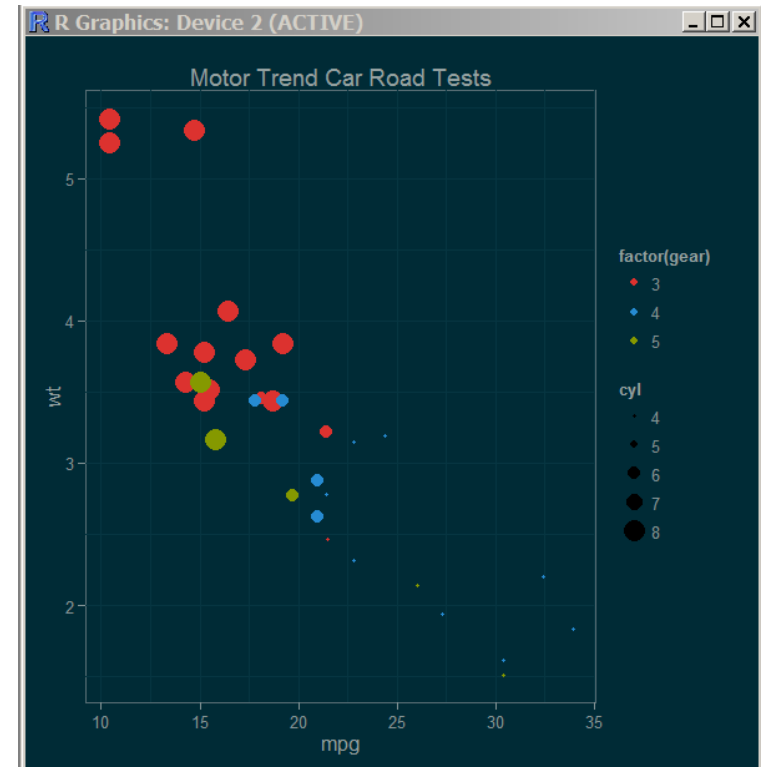


2013.2.8

例子

■ 深色风格

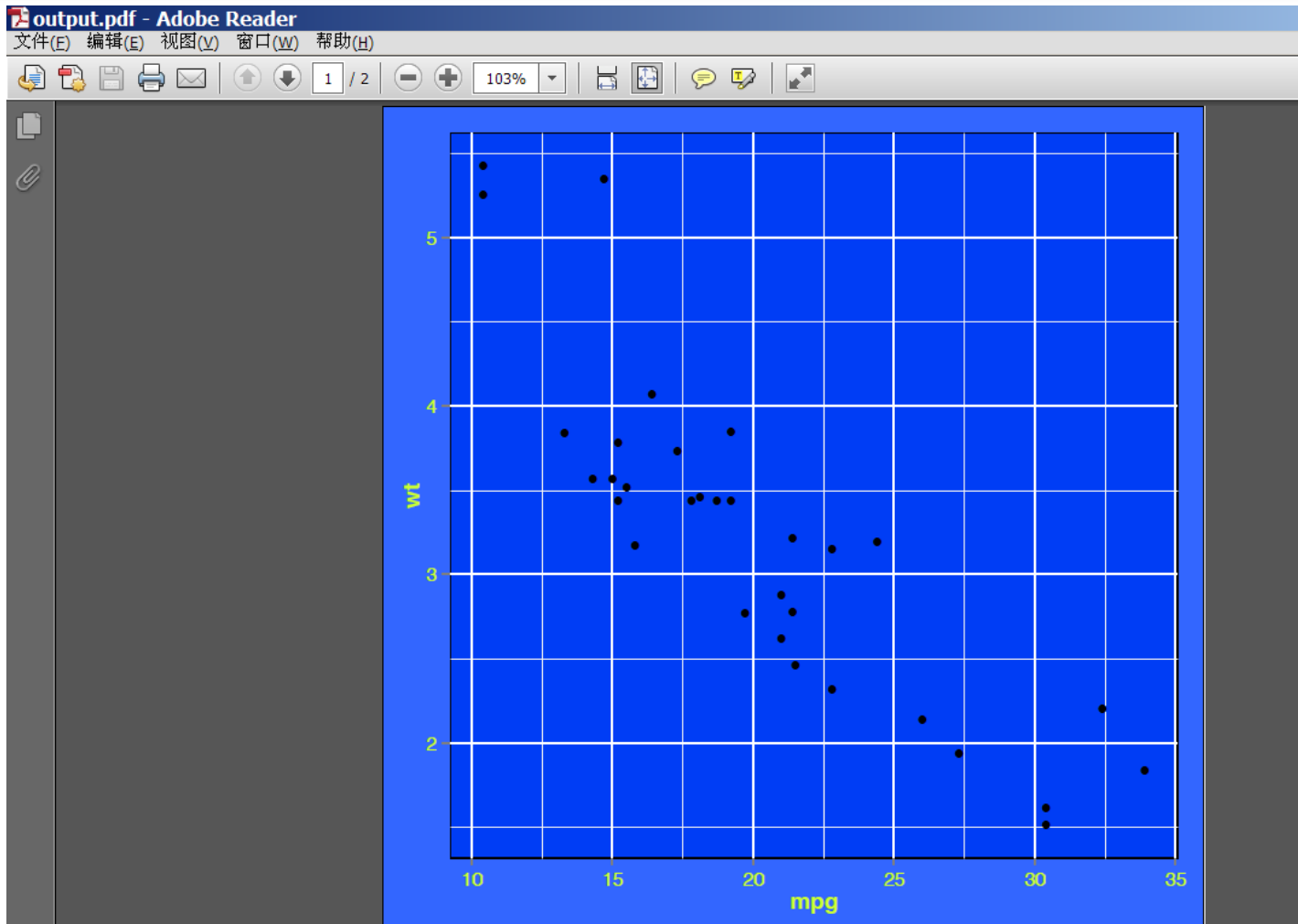
```
ggplot(mtcars,aes(x=mpg,y=wt,size=cyl,colour=factor(gear)))+  
geom_point()+  
theme_solarized(light=FALSE) +  
scale_colour_solarized("red")+  
ggtitle("Motor Trend Car Road Tests")
```



输出到文件ggsave()

```
qplot(mpg, wt, data = mtcars)
ggsave(file = "output.pdf")
pdf(file = "output.pdf", width = 6, height = 6)
# If inside a script, you will need to explicitly print() plots
qplot(mpg, wt, data = mtcars)
qplot(wt, mpg, data = mtcars)
dev.off()
```


结果



2013.2.8

在同一页面上画多幅图

■ 先产生三幅图

```
(a <- qplot(date, unemploy, data = economics, geom = "line"))
```

```
(b <- qplot(uempmed, unemploy, data = economics) +  
geom_smooth(se = F))
```

```
(c <- qplot(uempmed, unemploy, data = economics, geom="path"))
```

viewport()

Library(grid)

A viewport that takes up the entire plot device

```
vp1 <- viewport(width = 1, height = 1, x = 0.5, y = 0.5)
```

```
vp1 <- viewport()
```

A viewport that takes up half the width and half the height,

located in the middle of the plot.

```
vp2 <- viewport(width = 0.5, height = 0.5, x = 0.5, y = 0.5)
```

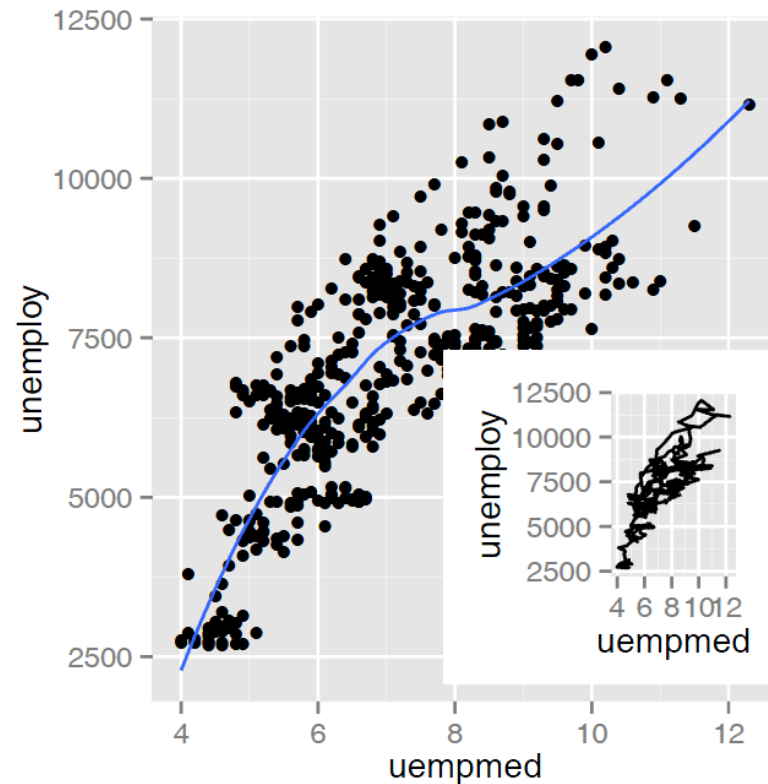
```
vp2 <- viewport(width = 0.5, height = 0.5)
```

A viewport that is 2cm x 3cm located in the center

```
vp3 <- viewport(width = unit(2, "cm"), height = unit(3, "cm"))
```

实验

```
pdf("polishing-subplot-1.pdf", width = 4, height = 4)
subvp <- viewport(width = 0.4, height = 0.4, x = 0.75, y = 0.35)
b
print(c, vp = subvp)
dev.off()
```

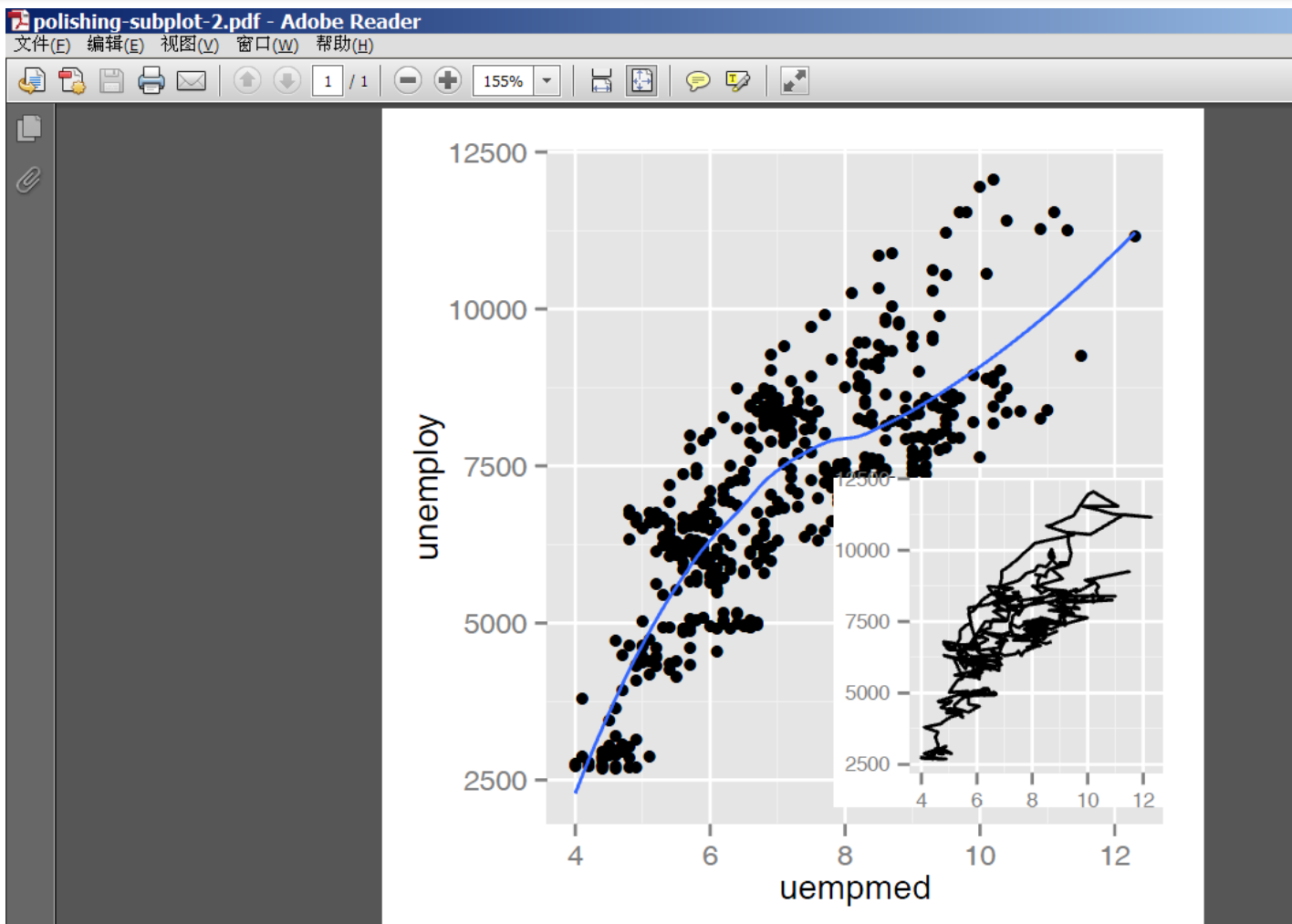


2013.2.8

改进

```
csmall <- c +  
theme_gray(9) +  
labs(x = NULL, y = NULL) +  
theme(plot.margin = unit(rep(0, 4), "lines"))  
pdf("polishing-subplot-2.pdf", width = 4, height = 4)  
b  
print(csmall, vp = subvp)  
dev.off()
```

结果

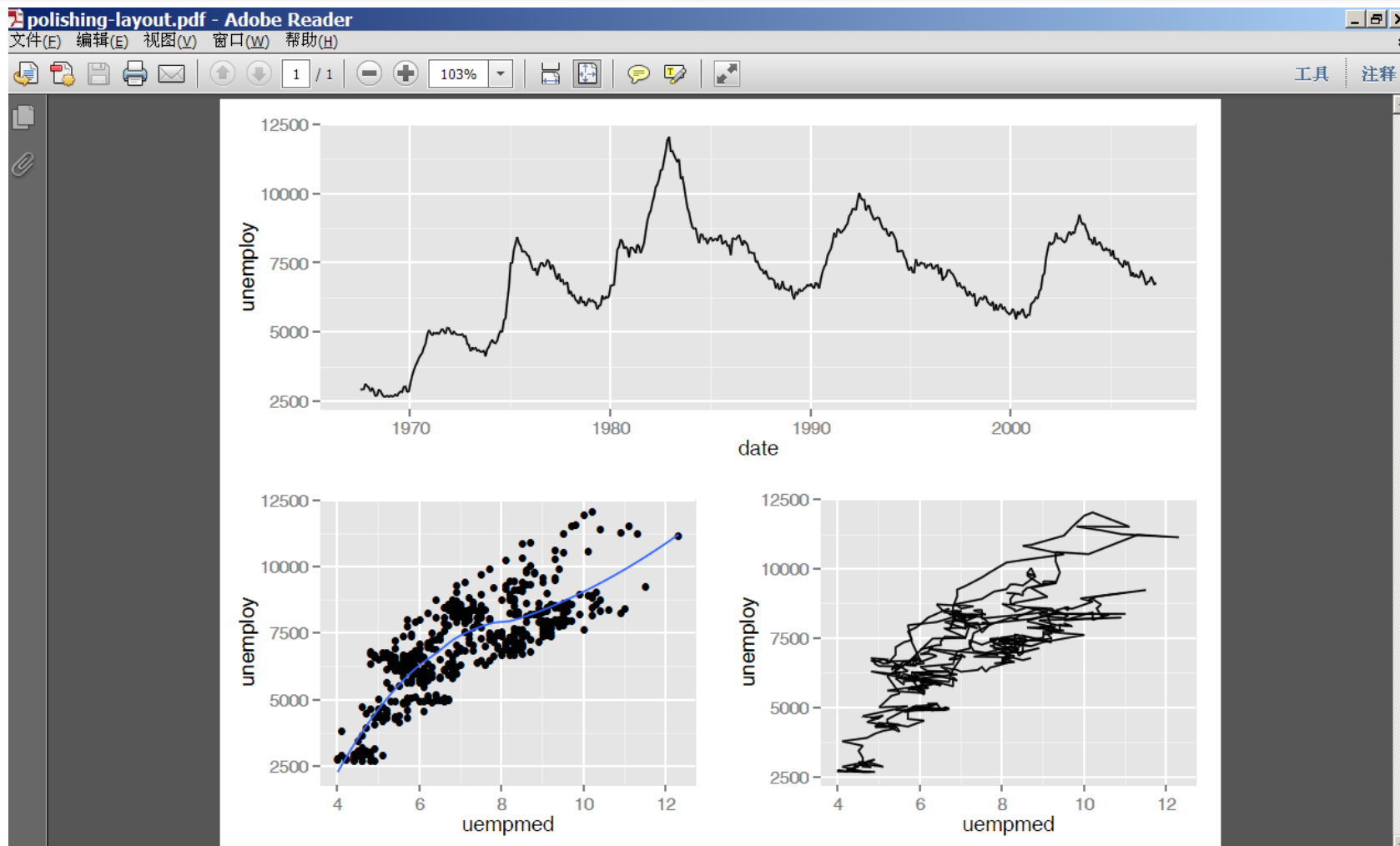


2013.2.8

指定比例

```
pdf("polishing-layout.pdf", width = 8, height = 6)
grid.newpage()
pushViewport(viewport(layout = grid.layout(2, 2)))
vlayout <- function(x, y)
viewport(layout.pos.row = x, layout.pos.col = y)
print(a, vp = vlayout(1, 1:2))
print(b, vp = vlayout(2, 1))
print(c, vp = vlayout(2, 2))
dev.off()
```

结果

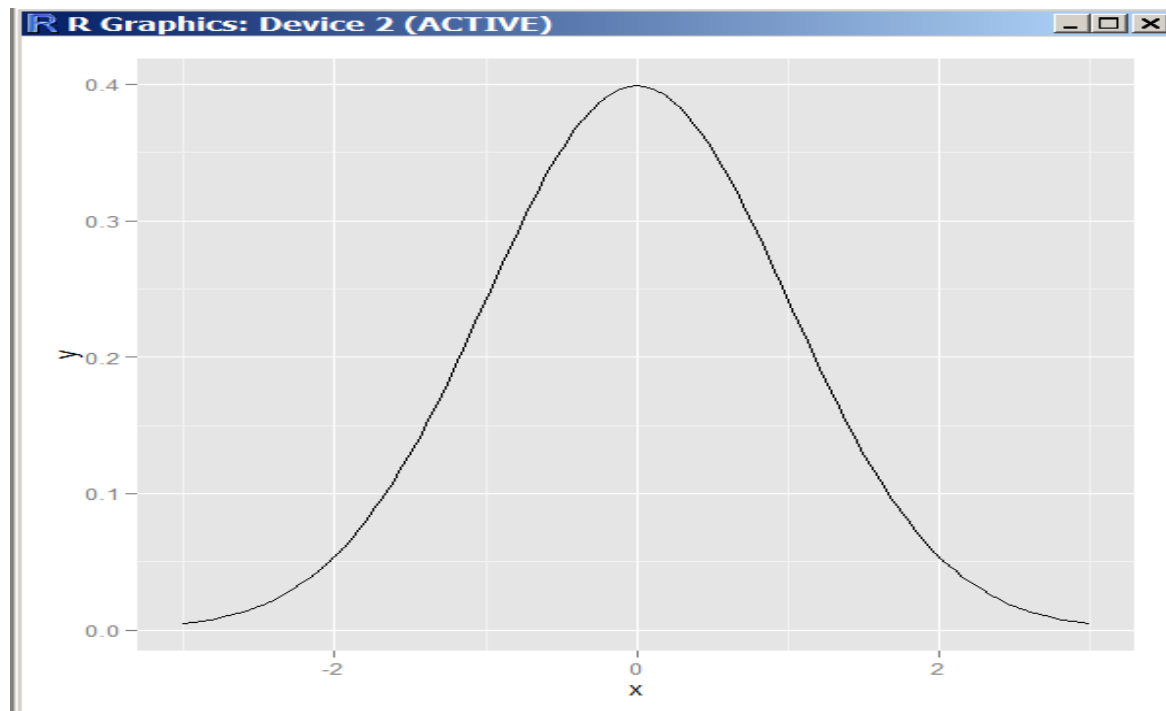


2013.2.8

画函数图像

```
p <- ggplot(data.frame(x=c(-3,3)), aes(x=x))
```

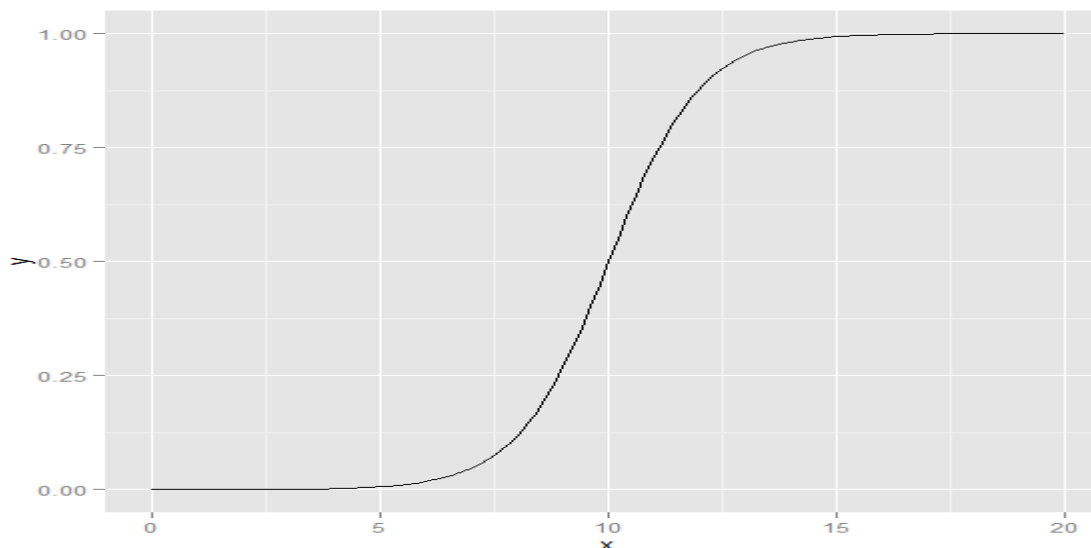
```
p + stat_function(fun = dnorm)
```



2013.2.8

画函数图像

```
myfun <- function(xvar) {  
  1/(1 + exp(-xvar + 10))  
}  
  
ggplot(data.frame(x=c(0, 20)), aes(x=x)) + stat_function(fun=myfun)
```



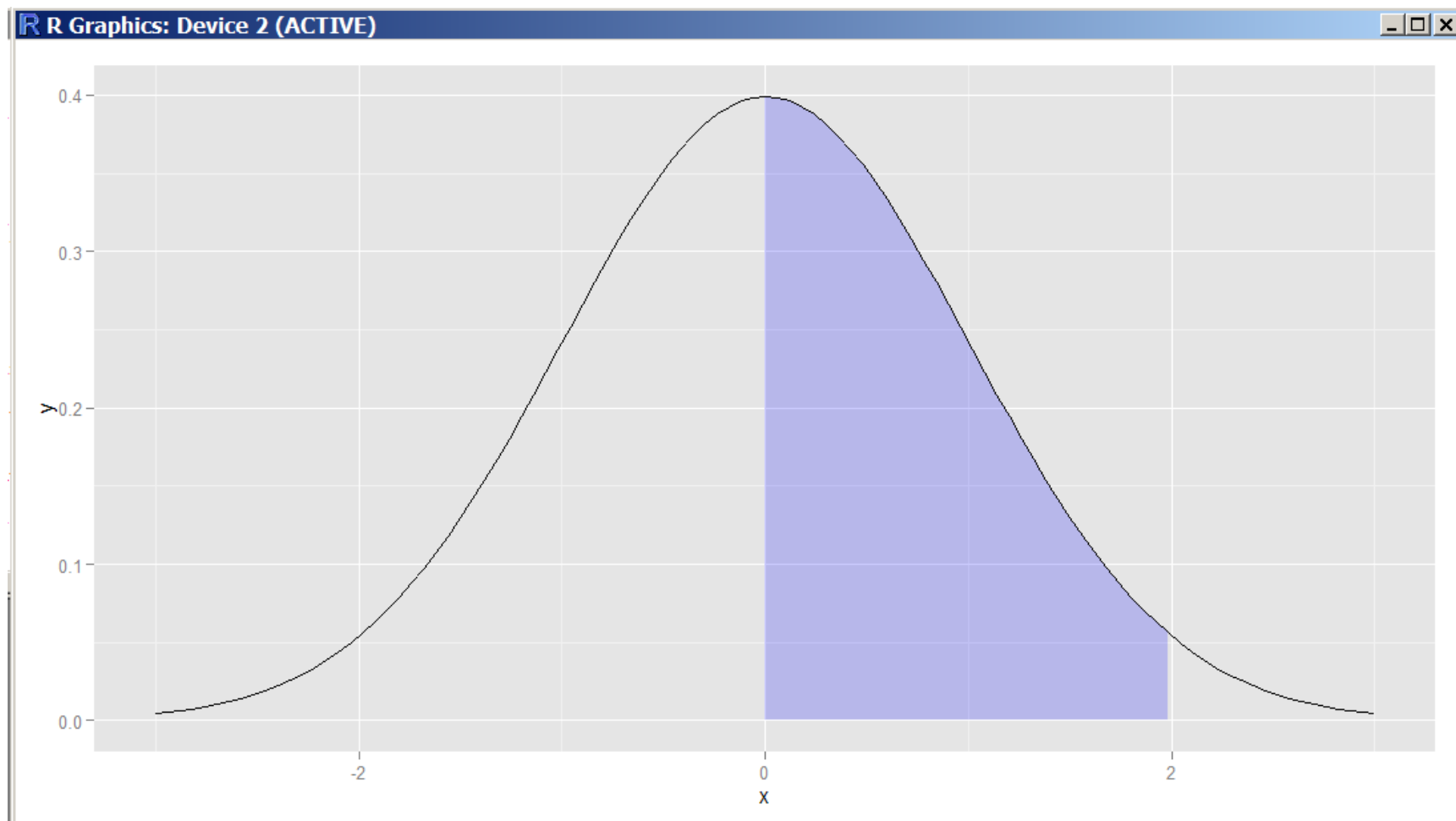
2013.2.8

画微积分中常见的曲边梯形

```
# Return dnorm(x) for 0 < x < 2, and NA for all other x
dnorm_limit <- function(x) {
  y <- dnorm(x)
  y[x < 0 | x > 2] <- NA
  return(y)
}

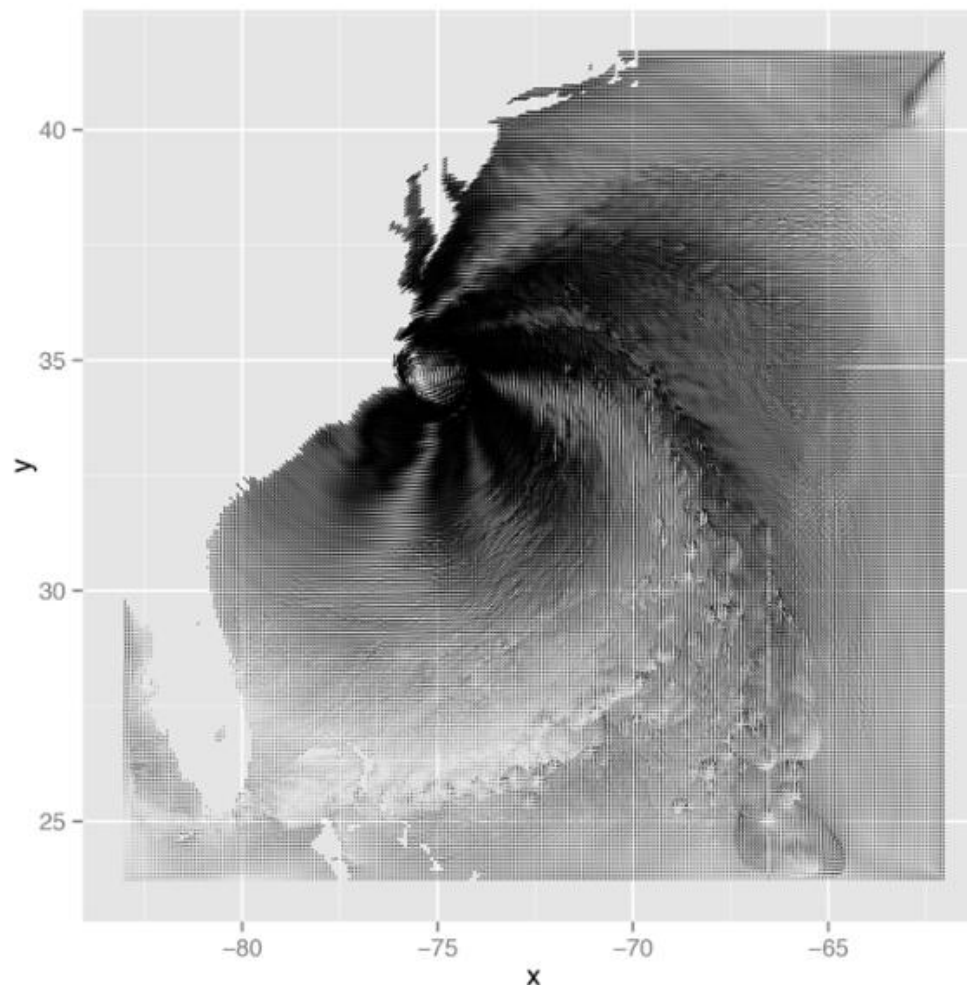
# ggplot() with dummy data
p <- ggplot(data.frame(x=c(-3, 3)), aes(x=x))
p + stat_function(fun=dnorm_limit, geom="area", fill="blue", alpha=0.2)
  +stat_function(fun=dnorm)
```

结果



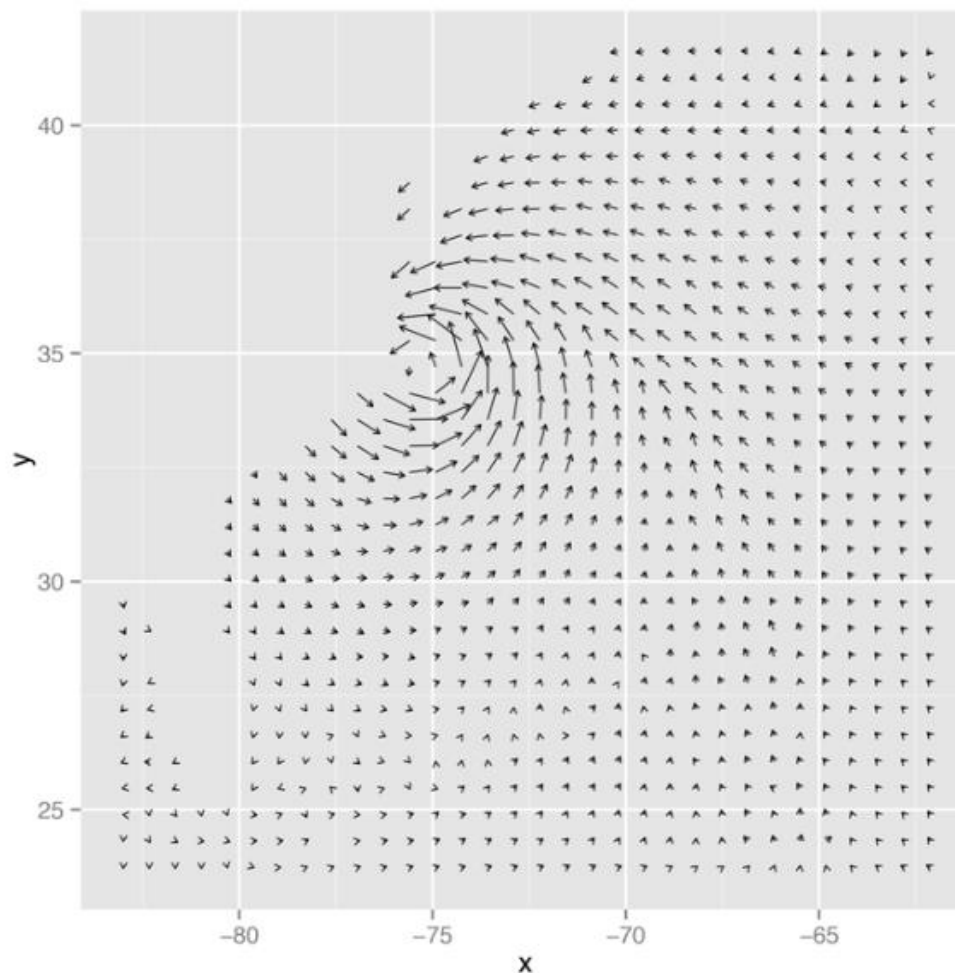
2013.2.8

ggplot2画向量场



2013.2.8

向量场



2013.2.8

向量场

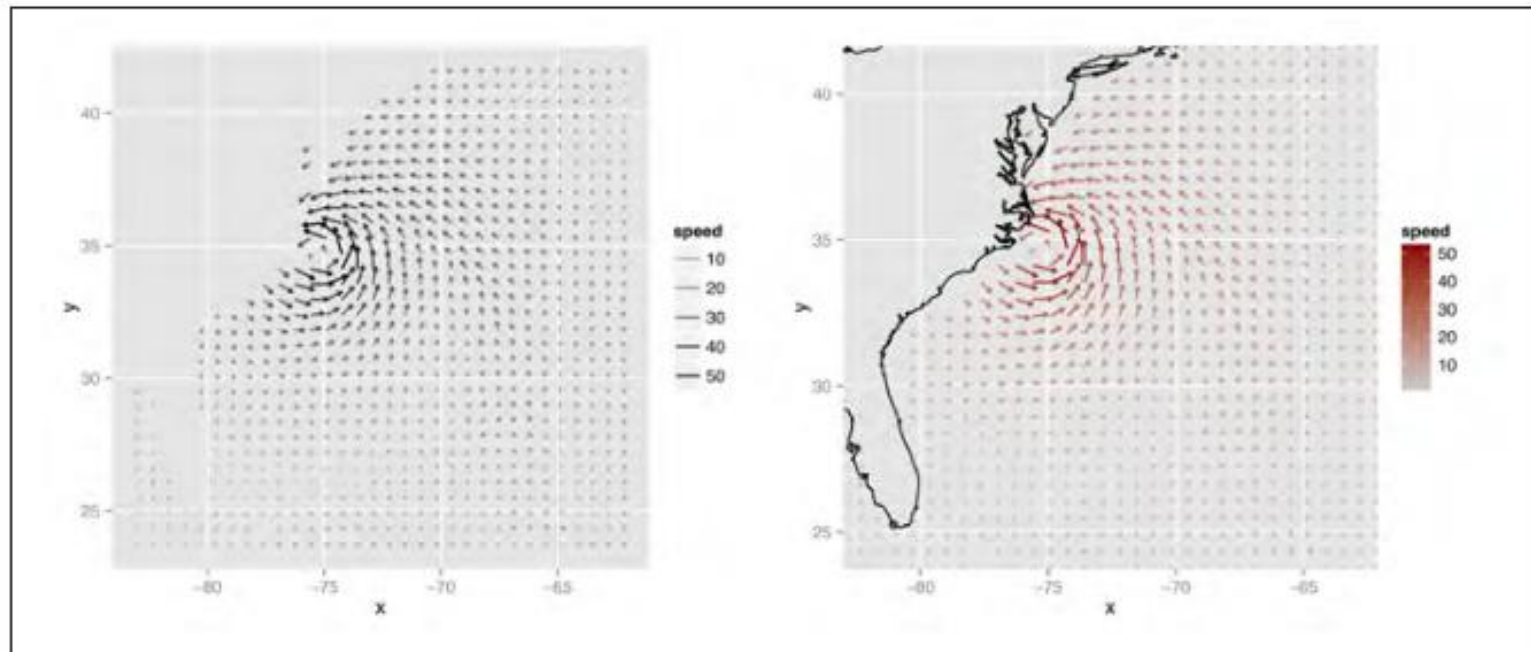


Figure 13-23. Left: vector field with speed mapped to alpha; right: with speed mapped to colour

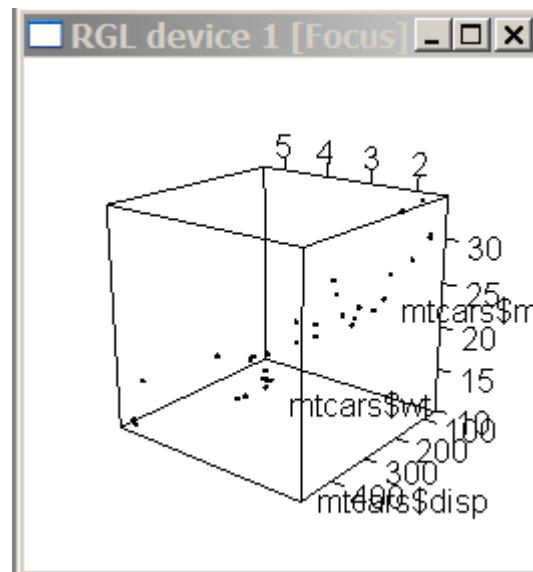
让图形动起来

- 并非ggplot2包的功能

```
library(rgl)
```

```
plot3d(mtcars$wt, mtcars$disp, mtcars$mpg, type="s", size=0.75, lit=FALSE)
```

```
play3d(spin3d())
```



2013.2.8

在R中实现动态气泡图

- <http://cos.name/2013/01/dynamic-bubble-plot-in-r/>
- 使用ggplot2包，animation包，ffmpeg软件
- 数据下载：<https://github.com/mbostock/bost.ocks.org/blob/gh-pages/mike/nations/nations.json>
- 代码下载：<http://blog.programet.org/wp-content/uploads/HansRosling.txt>



代码

#Import the Strange Dataset

```
rawdata=readLines("nations.json")
```

#Get the useful words

```
data=unlist(strsplit(rawdata,"\\[\\\\{\"|\":\\\\\"|\",|\"|\":\\\\\\\\[\\\\\\\\,\\\\\\\\[\\\\\\\\\\\\\\\\,\"|\\\\\\\\\\\\\\\\\\\\},\\\\\\\\{\"|\\\\\\\\\\\\\\\\\\\\}\\\\\\\\"))
```

```
data=data[which(data!="")]
```

```
n=length(data)
```

代码

```
#Set index
nameind=which(data=="name")
regind=which(data=="region")
incind=which(data=="income")
popind=which(data=="population")
lifind=which(data=="lifeExpectancy")
endpoint=nameind-1
endpoint=endpoint[2:180]
endpoint[180]=45999

name=data[nameind+1]
region=data[regind+1]
```

代码

```
#An empty data.frame type
initdfr=data.frame(name=rep(0,210),region=rep(0,210),year=1800:2009,income=rep(0,210),pop=rep(0,210),life=rep(0,210))

#An string split function only working for the comma
splf=function(x) return(as.numeric(unlist(strsplit(x,","))))

#Linear interpolation function, for those zero data
itpl=function(a)
{
  ind=which(a>0)
  if (ind[1]>1)
    a[1:(ind[1]-1)]=rep(a[ind[1]],ind[1]-1)
  n=length(ind)
  if (ind[n]<length(a))
    a[(ind[n]+1):length(a)]=rep(a[ind[n]],length(a)-ind[n])
  for (i in 1:(n-1))
    a[ind[i]:ind[i+1]]=rep(a[ind[i]],ind[i+1]-ind[i]+1)+(seq(ind[i],ind[i+1],1)-ind[i])*(a[ind[i+1]]-a[ind[i]])/(ind[i+1]-ind[i])
  return(a)
}
```

代码

```
#Drag data information from words
tbls=NULL
for (i in 1:180)
{
  dfr=initdfr
  dfr$name=rep(name[i],210)
  dfr$region=rep(region[i],210)
  for (j in (incind[i]+1):(popind[i]-1))
  {
    tmp=splt(data[j])
    ii=which(dfr$year==tmp[1])
    dfr$income[ii]=tmp[2]
  }
}
```

代码

```
for (j in (popind[i]+1):(lifind[i]-1))
{
    tmp=splt(data[j])
    ii=which(dfr$year==tmp[1])
    dfr$pop[ii]=tmp[2]
}
for (j in (lifind[i]+1):endpoint[i])
{
    tmp=splt(data[j])
    ii=which(dfr$year==tmp[1])
    dfr$life[ii]=tmp[2]
}

tbls=rbind(tbls,dfr)
}
```

代码

```
#Two country with only one record, meaningless  
ind=which(tbls$name=="Mayotte")  
tbls=tbls[-ind,]  
ind=which(tbls$name=="Tokelau")  
tbls=tbls[-ind,]  
name=name[c(-28,-177)]  
region=region[c(-28,-177)]
```

代码

```
#Linear interpolation
for (i in 1:178)
{
  ind=(210*(i-1)+1):(i*210)
  tbls$income[ind]=itpl(tbls$income[ind])
  tbls$pop[ind]=itpl(tbls$pop[ind])
  tbls$life[ind]=itpl(tbls$life[ind])
}
```


代码

```
require(ggplot2)

#Draw function with ggplot2
drawit=function(yr,scl=15)
{
  ind=which(tbls$year==yr)
  d.f=data.frame(yr=yr)
  p=ggplot(aes(x=log(income),y=life,size=pop,colour=as.factor(region)),pch=21,data=tbls[ind,])
  p+geom_point(show_guide = FALSE)+
    geom_point(shape = 1,colour = "black",show_guide = FALSE)+
    xlim(5.5,11.7)+ylim(10,83)+scale_area(range = c(1, scl))+
    annotate("text", x=10, y=15, label = yr,size=30,color="grey")
}

#drawit(1800)
```

代码

```
#Automatically repeat the drawing procedure  
finaldraw=function(a,b)  
{  
  for (i in 1:10)  
    print(drawit(a))  
  for (i in a:b)  
    print(drawit(i))  
  for (i in 1:10)  
    print(drawit(b))  
}
```

代码

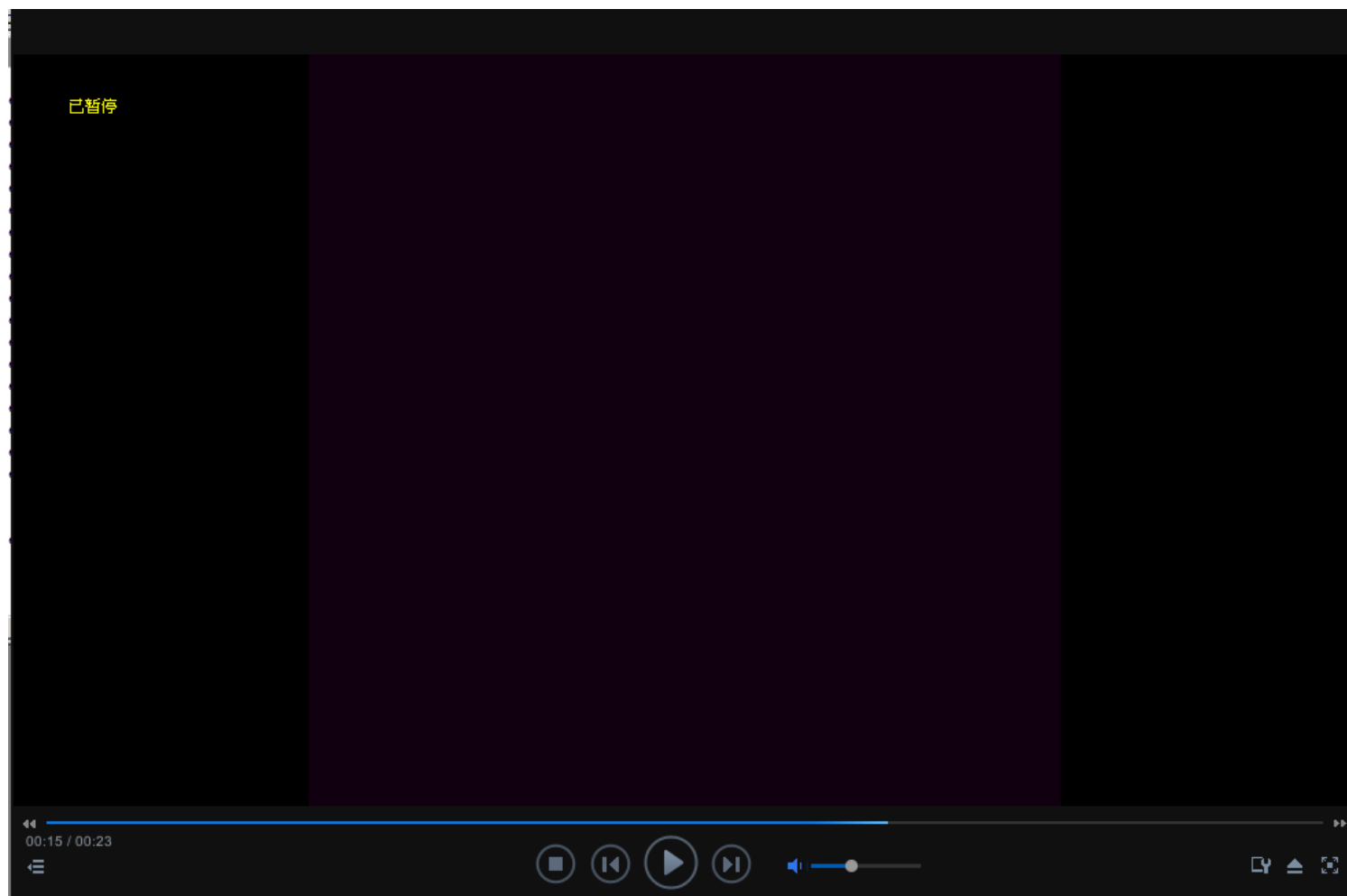
```
#finaldraw(1800,2009)

require(animation)

#sett ffmpeg in Windows = =||
oopts = ani.options(ffmpeg = "D:/ffmpeg/bin/ffmpeg.exe")

#Use the function from animation to make the final movie
saveVideo({
  finaldraw(1800,2009)
  ani.options(interval = 0.1, nmax = 230)
}, video.name = "HansRosling.mp4", other.opts = "-b 500k")
```

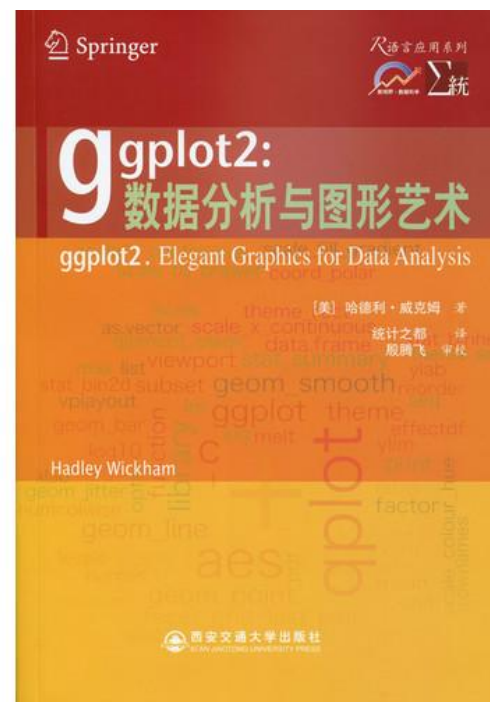
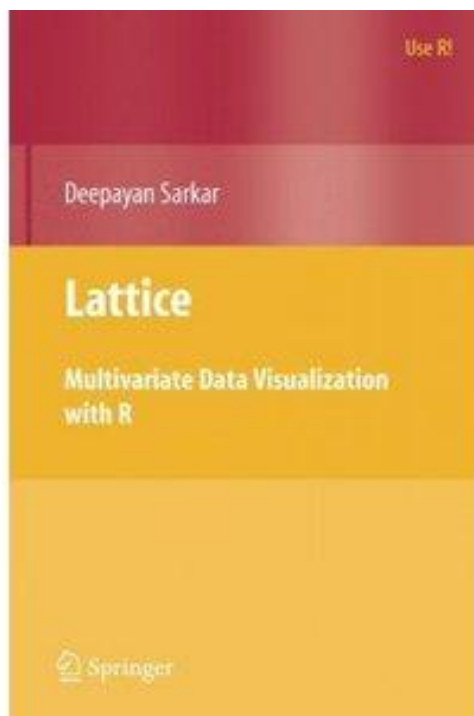
输出动画



2013.2.8

其它作图包

- Lattice
- Grid



2013.2.8

炼数成金逆向收费式网络课程

- 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