



语言

数据可视化

学习内容

- 一 R绘图基础
- 二 各类图形绘制
- 三 习题和资料

为什么要学习R语言

➤ R是

- 一个开放(GPL)的统计编程环境
- 一种语言，是S语言(由AT&T Bell实验室的Rick Becker, John Chambers, Allan Wilks开发)的一种方言(dialect)之一，另一则为S-plus.
- 一种软件，是集统计分析 with 图形直观显示于一体的统计分析

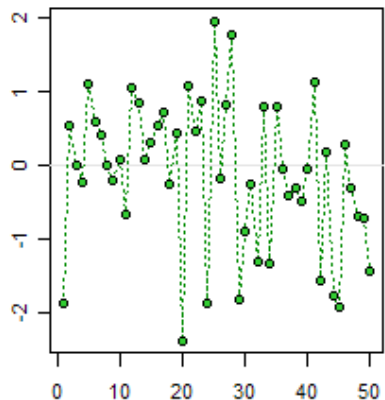
➤ R作为一个计划(project)，最早(1995年)是由Auckland大学统计系的Robert Gentleman和Ross Ihaka开始编制，目前由R核心开发小组(R Development Core Team-以后用R DCT表示)维护，他们完全自愿、工作努力负责，并将全球优秀的统计应用软件打包提供给我们。我们可以通过R计划的网站(<http://www.r-project.org>)了解有关R的最新信息和使用说明，得到最新版本的R软件和基于R的应用统计软件包.

为什么要学习R语言

- R是完全免费的！！而S-Plus尽管是非常优秀的统计分析软件，可是你需要支付一笔\$US。
- R可以在运行于UNIX, Windows和Macintosh的操作系统上。
- R嵌入了一个非常实用的帮助系统。
- R具有很强的作图能力。
- 我们将R程序容易地移植到S-Plus程序中，反之S的许多过程直接或稍作修改用于R。
- 通过R语言的许多内嵌统计函数，很容易学习和掌握R语言语法。
- 我们可以编制自己的函数来扩展现有的R语言(这就是为什么它在不断等级完善!!)
-

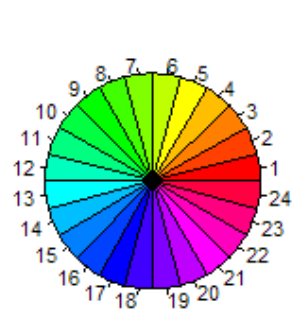
一 R绘图基础

Simple Use of Color In a Plot



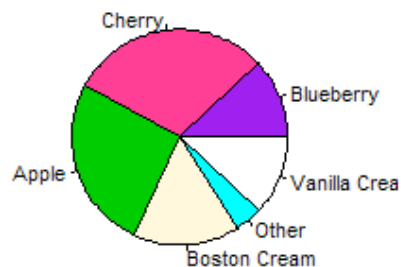
Just a Whisper of a Label

A Sample Color Wheel



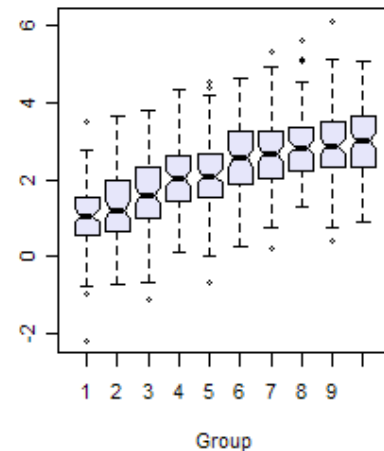
(Use this as a test of monitor linearity)

January Pie Sales

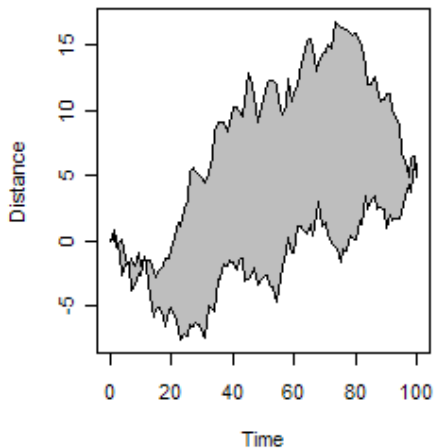


(Don't try this at home kids)

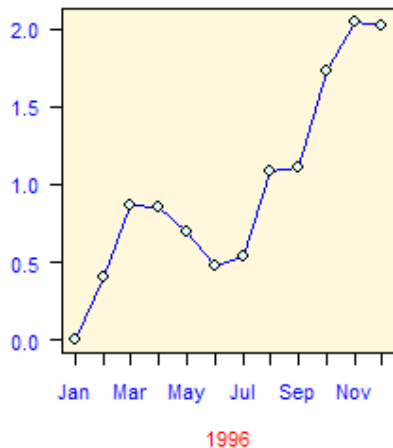
Notched Boxplots



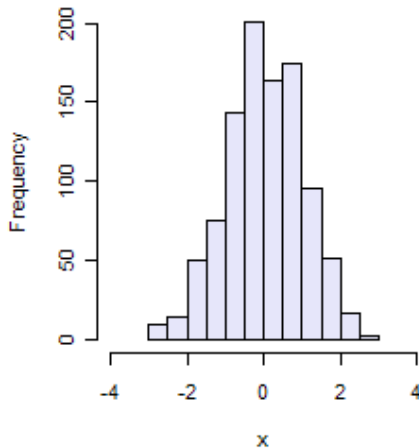
Distance Between Brownian Motions



The Level of Interest in R



1000 Normal Random Variates



A Topographic Map of Maunga Whau

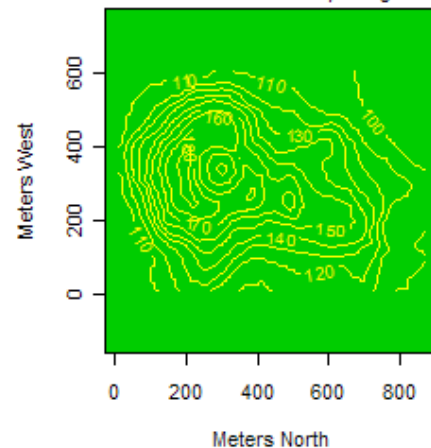
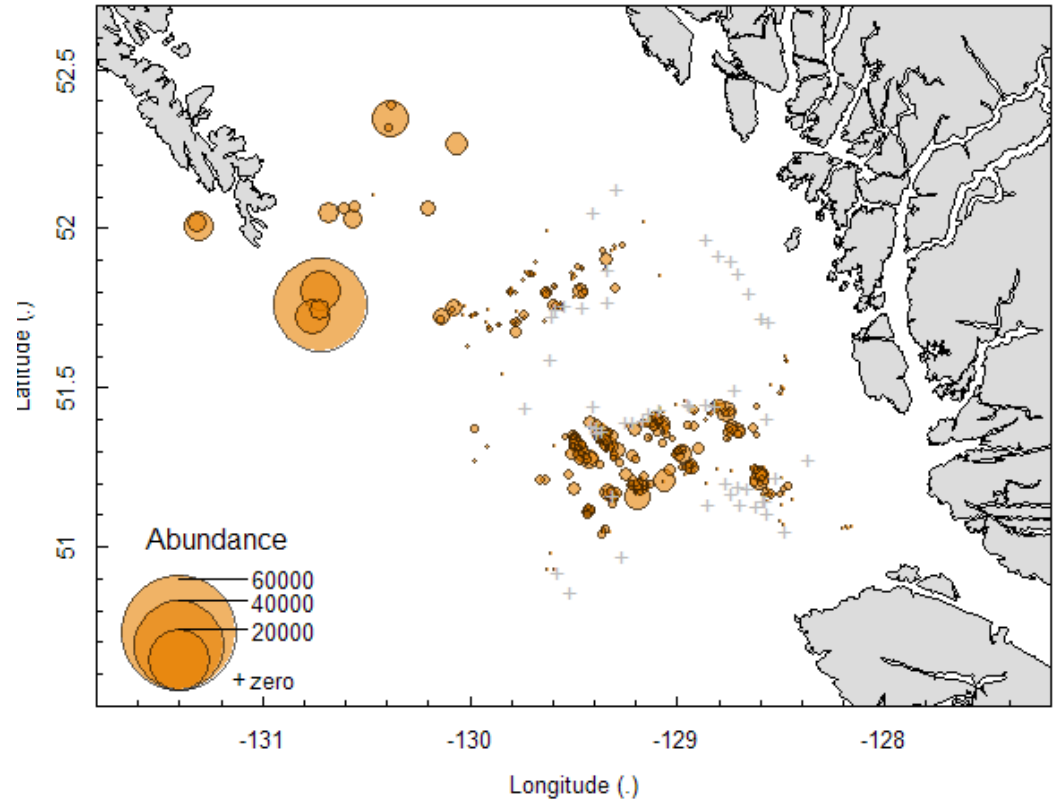
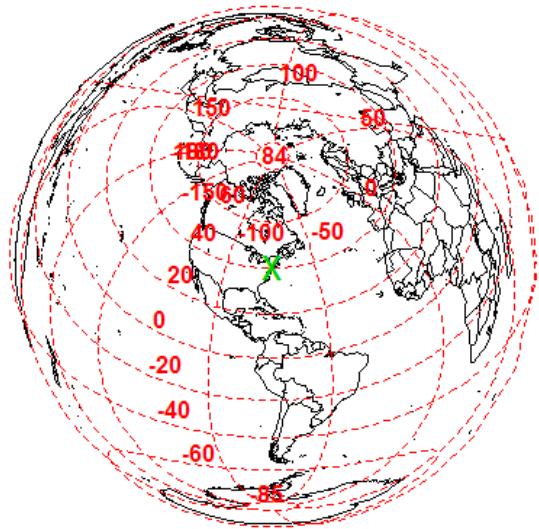


图 R绘制的图形

绘制地图

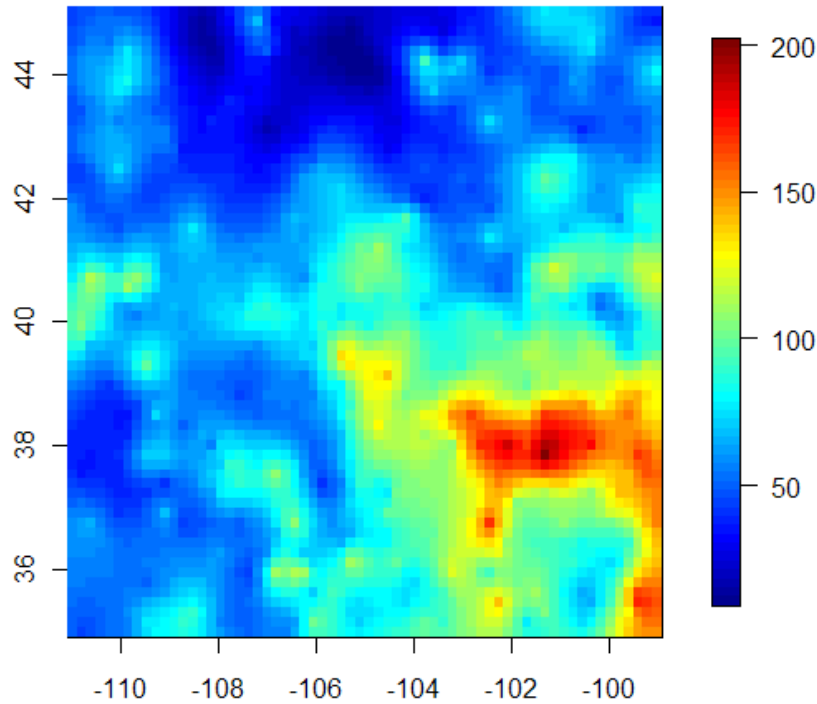


• 图 左图 maps包 map()

右图 PBSmapping包 addBubbles()

绘制地图

Plot of first (black) and second (red) nearest neighbours



• 图 `fields` 包实例

`spdep` 包实例

R绘图功能

- R具备卓越的绘图功能，通过参数设置对图形进行精确控制。绘制的图形能满足出版印刷的要求，可以输出Jpg、tiff、eps、emf、pdf、png等各种格式。
- 通过与GhostScript软件的结合，可以生成600dpi，1200dpi的等各种分辨率和尺寸的图形。
- 绘图是通过绘图函数结合相应的选项完成的。

R绘图功能

- `demo (graphics)` : 了解R绘图功能
- 绘图函数包括：
 - 高阶绘图函数 High-level Plotting Function
 - 产生一个新的图区，可能包括坐标轴、标签、标题等。
 - 低阶绘图函数 Low-level Plotting Function
 - 在已有的图上加更多的元素
- 绘图参数
 - 缺省值
 - `?par()`

高阶绘图函数

`plot(x)`、`plot(x,y)`

绘制散点图等多种图形,根据数据的类,调用相应的函数绘图

`pie(x)`

饼图

`boxplot(x)`

箱线图

`hist(x)`

频率直方图

`coplot(x~y|z)`

条件分割图

`Interaction.plot(f1,f2,y)`

交互效应图

`qqplot(x,y)`、`qqnorm(x)`

QQ图

`contour(x,y,z)`

等高线图

`persp(x,y,z)`

三维透视图

高阶绘图函数

<code>barplot(x)</code>	柱状图/条形图
<code>matplot(x,y)</code>	矩阵图
<code>mosaicplot(x)</code>	马赛克图
<code>pairs(x)</code>	散点图矩阵
<code>sunflowerplot(x,y)</code>	向日葵散点图
<code>stars(x)</code>	星状图
<code>stripchart(x)</code>	带状图
<code>dotchart(x)</code>	Cleveland点图
<code>filledcontour(x,y,z)</code>	颜色等高线图
<code>image(x,y,z)</code>	颜色图

低阶绘图函数

<code>points()</code>	添加点
<code>lines()</code>	添加线
<code>curve()</code>	添加曲线
<code>abline()</code>	添加给定斜率的线
<code>title()</code>	添加标题
<code>text()</code>	添加文字
<code>mtext()</code>	在图的边空添加文字
<code>axis()</code>	画坐标轴

低阶绘图函数

<code>legend()</code>	添加图例
<code>rug()</code>	添加刻度线
<code>rec()</code>	绘制长方形
<code>polygon()</code>	绘制多边形
<code>segments()</code>	添加线段
<code>arrows()</code>	画箭头
<code>box()</code>	添加外框

绘图参数

参数用在函数内部，在没有设定值时使用缺省值。

`font` = 字体, `lty` = 线类型, `lwd` = 线宽度, `pch` = 点的类型,

`xlab` = 横坐标, `ylab` = 纵坐标,

`xlim` = 横坐标范围,

`ylim` = 纵坐标范围,

也可以对整个要绘制图形的各种参数进行设定

参见 `? par()`

绘图参数

参数	含义	缺省	参数	含义	缺省
adj	字符相对位置	0.5	bty	边框样式	"o"
bg	图形背景色	"transparent"	cex	文本缩放倍数	1
col	图中符号颜色	"black"	family	字体族	""
fg	前景色	"black"	font	文本字体样式	1
lab	坐标轴刻度数目	5,5,7	las	坐标轴标签样式	0
lend	线条末端样式	"round"	lheight	图中文本行高	1
ljoin	线条相交样式	"round"	lty	线条样式	1
lwd	线条宽度	1	mex	坐标轴边界宽度	1
mar	图形边界空白	c(5,4,4,2)+0.1		缩放倍数	
mfrow	一页多图	c(1,1)	pch	点符号	1
mfcop	一页多图	c(1,1)	pty	作图区域形状	"m"
mgp	坐标轴边界宽度	c(3,1,0)	srt	字符串旋转角度	0
oma	外边界	c(0, 0, 0, 0)	xaxt	X 坐标轴样式	"s"
tck	坐标刻度线高度	NA	yaxt	Y 坐标轴样式	"s"
tcl	坐标刻度线高度	-0.5	xlog	X 坐标取对数	F
xpd	处理超出边界	F	ylog	Y 坐标取对数	F

一页多图

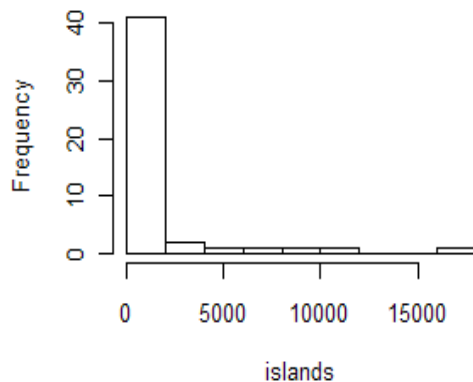
图 一页多图

```
par()
```

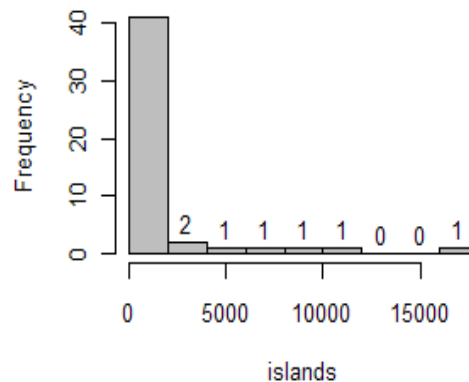
```
par(mfrow=c(2,2))
```

...

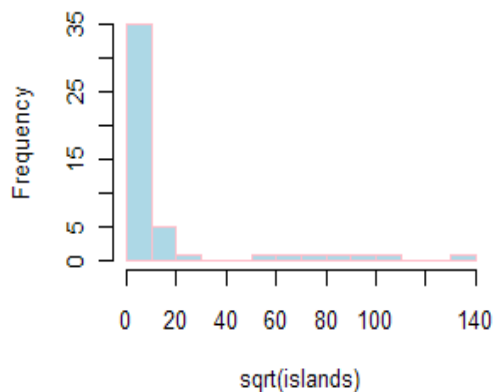
Histogram of islands



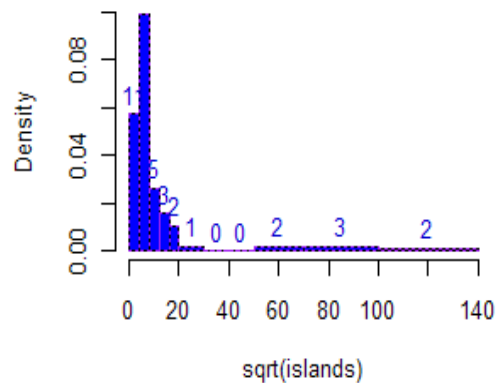
Histogram of islands



Histogram of sqrt(islands)



Histogram of sqrt(islands)



在原有图形上添加元素

举例：先执行`par(mfrow=c(2,2))`

```
x <- rnorm(100) # 生成随机数
```

```
hist(x, freq=F) # 绘制直方图
```

```
curve(dnorm(x), add=T) # 添加曲线
```

```
h <- hist(x, plot=F) # 绘制直方图
```

```
ylim <- range(0, h$density, dnorm(0)) # 设定纵轴的取值范围
```

```
hist(x, freq=F, ylim=ylim) # 绘制直方图
```

```
curve(dnorm(x), add=T, col="red") # 添加曲线
```

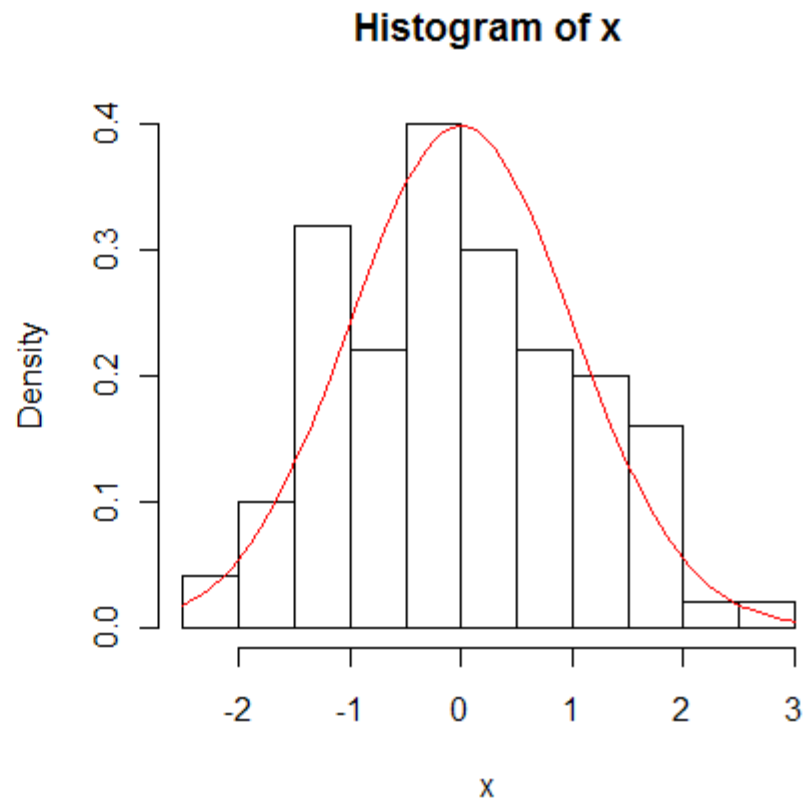
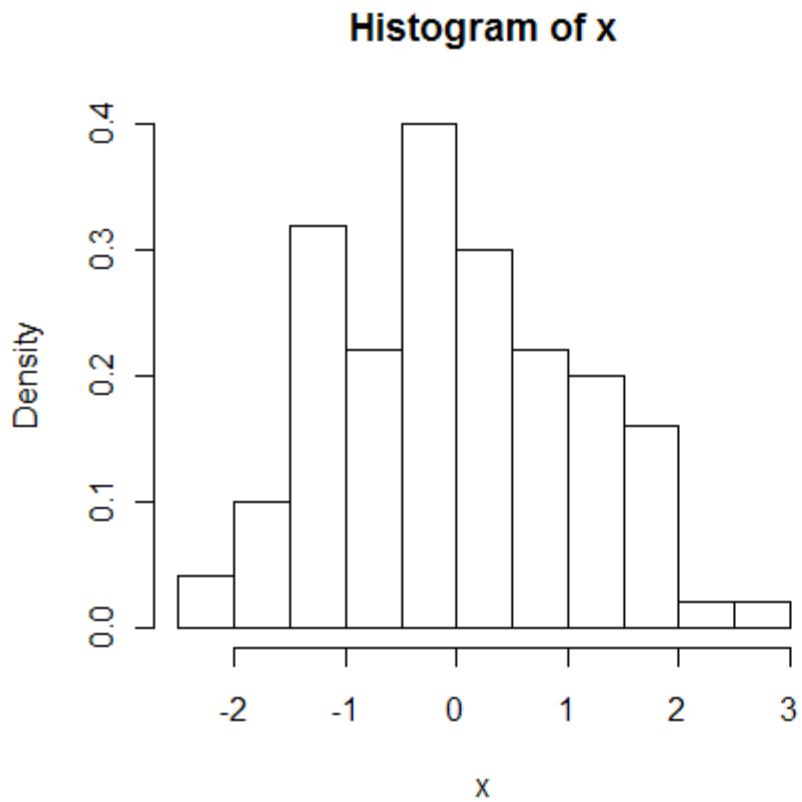


图 在原有直方图上添加曲线

二、各类图形绘制

1 基于比较的绘图

柱形图 .R、条形图 .R、折线图 .R、

2 基于关系的绘图

散点图 .R、气泡图 .R

3 基于组合的绘图

饼图 .R、

4 基于分布的绘图

直方图 .R、核密度图 .R、凹槽箱线图 .R、
箱线图 .R、小提琴图 .R

5 基于时间的绘图

LOESS 曲线图 .R、阶梯图 .R

6 基于空间的绘图

地图添加标记 .R、地图中画气泡 .R、地图
中画线 .R

7 多维信息可视化

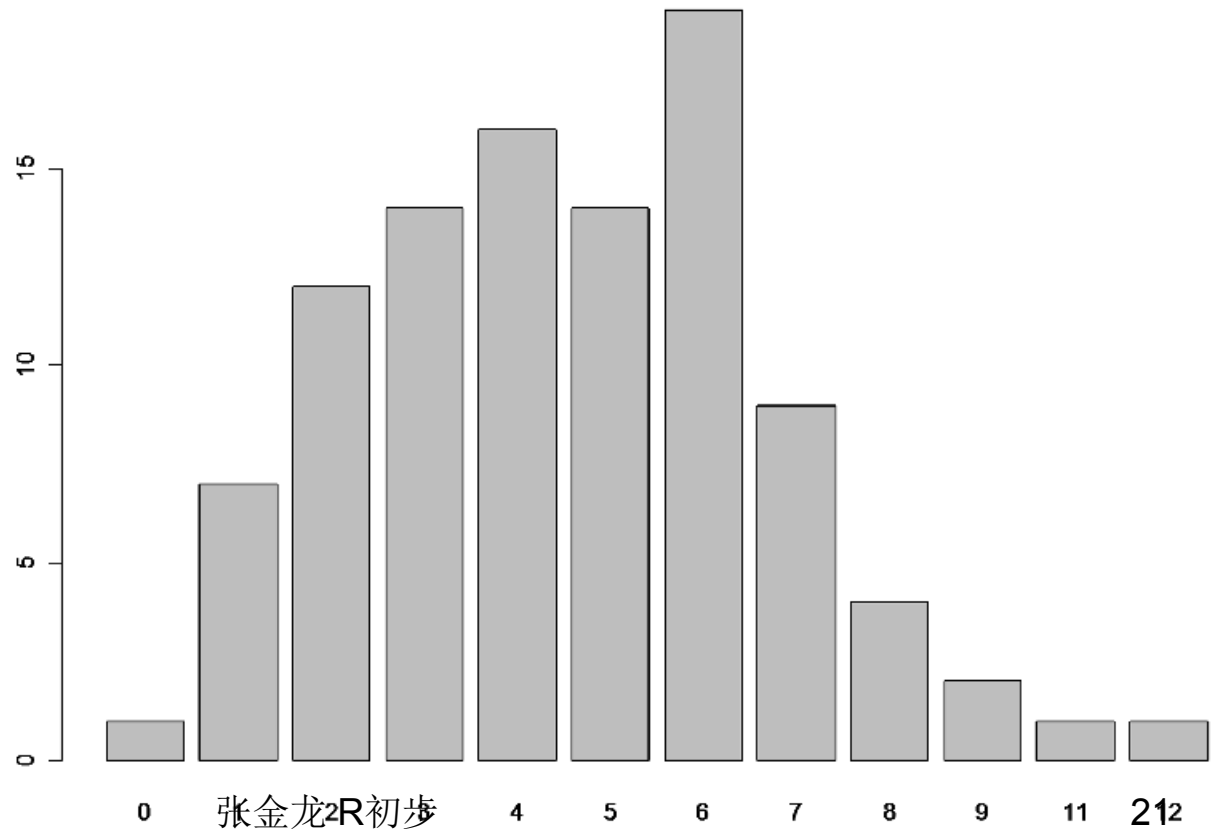
散点图矩阵 .R、线图 .R、平行坐标图 .R、
星状图 .R、堆叠柱状图 .R、热力图 .R、
马赛克图 .R、切尔诺夫脸谱图 .R

8 文本与文档可视化

词云-英文 .R、词云-中文 .R

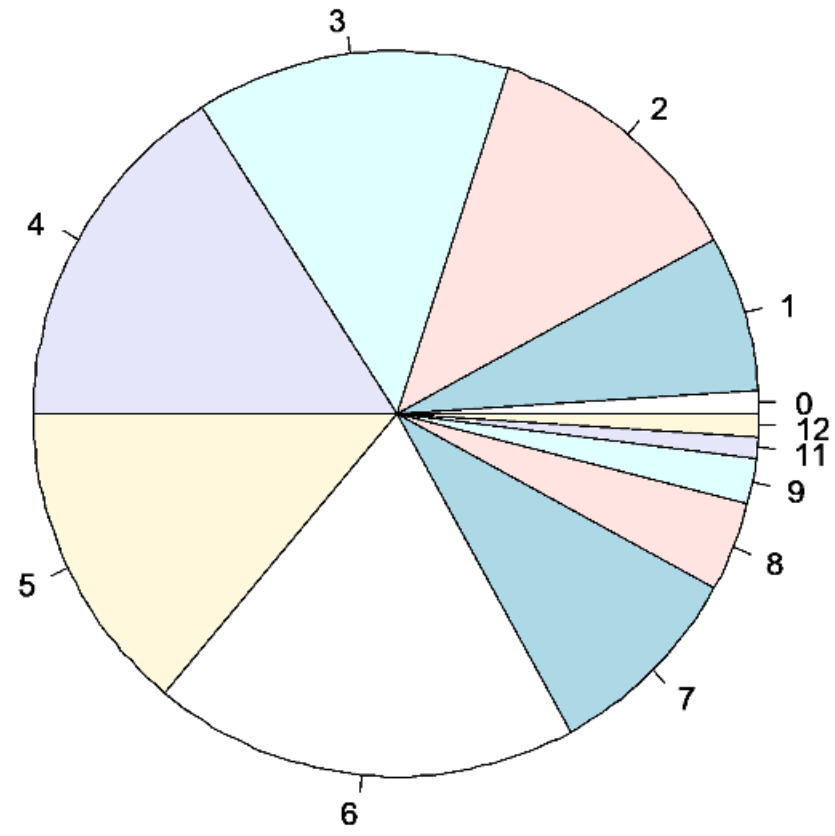
条形图

```
> tN <- table(Ni <- rpois(100, lambda=5))  
> tN  
> 0 1 2 3 4 5 6 7 8 9 11 12  
1 7 12 14 16 14 19 9 4 2 1 1  
> r <- barplot(tN, col='gray')
```



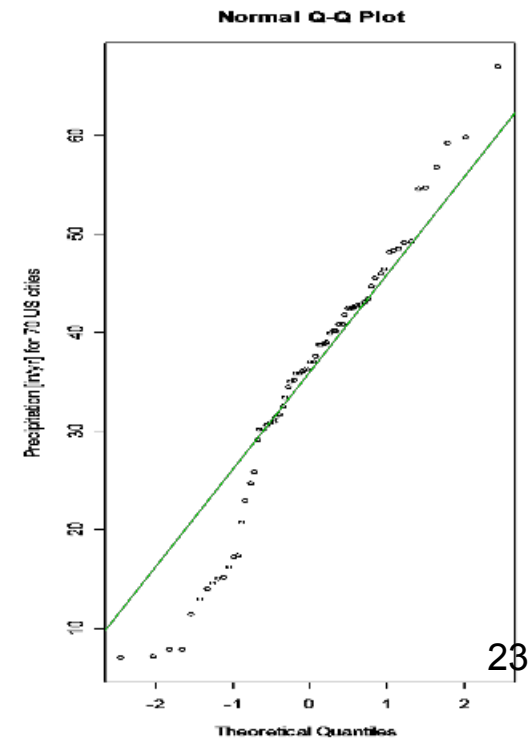
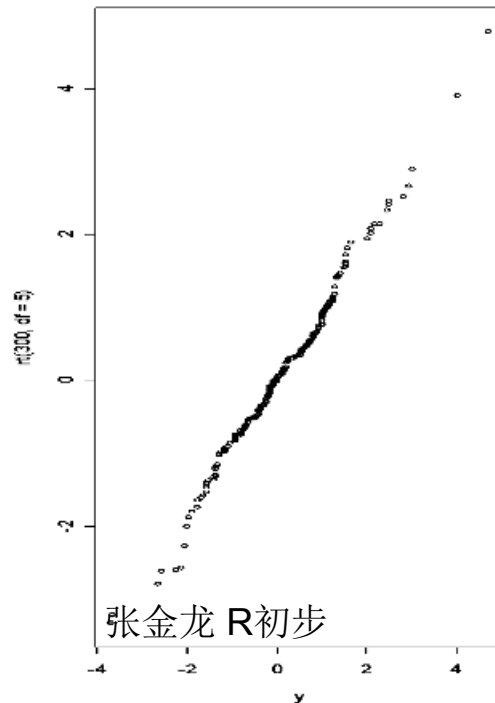
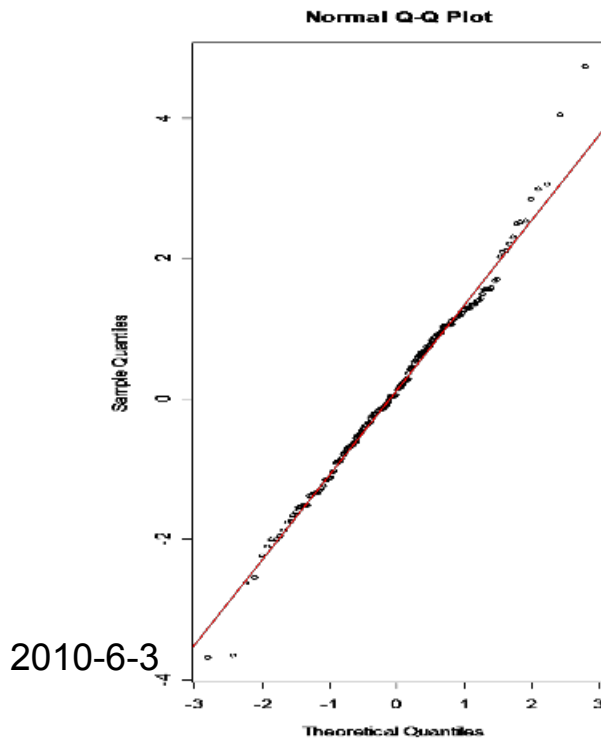
饼图

- pie(tN)



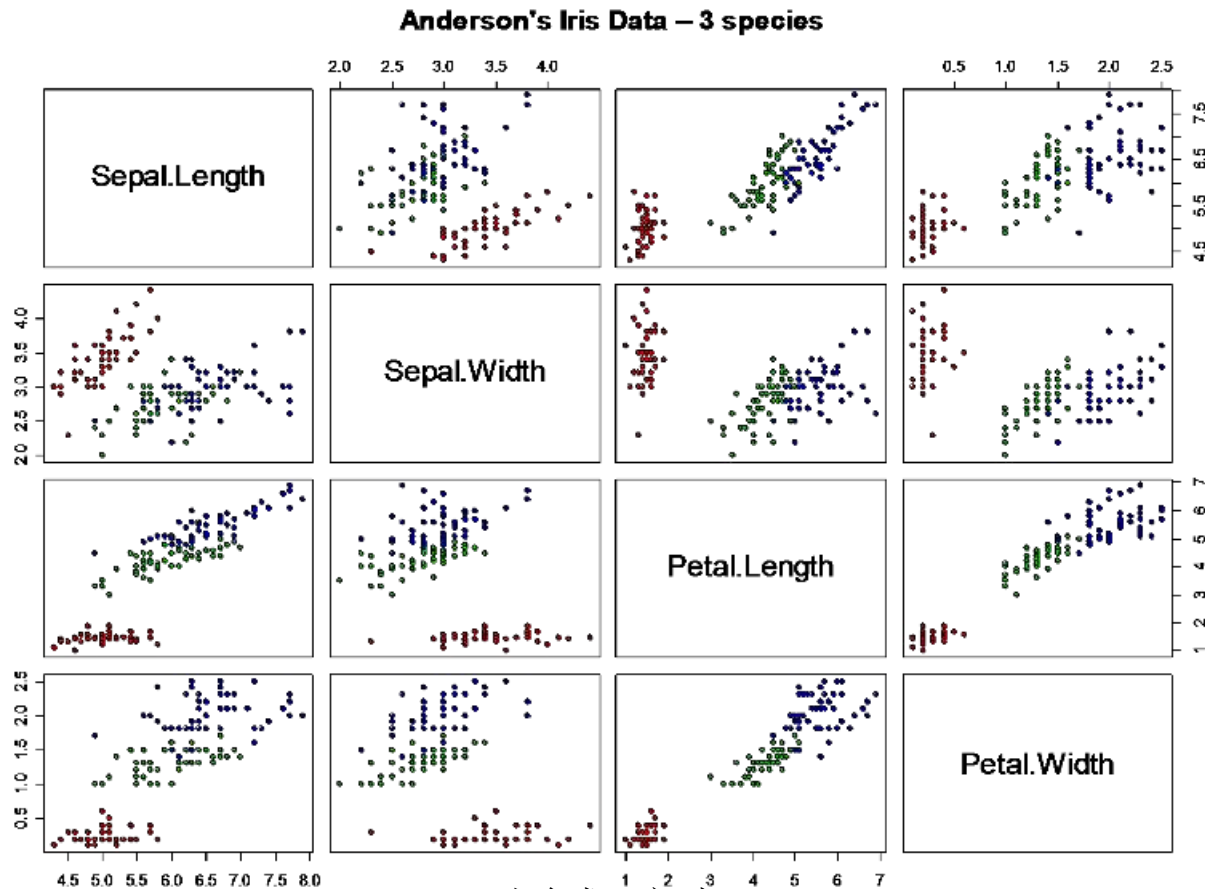


- > `par(mfrow=c(1,3))`
- > `y <- rt(200, df = 5)`
- > `qqnorm(y); qqline(y, col = 2) #和正态分布比`
- > `qqplot(y, rt(300, df = 5)); #和t(5)分布比`
- > `data(precip)`
- > `qqnorm(precip, ylab = "Precipitation [in/yr] for 70 US cities"); qqline(precip,col=3) #和正态分布比`
- > `par(mfrow=c(1,1))`



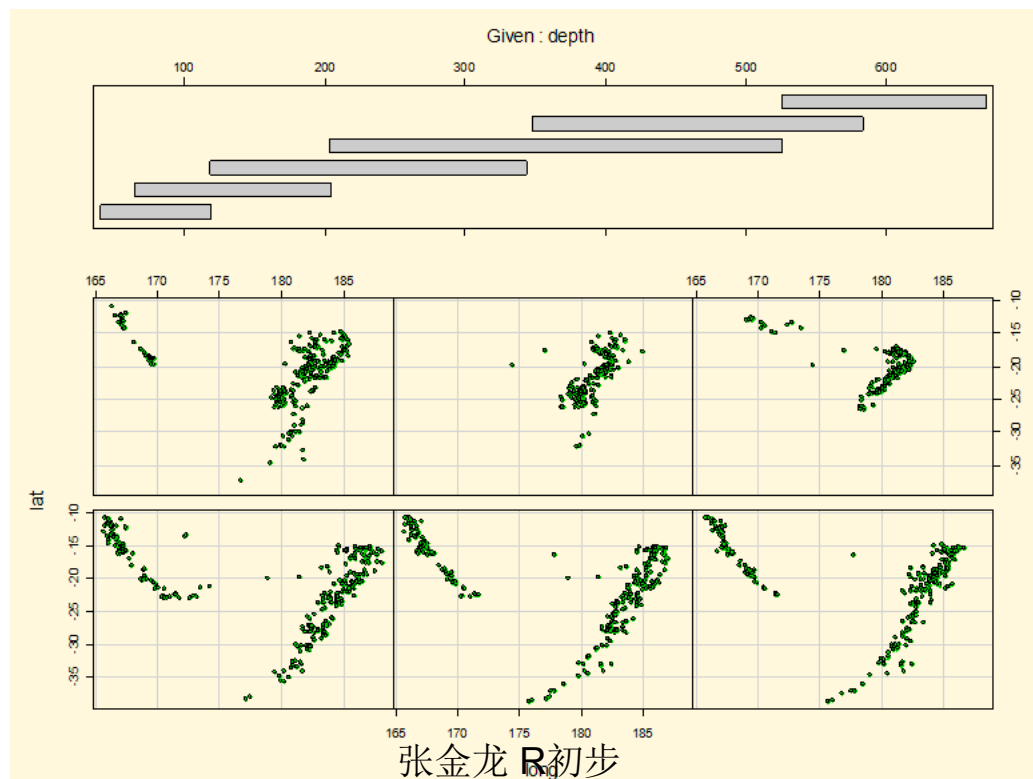
散点图矩阵

- > data(iris)
- > pairs(iris[1:4], main = "Anderson's Iris Data -- 3 species", pch = 21, bg = c("red", "green3", "blue")[codes(iris\$Species)])
- > #iris为150×5数据,这里是4个数量变量的点图(最后一个分类变量(iris\$Species))



条件分割图

- `par(bg = "cornsilk")`
- `data(quakes)`
- `coplot(lat ~ long | depth, data = quakes, pch = 21, bg = "green3")#Hit`
<Return> to see next plot:



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/3462410101010125>