ggplot2

4 марта 2015

Виноградова Светлана

Author

ggplot2 was developed by Hadley Wickham, assistant professor of statistics at Rice University, Houston. In July 2010 the latest stable release (Version 0.8.8) was published.

Hadley Wickham
Dobelman Family Junior Chair
Statistics, Rice University
6100 Main St MS#138
Houston TX 77005-1827

February 3, 2010

515 450 8171
hadley@rice.edu
http://had.co.nz

2008 Ph.D. (Statistics), Iowa State University, Ames, IA. “Practical tools for exploring data and models.”

2004 M.Sc. (Statistics), First Class Honours, The University of Auckland, Auckland, New Zealand.

2002 B.Sc. (Statistics, Computer Science), First Class Honours, The University of Auckland, Auckland, New Zealand.

1999 Bachelor of Human Biology, First Class Honours, The University of Auckland, Auckland, New Zealand.

http://www.ceb-institute.org/bbs/wp-content/uploads/2011/09/handout_ggplot2.pdf

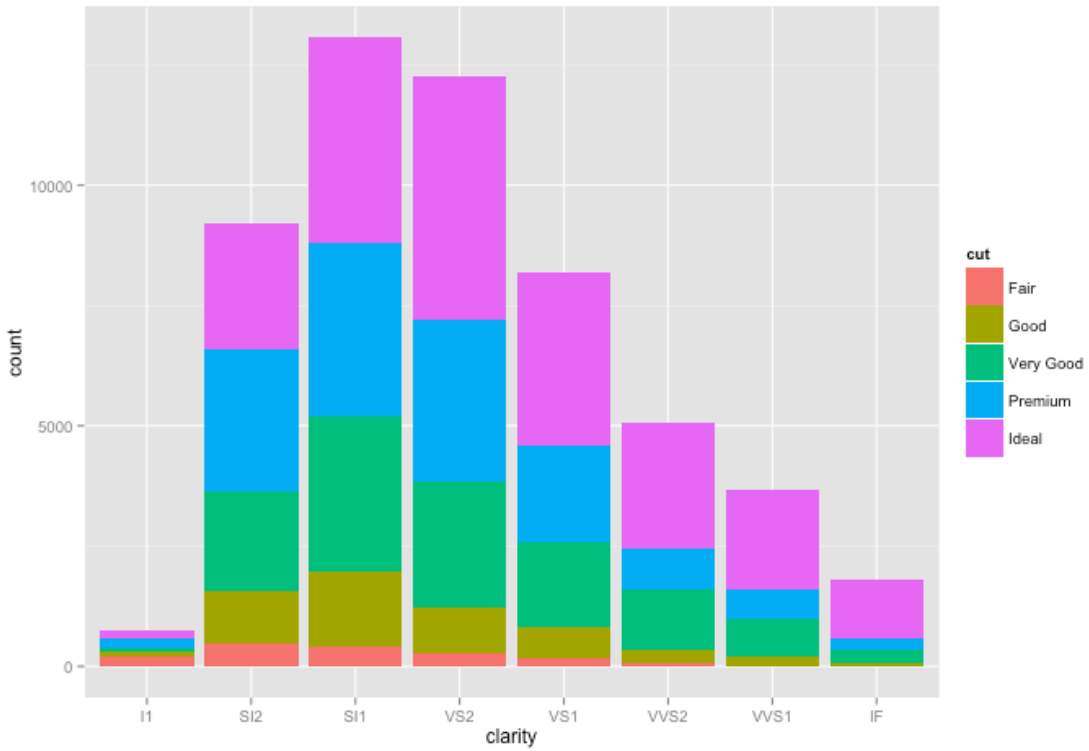
устанавливаем и загружаем пакет
install.package("ggplot2")
library("ggplot2")

данные, которые мы будем использовать для демонстрации
head(diamonds)
head(mtcars)

```
 carat       cut color clarity depth table price    x    y    z
1  0.23     Ideal     E     SI2  61.5    55   326 3.95 3.98 2.43
2  0.21   Premium     E     SI1  59.8    61   326 3.89 3.84 2.31
3  0.23      Good     E     VS1  56.9    65   327 4.05 4.07 2.31
4  0.29   Premium     I     VS2  62.4    58   334 4.20 4.23 2.63
5  0.31      Good     J     SI2  63.3    58   335 4.34 4.35 2.75
6  0.24 Very Good     J    VVS2  62.8    57   336 3.94 3.96 2.48
```

```
 mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

# сравним qplot и ggplot
# qplot histogram
qplot(clarity, data=diamonds, fill=cut, geom="bar")
# ggplot histogram -> same output
ggplot(diamonds, aes(clarity, fill=cut)) + geom_bar()



```
### как мы используем qplot
# scatterplot
qplot(wt, mpg, data=mtcars)
# можно менять данные прямо внутри
qplot(log(wt), mpg - 10, data=mtcars)
# добавим aesthetic mapping
qplot(wt, mpg, data=mtcars, color=qsec)
# поменяем размер точек
qplot(wt, mpg, data=mtcars, color=qsec, size=3)
# используем alpha blending
qplot(wt, mpg, data=mtcars, alpha=qsec)
```

```
qplot(wt, mpg, data=mtcars, colour=cyl)
levels(mtcars$cyl)
qplot(wt, mpg, data=mtcars, colour=factor(cyl))
# use different aesthetic mappings
qplot(wt, mpg, data=mtcars, shape=factor(cyl))
qplot(wt, mpg, data=mtcars, size=qsec)
# combine mappings (hint: hollow points, geom-concept, legend combination)
qplot(wt, mpg, data=mtcars, size=qsec, color=factor(carb))
qplot(wt, mpg, data=mtcars, size=qsec, color=factor(carb), shape=I(1))
qplot(wt, mpg, data=mtcars, size=qsec, shape=factor(cyl), geom="point")
qplot(wt, mpg, data=mtcars, size=factor(cyl), geom="point")
```

```
# bar-plot
qplot(factor(cyl), data=mtcars, geom="bar")
```

генерим данные и представляем их в удобном виде

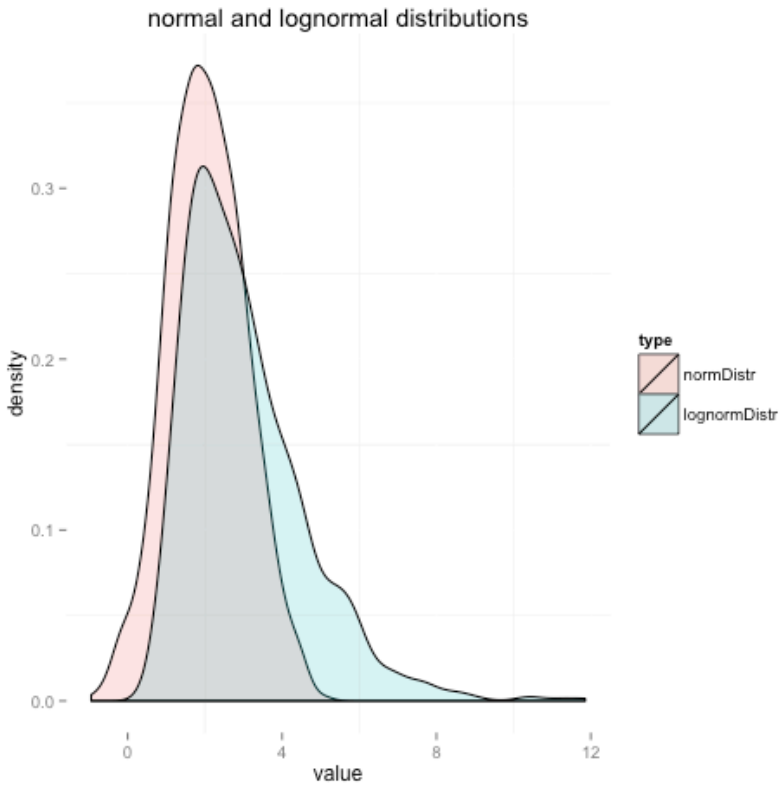
```
normDistr <- rnorm(1000, mean=2, sd=1)
lognormDistr <- rlnorm(1000, meanlog=1, sdlog=0.5)
x_data <- cbind(normDistr, lognormDistr)
```

```
 normDistr lognormDistr
[1,]  0.5533576     3.249032
[2,]  3.9941948     4.119568
[3,] -0.9461556     3.748958
[4,]  1.2034251     1.104947
[5,]  4.0532968     4.552704
[6,]  3.2065650     2.816817
```

```
library(reshape2)
x_data_melted <- melt(x_data)
colnames(x_data_melted) <- c("n", "type", "value")
```

```
  n      type      value
1 1 normDistr  0.5533576
2 2 normDistr  3.9941948
3 3 normDistr -0.9461556
4 4 normDistr  1.2034251
5 5 normDistr  4.0532968
6 6 normDistr  3.2065650
```

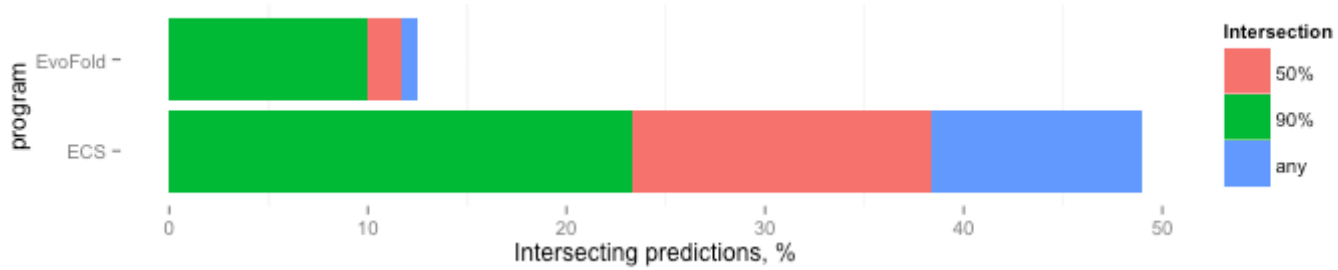
```
ggplot(x_data_melted, aes(value, fill = type)) + geom_density(alpha = 0.2) +
  theme(panel.background = element_rect(fill = 'white', colour = 'white')) +
  ggtitle("normal and lognormal distributions")
```

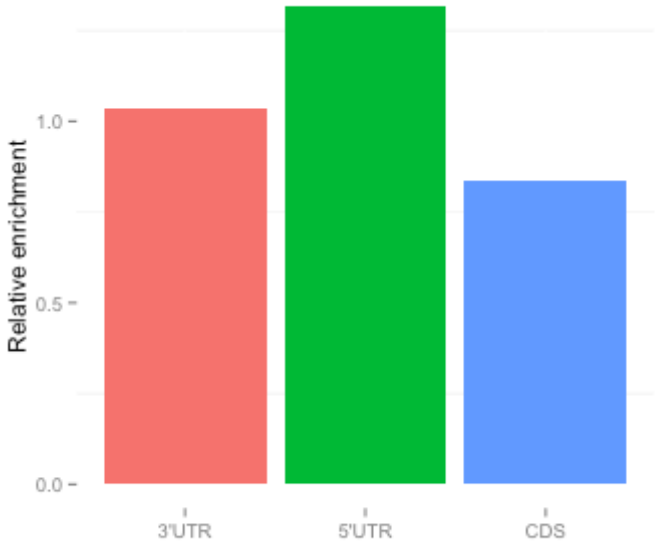


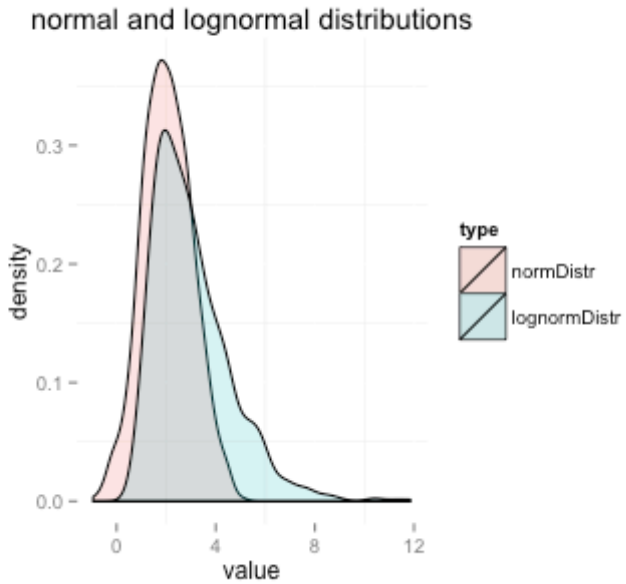
```
# в переменных p1, p2 и p3 у вас хранятся объекты типа ggplot
# тогда можно нарисовать сразу 2 графика
library(grid)
library(gridExtra)
grid.arrange(p1, p2, nrow=2)

# а можно все 3 и красиво их упорядочить

grid.newpage()
pushViewport(viewport(layout = grid.layout(3, 2)))
print(p1, vp = viewport(layout.pos.row = 1, layout.pos.col = 1:2))
print(p2, vp = viewport(layout.pos.row = 2:3, layout.pos.col = 1))
print(p3, vp = viewport(layout.pos.row = 2:3, layout.pos.col = 2:2))
```


program
EvoFold
ECS
0
10
20
30
40
50
Intersecting predictions, %
Intersection
50%
90%
any


Relative enrichment
1.0
0.5
0.0
3'UTR
5'UTR
CDS


normal and lognormal distributions
density
0.3
0.2
0.1
0.0
0
4
8
12
value
type
normDistr
lognormDistr