ggpubr 包系列学习教程

Xwyturbo

2022-11-03

1.R 包的安装及加载

```
#install.packages("ggpubr")
library(ggpubr)
```

2. 常用基本图形的绘制

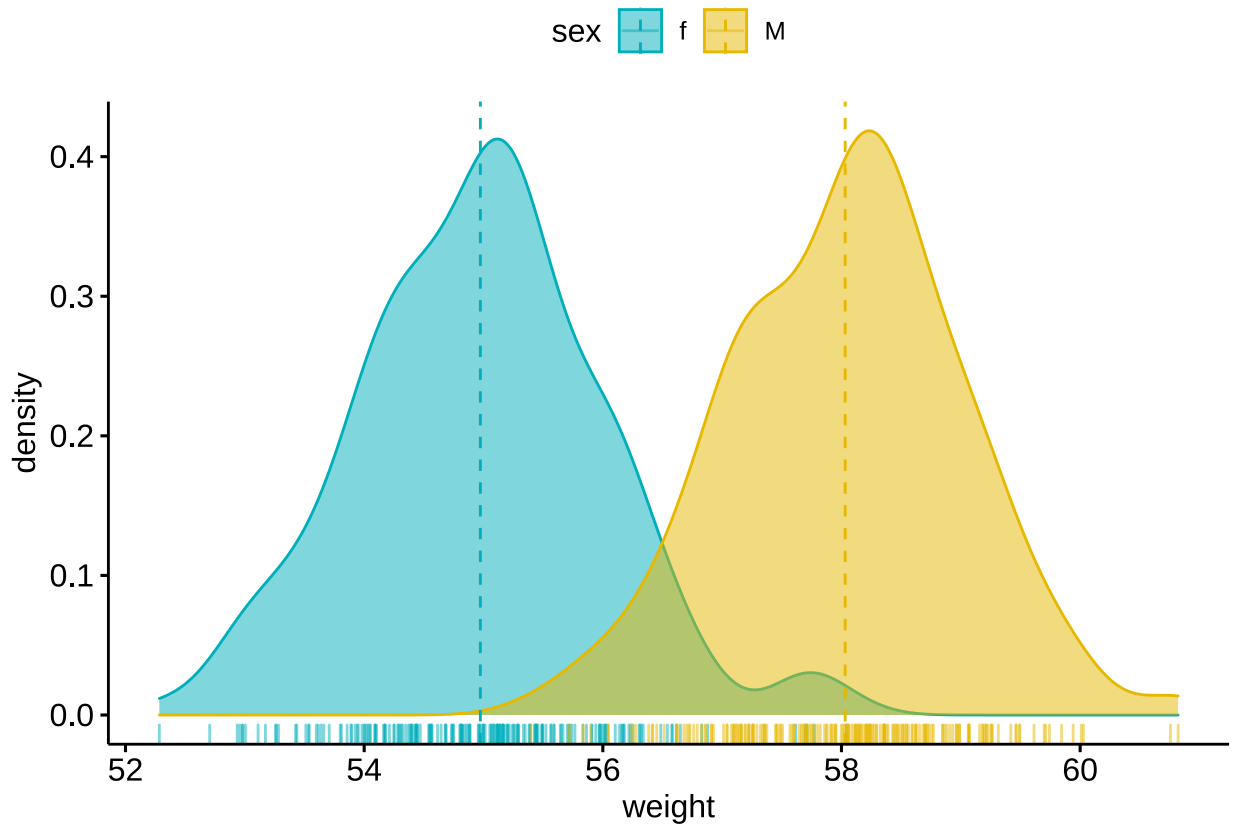
2.1 带有均值线和地毯线的密度图

```
library(ggpubr)
# 构建数据集 set.seed(1234)
df <- data.frame( sex=factor(rep(c("f", "M"), each=200)),
                  weight=c(rnorm(200, 55), rnorm(200, 58)))
# 预览数据格式
head(df)
```

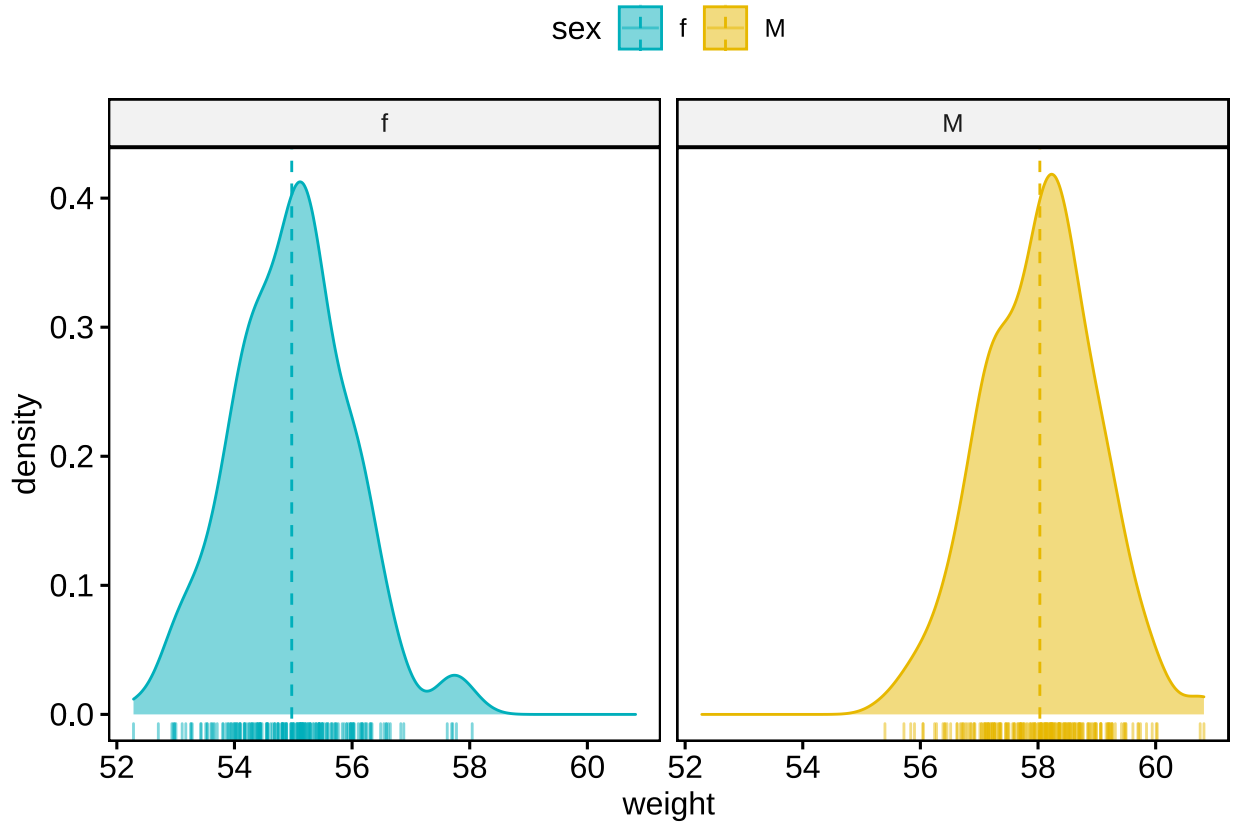
```
##   sex   weight
## 1   f 55.74066
## 2   f 56.10413
## 3   f 55.16629
## 4   f 56.14051
## 5   f 54.68209
## 6   f 55.34934
```

```
# 绘制密度图
# rug 参数添加地毯线,
# add 参数可以添加均值 mean 和中位数 median,
# 按性别" sex "分组标记边框线颜色和填充色, 使用 palette 参数自定义颜色
p1 <- ggdensity(df, x="weight", add = "mean", rug = TRUE, color = "sex",
                fill = "sex",palette = c("#00AFBB", "#E7B800"))
```

```
p1
```

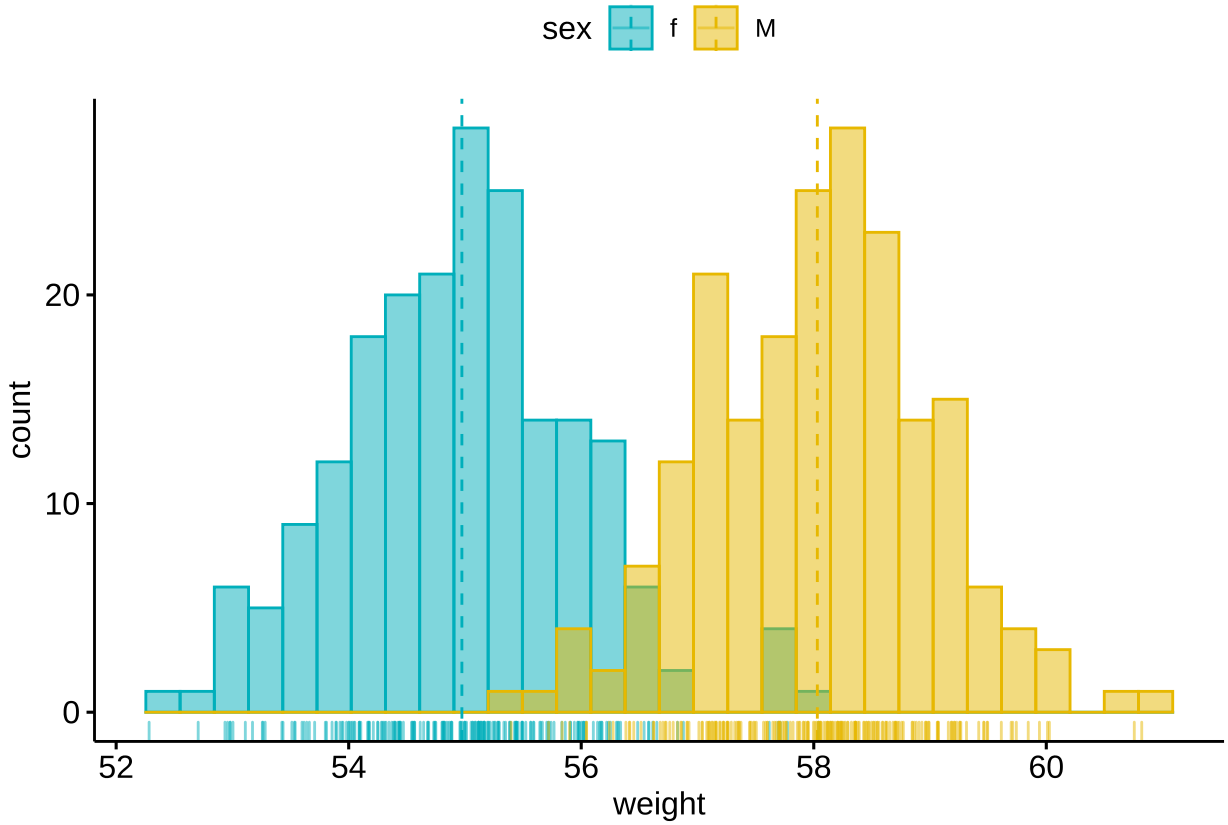



```
p11 <- ggdensity(df, x="weight",facet.by = "sex",
                 add = "mean", rug = TRUE, color = "sex",
                 fill = "sex",palette = c("#00AFBB", "#E7B800"))
p11
```

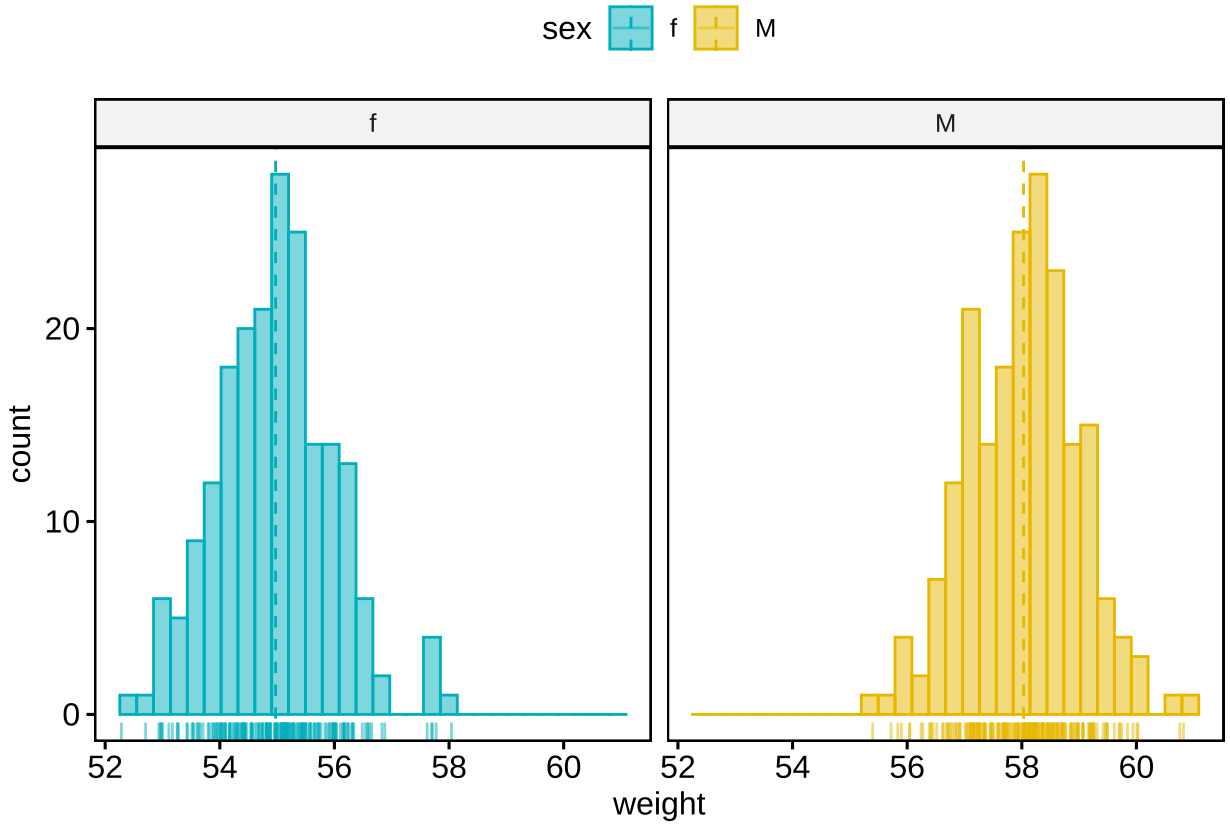


2.2 带有均值线和边际地毯线的直方图

```
library(ggpubr)
p2 <- gghistogram(df, x="weight",
                  add = "mean", rug = TRUE, color = "sex",
                  fill = "sex",palette = c("#00AFBB", "#E7B800"))
p2
```




```
p22 <- gghistogram(df, x="weight",facet.by = "sex",
                   add = "mean", rug = TRUE, color = "sex",
                   fill = "sex",palette = c("#00AFBB", "#E7B800"))
p22
```



2.3 箱线图 + 分组形状 + 统计

```
library(ggpubr)
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
```

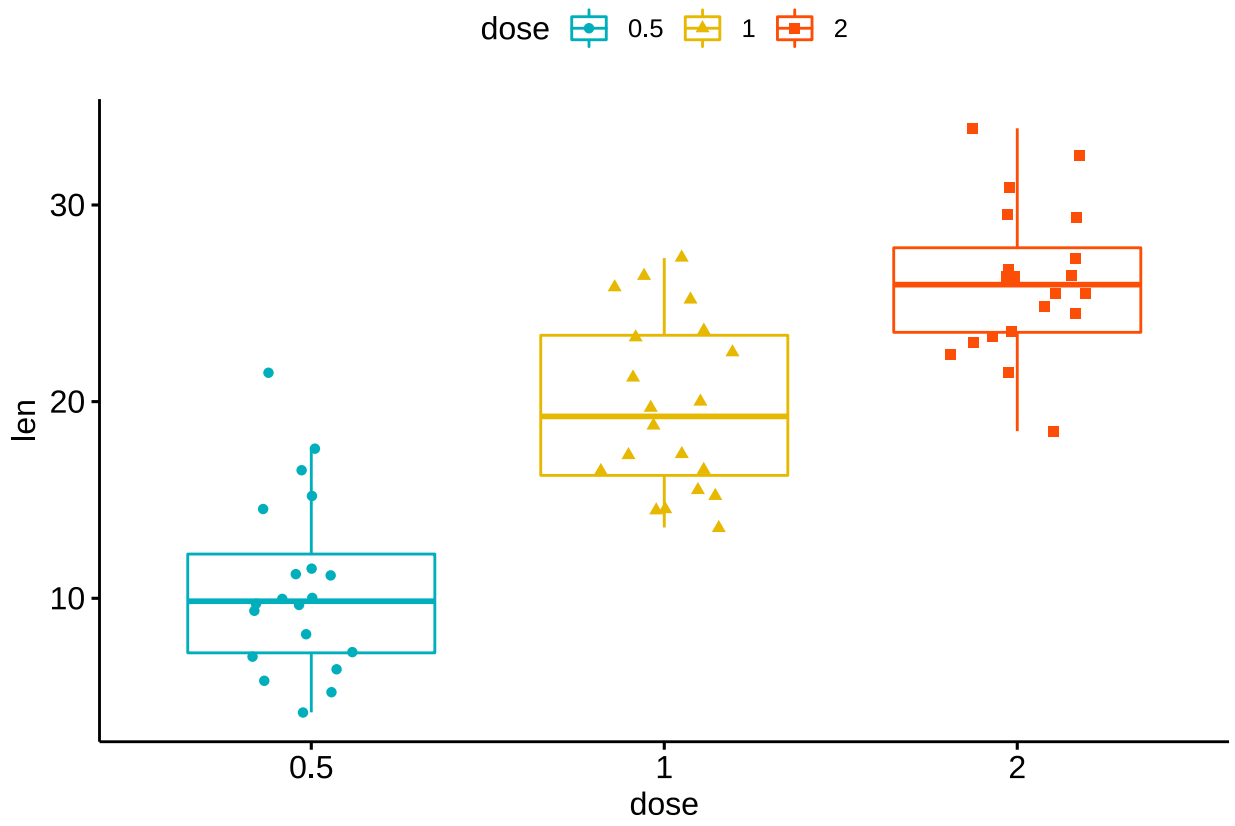
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
head(ToothGrowth)
```

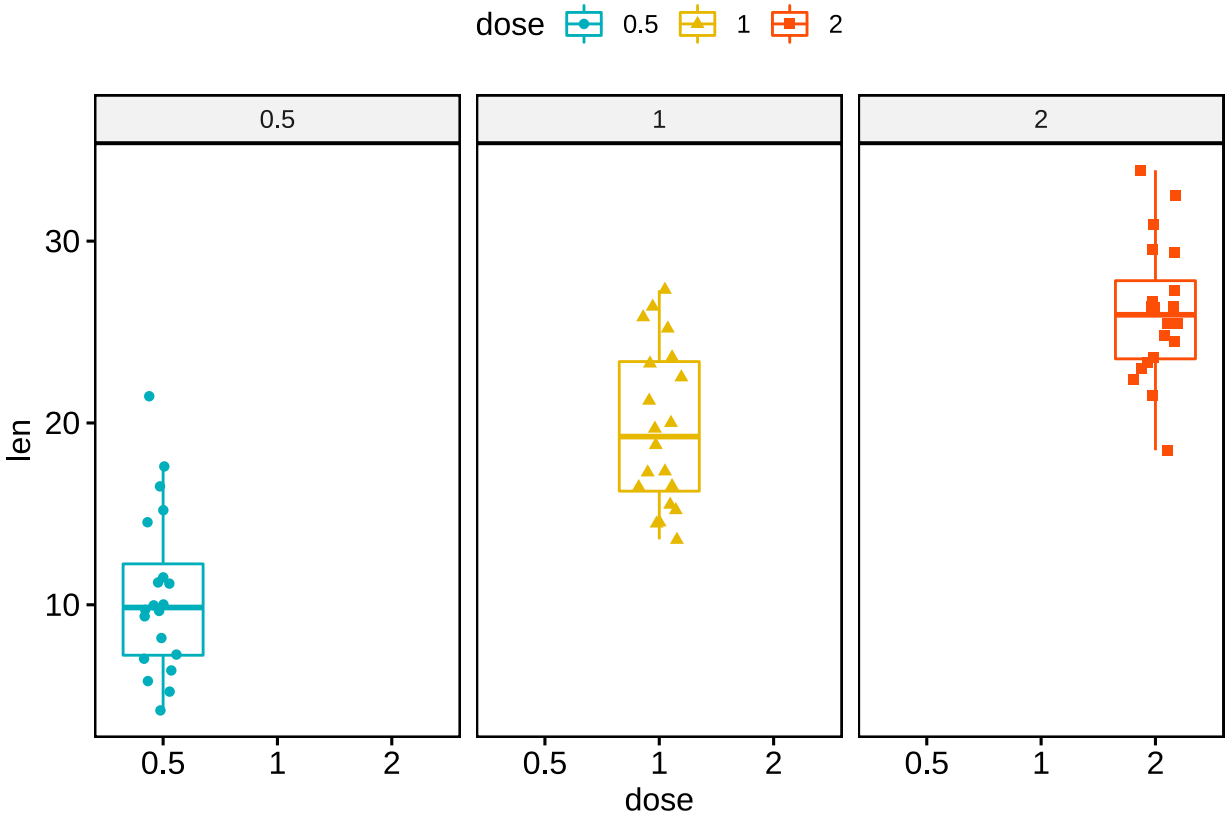
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
```

```
## 6 10.0   VC  0.5
```

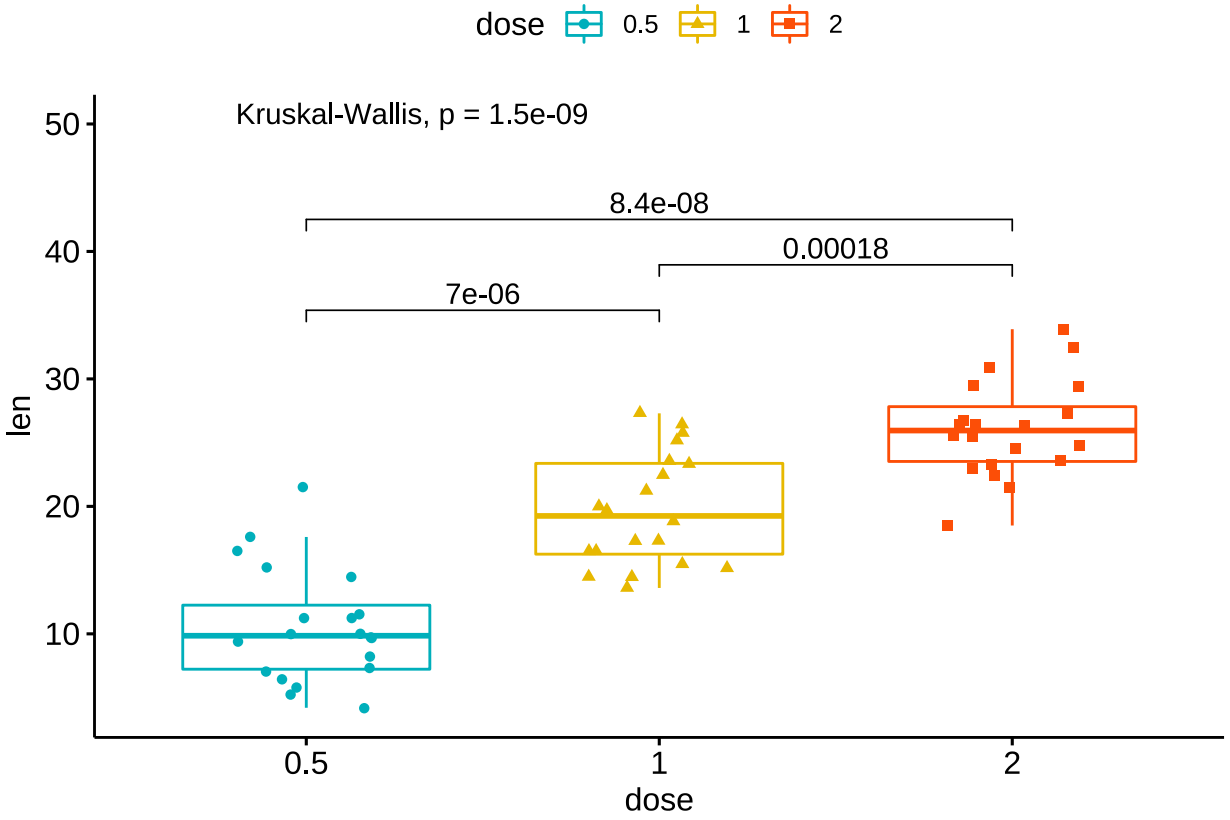
```
#jitter 参数添加扰动点，点的形状 shape 由 dose 变量映射
p3 <- ggboxplot(ToothGrowth, x="dose", y="len", color = "dose",
               palette = c("#00AFBB", "#E7B800", "#FC4E07"),
               add = "jitter", shape="dose")
p3
```




```
p32 <- ggboxplot(ToothGrowth, x="dose", y="len",facet.by = "dose",
                  color = "dose",
               palette = c("#00AFBB", "#E7B800", "#FC4E07"),
               add = "jitter", shape="dose")
p32
```




```
# stat_compare_means 参数比较不同组之间的均值，
# 并增加不同组间比较的 p-value 值，可以自定义需要标注的组间比较
my_comparisons <- list(c("0.5", "1"), c("1", "2"), c("0.5", "2"))
p4 <- p3 + stat_compare_means(comparisons = my_comparisons)+
      stat_compare_means(label.y = 50)
p4
```



2.4 内有箱线图的小提琴图 + 星标记

```
library(ggpubr)
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
```

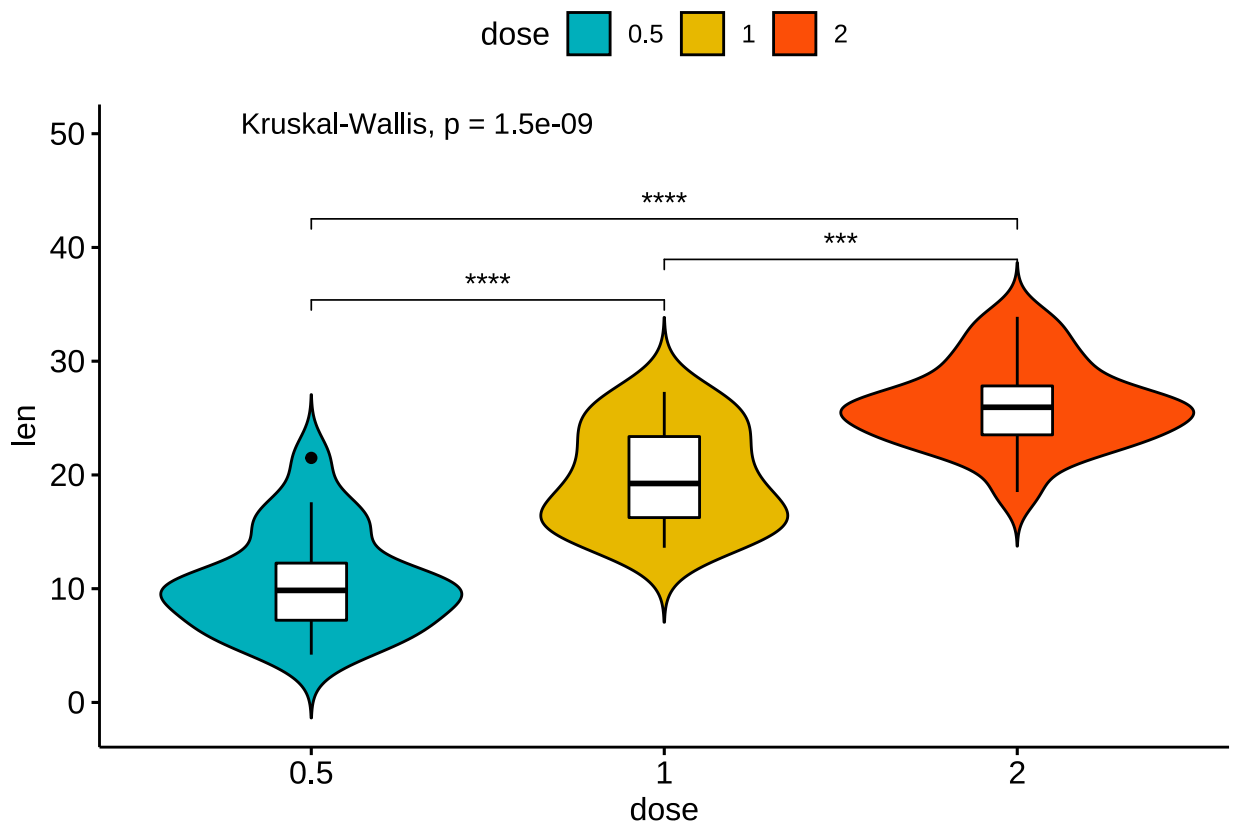
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
```

```
## 6 10.0   VC  0.5
```

```
# add = “boxplot”添加箱线图
# stat_compare_means 中设置 lable=”p.signif”,
# 即可添加星添加组间比较连线和统计 P 值按星分类
# add 添加箱线图，label 标注选择显著性标记（星号）
p5 <- ggviolin(ToothGrowth, x="dose", y="len", fill = "dose",
        palette = c("#00AFBB", "#E7B800", "#FC4E07"),
        add = "boxplot", add.params = list(fill="white"))+
        stat_compare_means(comparisons = my_comparisons, label = "p.signif") +
        stat_compare_means(label.y = 50)
 p5
```



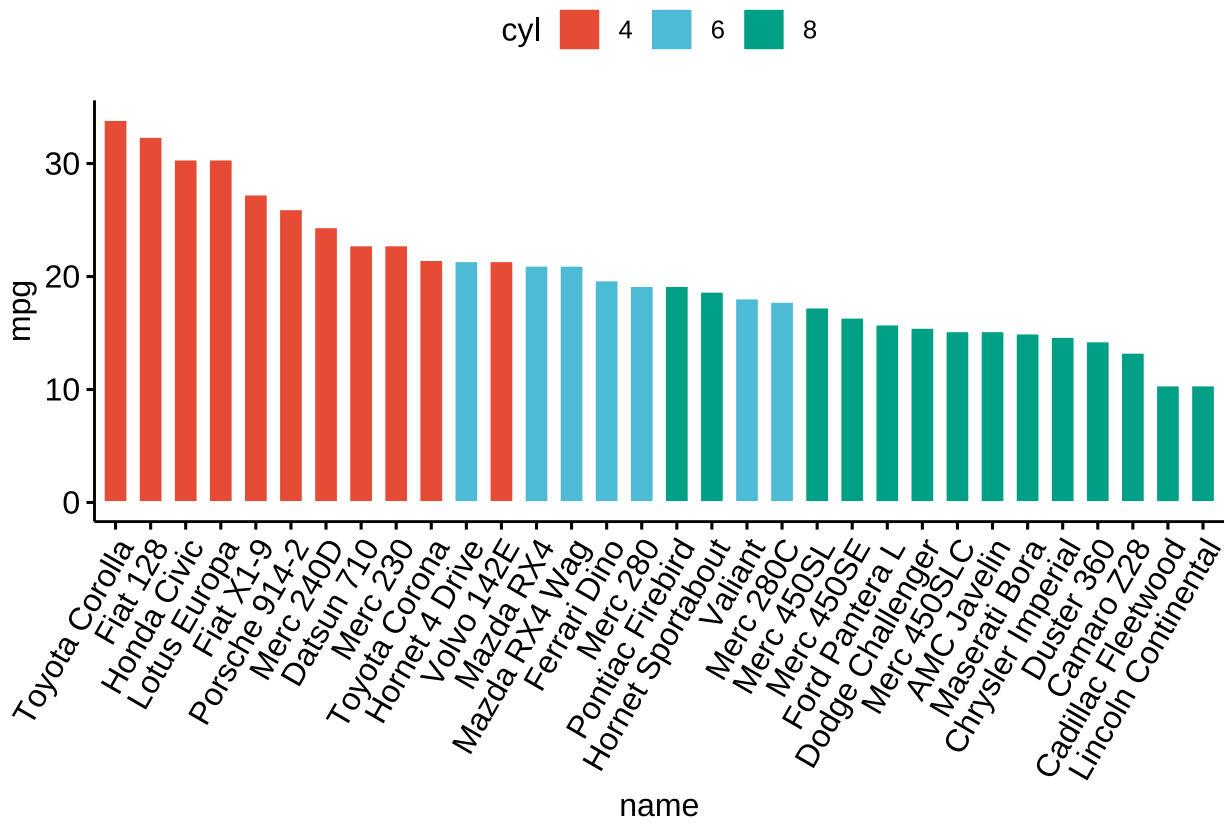
2.5 条形/柱状图绘制 (barplot)

```
library(ggpubr)
# 加载数据集
data("mtcars")
df2 <- mtcars
# 设置因子变量
```

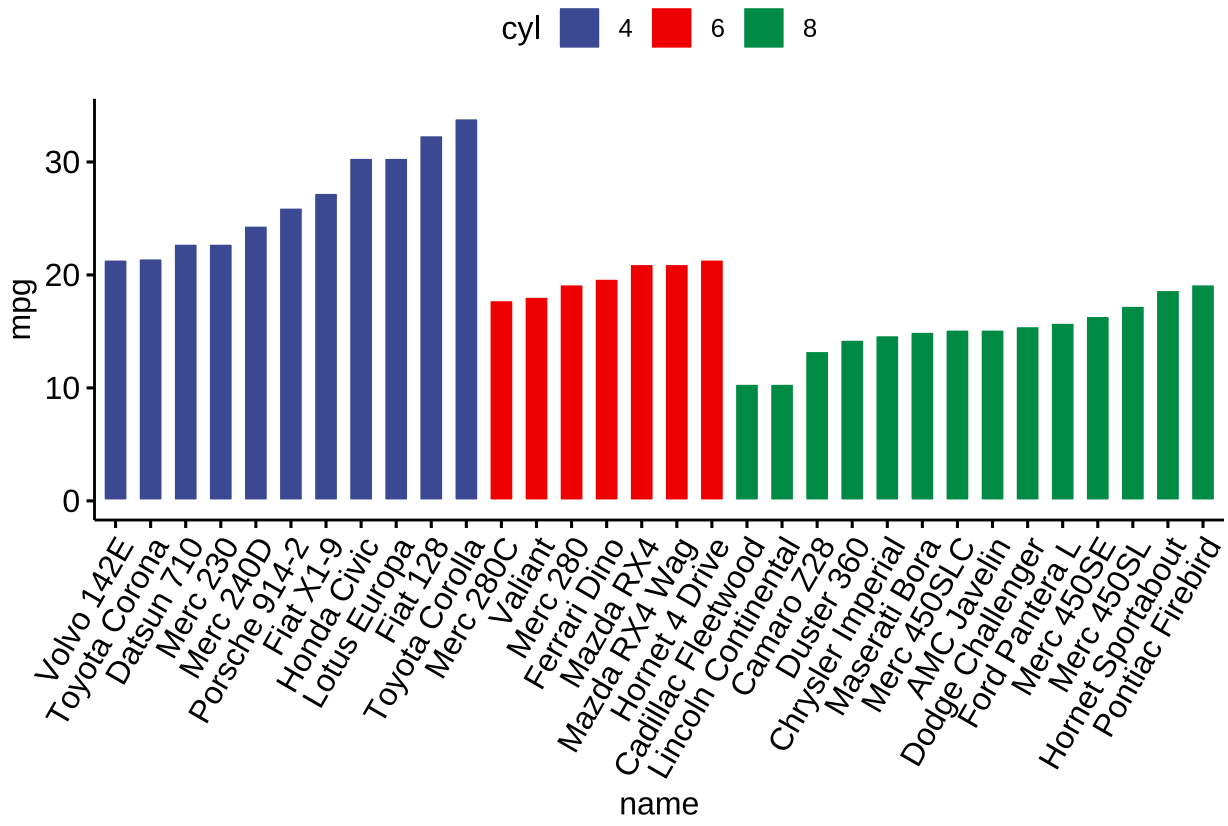
```r
df2$cyl <- factor(df2$cyl)
df2$name <- rownames(df2) # 添加一新列 name
head(df2[, c("name", "wt", "mpg", "cyl")])
```

```
##                                name    wt  mpg cyl
## Mazda RX4                  Mazda RX4 2.620 21.0   6
## Mazda RX4 Wag          Mazda RX4 Wag 2.875 21.0   6
## Datsun 710                Datsun 710 2.320 22.8   4
## Hornet 4 Drive        Hornet 4 Drive 3.215 21.4   6
## Hornet Sportabout Hornet Sportabout 3.440 18.7   8
## Valiant                      Valiant 3.460 18.1   6
```

```r
# 颜色按 nature 配色方法 (支持 ggsci 包中的本色方案，如："npg"，"aaas"，"lancet"，"jco"，"uc
 p6 <- ggbarplot(df2, x="name", y="mpg", fill = "cyl", color = "white",
          palette = "npg", # 杂志 nature 的配色
          sort.val = "desc", # 降序排序
          sort.by.groups=FALSE, # 不按组排序
          x.text.angle=60)
p6
```



```
# 按组进行排序
p7 <- ggbarplot(df2, x="name", y="mpg", fill = "cyl", color = "white",
          palette = "aaas", # 杂志 Science 的配色
          sort.val = "asc", # 升序排序，区别于 desc
          sort.by.groups=TRUE,x.text.angle=60)
# 按组排序 x.text.angl 设置 x 轴标签旋转角度
p7
```

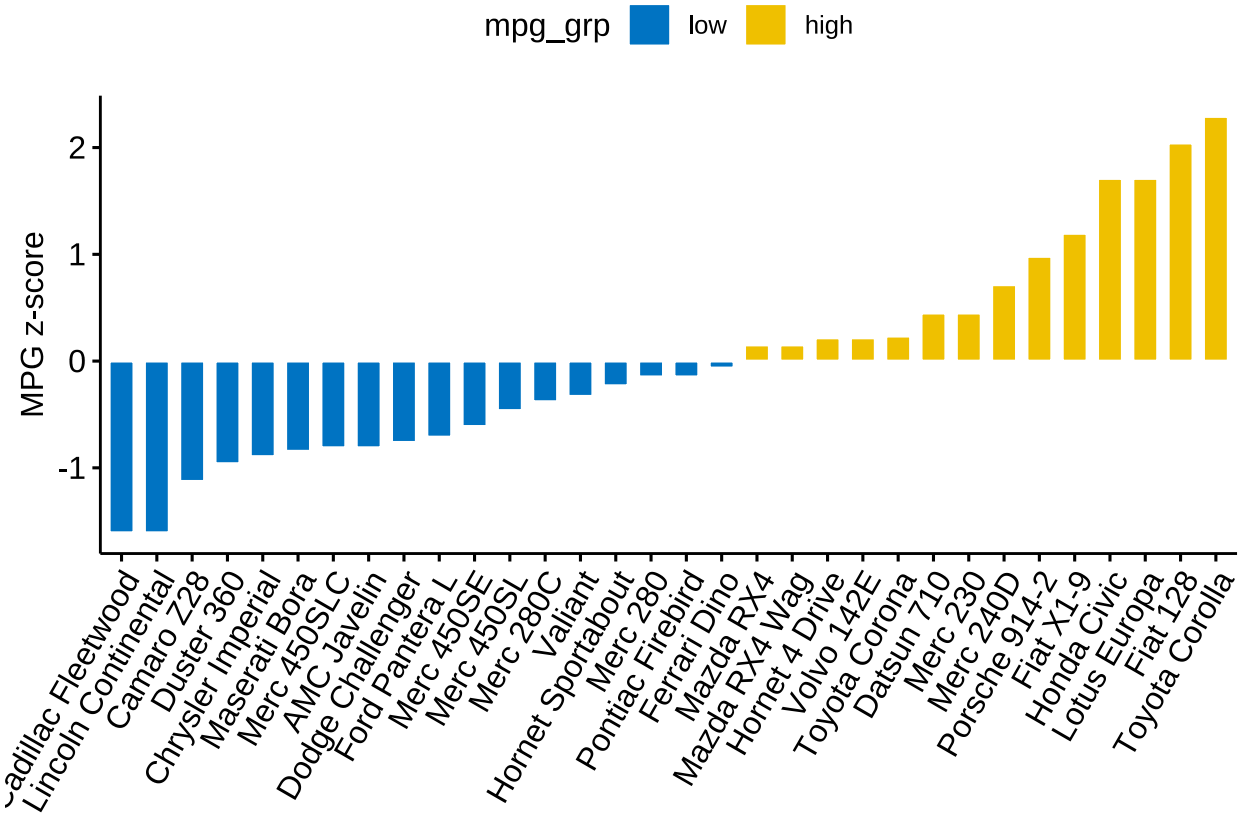



```
# 偏差图绘制 (Deviation graphs)，偏差图展示了与参考值之间的偏差。
df2$mpg_z <- (df2$mpg-mean(df2$mpg))/sd(df2$mpg)
# 相当于 Zscore 标准化，减均值，除标准差
df2$mpg_grp <- factor(ifelse(df2$mpg_z<0, "low", "high"),
                      levels = c("low", "high"))
# 设置分组因子
head(df2[, c("name", "wt", "mpg", "mpg_grp", "cyl")])
```

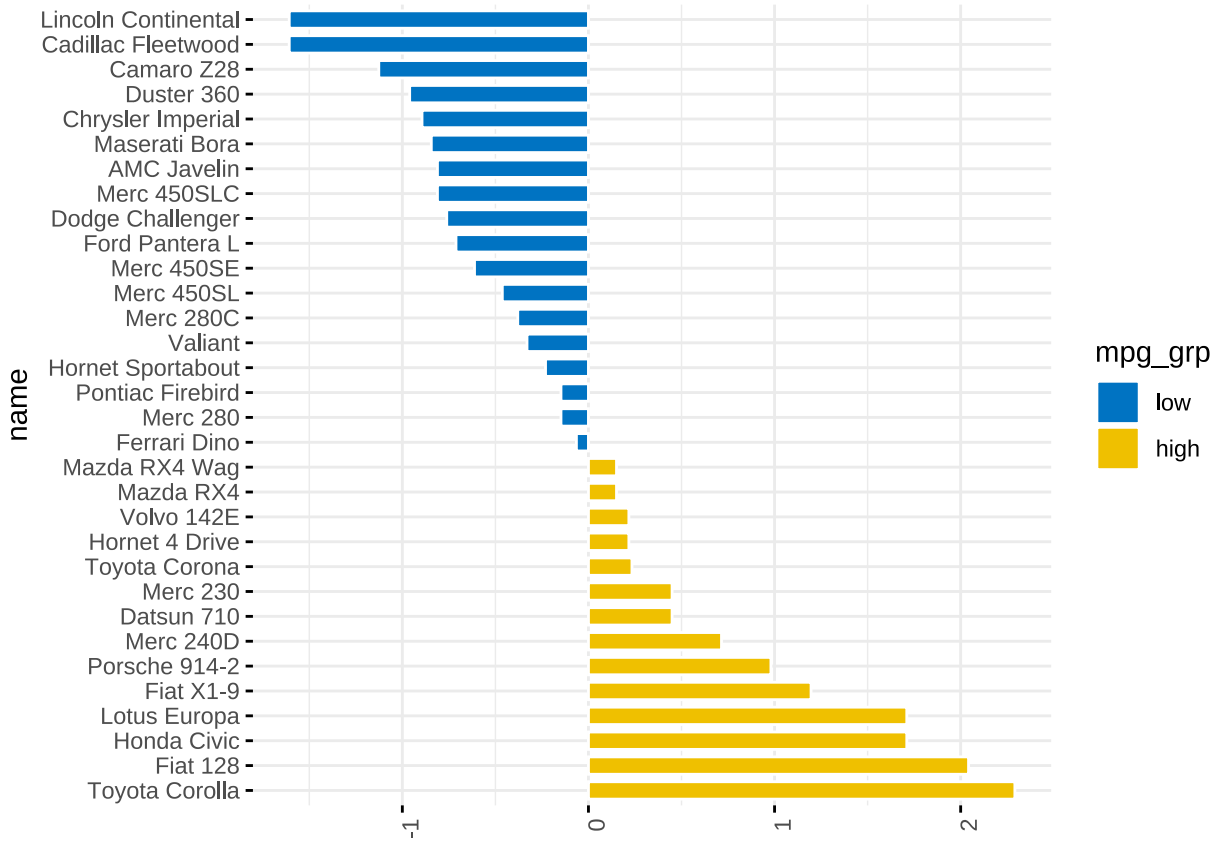
```
##                           name    wt  mpg mpg_grp cyl
## Mazda RX4                 Mazda RX4 2.620 21.0    high   6
## Mazda RX4 Wag         Mazda RX4 Wag 2.875 21.0    high   6
## Datsun 710               Datsun 710 2.320 22.8    high   4
```

```
## Hornet 4 Drive       Hornet 4 Drive 3.215 21.4    high   6
## Hornet Sportabout Hornet Sportabout 3.440 18.7     low   8
## Valiant                     Valiant 3.460 18.1     low   6
```

```
p8 <- ggbarplot(df2, x="name", y="mpg_z", fill = "mpg_grp", color = "white",
        palette = "jco", sort.val = "asc", sort.by.groups = FALSE,
        x.text.angle=60, ylab = "MPG z-score", xlab = FALSE)
p8
```

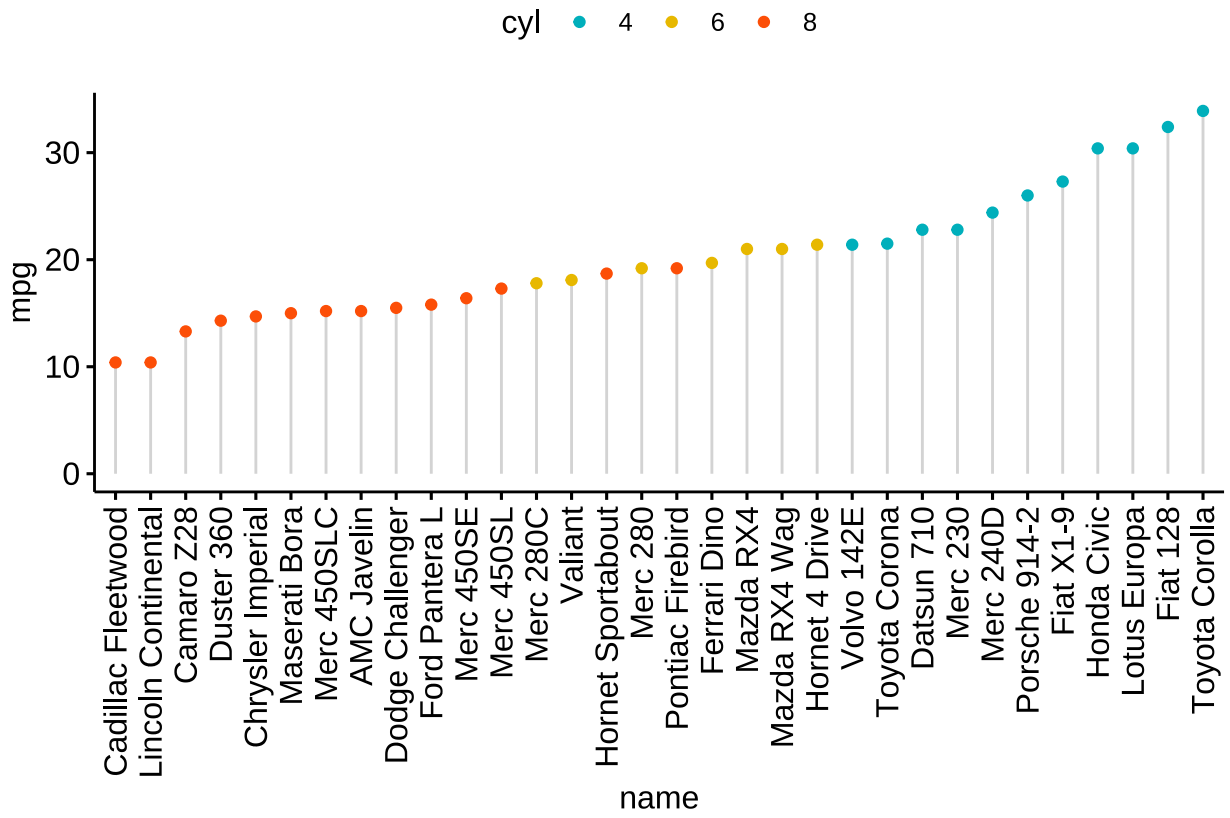



```
# rotate 设置 x/y 轴对换
p9 <- ggbarplot(df2, x="name", y="mpg_z", fill = "mpg_grp", color = "white",
         palette = "jco", sort.val = "desc", sort.by.groups = FALSE,
         x.text.angle=90, ylab = "MPG z-score", xlab = FALSE,
         rotate=TRUE, ggtheme = theme_minimal())
p9
```

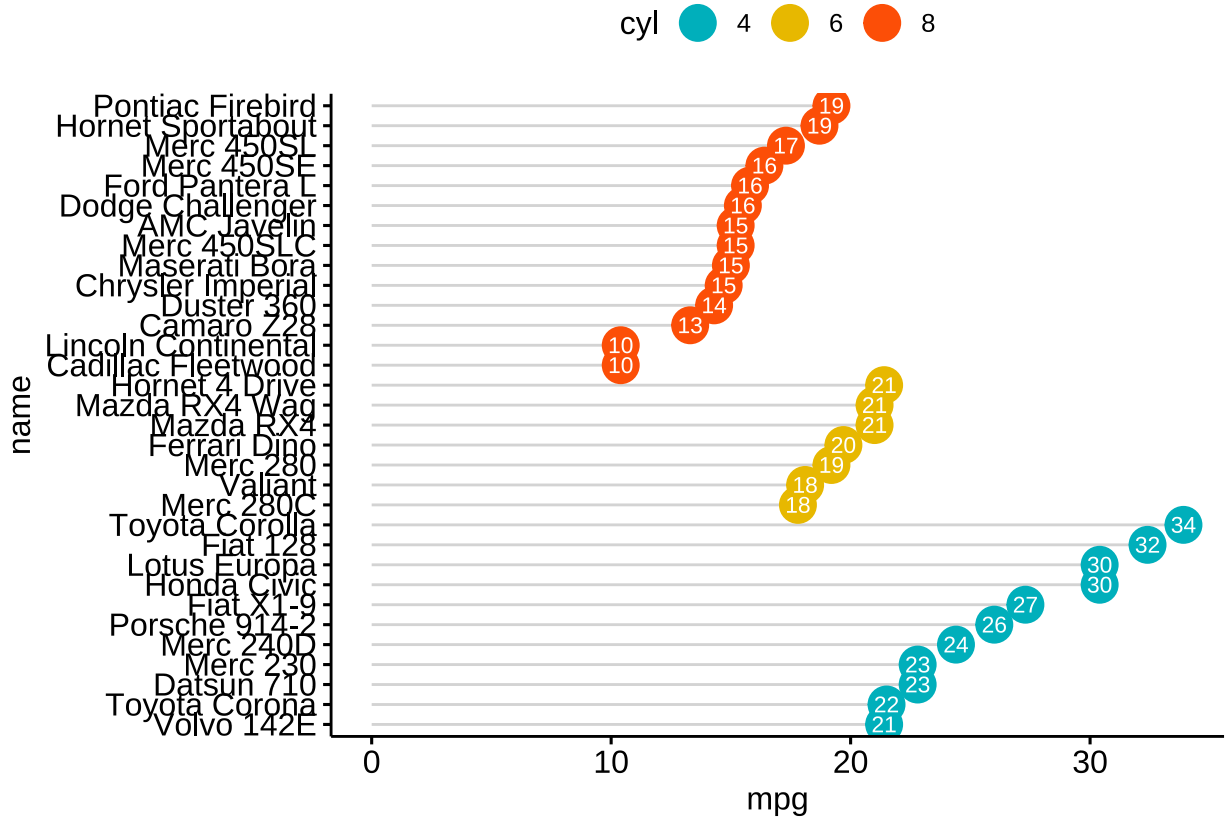


2.6 棒棒糖图绘制 (Lollipop chart), 棒棒图可以代替条形图展示数据

```
library(ggpubr)
library(ggpubr)
# 加载数据集
data("mtcars")
df2 <- mtcars
df2$cyl <- factor(df2$cyl)
df2$name <- rownames(df2) # 添加一新列 name
p10 <- ggdotchart(df2, x="name", y="mpg", color = "cyl",
          palette = c("#00AFBB", "#E7B800", "#FC4E07"),
          sorting = "ascending",
          add = "segments", ggtheme = theme_pubr())
p10
```

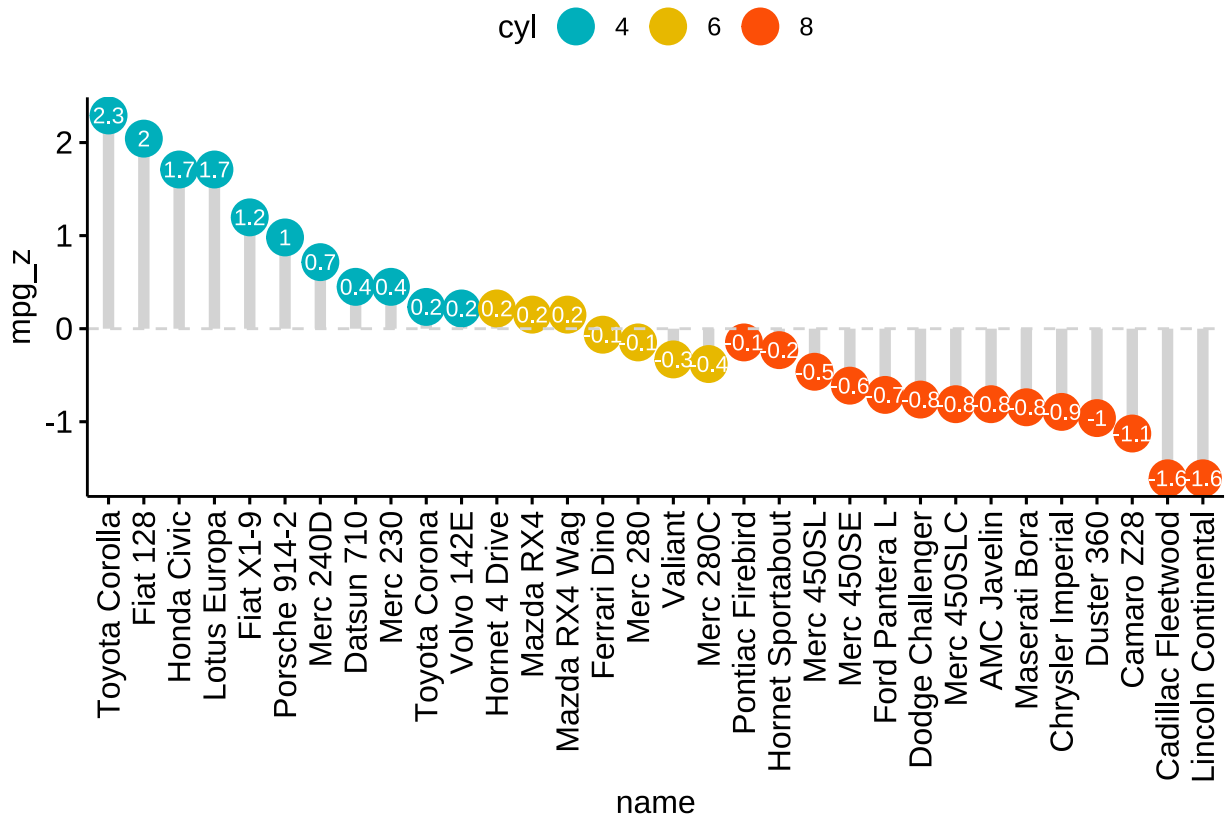



```
# 设置其他参数，dot.size = 6 调整糖的大小，添加 label 标签，设置字体样式和方向
p11 <- ggdotchart(df2, x="name", y="mpg", color = "cyl",
          palette = c("#00AFBB", "#E7B800", "#FC4E07"),
          sorting = "descending", add = "segments", rotate = TRUE,
          group = "cyl", dot.size = 6,
          label = round(df2$mpg), font.label = list(color="white",
          size=9, vjust=0.5), ggtheme = theme_pubr())
p11
```




```
# 偏差图绘制 (Deviation graphs), 偏差图展示了与参考值之间的偏差。
df2$mpg_z <- (df2$mpg-mean(df2$mpg))/sd(df2$mpg)
# 相当于 Zscore 标准化，减均值，除标准差
df2$mpg_grp <- factor(ifelse(df2$mpg_z<0, "low", "high"),
                      levels = c("low", "high"))
# 棒棒糖偏差图
p12 <- ggdotchart(df2, x = "name", y = "mpg_z",
                    color = "cyl", # Color by groups
                    palette = c("#00AFBB", "#E7B800", "#FC4E07"),
                  # Custom color palette
                    sorting = "descending", # Sort value in descending order
                    add = "segments", # Add segments from y = 0 to dots
                    add.params = list(color = "lightgray", size = 2),
                  # Change segment color and size
                    group = "cyl", # Order by groups
                    dot.size = 6, # Large dot size
                    label = round(df2$mpg_z,1),
                  # Add mpg values as dot labels, 设置一位小数
                    font.label = list(color = "white", size = 9, vjust = 0.5),
```

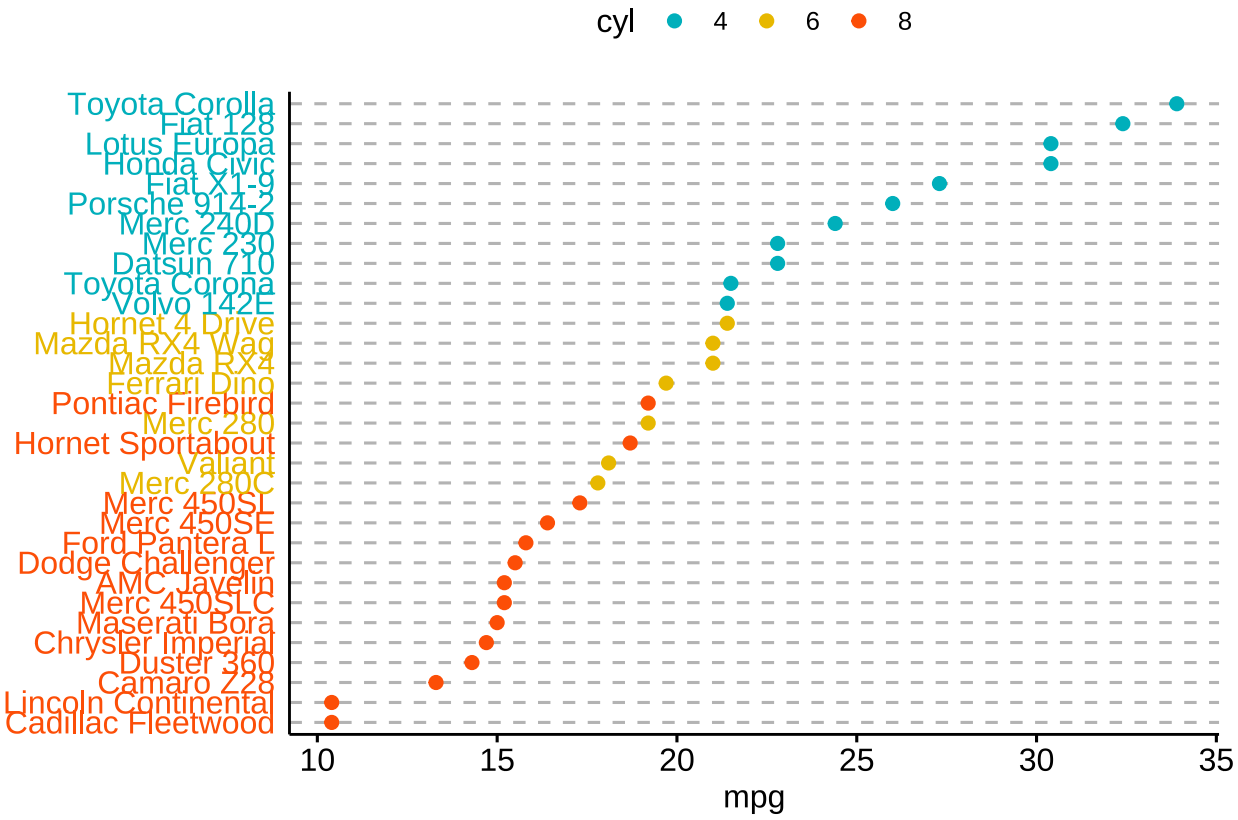
```
                  # Adjust label parameters
                    ggtheme = theme_pubr()) +
                    geom_hline(yintercept = 0, linetype = 2,
                               color = "lightgray")
p12
```



2.7Cleveland 点图绘制

```
library(ggpubr)
library(ggpubr)
# 加载数据集
data("mtcars")
df2 <- mtcars
df2$cyl <- factor(df2$cyl)
df2$name <- rownames(df2) # 添加一新列 name
# 偏差图绘制 (Deviation graphs), 偏差图展示了与参考值之间的偏差。
df2$mpg_z <- (df2$mpg-mean(df2$mpg))/sd(df2$mpg)
# 相当于 Zscore 标准化，减均值，除标准差
df2$mpg_grp <- factor(ifelse(df2$mpg_z<0, "low", "high"),
```

```
                        levels = c("low", "high"))
# theme_cleveland() 主题可设置为 Cleveland 点图样式
p13 <- ggdotchart(df2, x = "name", y = "mpg",
                       color = "cyl", # Color by groups
                       palette = c("#00AFBB", "#E7B800", "#FC4E07"), # Custom color palette
                       sorting = "descending", # Sort value in descending order
                       rotate = TRUE, # Rotate vertically
                       dot.size = 2, # Large dot size
                       y.text.col = TRUE, # Color y text by groups
                       ggtheme = theme_pubr() # ggplot2 theme
                       ) +
                       theme_cleveland() # Add dashed grids
p13
```



3. 常用基本绘图函数及参数

3.1 基本绘图函数

```
#gghistogram        Histogram plot # 绘制直方图
#ggdensity        Density plot # 绘制密度图
#ggdotplot        Dot plot # 绘制点图
#ggdotchart        Cleveland's Dot Plots # 绘制 Cleveland 点图
#ggline        Line plot # 绘制线图
#ggbarplot        Bar plot # 绘制条形/柱状图
#ggerrorplot        Visualizing Error # 绘制误差棒图
#ggstripchart        Stripcharts # 绘制线带图
#ggboxplot        Box plot # 绘制箱线图
#ggviolin        Violin plot # 绘制小提琴图
#ggpie        Pie chart # 绘制饼图
#ggqqplot        QQ Plots # 绘制 QQ 图
#ggscatter        Scatter plot # 绘制散点图
#ggmaplot        MA-plot from means and log fold changes # 绘制 M-A 图
#ggpaired        Plot Paired Data # 绘制配对数据
#ggecdf        Empirical cumulative density function  # 绘制经验累积密度分布图
```

3.2 基本参数

```
# ggtext        Text # 添加文本
# border        Set ggplot Panel Border Line # 设置画布边框线
# grids        Add Grids to a ggplot # 添加网格线
# font        Change the Appearance of Titles and Axis Labels # 设置字体类型
# bgcolor        Change ggplot Panel Background Color # 更改画布背景颜色
# background_image        Add Background Image to ggplot2 # 添加背景图片
# facet        Facet a ggplot into Multiple Panels # 设置分面
# ggpar        Graphical parameters # 添加画图参数
# ggparagraph        Draw a Paragraph of Text # 添加文本段落
# ggtexttable        Draw a Textual Table # 添加文本表格
# ggadd        Add Summary Statistics or a Geom onto a ggplot # 添加基本统计结果或其他几何图形
# ggarrange        Arrange Multiple ggplots # 排版多个图形
# annotate_figure        Annotate Arranged Figure # 添加注释信息
# gradient_color        Set Gradient Color # 设置连续型颜色
# xscale        Change Axis Scale: log2, log10 and more # 更改坐标轴的标度
# add_summary        Add Summary Statistics onto a ggplot # 添加基本统计结果
# set_palette        Set Color Palette # 设置画板颜色
```

```
# rotate        Rotate a ggplot Horizontally # 设置图形旋转
# rotate_axis_text        Rotate Axes Text # 旋转坐标轴文本
# stat_stars        Add Stars to a Scatter Plot # 添加散点图星标
# stat_cor        Add Correlation Coefficients with P-values to a Scatter Plot # 添加相关系数
# stat_compare_means        Add Mean Comparison P-values to a ggplot # 添加平均值比较的 P 值
# diff_express      Differential gene expression analysis results # 内置差异分析结果数据集
# ggexport    Export ggplots # 导出图片
# theme_transparent        Create a ggplot with Transparent Background # 设置透明背景
# theme_pubr        Publication ready theme # 设置出版物主题
```

4. 参考来源

[1]: https://www.rdocumentation.org/packages/ggpubr/versions/0.1.4

[2]: https://mp.weixin.qq.com/s/ZKxzKZ4NBTcsJ6vFimxoGA

[3]: http://blog.sciencenet.cn/blog-3334560-1091714.html

[4]: https://mp.weixin.qq.com/s/ZR2sfhVnqxHwDydz7iCGRw