

```

1 library(deplyr)
2 library(psych) # 本例中用到describe函数(描述统计)
3 library(lm.beta) # 用于在回归分析中输出标准化回归系数 $\beta$ 
4 library(ggplot2) # 绘图必备
5 library(gridExtra) # 多图组合
6 library(PerformanceAnalytics) # 本例中用到chart.Correlation函数
7
8 #描述统计
9 data=get(data("anscombe")) # 统计学经典四组数据，具有相同的统计量(M, SD, r)，但本质不同
10 describe(data)
11 chart.Correlation(data) # 绘制变量分布及相关关系图
12
13 # 线性回归：以下4种写法在作用上等价(建模、赋值、报告结果) -----
14 # y1 = b0 + b1*x1
15 model.1=lm(y1 ~ x1, data)
16 summary(lm.beta(model.1))
17 # y2 = b0 + b1*x2
18 model.2=lm(y2 ~ x2, data)
19 model.2 %>% lm.beta() %>% summary()
20 # y3 = b0 + b1*x3
21 summary(lm.beta( (model.3=lm(y3 ~ x3, data)) ))
22 # y4 = b0 + b1*x4
23 (model.4=lm(y4 ~ x4, data)) %>% lm.beta() %>% summary()
24
25 # 散点图 -----
26 p1=ggplot(data=data, aes(x=x1, y=y1)) + geom_point() + geom_smooth(method="lm") + lab
27 p2=ggplot(data=data, aes(x=x2, y=y2)) + geom_point() + geom_smooth(method="lm") + lab
28 p3=ggplot(data=data, aes(x=x3, y=y3)) + geom_point() + geom_smooth(method="lm") + lab
29 p4=ggplot(data=data, aes(x=x4, y=y4)) + geom_point() + geom_smooth(method="lm") + lab
30 grid=grid.arrange(grobs=list(p1, p2, p3, p4), ncol=2, nrow=2)
31 ggsave("Plot.pdf", grid, width=8, height=6) # PDF保存的是矢量图，任意放缩不会模糊，推荐使
32 ggsave("Plot.png", grid, width=8, height=6, dpi=300) # 科研绘图一般要求分辨率至少达到 300

```

factoextra是一个R软件包，可以轻松提取和可视化探索性多变量数据分析的输出，其中包括：主成分分析（PCA），用于通过在不丢失重要信息的情况下减少数据的维度来总结连续（即定量）多变量数据中包含的信息。对应分析（CA），它是适用于分析由两个定性变量（或分类数据）形成的大型列联表的主成分分析的扩展。多重对应分析（MCA），它是将CA改编为包含两个以上分类变量的数据表格。多因素分析（MFA）专用于数据集，其中变量按组（定性和/或定量变量）组织。分层多因素分析（HMFA）：在数据组织为分层结构的情况下，MFA的扩展。