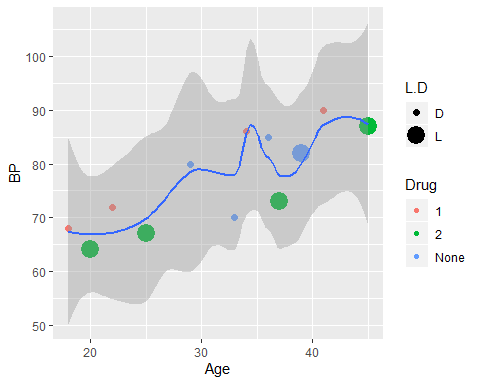
BIOS6621-Homework5-Randy-20191008

Randy

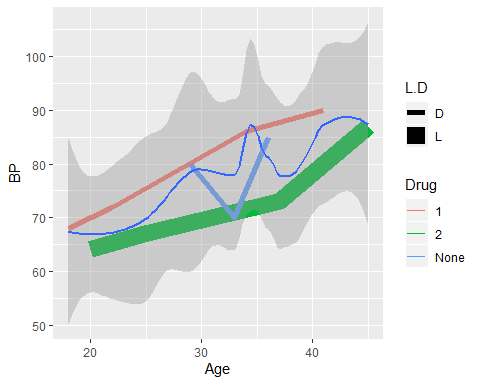
10/8/2019

##### Question1a

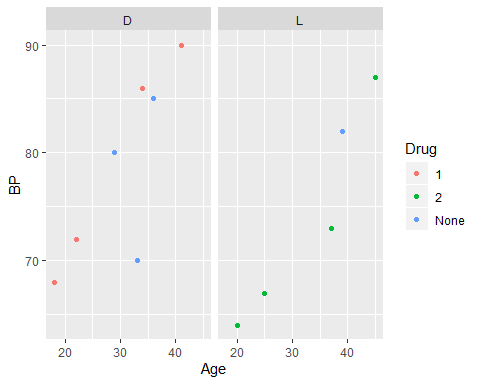
library(ggplot2)  
BIOS6621 <- read.csv("C:/Users/Goodgolden5/Desktop/BIOS6621.txt", sep="")  
View(BIOS6621)  
ggplot(BIOS6621, aes(Age, BP))+  
 geom\_point(aes(color = Drug, size = L.D))+  
 geom\_smooth(span = 0.5)



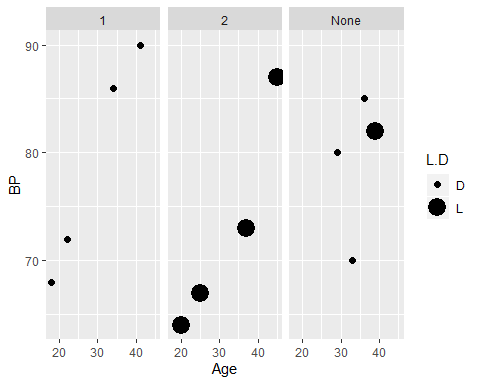
ggplot(BIOS6621, aes(Age, BP))+  
 geom\_path(aes(color = Drug, size = L.D))+  
 geom\_smooth(span = 0.5)



ggplot(BIOS6621, aes(Age, BP))+  
 geom\_point(aes(color = Drug))+  
 facet\_wrap(~L.D)



ggplot(BIOS6621, aes(Age, BP))+  
 geom\_point(aes(size = L.D))+  
 facet\_wrap(~Drug)



1. The blood pressure is positively correlated with age increasing.
2. This kind of correlation is shown in both Treatment Drug1 and Drug2 groups. But not so in Untreated groups.
3. Generally, in the same age group the Drug1 group has a higher blood pressure than Drug2 group patients.
4. The dead patients normally have higher blood pressure than the lived patients in the same age group.
5. More patients with Drug1 and Untreated groups are dead; more people are alive with Drug2 treatment.

##### Question1b

With the increasing of age, the patient get higher blood pressure. Higher blood pressure contributes to the higher mortality.

Just based on the data, we cannot conclude the effects of the drugs on the blood pressures. We can only see that the Drug1 was more often applied to higher blood pressure patients comparing to the same age group of Drug2.

Again we cannot conclude the direct relation between the drug effects and the mortality. Because the higher blood pressure patients faced more risk of high mortality. Without more details and evidence, we cannot conclude, although there are higher alive rates for Drug2 group. We do not know if the results are caused by the drugs or not.

##### Question2

**Figure1** MPG(City) or MPG(Highway) is a numerical variable on the horizontal axis.

Horsepower or EnginSize is a numerical variable on the vertical axis.

**Figure3b** Both line and scatterplot are applied for the graph.

Weight is a continuous numerical variable on the horizontal axis.

Height is a continuous numerical variable on the vertical axis.

Sex is a categorical variable coded by color.

**Figure8** Bar graph was applied to represent the data.

Two different panels to distinguish 2007 and 1997.

Population of Death is a continuous numerical variable on the horizontal axis.

Type of Cancer is a categorical variable on the vertical axis.

Sex is a categorical variable coded by color.

Type of Cases is a categorical variable coded by different shades and bar sizes.

**Figure9** The Line or path graph was applied to represent the data.

Date is a numerical variable on the horizontal axis.

Closing price is a categorical variable on the vertical axis.

Stock is a categorical variable coded by color.

**Figure10** The same as the Figure9.

Date is a numerical variable on the horizontal axis.

Closing price is a categorical variable on the vertical axis.

Stock is a categorical variable presented in different panels.

**Figure11** Product is a categorical variable on the horizontal axis.

Sales value is a continuous variable on the vertical axis.

Sales is a categorical variable presented in different colors and size of the bars.

Season of the year is a categorical variable presented in different panels.

**Figure13** Product Brands is a categorical variable on the horizontal axis.

Sales value is a continuous variable on the vertical axis.

Sales of different types is a categorical variable presented in different colors and size of the bars.

Season of the year is a categorical variable presented in different panels.

**Figure16** Cholesterol level is a continuous variable on the horizontal axis.

The horizontal box plot does not show its vertical axis.

Ages and Genders are categorical variables presented in different panels.

**Figure17** Age is a numerical variable on the horizontal axis.

Weight or Height is a continuous numerical variable on the vertical axis.

Sex is a categorical variable presented in different colors and type of lines.

The attributions of Weight or Height is a categorical variable presented in different panels.

**Figure18** It is a scatter plot matrix pairwised with different factors.

The Sepal Width, Sepal Length, Petal Width, and Petal Length are the continuous numerical variables pairwised as the horizontal and vertical axes.

Iris Species is the categorical variable coded with different colors and types.