ggplot-Basics

Five named plots

1. scatter: geom_point()
2. line : geom_line()
3. bar count: geom_bar()
4. column: geom_col()
5. frequency: geom_histogram()

Data

```
girth <- c(40,45,53,55,63,65,67)
height <- c(56.2,33.3,61.2,38.4,47.4,30.0,40.4)
age <- c(15,18,28,40,55,65,76)
sex <- c('M','F','M','F','M','M','F')
df <- data.frame(girth,height,age,sex)
df
```

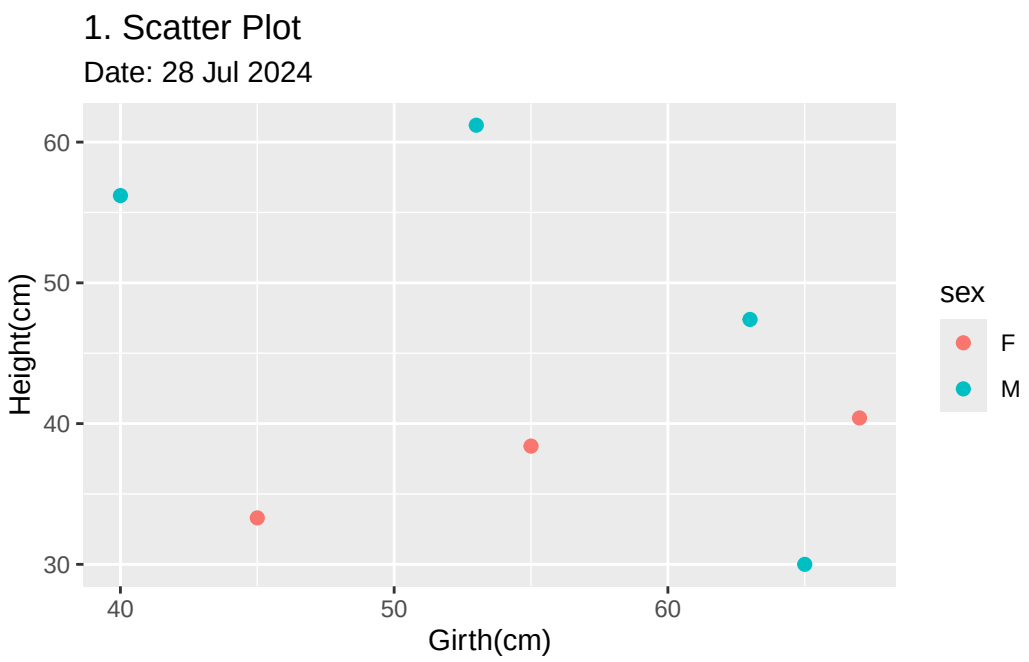
```
  girth height age sex
1    40   56.2  15   M
2    45   33.3  18   F
3    53   61.2  28   M
4    55   38.4  40   F
5    63   47.4  55   M
6    65   30.0  65   M
7    67   40.4  76   F
```

---

1. Scatter Plot

- geom_point = Scatter Plot
- `aes(color=sex)`
- `lab(title,subtitle,xlabel,ylabel)`

```
# 1. Scatter
library(tidyverse)
ggplot(df, aes(x=girth, y=height, color=sex))+
  geom_point(size=2)+
  labs(
    title="1. Scatter Plot",
    subtitle="Date: 28 Jul 2024",
    x = "Girth(cm)",
    y = "Height(cm)"
  )
```



---

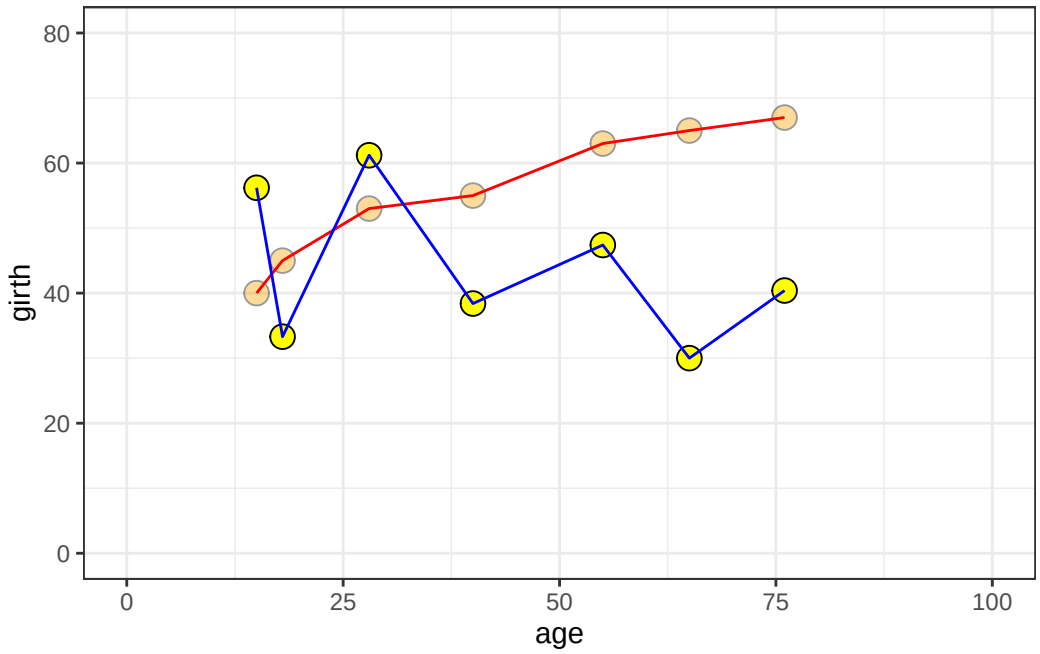
2. ScatterLines

- `xlims, ylims`
- `marker (attributes from geom_point())`

```
# 2. ScatterLines
ggplot(df)+
  xlim(0,100)+
  ylim(0,80)+

  geom_point(x=age,y=girth,shape=21, size=4,fill="orange",alpha=0.4)+
  geom_line(aes(x=age,y=girth,color="girth"),show.legend=FALSE)+
  geom_point(x=age,y=height,shape=21, size=4,fill="yellow")+
  geom_line(aes(x=age,y=height,color="height"),show.legend=FALSE)+
  scale_color_manual(values=c("red","blue"))+

  theme_bw()
```



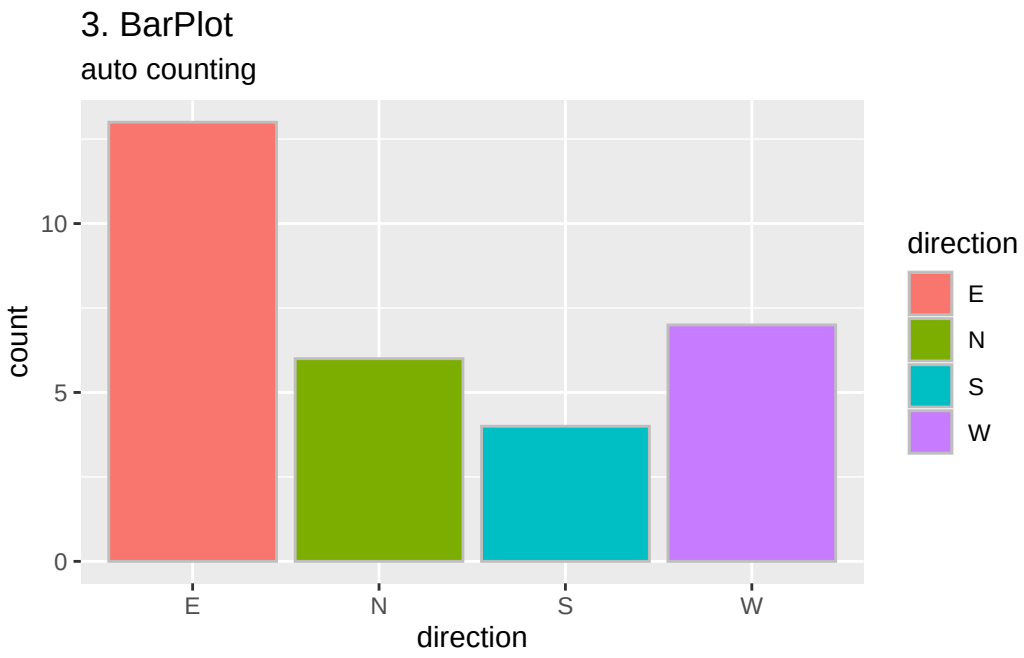
3. BarPlot

- counts categories
- runif()

```
# barplot
direction = c('N','S','E','W','W','E','W','E','E','E','E','S','S','N','N','W','E','E','N','N
speed = rnorm(30, 5, 0.2)
dfWind <- data.frame(direction,speed)
str(dfWind)
```

```
'data.frame':   30 obs. of  2 variables:
 $ direction: chr  "N" "S" "E" "W" ...
 $ speed    : num  5.14 4.97 5.18 4.74 4.9 ...
```

```
ggplot(dfWind,aes(x=direction, fill=direction))+
  geom_bar(color="gray")+
  labs(
    title="3. BarPlot",
    subtitle="auto counting"
  )
```



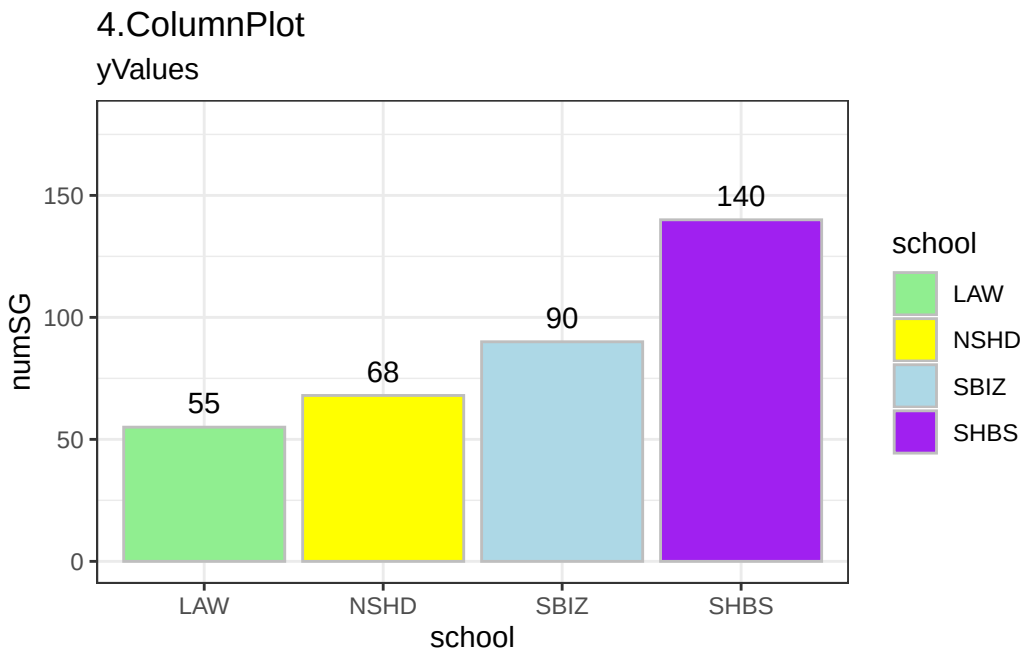
4. ColumnPlot

- annotate values : geom_text()

```
school <- c("NSHD","SHBS","SBIZ","LAW")
numSG <- c(68,140,90,55)
df2 <- data.frame(school,numSG)
df2
```

```
  school numSG
1   NSHD    68
2   SHBS   140
3   SBIZ    90
4    LAW    55
```

```
ggplot(df2, aes(x=school, y=numSG, fill=school))+
  labs(
    title="4.ColumnPlot",
   subtitle="yValues")+
  geom_col(color="gray")+
  scale_fill_manual(values=c("lightgreen","yellow","lightblue","purple"))+
  geom_text(aes(label=numSG),nudge_y=10)+
  ylim(0,180)+
  theme_bw()
```



5. Histogram

- geom_histogram
- bin

```
str(dfWind)
```

```
'data.frame':   30 obs. of  2 variables:
 $ direction: chr  "N" "S" "E" "W" ...
 $ speed    : num  5.14 4.97 5.18 4.74 4.9 ...
```

```
ggplot(dfWind, aes(x=speed))+
  geom_histogram(bins = 8, fill="lightgray",color="black")+
  labs(
    title="5. Histogram",
    subtitle = "bins=8",
    xlab = "Wind Speed",
    ylab = "Frequency"
  )
```