NoteDay3

Diamond

2020/7/8

chapter 3 Visuation:ggplot2

install first package, ggplot2, by installing tidyverse

Test work place simply.

```
getwd()
```

```
## [1] "D:/zju/ /  /  /DataScienceAndApplications"
```

1.Load data

```
#National Parks in California
ca <- read_csv("data/ca.csv")
```

```
## Parsed with column specification:
## cols(
##   region = col_character(),
##   state = col_character(),
##   code = col_character(),
##   park_name = col_character(),
##   type = col_character(),
##   visitors = col_double(),
##   year = col_double()
## )
```

```
#Acadia National Park
acadia <- read_csv("data/acadia.csv")
```

```
## Parsed with column specification:
## cols(
##   region = col_character(),
##   state = col_character(),
##   code = col_character(),
##   park_name = col_character(),
##   type = col_character(),
##   visitors = col_double(),
##   year = col_double()
## )
```

```
#Southeast US National Parks
se <- read_csv("data/se.csv")
```

```
## Parsed with column specification:
## cols(
##   region = col_character(),
##   state = col_character(),
##   code = col_character(),
##   park_name = col_character(),
##   type = col_character(),
##   visitors = col_double(),
##   year = col_double()
## )
```

```
#2016 Visitation for all Pacific West National Parks
visit_16 <- read_csv("data/visit_16.csv")
```

```
## Parsed with column specification:
## cols(
##   region = col_character(),
##   state = col_character(),
##   code = col_character(),
##   park_name = col_character(),
##   type = col_character(),
##   visitors = col_double(),
##   year = col_double()
## )
```

```
#All Nationally designated sites in Massachusetts
mass <- read_csv("data/mass.csv")
```

```
## Parsed with column specification:
## cols(
##   region = col_character(),
##   state = col_character(),
##   code = col_character(),
##   park_name = col_character(),
##   type = col_character(),
##   visitors = col_double(),
##   year = col_double()
## )
```

A Grammar of Graphics!

ggplot(data = <DATA>) + <GEOM_FUNCTION>(mapping = aes(<MAPPINGS>),stat = <STAT>,position = <POSITION>) + <COORDINATE_FUNCTION> + <FACET_FUNCTION>

You can uniquely describe any plot as a combination of these 7 parameters.

```
head(ca)
```

```
## # A tibble: 6 x 7
##   region state code  park_name                     type          visitors  year
##   <chr>  <chr> <chr> <chr>                         <chr>            <dbl> <dbl>
## 1 PW     CA    CHIS  Channel Islands National Park National Park     1200  1963
## 2 PW     CA    CHIS  Channel Islands National Park National Park     1500  1964
## 3 PW     CA    CHIS  Channel Islands National Park National Park     1600  1965
## 4 PW     CA    CHIS  Channel Islands National Park National Park      300  1966
## 5 PW     CA    CHIS  Channel Islands National Park National Park    15700  1967
## 6 PW     CA    CHIS  Channel Islands National Park National Park    31000  1968
```

```
#view(ca) other worksheet will come out
```

Hello,Li YinHe