2017년 1학기 기말고사

양승훈

2017 6 20

## 인사말

한 학기 동안 엑셀과 R을 배우느라 고생이 많았습니다. 이번 기말고사를 통해 자신의 늘어난 실력과 자신감을 확인할 수 있길 바랍니다.

## 시험 안내

이번 시험은 아래의 1) 기본형 10문제(각 5점)와 2) 심화형 5문제(각 10점)으로 구성되어 있습니다. 주어진 시간은 16:00~17:00(1시간)입니다. 답안은 R코드로 "정민주\_기말고사.r" 방식으로 저장해서 제출하면 됩니다. 답안 제출을 위해서는 **tidyverse** 패키지를 설치하고 로드해야 합니다. 단답식 1-3번을 제외한 모든 답안은 그대로 실행이 가능해야 합니다. 그렇기 때문에 반드시 코드 형식으로 답안을 작성할 것. 모든 답은 부분 점수가 있으니 포기하지 마세요!

### 답안 작성 예시

### 1 번 문제

head(mpg)

###2 번 문제

library(tidyverse)

## 기본형(단답식) 문제

1. R에서 텍스트 파일을 로드하는 명령어는? (대소문자 주의)
2. R에서 내장 데이터를 로드하는 명령어는?
3. R에서 점을 그릴 수 있는 기본 함수(ggplot 아님)와 인자를 입력하시오. **cars**의 dist 변수를 x로, speed 변수를 y로 할 것. (2가지 방법 모두 입력할 것)
4. 아래 **mpg** 데이터 모든 변수의 기초 통계값을 확인하는 명령어를 입력하시오. () (만약 **cars** 를 확인한다면 명령어(**cars**) 방식으로 입력하시오. tidyverse 패키지를 로드하여야 mpg 데이터 사용 가능.)

library(tidyverse)

## Loading tidyverse: ggplot2  
## Loading tidyverse: tibble  
## Loading tidyverse: tidyr  
## Loading tidyverse: readr  
## Loading tidyverse: purrr  
## Loading tidyverse: dplyr

## Conflicts with tidy packages ----------------------------------------------

## filter(): dplyr, stats  
## lag(): dplyr, stats

head(mpg)

## # A tibble: 6 x 11  
## manufacturer model displ year cyl trans drv cty hwy fl  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int> <chr>  
## 1 audi a4 1.8 1999 4 auto(l5) f 18 29 p  
## 2 audi a4 1.8 1999 4 manual(m5) f 21 29 p  
## 3 audi a4 2.0 2008 4 manual(m6) f 20 31 p  
## 4 audi a4 2.0 2008 4 auto(av) f 21 30 p  
## 5 audi a4 2.8 1999 6 auto(l5) f 16 26 p  
## 6 audi a4 2.8 1999 6 manual(m5) f 18 26 p  
## # ... with 1 more variables: class <chr>

1. **mpg** 데이터 중 cyl 변수 기준으로 cty 변수의 합을 표로 구현하는 명령어와 인자를 입력해 실행 가능하게 하시오.
2. **mpg** 데이터 class 변수의 값에 따른 개수를 표로 구현하는 명령어와 인자를 입력해 실행 가능하게 하시오.
3. **mpg** 데이터의 hwy 변수를 자동차 제작사에 따라 박스 플롯으로 그려내는 명령어와 인자를 입력해 실행 가능하게 하시오. (ggplot이 아닌 기본 명령어로 할 것)
4. 아래의 **iris** 데이터 중 Species가 'setosa'만 추려내는 명령어는? (아무 방법이나 상관 없음)

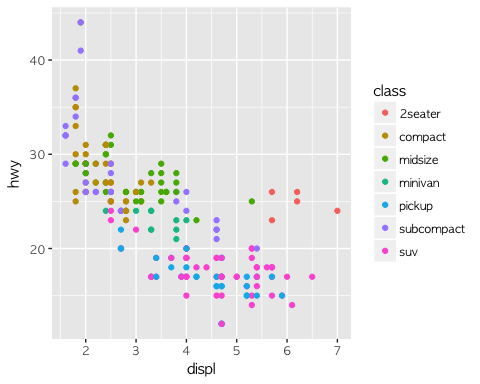
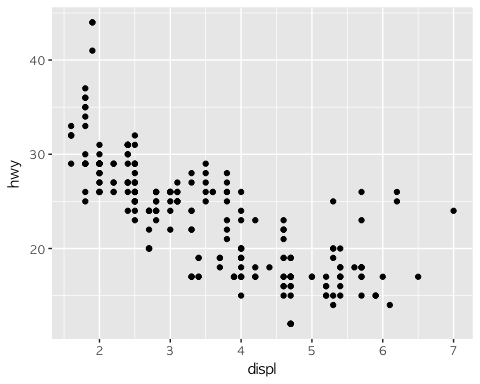
data(iris)  
head(iris)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 1 5.1 3.5 1.4 0.2 setosa  
## 2 4.9 3.0 1.4 0.2 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 5 5.0 3.6 1.4 0.2 setosa  
## 6 5.4 3.9 1.7 0.4 setosa

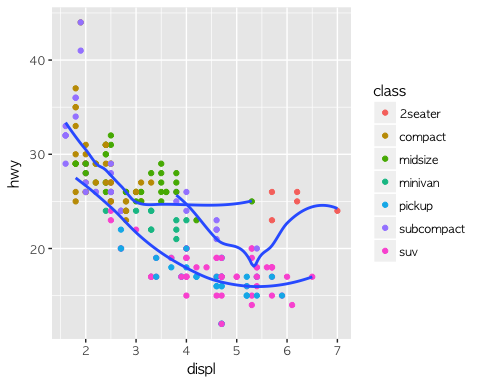
1. **iris** 데이터 각 변수의 '속성'을 파악하는 명령어를 입력하시오.
2. **iris** 데이터 마지막 6행을 보여주는 명령어를 입력하시오.

## 심화 문제 (ggplot 그래프 그리기)

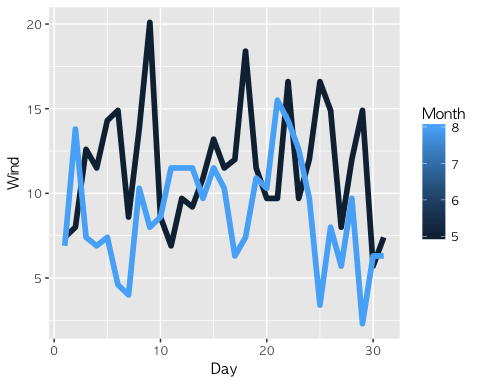
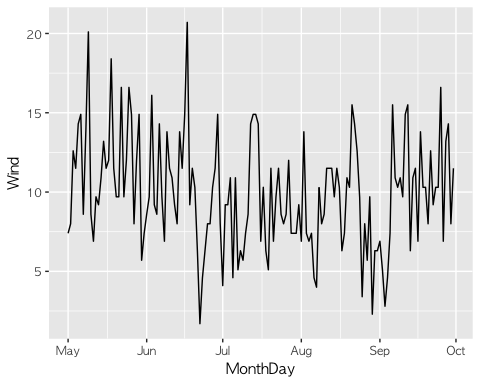
11-13. 아래 3개의 그래프와 같은 모양이 나오도록 **mpg** 데이터를 활용해 그래프를 만드는 ggplot 코드를 각각 작성하시오. (각 10점, 힌트 : 참조 변수는 displ, hwy, class)



## `geom\_smooth()` using method = 'loess'



14-15 아래 2개의 그래프와 같은 모양이 나오도록 **airquality** 데이터를 활용해 날짜별 바람의 정도를 파악할 수 있는 선 그래프를 그리는 ggplot 코드를 각각 작성하시오. (각 10점, 힌트 : MonthDay 변수를 만들고 나서 작업하시오. 참조 변수는 Month, Day, Wind. 두 번째 그래프는 5월과 8월만 포함하여 Day를 변수로 그림.)



## 보너스 문제 (이 문제를 풀어서 함께 제출하면 중간고사 점수에 20점 추가)

16-17. 아래의 그래프와 같은 모양이 나오도록 함께 첨부한 **seoul\_dust.csv** 데이터를 활용하여 서울의 미세먼지 정도를 측정할 수 있는 두 개의 그래프를 그리는 ggplot 코드를 각각 작성하시오. 첫 번째 그래프는 강남구, 마포구, 중랑구 장소만 추려서 선그래프를 그린 것이고, 두 번째 그래프는 tinydust가 100초과인 것만 추려서 그린 선그래프임. (각 10점. seoul\_dust.csv는 seoul 데이터 프레임에 저장할 것.)

## Parsed with column specification:  
## cols(  
## date = col\_integer(),  
## place = col\_character(),  
## CO2 = col\_double(),  
## O3 = col\_double(),  
## CO = col\_double(),  
## SO2 = col\_double(),  
## TinyDust = col\_integer(),  
## MicroDust = col\_integer(),  
## Year = col\_integer(),  
## Month = col\_character(),  
## Day = col\_character(),  
## date2 = col\_date(format = "")  
## )

