

02 환경 설정 및 변수 및 자료형

- R 스튜디오의 두 종류의 환경 설정
 - 글로벌 옵션(Global Options)
 - 프로젝트 옵션(Project Options)

1-1 Soft-wrap 자동 줄바꿈 옵션 설정하기 ¶

- (1) [Tools -> Global Options] 선택
- (2) [Code] 탭 선택
- (3) [Soft - wrap R source File 항목 체크]

1-2 인코딩 방식 설정하기

- 스크립트 내용 중 한글 부분이 깨져 나오는 경우가 있다. 이는 대부분 인코딩 설정이 달라 생기는 문제
- (1) [Tools -> Project Options] 선택
- (2) [Code Editing] 탭 선택
- (3) 'Text encoding'을 UTF-8로 변경

In [48]:

```
sessionInfo()
```

R version 3.5.1 (2018-07-02)

Platform: x86_64-w64-mingw32/x64 (64-bit)

Running under: Windows 10 x64 (build 17134)

Matrix products: default

locale:

[1] LC_COLLATE=Korean_Korea.949 LC_CTYPE=Korean_Korea.949

[3] LC_MONETARY=Korean_Korea.949 LC_NUMERIC=C

[5] LC_TIME=Korean_Korea.949

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] ggplot2_3.0.0 RevUtils_11.0.1 RevUtilsMath_11.0.0

loaded via a namespace (and not attached):

[1] Rcpp_0.12.18 compiler_3.5.1 pillar_1.3.0 plyr_1.8.4

[5] bindr_0.1.1 base64enc_0.1-3 tools_3.5.1 digest_0.6.15

[9] uuid_0.1-2 jsonlite_1.5 evaluate_0.11 tibble_1.4.2

[13] gtable_0.2.0 pkgconfig_2.0.1 rlang_0.2.1 IRdisplay_0.5.0

[17] IRkernel_0.8.11 bindrcpp_0.2.2 withr_2.1.2 repr_0.15.0

[21] stringr_1.3.1 dplyr_0.7.6 grid_3.5.1 tidyselect_0.2.4

[25] glue_1.3.0 R6_2.2.2 pbdZMQ_0.3-3 purrr_0.2.5

[29] magrittr_1.5 scales_0.5.0 htmltools_0.3.6 assertthat_0.2.0

[33] colorspace_1.3-2 labeling_0.3 stringi_1.1.7 lazyeval_0.2.1

[37] munsell_0.5.0 crayon_1.3.4

2. 변수, 함수, 패키지

- 변수 : 변하는 수.
- 상수 : 고정된 값을 가지고 있음.
- 변수 만들기 : 변수명 <- 값 또는 식

In [1]:

```
a <- 1  
a
```

1

In [2]:

```
d <- 3.5  
d
```

3.5

In [3]:

```
a + d
```

4.5

여러개의 값을 갖는 변수 만들어 보기

- c() 함수를 이용하여 여러개의 값을 만든다.
- c는 합치는 의미 : Combine의 머리글자

In [4]:

```
var1 <- c(1,2,3,4,5)  
var1
```

1 2 3 4 5

In [7]:

```
var2 <- c(1:5)  
var2
```

1 2 3 4 5

In [8]:

```
var1 + var2
```

2 4 6 8 10

(실습) 만약 var1과 var2의 길이가 다르다면 어떻게 될까?

seq()를 활용해 보기

- seq(시작, 끝, by=증가)

In [10]:

```
var4 <- seq(1, 15, by=3) # 1~15까지 3간격으로 값을 증가시키고 15를 넘어가면 종료  
var4
```

1 4 7 10 13

In [11]:

```
var4 + 5 # 각각의 원소에 5를 더해준다.
```

6 9 12 15 18

In [12]:

```
### 문자로된 변수 만들기  
str1 <- "Hello World"  
str1
```

'Hello World'

In [13]:

```
str2 <- c("Hello", "World", "Today")  
str2
```

'Hello' 'World' 'Today'

(해보기) 문자로된 변수를 +1 하면 어떻게 될까?

In []:

```
str1 + 3
```

2-2 함수

- 데이터 분석은 함수로 시작해서 함수로 끝난다.
 - 함수를 얼마나 잘 이용할 수 있는가? 이에 대한 조작방법을 익히는 것.

In [18]:

```
### 변수 만들기  
x <- c(1,2,3,4,5)  
x
```

1 2 3 4 5

In [19]:

```
mean(x)
```

3

In [20]:

```
print(min(x))  
print(max(x))
```

[1] 1
[1] 5

In [22]:

```
print(median(x))
```

[1] 3

문자열 처리 함수

In [23]:

```
str5 = c("Hello", "World", "toto")  
paste(str5, collapse = ",")
```

'Hello,World,toto'

In [24]:

```
paste(str5, collapse = " ")
```

'Hello World toto'

결과를 저장하여 변수 만들기

In [25]:

```
result = paste(str5, collapse = " ")  
result
```

'Hello World toto'

2-3 패키지 이해하기

- 패키지는 함수들을 모아둔 상자들이 여러개 들어 있는 상자 꾸러미.
- 하나의 패키지에 다양한 함수가 들어 있다.

패키지 설치 후, 로드 그리고 사용하기

- ggplot2 패키지 설치 - 패키지는 설치 시 다운로드를 해야 하므로 접속이 되어 있어야.
 - `install.packages("ggplot2")`

In [27]:

```
install.packages("ggplot2")
```

Warning message:

"unable to access index for repository <http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/3.5>: (<http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/3.5>:)
URL 'http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/3.5/PACKAGES'를 열 수 없습니다"

package 'ggplot2' successfully unpacked and MD5 sums checked

The downloaded binary packages are in

C:\Users\WWITHJS\AppData\Local\Temp\WRtmp\zp6DEW\downloaded_packages

In [28]:

```
### 불러오기  
library(ggplot2)
```

In [29]:

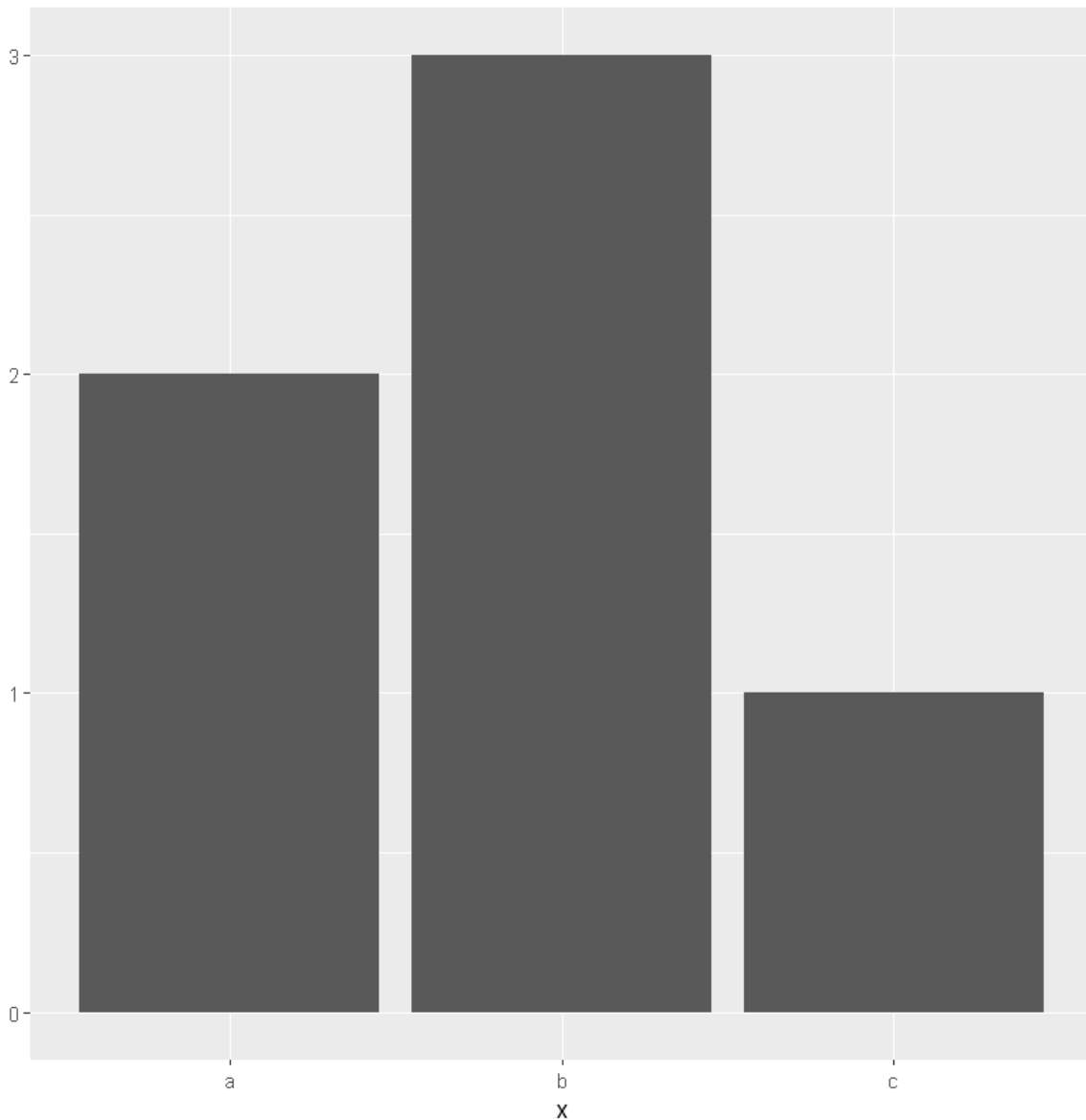
```
### 확인  
search()
```

```
'.GlobalEnv' 'package:ggplot2' 'jupyter:irkernel' 'package:RevoUtils' 'package:stats'  
'package:graphics' 'package:grDevices' 'package:utils' 'package:datasets'  
'package:RevoUtilsMath' 'package:methods' 'Autoloads' 'package:base'
```

In [30]:

```
### 함수 사용
```

```
x <- c('a', 'a', 'b', 'b', 'b', 'c')  
qplot(x)
```



ggplot2에 내장된 mpg 데이터 셋 이용해 보기

In [34]:

```
dim(mpg)
```

```
234 11
```

In [35]:

```
is(mpg)
```

```
'tbl_df' 'tbl' 'data.frame' 'list' 'oldClass' 'vector'
```

In [36]:

```
names(mpg)
```

```
'manufacturer' 'model' 'displ' 'year' 'cyl' 'trans' 'drv' 'cty' 'hwy' 'fl' 'class'
```

In [39]:

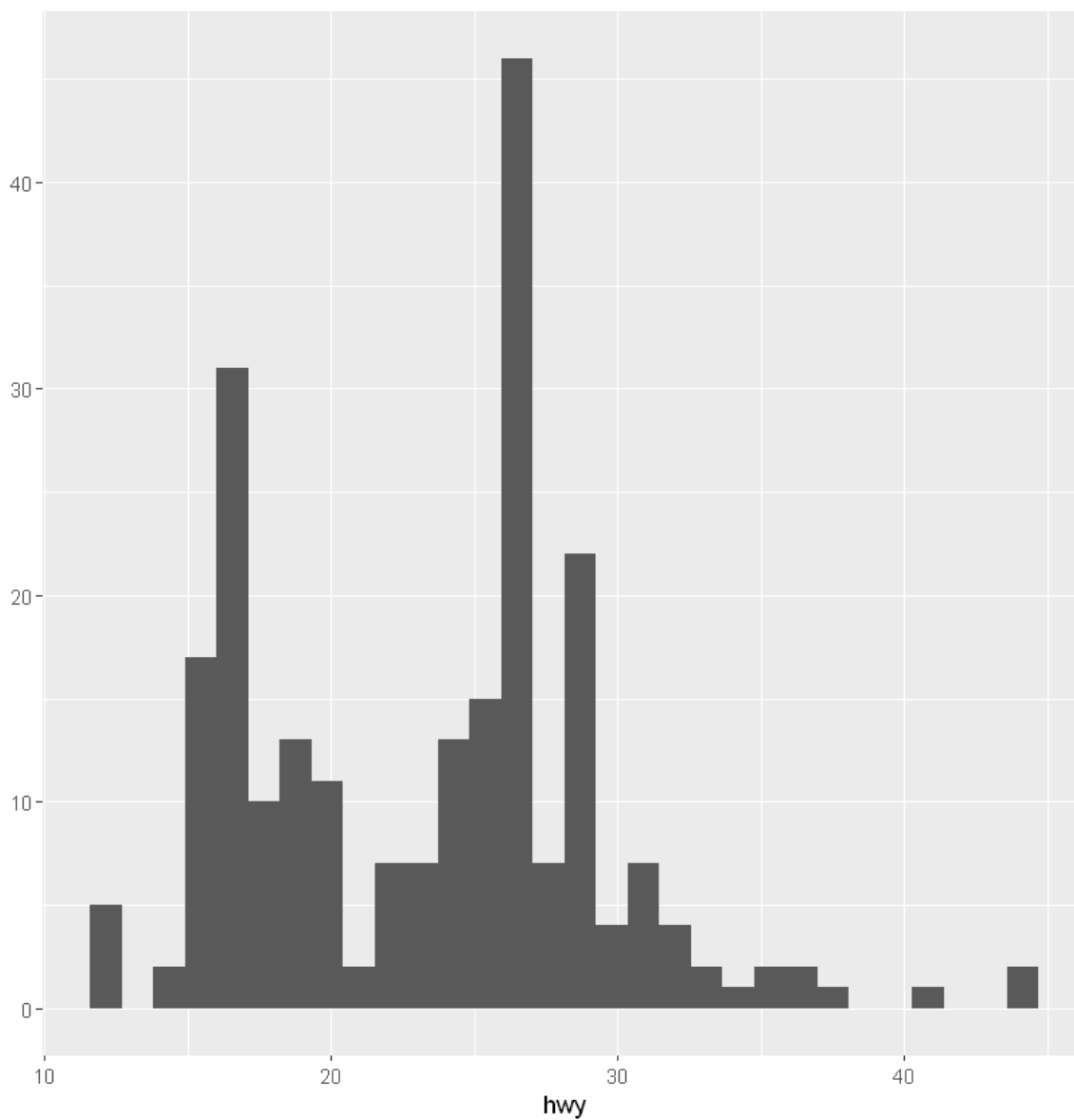
```
colnames(mpg)
```

```
'manufacturer' 'model' 'displ' 'year' 'cyl' 'trans' 'drv' 'cty' 'hwy' 'fl' 'class'
```

In [33]:

```
library(ggplot2)  
qplot(data=mpg, x=hwy)
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



In [41]:

```
mpg
```

manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p	compact
audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p	compact
audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact
audi	a4 quattro	2.8	1999	6	auto(l5)	4	15	25	p	compact
audi	a4 quattro	2.8	1999	6	manual(m5)	4	17	25	p	compact
audi	a4 quattro	3.1	2008	6	auto(s6)	4	17	25	p	compact
audi	a4 quattro	3.1	2008	6	manual(m6)	4	15	25	p	compact
audi	a6 quattro	2.8	1999	6	auto(l5)	4	15	24	p	midsize
audi	a6 quattro	3.1	2008	6	auto(s6)	4	17	25	p	midsize
audi	a6 quattro	4.2	2008	8	auto(s6)	4	16	23	p	midsize
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14	20	r	suv
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	11	15	e	suv
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14	20	r	suv
chevrolet	c1500 suburban 2wd	5.7	1999	8	auto(l4)	r	13	17	r	suv
chevrolet	c1500 suburban 2wd	6.0	2008	8	auto(l4)	r	12	17	r	suv
chevrolet	corvette	5.7	1999	8	manual(m6)	r	16	26	p	2seater
chevrolet	corvette	5.7	1999	8	auto(l4)	r	15	23	p	2seater
chevrolet	corvette	6.2	2008	8	manual(m6)	r	16	26	p	2seater
chevrolet	corvette	6.2	2008	8	auto(s6)	r	15	25	p	2seater
chevrolet	corvette	7.0	2008	8	manual(m6)	r	15	24	p	2seater
chevrolet	k1500 tahoe 4wd	5.3	2008	8	auto(l4)	4	14	19	r	suv
chevrolet	k1500 tahoe 4wd	5.3	2008	8	auto(l4)	4	11	14	e	suv
...
toyota	toyota tacoma 4wd	3.4	1999	6	auto(l4)	4	15	19	r	pickup
toyota	toyota tacoma 4wd	4.0	2008	6	manual(m6)	4	15	18	r	pickup

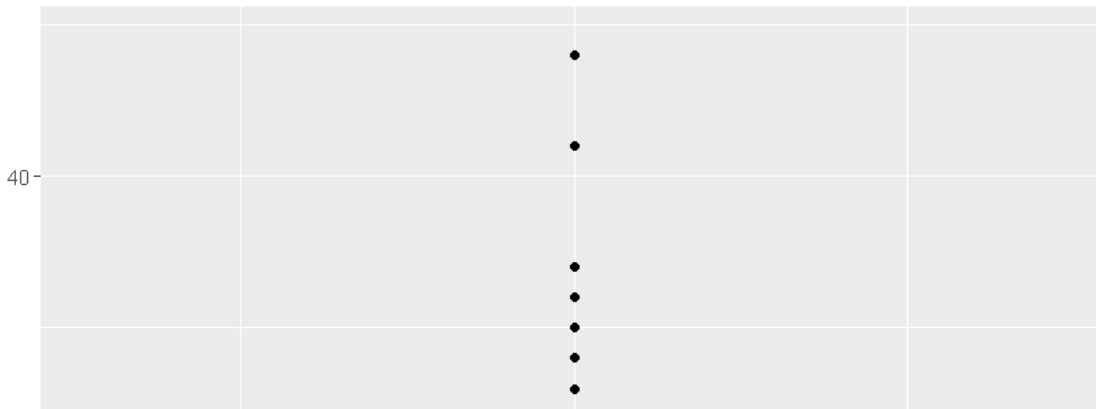
manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
toyota	toyota tacoma 4wd	4.0	2008	6	auto(l5)	4	16	20	r	pickup
volkswagen	gti	2.0	1999	4	manual(m5)	f	21	29	r	compact
volkswagen	gti	2.0	1999	4	auto(l4)	f	19	26	r	compact
volkswagen	gti	2.0	2008	4	manual(m6)	f	21	29	p	compact
volkswagen	gti	2.0	2008	4	auto(s6)	f	22	29	p	compact
volkswagen	gti	2.8	1999	6	manual(m5)	f	17	24	r	compact
volkswagen	jetta	1.9	1999	4	manual(m5)	f	33	44	d	compact
volkswagen	jetta	2.0	1999	4	manual(m5)	f	21	29	r	compact
volkswagen	jetta	2.0	1999	4	auto(l4)	f	19	26	r	compact
volkswagen	jetta	2.0	2008	4	auto(s6)	f	22	29	p	compact
volkswagen	jetta	2.0	2008	4	manual(m6)	f	21	29	p	compact
volkswagen	jetta	2.5	2008	5	auto(s6)	f	21	29	r	compact
volkswagen	jetta	2.5	2008	5	manual(m5)	f	21	29	r	compact
volkswagen	jetta	2.8	1999	6	auto(l4)	f	16	23	r	compact
volkswagen	jetta	2.8	1999	6	manual(m5)	f	17	24	r	compact
volkswagen	new beetle	1.9	1999	4	manual(m5)	f	35	44	d	subcompact
volkswagen	new beetle	1.9	1999	4	auto(l4)	f	29	41	d	subcompact
volkswagen	new beetle	2.0	1999	4	manual(m5)	f	21	29	r	subcompact
volkswagen	new beetle	2.0	1999	4	auto(l4)	f	19	26	r	subcompact
volkswagen	new beetle	2.5	2008	5	manual(m5)	f	20	28	r	subcompact
volkswagen	new beetle	2.5	2008	5	auto(s6)	f	20	29	r	subcompact
volkswagen	passat	1.8	1999	4	manual(m5)	f	21	29	p	midsize
volkswagen	passat	1.8	1999	4	auto(l5)	f	18	29	p	midsize
volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	p	midsize
volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	p	midsize
volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	p	midsize
volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	p	midsize
volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	p	midsize

In [45]:

```
help(mpg) # 또는 ?qplot
```

In [47]:

```
# 연료와 city miles per gallon
qplot(data = mpg, x = cty)
# mpg(연료)와 drv, hwy
qplot(data=mpg, x=drv, y=hwy)
# mpg(연료)와 drv, hwy
# f = front-wheel drive, r = rear wheel drive, 4 = 4wd
# highway miles per gallon
qplot(data=mpg, x=drv, y=hwy, geom="boxplot")
qplot(data=mpg, x=drv, y=hwy, geom="boxplot", color=drv)
```



(책) 혼자서 해보기

- (1) 시험점수 만들고 출력하기(p77)
- (2) 전체 평균 구하기 - 학생들의 전체 평균 점수를 구해보기
- (3) 전체 평균 만들고 출력하기 - 전체 변수를 담고 있는 새 변수를 만들기.

In []: