
[목차]

29장 → 28장

30장 → 29장

31장 → 30장

[p.103: 2행]

하나의 열만 반환하므로 벡터를 반환

→

이 경우에는 하나의 열의 값을 하나의 (팩터형) 벡터의 형태로 반환한다.

[p.117]

```
theUrl <-http://www.jaredlander.com/data/TomatoFirst.csv
```

```
->
```

```
theUrl <- "http://www.jaredlander.com/data/TomatoFirst.csv"
```

[p.153 : 그림 7.9 캡션]

컷은 수평으로, 투명도는 수평으로 정렬돼 있는 것을 주목한다.

→

컷은 수직으로, 투명도는 수평으로 정렬돼 있는 것을 주목한다.

[p.180 : 코드 아래에서 8행]

순화하면서 → 순회하면서

[p.188 : 코드 아래에서 9행]

행의 길이를 더한다. → 행의 길이를 더한다.

[p.189 : 11.2 집계 6행]

포물러는 틸드(~)를 사이를 두고 → 포물러는 틸트(~)를 사이에 두고

[p.209 : 1행]

aggregate함수와는 각 열에 대해 여러 가지 수치들을 한꺼번에 계산할 수 있다.

→

aggregate함수와는 달리 열에 대해 여러 가지 수치들을 한꺼번에 계산할 수 있다.

[p.211 : 아래에서 5행]

먼저 plyr 패키지를 먼저 로딩하고 → 먼저 plyr 패키지를 로딩하고

[p.225 : 아래]

복잡한 필터링 조건은 → 복잡한 **필터링** 조건은

[p.240 : 첫 번째 명령어]

topN 함수를 두 번째 인자를 3으로 지정해 적용

→

topN **함수**의 두 번째 인자를 3으로 지정해 적용

[p.245 : 5행]

어떤 함수의 행동이 입력된 → 어떤 함수의 **행동**이 입력된

[p.256 : 결론 1행]

리스트 순회하는 문제는 → **리스트**를 순회하는 문제는

[p.259 : 3행]

파일들의 리스트는 dir 함수로 확인할 수 있고, → 파일들의 리스트는 dir 함수**로** 확인할 수 있고,

[p.259 : 두 번째 명령어]

데이터를 할당한 좋은 이름을 만든다. → 데이터를 할당**할** 좋은 이름을 만든다.

[p.264 : 1행]

모든 데이터 프레임이 → 모든 데이터 프레임**들이**

[p.264 : 2행]

하면서 모든 요소들을 하나로 조인할 수 있다 → 하면서 모든 요소들을 하나로 조인할 수 **있다**

[p.266 : 14.3 3행]

R에서는 흔히 있는 일지만 → R에서는 흔히 있는 **일이지만**

[p.268 : 코드 마지막행]

```
+ scale_y_continuous(labels = multiple_format(extra=dollar, multiple="B"))
->
+   scale_y_continuous(labels=multiple_format(extra=dollar,
+                                             multiple="B"))
```

[p.270 : 마지막행]

이 장은 11장에서 → 이 장은 **11장**의

[p.272 : 코드 4단락]

```
> # 행에서 데이터를 지정하는 방식으로 새로운 티블 만들.  
> trophies2 <- tibble(  
->  
> # 행 방향으로 만들기 쉽도록 트리블을 이용한 또다른 티블을 만들  
> trophies2 <- tribble(  

```

[p.273 : 3행]

다이아몬드 → 다이아몬드

[p.274 : 아래 단락 코드]

```
('color' = "Color")  
→  
(color = 'Color')
```

[p.275 : 5행]

책의 레이아웃이 → 책의 레이아웃이

[p.276 : 아래에서 3행]

양쪽 데이터 → 양쪽 테이블

[p.277 : 마지막 단락]

```
> semi_join(diamondColors, diamonds, by = c('color' = 'Color'))  
->  
> semi_join(diamondColors, diamonds, by = c('Color' = 'color'))
```

[p.278 : 두번째 코드 단락]

```
> anti_join(diamondColors, diamonds, by = c('color' = 'Color'))  
->  
> anti_join(diamondColors, diamonds, by = c('Color' = 'color'))
```

[p.287 : 2행]

정확한 위치로 쉽표의 놓아야 → 정확한 위치에 쉽표를 놓아야

[p.287 : 아래에서 5행]

벡터를 조합해 → 벡터를 조합해

[p.293: 마지막 행]

구분하기 때문에 → 구분하기 **때**문에

[p.294 : 아래에서 3행]

어떤 것을 달과 날짜까지 → 어떤 **것은** 달과 날짜까지

[p.295 : 2행]

위키피디아의 → 위키피**디**아의

[p.295 : 아래에서 3행]

사용하는 것이 아니면 → 사용하는 것이 **아**니라

[p.295 : 아래에서 2행]

이러 그룹핑은 → 이**런** 그룹핑은

[p.301 : 두 번째 코드]

텍스틀 → 텍**스트**를

[p.309 : 아래에서 8행]

난수를 생성하는 것은 아니라, → 난수를 생성하는 **것**이 아니라,

[p.309 : 아래에서 6행]

성공할 회수를 → 성공할 **횟**수를

[p.309 : 아래에서 5행]

size=10(시행 '회수') → size=10(시행 '**횟**수')

[p.311 : 코드 1행]

10,000회 관찰한 것은 → 10,000회 관찰한 **것**을

[p.315 : 3행]

포아송 분포는 정규분포에 가까와진다. → 포아송 분포는 정규분포에 가까**워**진다.

[p.324]

중앙값은 숫자들을 숫서대로 → 중앙값은 숫자들을 **순**서대로

[p.327 : 2행]

여기서는 ggplot2 패키지를 기초를 둔 → 여기서는 ggplot2 패키지**에** 기초를 둔

[p.328 : 7행]

추가 기능을 이해 scales 패키지 로딩 → 추가 기능을 **위해** scales 패키지 로딩

[p.333 : 3행]

두 사이에 인과관계가 → 두 **변수** 사이에 인과관계가

[p.333 : 3행]

웹 톱 552번에 → **웹툰** 552번에

[p.333 : 주석 1]

랜덤 번호의 웹 톱이다. → 랜덤 번호의 **웹툰**이다.

[p.336 : 7행]

분포에서 외곽으로 떨어진 곳에 있어서 우리는 모집단의 평균이 \$2.50가 같지

→

분포에서 외곽으로 떨어진 곳에 있어서 우리는 모집단의 평균이 \$2.50**와** 같지

[p.339 : 아래에서 3행]

검정이 실패했으므로 → 검정이 **실패**했으므로

[p.340 : (코드) 마지막행]

`sqrt(NROW(tip)),`

→

`sqrt(NROW(tip)))`

[p.345 : 그림 18.8 캡션]

다른다는 것을 알 수 있다. → **다르**다는 것을 알 수 있다.

[p.349 : 1행]

```
theUrl <-http://www.jaredlander.com/data/TomatoFirst.csv
```

```
->
```

```
theUrl <- "http://www.jaredlander.com/data/TomatoFirst.csv"
```

[p.351]

변수 day의 각 레벨에 대해 자동으로 → 변수 day의 각 **레벨**에 대해 자동으로

[p.356 : 6행]

header로 첫 행이 열 이름으로 쓴다는 것을 지정한다. stringsAsFactors는 FALSE를 지정하면 문자열로 자동으로 팩터로 변화시키는 것을 막을 수 있다.

→

header로 첫 행을 열 이름으로 **쓰**를 지정한다. stringsAsFactors의 **인자**를 FALSE로 지정하면 문자열을 자동으로 팩터로 변화시키는 것을 막을 수 있다.

[p.357 : 1행]

평방피트당 가격value for square foot → 평방피트당 가격value **per** square foot

[p.357 : 2행]

콘도를 분석과는 다소 동떨어진 → **콘도**의 분석과는 다소 동떨어진

[p.358 : 그림 19.4 캡션]

이것은 보면 브루클린과 퀸즈가 하나의 봉우리를, → 이것은 보면 브루클린과 **퀸즈**가 하나의 봉우리를,

[p.358 : 그림 19.4 캡션]

브롱크스와 스테튼아일랜드는 → 브롱크스와 **스테튼아일랜드**는

[p.361 : 그림 19.6 캡션]

평방피드 당 → **평방피트** 당

[p.365 : 9행]

판단하는 것을 중요하기 때문이다. → 판단하는 것을 **중요시**하기 때문이다.

[p.370 : 그림 19.2 캡션]

이것은 보면 → 이것을 **보**면

[p.370 : 아래에서 3행]

그림 19.2에서 → 그림 **19.12**에서

[p.371 : 1행]

하나의 변수로 **높**고 → 하나의 변수로 **놓**고

[p. 375 : 8행]

미국 커뮤니티 조사 → 미국 **커뮤니티** 조사

[p.380 : 3행]

앞에서 사용했더는 → 앞에서 **사용했**던

[p.384 : 아래에서 2행]

제공근 → **제공근**

[p.385 : 7행]

데이터 세트과 다르다. → 데이터 세트**와** 다르다.

[p.385 : 8행]

피실험자에서 어떤 일이 발생했는지 → 피실험자**에게** 어떤 일이 발생했는지

[p.385 : 아래에서 2행]

이전에 발행했는지 여부를 → 이전에 **발생**했는지 여부를

[p.388 : 1행]

자연스럽 방법이다. → 자연스**런** 방법이다.

[p.393 : 2행]

변수의 변형 시키는 일 등과 → 변수를 **변형**하는 일 등과

[p.398 : 그림 21.3 캡션]

점들이 x축과 평할 곡선에 위에 그려진 것을 확인하자.

→

점들이 x축과 평할 곡선 **위에** 그려진 것을 확인하자.

[p.399 : 1행]

모형의 적합도가 → 모형의 **적합**도가

[p.399 : 그림 21.5 캡션]

최적 적합을 갖고 있다고 보기 어렵다. → 최적 **적합**을 갖고 있다고 보기 어렵다.

[p.400 : 아래에서 2행]

왜냐하면, 적합도라는 것이 → 왜냐하면, **적합**도라는 것이

[p.405 : 아래에서 3행]

브리안 리플리(Brain Ripley) → **브라이언** 리플리(Brian Ripley)

[p.406 : 2행]

정확도를 평가는 것으로 → 정확도를 **평가**하는 것으로

[p.406 : 아래쪽 코드]

glm으로 모형 적합을 다시 시행 → glm으로 모형 **적합**을 다시 시행

[p.407 : 코드 아래에서 3행]

NOVA → **ANOVA**

[p.409 : 그림 21.8 캡션]

그 모양을 비슷하다. → 그 **모양은** 비슷하다.

[p.413]

원래의 척도와 바이어스와 → **원래의** 척도와 바이어스와

[p.416 : 1행]

더하기만 할지, 아니면 → 더하기만 할지, **아니면**

[p.418 : 3행]

논란을 있지만 라쏘 회귀 → **논란은** 있지만 라쏘 회귀

[p.419 : 1행]

다루야 하는 상황이 → **다뤄야** 하는 상황이

[p.419 : 3행]

불가능할 경우도 있지만, → **불가능한** 경우도 있지만,

[p.420 : 7행]

획기적인 논문을 발표한 → **획기적인** 논문을 발표한

[p.429 : 1행]

a값으로 0으로 지정하면 → **a값을** 0으로 지정하면

[p.432 : 8행]

여러 개의 인자들 취하므로 .multicombine을 TRUE로 설정해 속도를 올릴 수 있다.

→

여러 개의 인자**를** 취하므로 .multicombine을 TRUE로 설정해 속도를 올릴 수 있다.