

대한민국 No.1 뷰티 스토어 ‘올리브영’

뷰티 관련 제품은 없는게 없는 ‘올리브영’에서는 도대체 무슨 뷰티 제품들을 팔고 있을까?

요즘 20대 여성들은 “이니스프리”, “에뛰드 하우스”처럼 독립적인 오프라인 화장품 가게에 가는 횟수가 줄었다.

올리브영에서 수많은 브랜드의 제품을 제공하여 각 소비자들의 입맛에 맞게 구매할 수 있도록 해주기 때문이다.

여러 브랜드의 색상, 품질, 가격 비교 등은 바로 이 ‘올리브영’에서 쉽게 할 수 있다.

그렇다면 과연 올리브영에 입점해 있는 뷰티 관련 브랜드들에는 어떤 것들이 있을까?

목차

1. 올리브영, 대체 어떤 뷰티제품을 파는 가게인가?
2. 올리브영에 입점한 브랜드들을 아이템별로 알아보기.
3. 어떤 아이템의 조합에서 중복되는 브랜드가 많을까?
4. 올리브영에 크게 자리잡은 브랜드들! 그 브랜드들이 궁금하다.
5. 결론

1. 올리브영, 대체 어떤 뷰티제품을 파는 가게인가?

- 웹 크롤링으로 ‘올리브영 공식 홈페이지’에서 item별, 그리고 그 item에 속한 category별로 판매하는 제품의 수를 가져와서 표 만들기
- 그 후 파일로 저장

```
#install.packages("rvest")
library(rvest)
```

```
## Loading required package: xml2
```

```
total<-NULL
category<-NULL
item<-NULL

#크롤링해오는 함수1 만들기
crawl_func1<-function(x,url){
  for (i in x){
    url2<-paste(url,i,sep = "")
    htxt<-read_html(url2)
    category<-append(category,html_nodes(htxt,"div.titBox h1")%>%html_text())
    total<-append(total,html_nodes(htxt,'p.cate_info_tx span')%>%html_text())
    items<-html_nodes(htxt,"div.history_cate_box ul li a.on")%>%html_text()%>%as.list()
    item<-append(item,unlist(items[1]))
  }
}

crawl_func1(c(1,2,6,7,4),"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010001000")
crawl_func1(c(1,8,6,7),"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010002000")
crawl_func1(4:9,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010003000")
crawl_func1(c(8,7,6,5,4,9),"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010004000")
crawl_func1(c(3,4,2),"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010005000")
crawl_func1(1:5,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010006000")
crawl_func1(7:9,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010007000")
crawl_func1(10:14,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=10000010007000")

#아이템, 카테고리, 개수를 크롤링해 온 데이터들을 item_consist라는 변수에 데이터프레임으로 저장하기
item_consist<-data.frame(item,category,total)
item_consist<-item_consist[-5,]
write.csv(item_consist,"item_consist.csv")

#item_consist 일부를 보기
head(item_consist)
```

```
##      item      category    total
## 1 스킨케어    페이스케어  1,307
## 2 스킨케어    마스크팩    564
## 3 스킨케어    클렌징     602
## 4 스킨케어    선케어     247
## 6 메이크업    베이스     478
## 7 메이크업    네일메이크업 591
```

- 크롤링한 데이터를 통해 올리브영 공식 홈페이지의 item에는 어떤 것들이 있는지, category에는 어떤 것들이 있는지, 그리고 이 item, category에 속한 제품의 수는 몇 개인지 알 수 있다.

-임의적으로 예를 들어 '스킨케어'를 item으로 이름짓고 '스킨케어'에 포함되는 '마스크팩'등과 같은 것을 category로 이름 지었다.

- 위의 데이터를 분석해 유의미한 데이터를 얻어보기

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## √ ggplot2 3.3.2    √ purrr  0.3.4
## √ tibble  3.0.3    √ dplyr  1.0.2
## √ tidyr   1.1.2    √ stringr 1.4.0
## √ readr   1.3.1    √ forcats 0.5.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter()      masks stats::filter()
## x readr::guess_encoding() masks rvest::guess_encoding()
## x dplyr::lag()         masks stats::lag()
## x purrr::pluck()       masks rvest::pluck()
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##      combine
```

```
#아이템별 제품수량 데이터 가져오기
category_quantity<-read.csv("item_consist.csv",stringsAsFactors = F)

#데이터 구조 알아보기
str(category_quantity)
```

```
## 'data.frame':    36 obs. of  4 variables:
## $ X      : int  1 2 3 4 6 7 8 9 10 11 ...
## $ item    : chr  "스킨케어" "스킨케어" "스킨케어" "스킨케어" ...
## $ category: chr  "페이스케어" "마스크팩" "클렌징" "선케어" ...
## $ total   : chr  "1,307" "564" "602" "247" ...
```

```
#total 변수 전처리하기
category_quantity$total<-as.numeric(gsub(",", "", category_quantity$total))

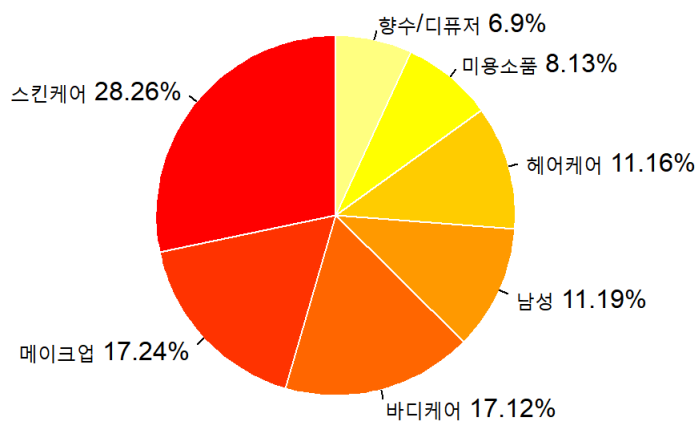
#item 별로 얼마나 많은 제품이 있는지 확인해보기
#오름차순으로 제품 수가 많은 item부터 제품 수가 적은 item 순으로 정렬하기
#그 후 총 제품의 수를 합친 'total_number' 변수 추가하기
#전체 제품 수에 대해 item별 제품 수의 비율 구하기

item_quantity<-category_quantity%>%
  group_by(item)%>%
  summarise(number=sum(total), .groups = 'drop')%>%
  arrange(desc(number))%>%
  mutate(total_number=sum(number))%>%
  mutate(item_percentage=(number/total_number)*100)

#올리브영에서 판매하는 제품의 수는 item 별로 어떻게 분포되어있나를 한 눈에 확인하기 위해
#위의 데이터를 시각화하기- pie chart 선택

slices<-round(item_quantity$item_percentage,2)
lbls<-item_quantity$item
lbls<-paste(lbls,slices)
lbls<-paste(lbls,"%",sep="")
pie(slices,init.angle=90,labels = lbls,col=heat.colors(length(lbls)),border="white",main="올리브영 아이템 구성"
)
```

올리브영 아이템 구성



- 올리브영의 뷰티 아이템 구성을 시각화해서 보니 스킨케어 제품이 상당한 수를 이루고 있음을 알 수 있다.
- 뷰티 아이템 중 무려 약 28.21% 정도를 차지하고 있고 스킨케어의 뒤를 이어 메이크업 제품이 약 17.25%를 차지한다.
- 남성만을 위한 아이템에 관한 제품들도 11.22%나 차지하고 있는 것으로 보아 꼭 올리브영이 여성을 위한 회사만은 아님을 알 수 있다.
- 그렇다면 이제 각각의 아이템에 어떤 카테고리들이 구성되어 있는지 알아보자.
- 메이크업과 스킨케어가 같은 것 아닌가?라고 생각하는 사람이 있을 수 있기 때문에 올리브영 내에서 메이크업 아이템 내에는 어떤 카테고리들이 있고, 스킨케어 아이템에는 어떤 카테고리들이 있는지 알아보자.
- bar그래프로 시각화해서 한 눈에 살펴보도록 하자.

#스킨케어의 구성과 메이크업의 구성 bar 그래프로 한 눈에 보기

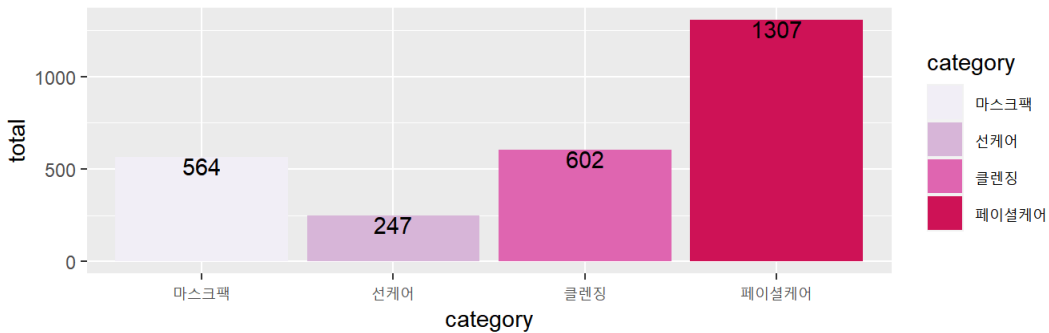
#스킨케어의 구성

```
skin<-category_quantity%>%filter(item=="스킨케어")
skin_composition<-ggplot(skin,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "PuRd")+
  labs(title = "스킨케어의 구성",x="category")+
  geom_text(vjust=1)
```

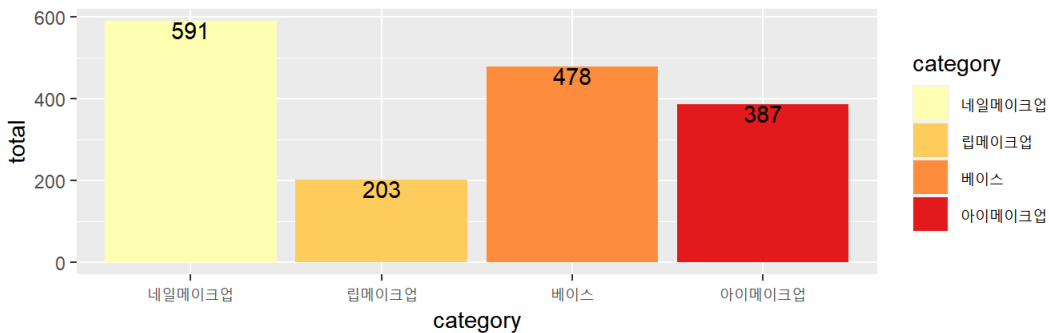
#메이크업의 구성

```
makeup<-category_quantity%>%filter(item=="메이크업")
makeup_composition<-ggplot(makeup,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "YlOrRd")+
  labs(title = "메이크업의 구성",x="category")+
  geom_text(vjust=1)
grid.arrange(skin_composition,makeup_composition)
```

스킨케어의 구성



메이크업의 구성



- 스킨케어의 카테고리에는 어떤 것들이 포함되는 지 살펴보자. 스킨, 로션과 같은 페이스케어에 관한 제품이 주를 이루며 마스크팩, 클렌징 용품도 스킨케어에 포함되는 것을 알 수 있다.
- 메이크업에서는 네일메이크업이라는 카테고리의 수가 가장 많으며 그 다음으로는 팩트, 비비와 같은 베이스 제품이 많다.

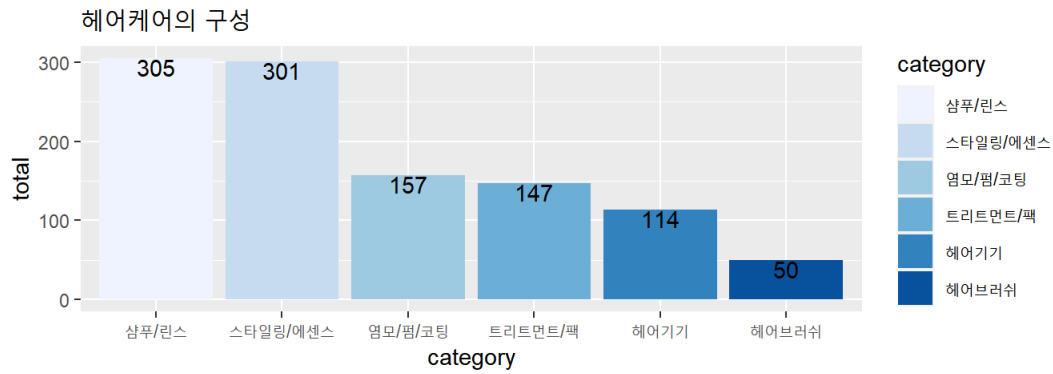
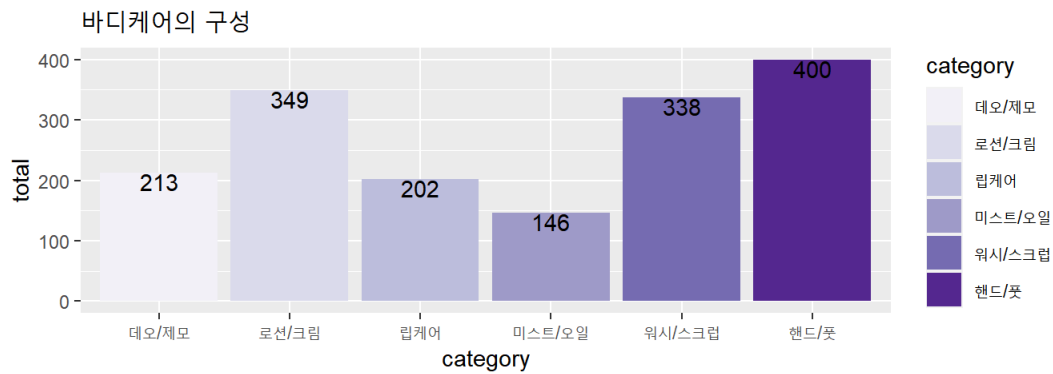
#바디케어의 구성과 헤어케어의 구성 bar 그래프로 한 눈에 보기

#바디케어의 구성

```
body<-category_quantity%>%filter(item=="바디케어")
body_composition<-ggplot(body,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "Purples")+
  labs(title = "바디케어의 구성",x="category")+
  geom_text(vjust=1)
```

#헤어케어

```
hair<-category_quantity%>%filter(item=="헤어케어")
hair_composition<-ggplot(hair,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "Blues")+
  labs(title = "헤어케어의 구성",x="category")+
  geom_text(vjust=1)
grid.arrange(body_composition,hair_composition)
```



- 바디케어에는 핸드/풋 즉, 손발을 케어해주는 제품이 가장 많다.
- 헤어케어에는 샴푸/린스 제품이 가장 큰 비율을 차지한다.

#향수/디퓨저의 구성과 미용소품의 구성 bar 그래프로 한 눈에 보기

#향수/디퓨저

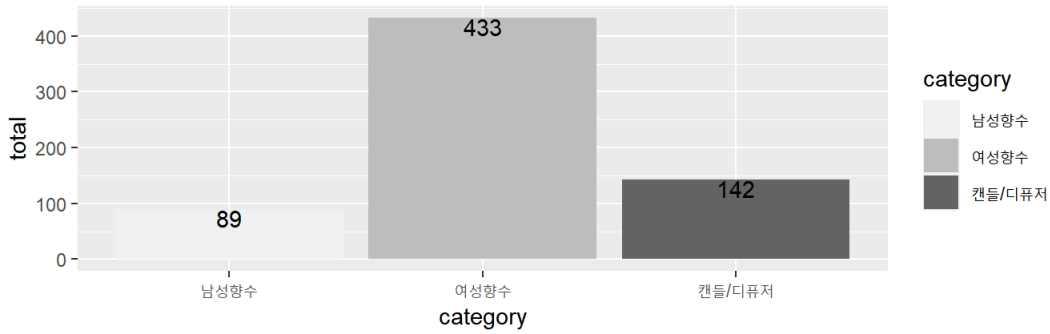
```
perfume<-category_quantity%>%filter(item=="향수/디퓨저")
perfume_composition<-ggplot(perfume,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "Greys")+
  labs(title = "향수/디퓨저의 구성",x="category")+
  geom_text(vjust=1)
```

#미용소품

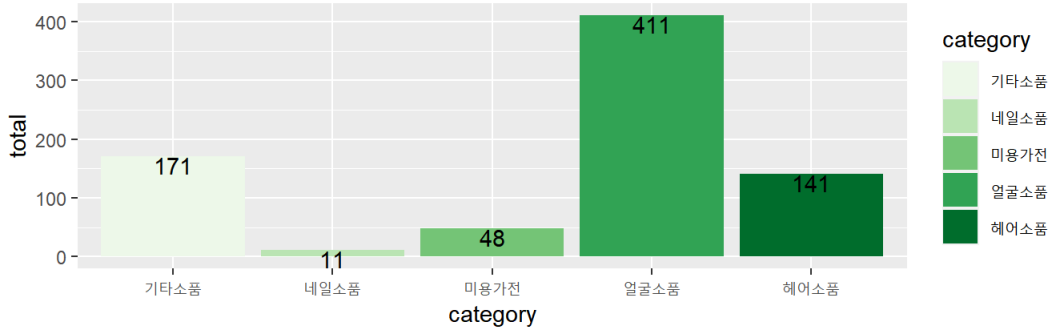
```
beauty_item<-category_quantity%>%filter(item=="미용소품")
beauty_item_composition<-ggplot(beauty_item,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "Greens")+
  labs(title = "미용소품의 구성",x="category")+
  geom_text(vjust=1)
```

```
grid.arrange(perfume_composition,beauty_item_composition)
```

향수/디퓨저의 구성



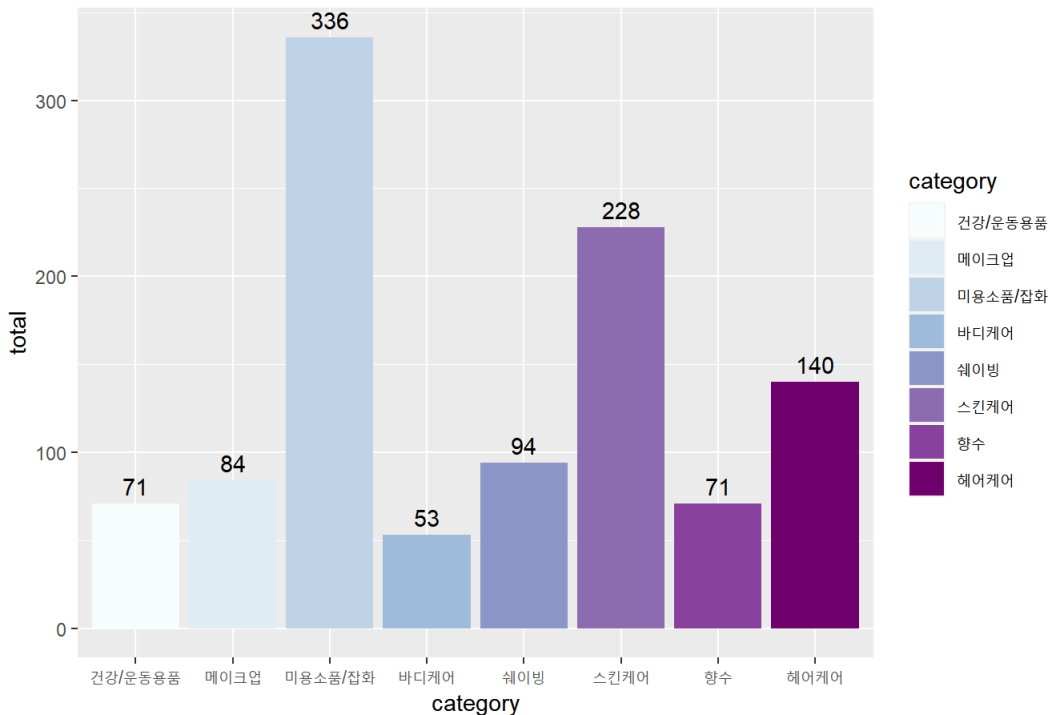
미용소품의 구성



- 향수/디퓨저에는 향수가 월등히 많고 그 중에서도 여성 향수 카테고리의 제품 수가 남성 향수에 비해 훨씬 많다.
- 미용소품에는 얼굴에 관련한 소품 카테고리의 제품 수가 가장 많다.

```
#남성용품의 구성요소 bar 그래프로 한눈에 보기
for_man<-category_quantity%>%filter(item=="남성")
for_man_composition<-ggplot(for_man,aes(x=category,y=total,fill=category,label=total))+
  geom_bar(stat = "identity")+
  scale_fill_brewer(palette = "BuPu")+
  labs(title = "남성제품의 구성",x="category")+
  geom_text(vjust=-0.5)
for_man_composition
```

남성제품의 구성



- 남성과 관련한 용품에는 스킨케어나 메이크업에 관련한 제품보다는 미용소품/잡화가 주를 이룬다.
- 여기서 하나의 인사이트를 도출해보자면, 여성을 위한 아이템에서는 스킨케어와 메이크업이 주를 이루었던 반면, 남성을 위한 아이템 들에는 소품/잡화가 많은 비중을 차지하는 것을 발견할 수 있다.

2. 올리브영에 입점한 '브랜드'들을 '카테고리별'로 알아보자.

```
#아이템별로 제품들 가져오기

# '스킨케어' 제품들 크롤링하기, 파일 만들기
brand<-NULL
goods<-NULL
price<-NULL
crwal_func2<-function(x,url){
  for (i in 1:x){
    url2<-paste(url,i,sep="")
    htxt<-read_html(url2)
    brand<-append(brand,html_nodes(htxt,"a.goodsList span.tx_brand")%>%html_text())
    goods<-append(goods,html_nodes(htxt,"p.tx_name")%>%html_text())
    price<-append(price,html_nodes(htxt,"span.tx_cur span.tx_num")%>%html_text())
  }
}

crwal_func2(51,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100010001&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(24,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100010002&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(25,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100010006&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(11,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100010007&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(14,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100010004&fltDispCatNo=&prdSort=03&pageIdx=")

skin_care_goods<-data.frame(brand,goods,price)
head(skin_care_goods)
```

```
##          brand
## 1   크리니크
## 2 디오디너리
## 3 디오디너리
## 4 리얼베리어
## 5   닥터지
## 6   마녀공장
##
##                                goods  price
## 1   [1+1한정기획] 크리니크 핑크수분크림 (72-아워) 39,000
## 2   나이아신아마이드 10%+징크 1% 30mL  5,900
## 3   글리코릭 애시드 7% 토닝 솔루션  8,700
## 4   [한정기획] 리얼베리어 익스트림 크림 50ml 기획(+25ml*2ea+1ml*3ea) 25,180
## 5   닥터지 레드블레미쉬클리어수딩크림 증량 UP 기획세트(70ml+30ml) 25,200
## 6   [한정기획] 마녀공장 비피다 바이옴 앰플 토너 400ml+각질패드 80매 증정 17,500
```

```
write.csv(skin_care_goods,"skin_care_goods.csv")

# '메이크업' 제품 크롤링하기, 파일 만들기

brand<-NULL
goods<-NULL
price<-NULL

crwal_func2(18,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100020001&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(25,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100020008&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(8,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100020008&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(15,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100020007&fltDispCatNo=&prdSort=03&pageIdx=")

makeup_goods<-data.frame(brand,goods,price)
write.csv(makeup_goods,"makeup_goods.csv")

# '바디케어' 제품 크롤링하기, 파일 저장하기
brand<-NULL
goods<-NULL
```

```

price<-NULL

crwal_func2(13,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030004&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(12,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030005&fltDispCatNo=&prdSort=01&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030006&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(15,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030007&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(9,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030008&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(10,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100030009&fltDispCatNo=&prdSort=03&pageIdx=")

body_care_goods<-data.frame(brand,goods,price)
write.csv(body_care_goods,"body_care_goods.csv")

# '헤어케어' 제품 크롤링하기, 파일 저장하기

brand<-NULL
goods<-NULL
price<-NULL

crwal_func2(12,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040008&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040007&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(13,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040006&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(7,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040005&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(5,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040004&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(3,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100040009&fltDispCatNo=&prdSort=03&pageIdx=")

hair_care_goods<-data.frame(brand,goods,price)
write.csv(hair_care_goods,"hair_care_goods.csv")

# '향수/디퓨저' 제품 크롤링하기, 파일 저장하기

brand<-NULL
goods<-NULL
price<-NULL

crwal_func2(17,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100050003&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(4,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100050004&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100050002&fltDispCatNo=&prdSort=03&pageIdx=")

perfume_goods<-data.frame(brand,goods,price)
write.csv(perfume_goods,"perfume_goods.csv")

# '미용 소품' 제품 크롤링하기, 파일 저장하기

brand<-NULL
goods<-NULL
price<-NULL

crwal_func2(15,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100060001&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100060002&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(1,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100060003&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(2,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100060004&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100060005&fltDispCatNo=&prdSort=03&pageIdx=")

```



```
ispCatNo=&prdSort=03&pageIdx=")
```

```
beauty_item<-data.frame(brand,goods,price)
write.csv(beauty_item,"beauty_item.csv")
```

```
# '남성 용품' 제품 크롤링하기, 파일 저장하기
```

```
brand<-NULL
goods<-NULL
price<-NULL
```

```
crwal_func2(11,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070007&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(4,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070008&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(6,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070009&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(5,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070010&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(3,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070011&fltDispCatNo=&prdSort=03&pageIdx=")
crwal_func2(4,"https://www.oliveyoung.co.kr/store/display/getMCategoryList.do?dispCatNo=100000100070012&fltDispCatNo=&prdSort=03&pageIdx=")
```

```
for_man_goods<-data.frame(brand,goods,price)
write.csv(for_man_goods,"for_man_goods.csv")
```

```
#아이템별 많이 나타나는 브랜드 살펴보기
```

```
library(tidyverse)
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
#스킨케어
```

```
#install.packages("wordcloud")
skin_care_goods<-read.csv("skin_care_goods.csv",stringsAsFactors = F)
```

```
#brand 빈도수를 table함수를 사용해서 알아보기
head(table(skin_care_goods$brand))
```

```
##
## 23 years old      AHC      DMCK      SRB      XTM      가스비
##              5         29         2         1         8         5
```

```
#더 보기 좋게 brand 빈도수 알아보기
```

```
skin_care<-skin_care_goods%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')
```

```
#나타나는 브랜드 이름을 가지고 워드 클라우드 그려보기
```

```
library(RColorBrewer)
palette<-brewer.pal(9,"Spectral")
wordcloud(skin_care$brand,freq = skin_care$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```

```
## Warning in wordcloud(skin_care$brand, freq = skin_care$freq, min.freq = 3, : 엠
## 브리올리스 could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(skin_care$brand, freq = skin_care$freq, min.freq = 3, : 비
## 온드 옴므 could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(skin_care$brand, freq = skin_care$freq, min.freq = 3, : 스
## 웨디쉬 could not be fit on page. It will not be plotted.
```

- 스킨케어 아이템에서는 ‘메디힐’, ‘차앤박’, ‘닥터자르트’, ‘바이오더마’ 등의 많은 브랜드가 입점해있다.

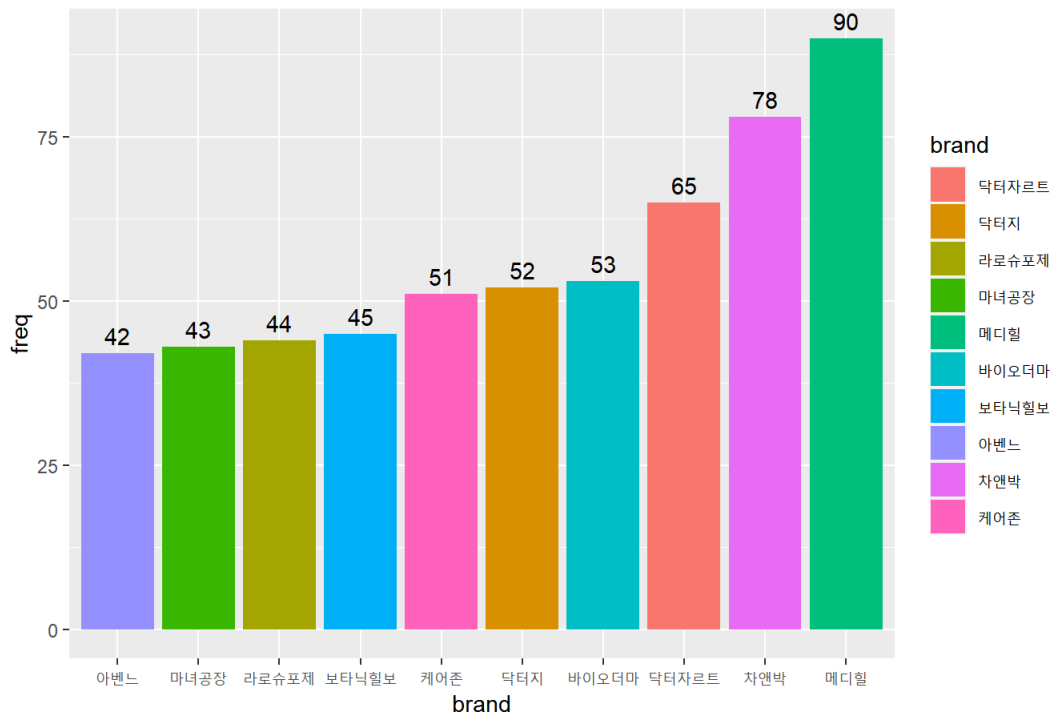
#빈도수를 기준으로 상위 10개의 브랜드 10개 뽑아보기

```
top_skin_care_brand<-skin_care%>%arrange(desc(freq))%>%head(.,10)
```

#스킨케어 아이템에 많이 입점한 브랜드를 빈도수 별로 그래프 그리기

```
top_skin_care_brand_bar<-ggplot(top_skin_care_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))
+
  geom_bar(stat = "identity")+
  labs(title = "스킨케어 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_skin_care_brand_bar
```

스킨케어 브랜드 입점



- '메디힐'은 90개의 제품을 올리브영 스킨케어 항목에 판매하고 있다.

#메이크업

```
makeup_goods<-read.csv("makeup_goods.csv",stringsAsFactors = F)
```

#brand 빈도수 알아보기

```
maekup<-makeup_goods%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')
```

```
library(RColorBrewer)
```

```
palette<-brewer.pal(9,"Spectral")
```

#나타나는 브랜드 이름을 가지고 워드 클라우드 그려보기

```
wordcloud(maekup$brand,freq = maekup$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```


- 따라서 메이크업 아이템에는 네일 관련한 제품이 압도적으로 많이 배치되어 있음을 알 수 있다.
- 네일 제품을 파는 브랜드를 제외하고는 메이크업의 다른 브랜드들은 비슷한 제품의 수로 올리브영 '메이크업'에 입점해 있음을 확인할 수 있다.

#바디케어

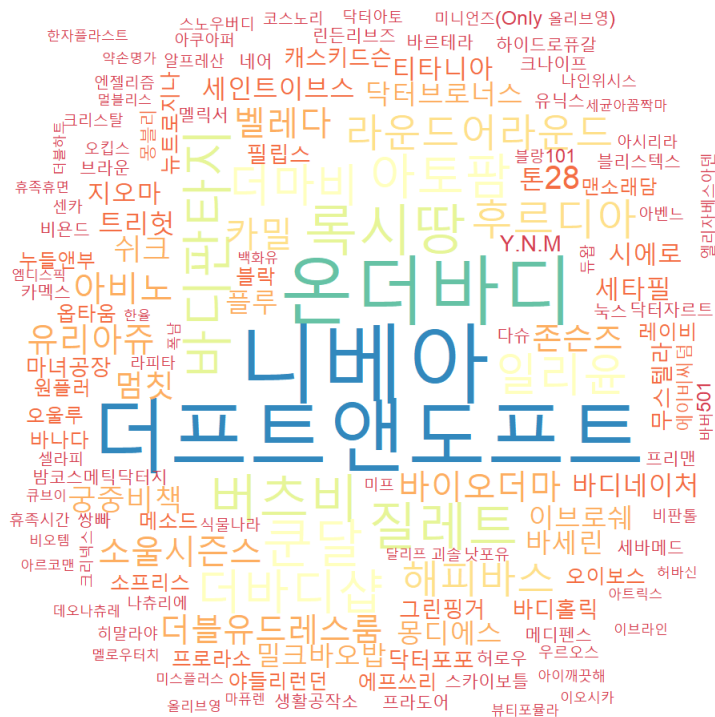
```
body_care_goods<-read.csv("body_care_goods.csv",stringsAsFactors = F)
```

```
body_care<-body_care_goods%>%group_by(brand)%>%summarize(freq=n(),.groups = 'drop')
```

```
library(RColorBrewer)
```

```
palette<-brewer.pal(9,"Spectral")
```

```
wordcloud(body_care$brand,freq = body_care$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```

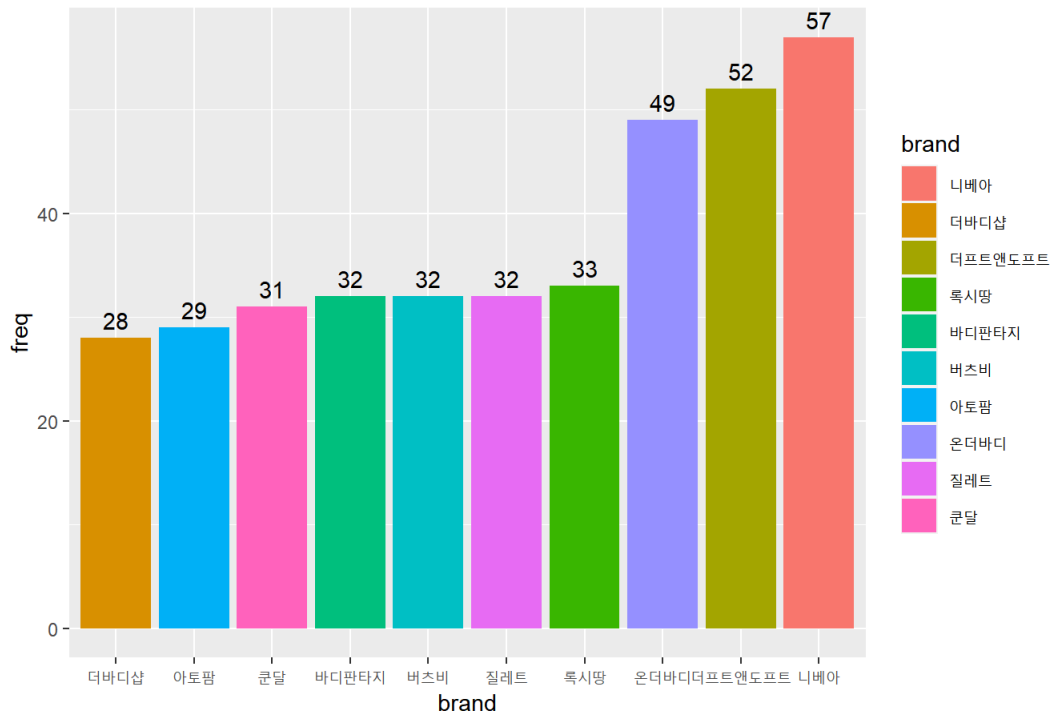


- 바디 케어에서는 ‘니베아’, ‘더프트앤도프트’, ‘온더바디’ 등의 브랜드들이 눈에 띈다.

```
top_body_care_brand<-body_care%>%arrange(desc(freq))%>%head(.,10)
```

```
top_body_care_brand_bar<-ggplot(top_body_care_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))
+
  geom_bar(stat = "identity")+
  labs(title = "바디케어 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_body_care_brand_bar
```

바디케어 브랜드 입점



- '니베아'가 바디케어에서 가장 많은 제품의 수를 판매하고 있다는 것을 알 수 있다.

#헤어케어

```
hair_care_goods<-read.csv("hair_care_goods.csv",stringsAsFactors = F)
```

#brand 빈도수 알아보기

```
hair_care<-hair_care_goods%>%group_by(brand)%>%summarise(freq=n()),.groups = 'drop')
```

```
library(RColorBrewer)
```

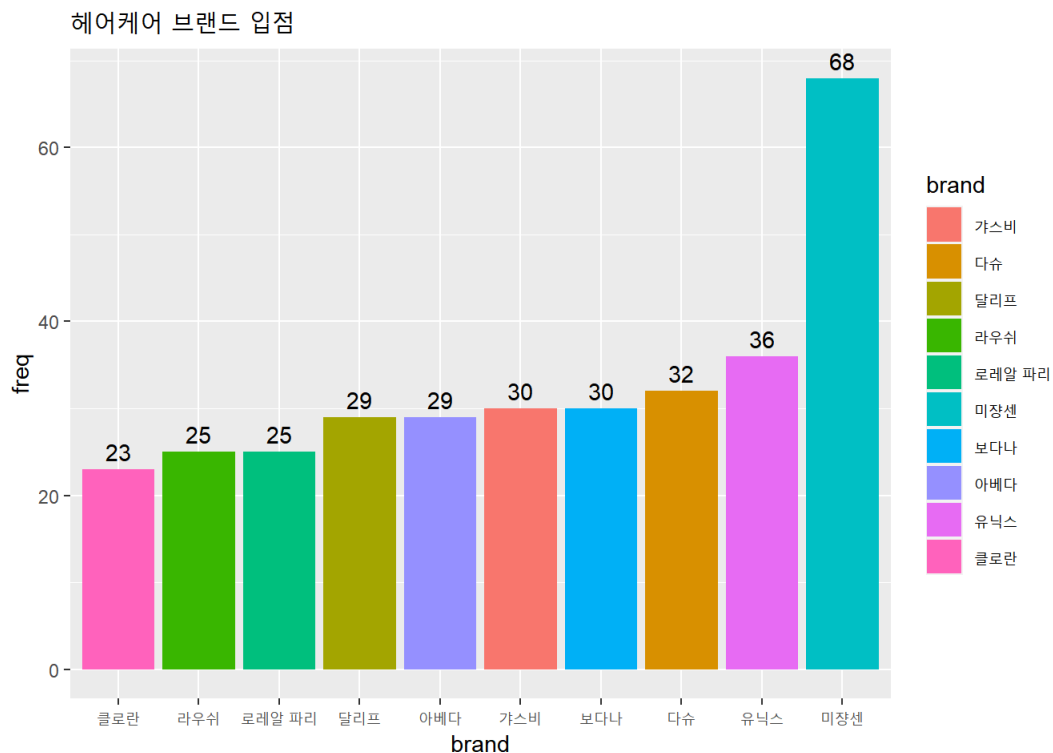
```
palette<-brewer.pal(9,"Spectral")
```

```
wordcloud(hair_care$brand,freq = hair_care$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```



```
top_hair_care_brand<-hair_care%>%arrange(desc(freq))%>%head(.,10)

top_hair_care_brand_bar<-ggplot(top_hair_care_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))
+
  geom_bar(stat = "identity")+
  labs(title = "헤어케어 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_hair_care_brand_bar
```

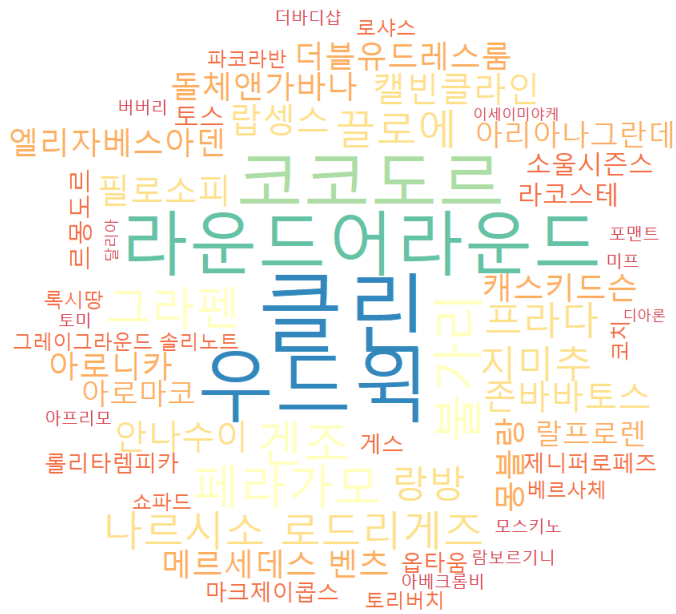


- '미장센'이 압도적으로 많이 올리브영 헤어케어 아이템에 입점해 있음을 알 수 있다.

```
#향수
perfume_goods<-read.csv("perfume_goods.csv",stringsAsFactors = F)

#brand 빈도수 알아보기
perfume<-perfume_goods%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')

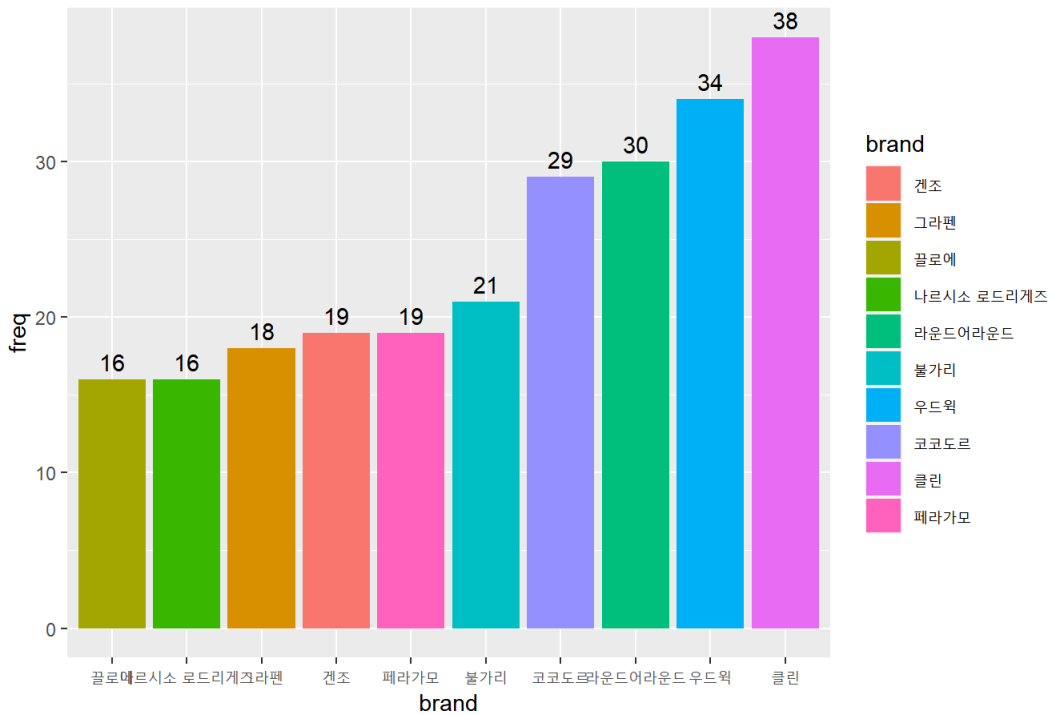
library(RColorBrewer)
palette<-brewer.pal(9,"Spectral")
wordcloud(perfume$brand,freq = perfume$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```



```
top_perfume_brand<-perfume%>%arrange(desc(freq))%>%head(.,10)
```

```
top_perfume_brand_bar<-ggplot(top_perfume_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))+
  geom_bar(stat = "identity")+
  labs(title = "향수/디퓨저 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_perfume_brand_bar
```

향수/디퓨저 브랜드 입점



- ‘클린’, 우드웍’ 두 브랜드가 향수 아이템에 비슷하게 많이 입점해있다.


```
#미용소품

beauty_item<-read.csv("beauty_item.csv",stringsAsFactors = F)
#brand 빈도수 알아보기
beauty<-beauty_item%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')

library(RColorBrewer)
palette<-brewer.pal(9,"Spectral")
wordcloud(beauty$brand,freq = beauty$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```

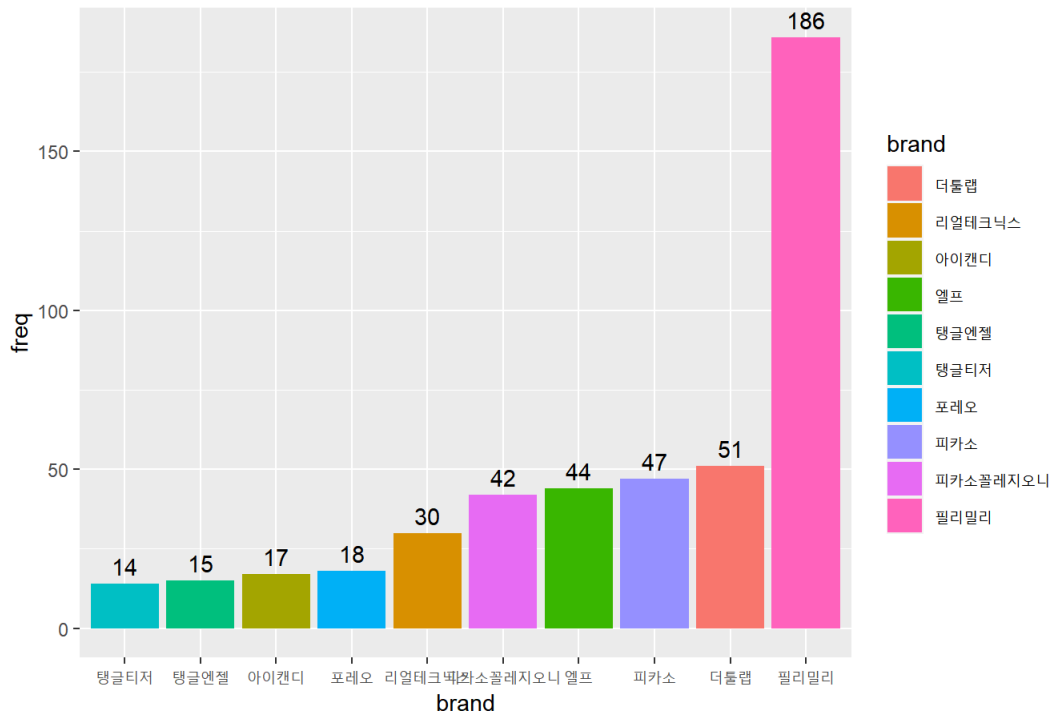


- '필리밀리'라는 브랜드가 미용소품에서 압도적으로 많이 입점해있다.

```
top_beauty_item_brand<-beauty%>%arrange(desc(freq))%>%head(.,10)

top_beauty_item__brand_bar<-ggplot(top_beauty_item_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))+
  geom_bar(stat = "identity")+
  labs(title = "미용소품 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_beauty_item__brand_bar
```

미용소품 브랜드 입점



- 실제로 '필리밀리'브랜드가 2위의 브랜드 '더툴랩'보다 3배 이상 더 많은 제품을 올리브영 미용소품 아이템에 팔고 있다.
- 올리브영 미용소품 아이템에서는 '필리밀리'의 영향력이 크다고 할 수 있다.

#남자

```
for_man_goods<-read.csv("for_man_goods.csv",stringsAsFactors = F)

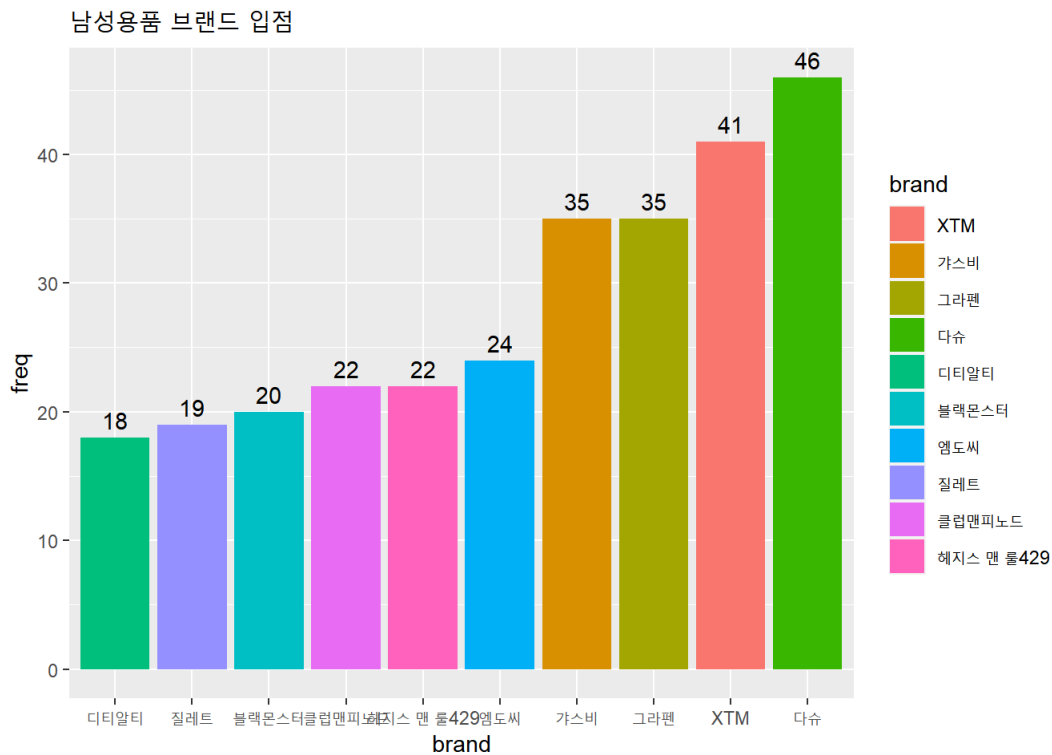
#brand 빈도수 알아보기
man<-for_man_goods%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')

library(RColorBrewer)
palette<-brewer.pal(9,"Spectral")
wordcloud(man$brand,freq = man$freq,min.freq = 3,
          random.order = F,random.color = F,colors = palette)
```



```
top_for_man_brand<-man%>%arrange(desc(freq))%>%head(.,10)

top_for_man_brand_bar<-ggplot(top_for_man_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))+
  geom_bar(stat = "identity")+
  labs(title = "남성용품 브랜드 입점",x="brand")+
  geom_text(vjust=-0.5)
top_for_man_brand_bar
```



- 브랜드 이름들을 살펴보니, 확실히 이때까지 나온 여성을 위한 제품의 브랜드와는 다른 브랜드들이 입점해 있는 것을 볼 수 있다.

3. 어떤 아이템의 조합에서 중복되는 브랜드가 많을까?

- 특정 브랜드가 꼭 하나의 아이템(ex. 스킨케어, 메이크업, 바디케어, 헤어케어)에 속하라는 법은 없다.
- 쉽게 예상해볼 수 있듯이, 립스틱, 섀도우와 같은 색조 화장에 주력하는 회사도 스킨 제품을 낼 수 있기 때문이다.
- 그렇다면 과연 어떤 조합에서 중복되는 브랜드가 많이 나타날까?
- 즉, 올리브영에 입점해 있는 브랜드는 어떤 조합에서 가장 많이 중복 입점해 있는 걸까?

```
library(tidyverse)

#아이템을 알려줄 새로운 변수 추가하기
skin_care_goods<-skin_care_goods%>%mutate(category="스킨케어")
makeup_goods<-makeup_goods%>%mutate(category="메이크업")
body_care_goods<-body_care_goods%>%mutate(category="바디케어")
hair_care_goods<-hair_care_goods%>%mutate(category="헤어케어")
perfume_goods<-perfume_goods%>%mutate(category="향수/디퓨저")
beauty_item<-beauty_item%>%mutate(category="미용소품")
for_man_goods<-for_man_goods%>%mutate(category="남성")

#카테고리별로 브랜드의 개수를 파악하기

#스킨케어 카테고리에 입점한 브랜드의 개수
skin_care_goods%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
```

```
## # A tibble: 265 x 2
##   brand      num
##   <chr>    <int>
## 1 23 years old     5
## 2 AHC             29
## 3 DMCK            2
## 4 SRB             1
## 5 XTM             8
## 6 가스비          5
## 7 구달            26
## 8 궁중비책        4
## 9 그라펜          6
## 10 그레이그라운드  1
## # ... with 255 more rows
```

```
skin_brand_num<-length(skin_care_goods$brand)
```

#메이크업 카테고리에 입점한 브랜드의 개수

```
makeup_goods%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
```

```
## # A tibble: 105 x 2
##   brand      num
##   <chr>    <int>
## 1 16브랜드      1
## 2 3CE           24
## 3 AHC           2
## 4 XTM           16
## 5 그라펜       11
## 6 나인위시스    2
## 7 다슈          9
## 8 닥터원더      5
## 9 닥터자르트    7
## 10 닥터자르트포맨 1
## # ... with 95 more rows
```

```
makeup_brand_num<-length(makeup_goods$brand)
```

#바디케어 카테고리에 입점한 브랜드의 개수

```
body_care_goods%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
```

```
## # A tibble: 232 x 2
##   brand      num
##   <chr>    <int>
## 1 8x4        2
## 2 ms.44     2
## 3 XTM        2
## 4 Y.N.M      6
## 5 괴슬       4
## 6 궁중비책   15
## 7 그라펜     1
## 8 그레이그라운드 1
## 9 그린핑거   10
## 10 나인위시스   4
## # ... with 222 more rows
```

```
body_brand_num<-length(body_care_goods$brand)
```

#헤어케어 카테고리에 입점한 브랜드의 개수

```
hair_care_goods%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
```

```
## # A tibble: 126 x 2
##   brand      num
##   <chr>      <int>
## 1 AZH         3
## 2 OGX         8
## 3 XTM         9
## 4 가스비      30
## 5 그라펜      10
## 6 그레이그라운드 1
## 7 그린핑거     1
## 8 꽃을튼남자   3
## 9 니심        2
## 10 다슈       32
## # ... with 116 more rows
```

```
hair_brand_num<-length(hair_care_goods$brand)
```

- 아이템 별로 조합하기
- 여기서는 스킨케어, 메이크업, 바디케어, 헤어케어 이 네 가지 아이템들의 조합을 살펴보았다.
- 미용 소품, 향수는 소품, 향수 등을 전문적으로 하는 회사들이 많았기 때문에 제외하였다.

#스킨과 메이크업 아이템에 입점한 브랜드의 개수

```
skin_makeup<-inner_join(skin_care_goods,makeup_goods,by="brand");head(skin_makeup)
```

```
##   X.x      brand                                goods.x price.x category.x
## 1    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
## 2    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
## 3    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
## 4    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
## 5    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
## 6    1  크리닉 [1+1한정기획] 크리닉 핑크수분크림 (72-아워)  39,000   스킨케어
##   X.y                                goods.y price.y
## 1   10                                크리닉 치크 팍  22,400
## 2  311      크리닉 비온드 퍼펙팅 파운데이션+컨실러 SPF19/PA++ 30ml  49,000
## 3  321      크리닉 이본 베타 플루이드-크림 메이크업 SPF15/PA++ 30ml  55,000
## 4  335      크리닉 수퍼 시티 블록 BB 쿠션 컴팩트 SPF50/PA+++  55,000
## 5  421  크리닉 이본베타 리프레쉬 하이드레이팅 앤 리페어링 파운데이션  51,000
## 6 1223      [1+1한정기획] 크리닉 래쉬 파워 마스크라  34,000
##   category.y
## 1   메이크업
## 2   메이크업
## 3   메이크업
## 4   메이크업
## 5   메이크업
## 6   메이크업
```

```
skin_makeup_brand<-skin_makeup%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
skin_makeup_brand_num<-length(skin_makeup_brand$brand)
```

#스킨과 바디케어 아이템에 입점한 브랜드의 개수

```
skin_body<-inner_join(skin_care_goods,body_care_goods,by="brand")
skin_body_brand<-skin_body%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
skin_body_brand_num<-length(skin_body_brand$brand)
```

#스킨과 헤어케어 아이템에 입점한 브랜드의 개수

```
skin_hair<-inner_join(skin_care_goods,hair_care_goods,by="brand")
skin_hair_brand<-skin_hair%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
skin_hair_brand_num<-length(skin_hair_brand$brand)
```

#메이크업과 바디케어 아이템에 입점한 브랜드의 개수

```
makeup_body<-inner_join(makeup_goods,body_care_goods,by="brand")
makeup_body_brand<-makeup_body%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
makeup_body_brand_num<-length(makeup_body_brand$brand)
```

#메이크업과 헤어케어 아이템에 입점한 브랜드의 개수

```
makeup_hair<-inner_join(makeup_goods,hair_care_goods,by="brand")
makeup_hair_brand<-makeup_hair%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
makeup_hair_brand_num<-length(makeup_hair_brand$brand)
```

#바디케어나 헤어케어 아이템에 입점한 브랜드의 개수

```
body_hair<-inner_join(body_care_goods,hair_care_goods,by="brand")
body_hair_brand<-body_hair%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')
body_hair_brand_num<-length(body_hair_brand$brand)
```

- 아이템별 상관관계 분석
- 어떤 조합에서 중복되는 브랜드가 가장 많이 나타나는지를 수치화하고 시각화 해보았다.

아이템별 상관관계 분석을 위해 데이터프레임 만들기

```
skin<-c(skin_brand_num,skin_makeup_brand_num,skin_body_brand_num,skin_hair_brand_num)
makeup<-c(skin_makeup_brand_num,makeup_brand_num,makeup_body_brand_num,makeup_hair_brand_num)
body<-c(skin_body_brand_num,makeup_body_brand_num,body_brand_num,body_hair_brand_num)
hair<-c(skin_hair_brand_num,makeup_hair_brand_num,body_hair_brand_num,hair_brand_num)
category_relation<-data.frame(skin,makeup,body,hair)
```

```
rownames(category_relation)<-c("skin","makeup","body","hair")
```

#이해하기 쉬운 표를 위해 이름을 따로 붙여주기

#상관관계 분석하기

#수치화

```
category_relation<-as.matrix(category_relation)
cor(category_relation)
```

```
##           skin      makeup      body      hair
## skin      1.0000000 -0.3131049 -0.2904344 -0.3436424
## makeup -0.3131049  1.0000000 -0.3626365 -0.3637390
## body  -0.2904344 -0.3626365  1.0000000 -0.3242460
## hair  -0.3436424 -0.3637390 -0.3242460  1.0000000
```

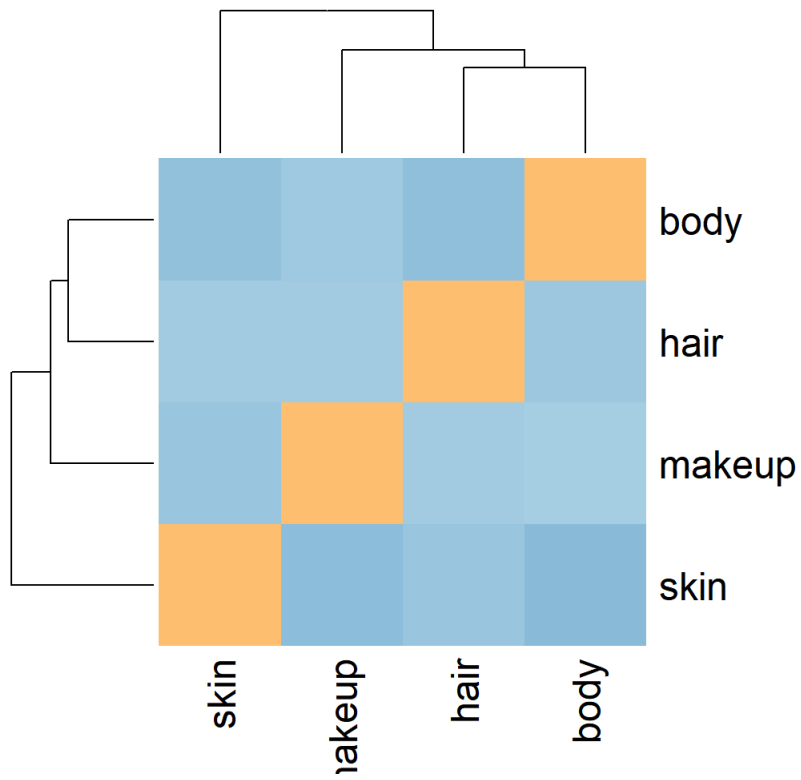
library (RColorBrewer)

```
col<-colorRampPalette(brewer.pal(7,"Paired"))(256)
```

#히트맵 그리기

#시각화

```
heatmap(category_relation,scale = "col",col=col)
```



- 분석 결과, 스킨케어 제품과 바디케어 제품에 공통으로 입점한 브랜드가 상대적으로 많았다.
- 아무래도 스킨케어 브랜드나 바디케어 브랜드 모두 사람의 피부를 주로 연구하는 브랜드이다 보니
- 스킨케어와 바디케어 모두에 입점한 브랜드가 상대적으로 많았다.
- 그에 비해, 메이크업 제품과 헤어케어 제품에 공동으로 입점한 브랜드는 상대적으로 적었다.

가설검증해보기

```
## 위의 히트맵에서 관찰할 수 있듯이 스킨케어와 바디케어는 중복되는 브랜드가 비교적 많이 입점해있다.

## 그렇다면 스킨케어와 바디케어에 중복해서 들어와있는 브랜드 중 브랜드 10개를 뽑아보자.
top_common_skin_body_brand<-head(skin_body_brand$brand,10)

#스킨케어에 입점해있는 브랜드 중 앞의 브랜드 10개에 해당하는 브랜드들의 제품 수의 합
top_skin_care_common_brand<-skin_care_goods%>filter(brand %in% top_common_skin_body_brand )%>group_by(brand)%>summarise(num=n(),.groups = 'drop')

#바디케어에 입점해있는 브랜드 중 앞의 브랜드 10개에 해당하는 브랜드들의 제품 수의 합
top_body_care_common_brand<-body_care_goods%>filter(brand %in% top_common_skin_body_brand )%>group_by(brand)%>summarise(num=n(),.groups = 'drop')

# 대립가설 : 브랜드 10개에서 스킨케어 제품의 수와 바디케어 제품의 수는 다를 것이다.

# 이 가설을 검증해보고 싶은 이유는 다음과 같다,

# 스킨케어와 바디케어에 동시에 입점하는 브랜드가 있다고 할 때, 과연 이 브랜드는 스킨케어와 바디케어에 비슷한 제품 수를 입점하게 될 것인지, 아니면 스킨을 더 많이 입점하는데 치중하는지, 바디에 더 많이 입점하는데 치중하는지 궁금했기 때문이다.

#brand를 기준으로 inner join하여 요약 표 만들기
each_skin_body_brand<-inner_join(top_skin_care_common_brand,top_body_care_common_brand,by="brand")

#raw 데이터로 바꿔줄 함수
rep.row<-function(x,n){
  m <- matrix(rep(x,each=n),nrow=n)
  return(m)
}

#rep.row 함수 적용하기
df <- data.frame(rbind(rep.row(c("XTM", "스킨"), 9),
  rep.row(c("XTM", "바디"), 2),
  rep.row(c("가스비", "스킨"), 5),
  rep.row(c("가스비", "바디"), 4),
  rep.row(c("궁중비책", "스킨"), 4),
  rep.row(c("궁중비책", "바디"), 19),
  rep.row(c("그라펜", "스킨"), 6),
  rep.row(c("그라펜", "바디"), 3),
  rep.row(c("그레이그라운드", "스킨"), 1),
  rep.row(c("그레이그라운드", "바디"), 1),
  rep.row(c("나인위시스", "스킨"), 20),
  rep.row(c("나인위시스", "바디"), 4),
  rep.row(c("눅스", "스킨"), 44),
  rep.row(c("눅스", "바디"), 6),
  rep.row(c("뉴트로지나", "스킨"), 22),
  rep.row(c("뉴트로지나", "바디"), 13),
  rep.row(c("니베아", "스킨"), 5),
  rep.row(c("니베아", "바디"), 64),
  rep.row(c("다슈", "스킨"), 4),
  rep.row(c("다슈", "바디"), 3)))

#열과 행 이름 문맥에 맞게 바꿔주기
names(df) <- c("브랜드", "스킨_바디")

#만든 raw 데이터를 통해 테이블 만들기
skin_body_brand_table<-table(df$브랜드,df$스킨_바디)

# 교차분석 진행하기
chisq.test(skin_body_brand_table, correct=FALSE)
```

```
## Warning in chisq.test(skin_body_brand_table, correct = FALSE): Chi-squared
## approximation may be incorrect
```

```
##
## Pearson's Chi-squared test
##
## data: skin_body_brand_table
## X-squared = 107.8, df = 9, p-value < 2.2e-16
```

- p-value가 매우 작은 것으로 나오므로 10개 브랜드는 스킨에 입점시키는 제품 수와 바디케어에 입점시키는 제품 수를 다르게 할 것이

라는 것을 알 수 있다.

- 즉, 스킨케어와 바디케어에 중복되는 10개의 브랜드의 제품 수에서는 스킨케어의 제품 수와 바디케어의 제품 수의 비율이 같지 않다.
- 다시 말해, 스킨케어와 바디케어에 동시에 입점하려고 하는 브랜드는, 스킨케어 제품 수 입점에 더 치중하거나 바디케어 제품 수 입점에 더 치중한다. 스킨케어, 바디케어 두 브랜드에 비슷하게 제품을 입점하지 않는다.
- 즉, 10개의 각 브랜드마다 스킨케어의 제품 수와 바디케어의 제품 수는 다르다.
- 이 검증을 한 궁극적인 이유는 다음과 같다, 특정 브랜드가 스킨케어와 바디케어에 비슷한 제품 수를 입점시킬 것인지, 아니면 스킨을 더 많이 입점하는데 치중하는지, 바디에 더 많이 입점하는데 치중하는지 궁금했기 때문이다.

#스킨케어와 바디케어 제품에 공동으로 입점한 브랜드를 워드클라우드를 그려 알아보기

```
#install.packages("wordcloud2")  
library(wordcloud2)  
skin_body_brand<-skin_body%>%group_by(brand)%>%summarise(num=n(),.groups = 'drop')  
wordcloud2(skin_body_brand,col="random-light")
```



- ‘니베아’, ‘닥터브로너스’, ‘닥터자르트’, ‘마녀공장’, ‘일리윤’ 등의 브랜드들이 올리브영의 스킨케어, 바디케어를 꽉 잡고 있음을 확인해 볼 수 있다.

4. 올리브영에 크게 자리잡은 브랜드들! 그 브랜드들이 궁금하다.

-아이템 별로 나누지 않고 결국 올리브영 전체에 가장 많이 입점한 브랜드가 궁금해서 이 챕터를 추가하였다.

#크롤링한 모든 제품을 합쳐서 전체 제품을 담은 데이터프레임 만들기

```
total_product<-skin_care_goods%>%rbind(.,makeup_goods)%>%rbind(.,body_care_goods)%>%rbind(.,hair_care_goods)%>%rbind(.,perfume_goods)%>%rbind(.,beauty_item)%>%rbind(.,for_man_goods)
```

#브랜드별로 그룹핑 한 다음 빈도수를 확인하고 내림차순으로 정렬하기, 그 후 10개만 추출하기

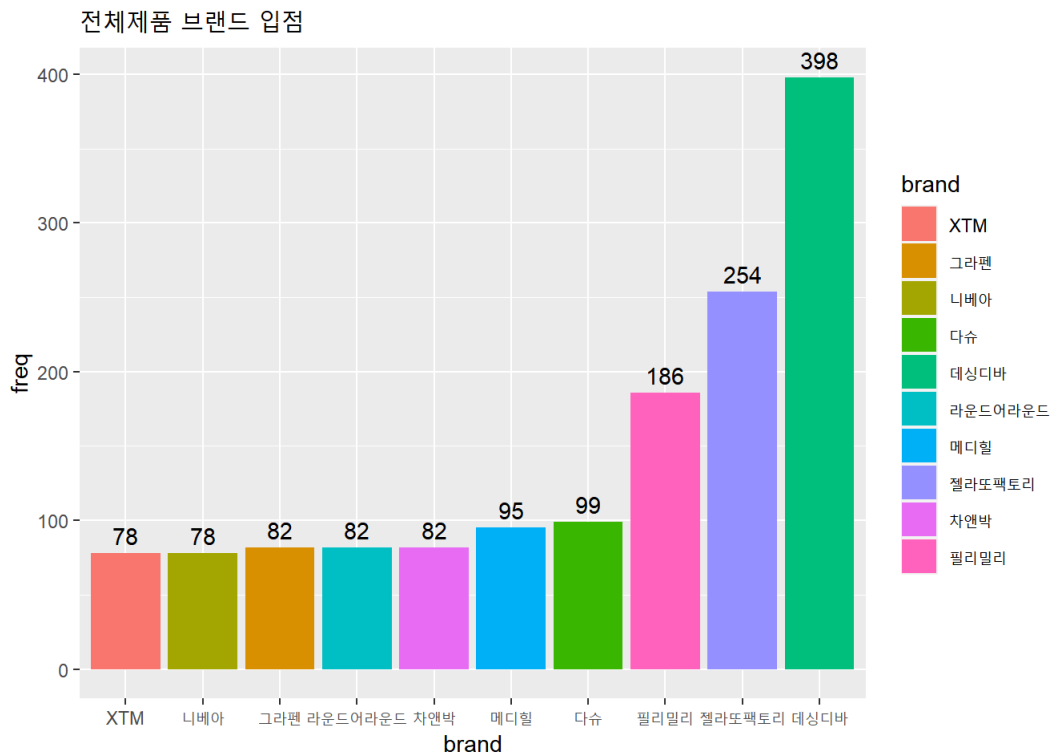
```
top_total_product_brand<-total_product%>%group_by(brand)%>%summarise(freq=n(),.groups = 'drop')%>%arrange(desc(freq))%>%head(.,10)
```

#앞서 뽑은 가장 많이 나타나는 10개의 브랜드인 브랜드만 추출하고 brand열과 item열을 추출하기, 그 후 브랜드별로 정렬하고 중복되는 값 없애기

```
total_top_brand_item<-total_product%>%filter(brand %in% top_total_product_brand$brand)%>%select(brand,category)%>%arrange(brand)%>%unique()
```

#시각화를 통해 어떤 브랜드가 올리브영에 가장 많이 입점하고 있는지 파악하기

```
ggplot(top_total_product_brand,aes(x=reorder(brand,freq),y=freq,fill=brand,label=freq))+  
  geom_bar(stat = "identity")+  
  labs(title = "전체제품 브랜드 입점",x="brand")+  
  geom_text(vjust=-0.5)
```



- '데싱디바', '젤라또팩토리'가 많은 수를 차지하는 것으로 보아 네일 관련 제품 수가 굉장히 많음을 알 수 있었다.
- '필리밀리'는 미용 소품을 판매하는 브랜드였던 것을 떠올려보면 올리브영은 화장품 뿐만 아니라 뷰티에 관련한 소품도 많이 팔고 있음을 알 수 있다.
- '다슈', '그라펜', 'XTM' 등 남성제품 브랜드들이 올리브영 내 전체 브랜드 입점순위에도 상위권에 랭크되어 있다.
- 이를 통해 남성들이 올리브영에 갔을 때 구매할 수 있는 제품 수가 상당히 많다는 것을 도출해낼 수 있다.
- 그런데 브랜드 이름만 보면 이 브랜드가 과연 어떤 아이템인지 알 수 없다.
- 즉, 필리밀리라는 브랜드의 제품 수가 많지만, 필리밀리 브랜드가 메이크업 아이템의 브랜드였는지, 헤어케어 아이템의 브랜드였는지, 미용 소품 아이템의 브랜드였는지 알 수 없다는 뜻이다.
- 각 브랜드의 해당 그룹을 찾으려면 다시 위로 올라가야 한다.
- 따라서 브랜드와 과연 이 브랜드가 어떤 아이템에 속하는지까지 한 눈에 볼 수 있는 network graph를 그려보았다.

#그래프를 그려 아이템과 많이 입점하고 있는 브랜드와 어떤 관계가 있는지 알아보기

```
#install.packages("igraph")
```

```
library(igraph)
```

```
##
```

```
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:dplyr':
##
##   as_data_frame, groups, union
```

```
## The following objects are masked from 'package:purrr':
##
##      compose, simplify
```

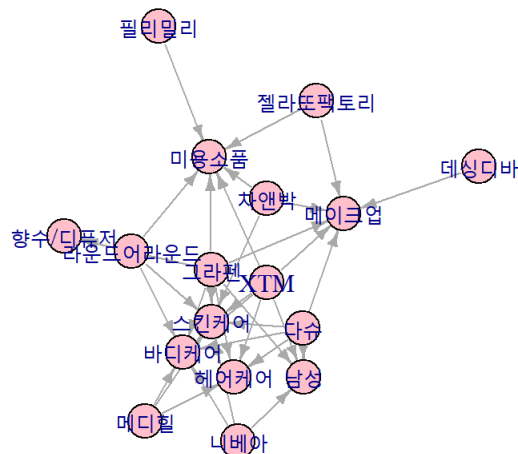
```
## The following object is masked from 'package:tidyr':
##
##   crossing
```

```
## The following object is masked from 'package:tibble':
##
##      as_data_frame
```

```
## The following objects are masked from 'package:stats':
##
##   decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##      union
```

```
g<-graph.data.frame(total_top_brand_item,directed = T)
plot(g,layout=layout.fruchterman.reingold,vertex.size=17,edge.arrow.size=0.5,vertex.color="pink")
```



- 올리브영에 가장 많이 입점해 있는 브랜드들은 ‘바디케어,’ ‘스킨케어’에 집중하는 브랜드임을 알 수 있다.
- ‘바디케어,’ ‘스킨케어’에 찌르는 브랜드들이 많기 때문이다.

결론

- 올리브영에는 뷰티 제품뿐만 아니라 반려견을 위한 제품, 건강식품, 일반식품, 아이용품 등등 더 다양한 제품들을 판매하고 있었다. 따라서 뷰티 제품들에만 한정하여 분석해본 것이 아쉬웠다.
- 올리브영에는 여러 브랜드들이 입점하는 기업이다보니 스킨케어, 메이크업, 바디케어, 헤어케어 등등 아이템별로 다 다른 브랜드들이 입점해있는 것이 놀라웠다.
- 여러 아이템들에 걸쳐서 제품을 제공하는 브랜드들도 있었지만 한 아이템에만 입점한 브랜드들도 많다는 것을 알 수 있었다.

- 또한 올리브영이 여성만을 위한 기업이라고 생각했는데 생각보다 남성을 위한 제품이 많았고 남성들도 필요한 뷰티 제품이 있으면 올리브영을 둘러도 좋을 것 같다.
- 다이소와 같이 올리브영에서도 여러 미용 소품들을 판매하고 있었고 제품 수가 적지 않음을 확인할 수 있었다. 따라서 다이소의 가격이 너무 저렴하여 미용 소품을 사기 꺼려진다면 올리브영에 둘러도 좋을 것 같다.
- 소규모의 뷰티 브랜드, 이제 막 시작하는 뷰티 브랜드 들도 많이 찾아볼 수 있었고 앞으로도 올리브영이 이런 브랜드들에게 기회를 주는 곳으로 남을 것 같다.
- 마지막으로 교차분석을 통해 스킨과 바디케어 제품에 들어오는 수는 다르다는 것을 검증해볼 수 있었다.