**HW2**

**Submit the homework named as HW1\_학번\_이름.doc, HW1\_학번\_이름.R (코드도 첨부)**

**데이터 설명 : BTS, zoom, 라면, 코로나에 대한 검색결과**

**search.csv**

일: 날짜 (YYYY-MM-DD)

라면 : 주어진 기간 가장 큰 검색량을 100으로 한다. 그 기준 날짜에 라면의 검색량.

Zoom : 주어진 기간 가장 큰 검색량을 100으로 한다. 그 기준 날짜에 zoom의 검색량.

코로나 : 주어진 기간 가장 큰 검색량을 100으로 한다. 그 기준 날짜에 코로나의 검색량.

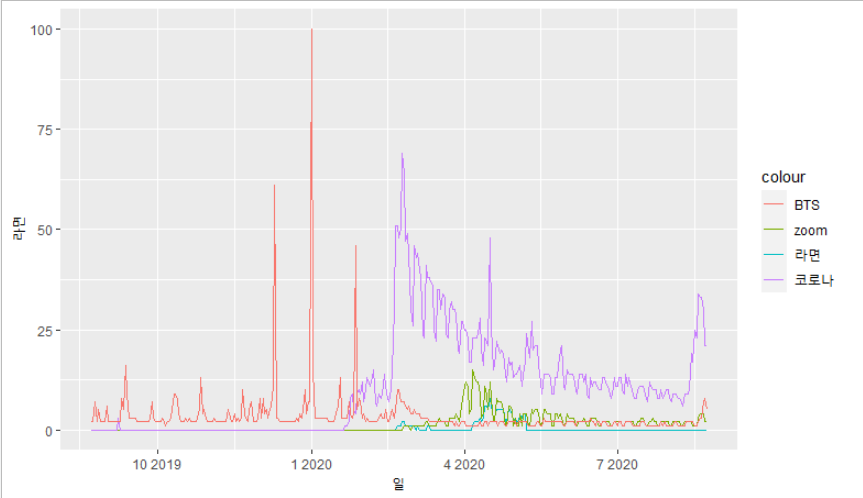
BTS : 주어진 기간 가장 큰 검색량을 100으로 한다. 그 기준 날짜에 BTS의 검색량.

**search\_gender.csv**

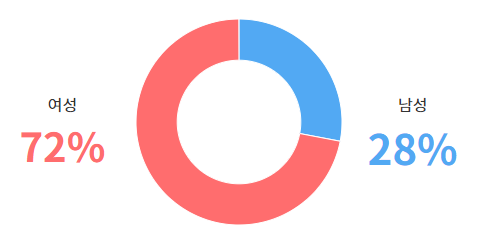
**search\_age.csv**

**search\_local.csv**

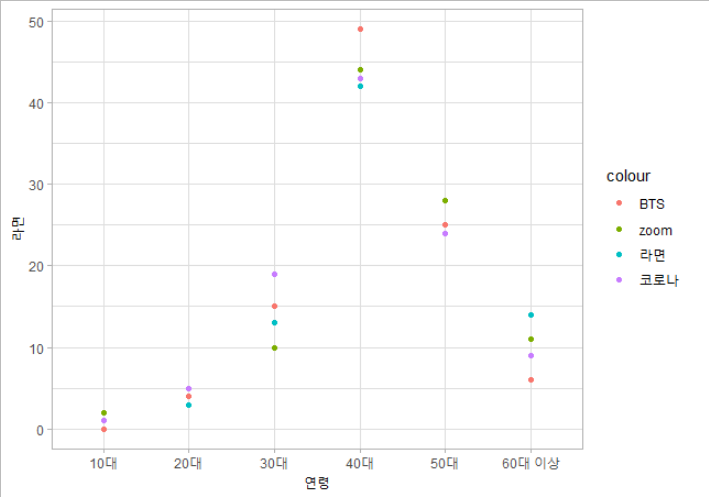
1. **상기 데이터에 대해 분석을 진행한다.**
   1. **주어진 데이터 search.csv, search\_gender.csv, search\_age.csv, search\_local.csv를 불러들이기 (search.csv를 data로 읽었다 하자. data$일 = as.Date(data$일) 을 적용하라.)**
   2. **검색량.csv 데이터를 geom\_line을 통해 다음과 같이 표현해라.**



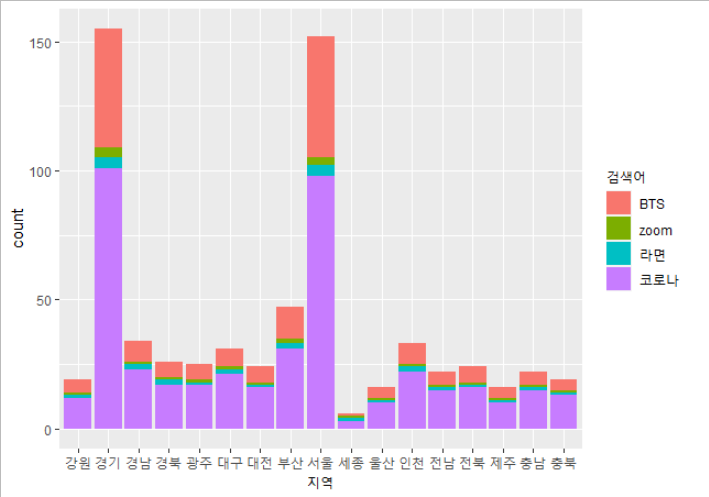
* 1. **검색량\_성별.csv데이터를 다음과 같이 pie chart로 표현해라**



* 1. **검색량\_연령.csv데이터를 다음과 같이 Scatter Plot으로 표현해라**



* 1. **검색량\_지역.csv데이터를 다음과 같이 Histogram으로 표현해라**



* 1. **위의 표현된 결과를 통해 각각 BTS, zoom, 라면, 코로나의 일자별 검색량의 변화를 해석하라.**

**데이터설명 : U.S SAT scores by state for 2010 (SAT\_2010.csv)**

A data.frame with 50 rows and 9 variables.

state : a factor with levels for each state

expenditure : average expenditure per student (in each state)

pupil\_teacher\_ratio : pupil to teacher ratio in that state

salary : teacher salary (in 2010 US $)

read : state average Reading SAT score

math state average Math SAT score

write state average Writing SAT score

total state average Total SAT score

sat\_pct percent of students taking SAT in that state

1. **상기 데이터에 대해 분석을 진행한다.**
   1. **“write”와 “math”에 대해 boxplot을 그리고, 분포에 대해 설명.**
   2. **산점도를 그려보고 가장 의미있다고 생각하는 1개을 제시하고 설명 (total변수 사용)**
   3. **ggplot을 이용하여 Salary가 50000이상인 데이터와 이하인 데이터에 대해 Sat\_pct와 total 을 이용해 Plot graph를 그리시오. 이 두 그래프를 비교하며 해석하시오.**