# 학력과 취업의 관계(실제 데이터 x)

# 변수: academic\_background(학력), employ(취업)

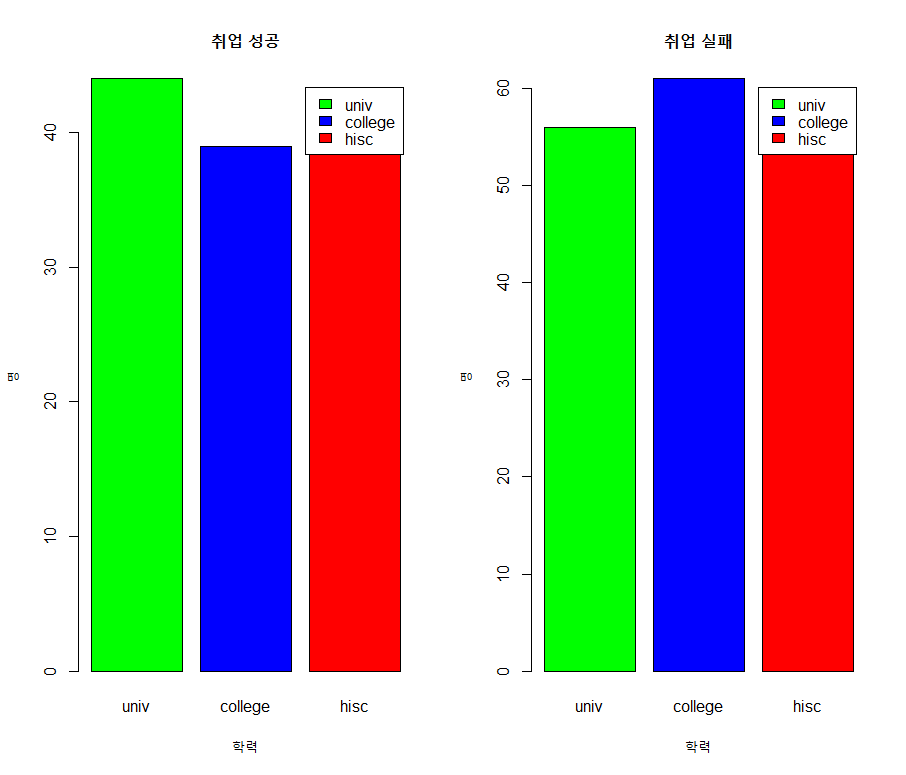
# academic\_background: 1(univ, 4년제), 2(college, 2년제), 3(hisc, 고졸)

# employ: 0(Failure), 1(success)

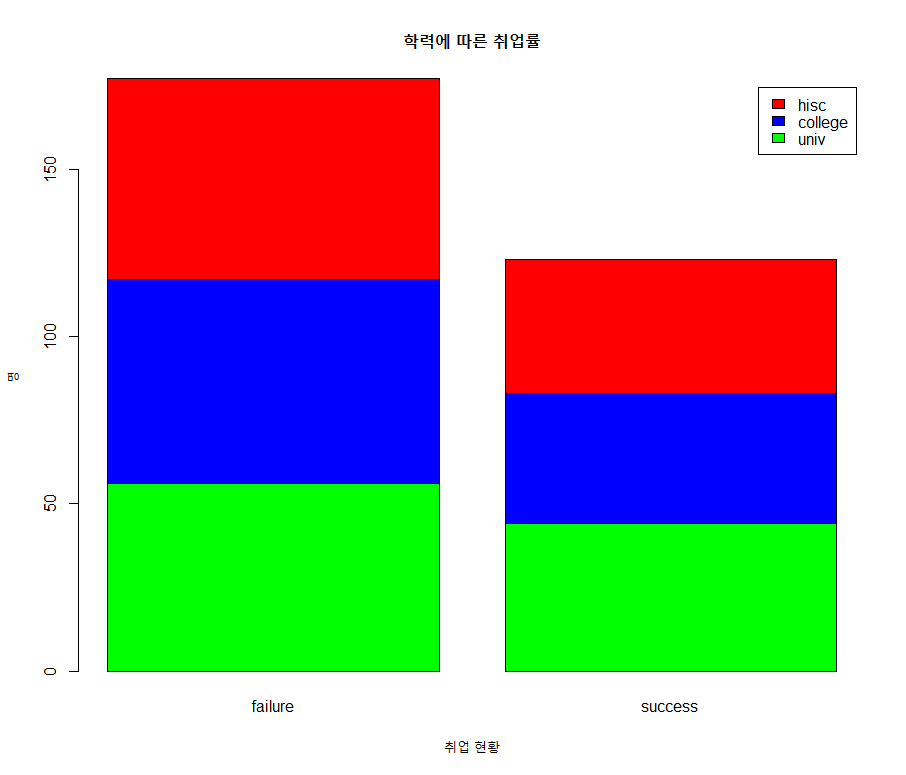
# 위의 정보를 통해 데이터 전처리를 해주시고, 아래의 그래프가 나오도록 실습해보세요.

\*아래의 그래프는 인강에서 나온 함수들과 함수속의 속성들을 전부 사용하여 만든 것입니다.

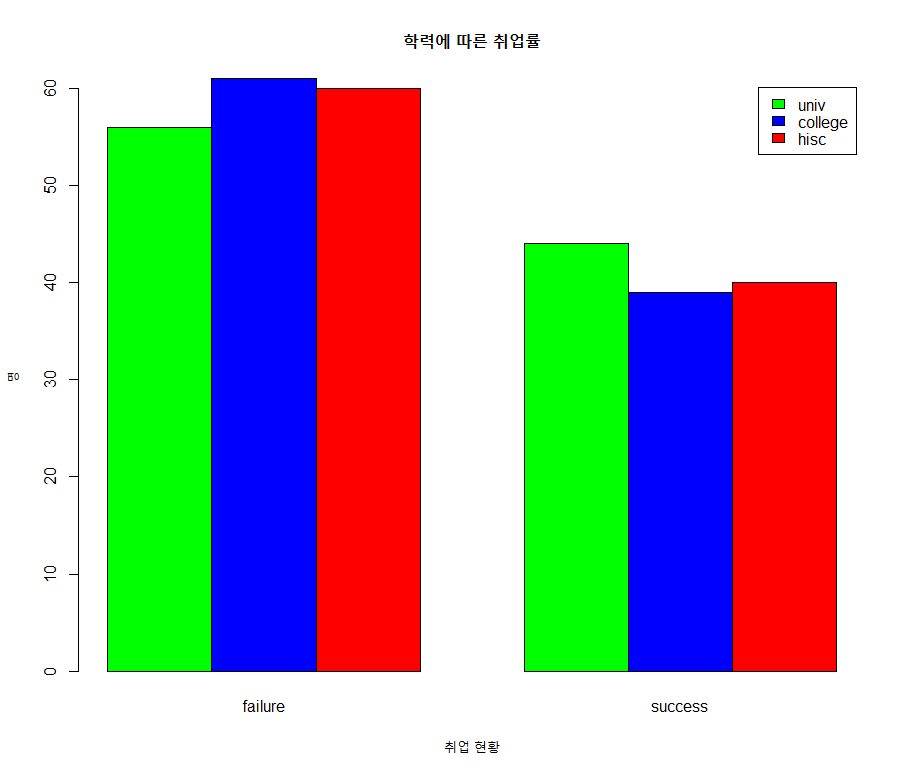
색깔이나 그래프 위치, 크기 등의 사소한 차이는 상관없습니다. 다만, 그래프에 사용한 데이터는 틀리지 않았으면 합니다.

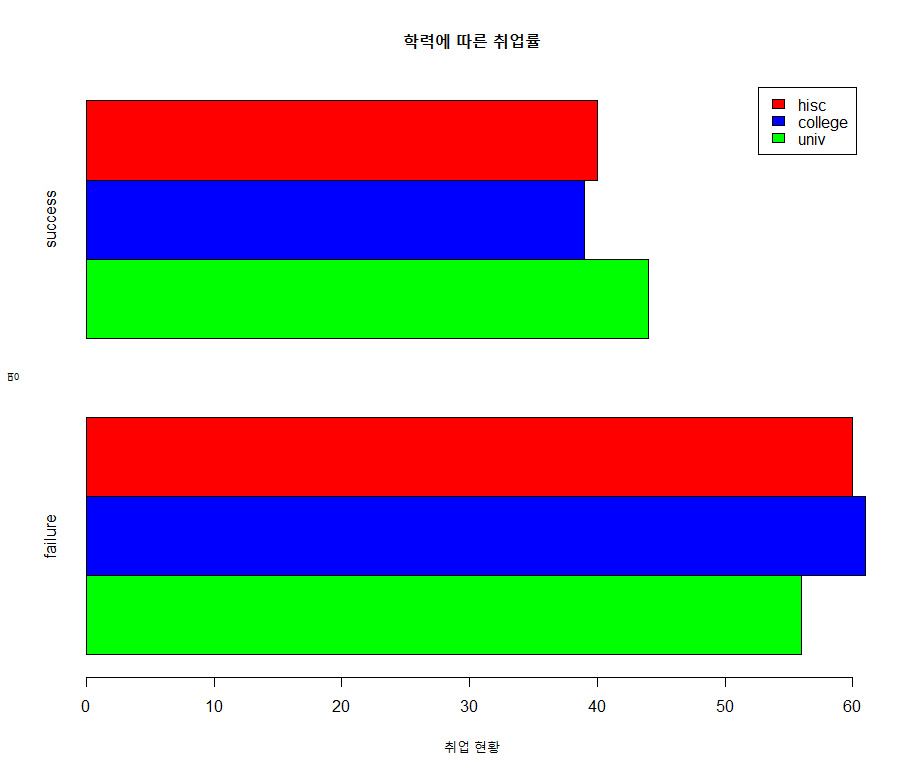


\*막대 그래프(스크린 2개), x축: 학력, y축: 사람 수(단위: 명)

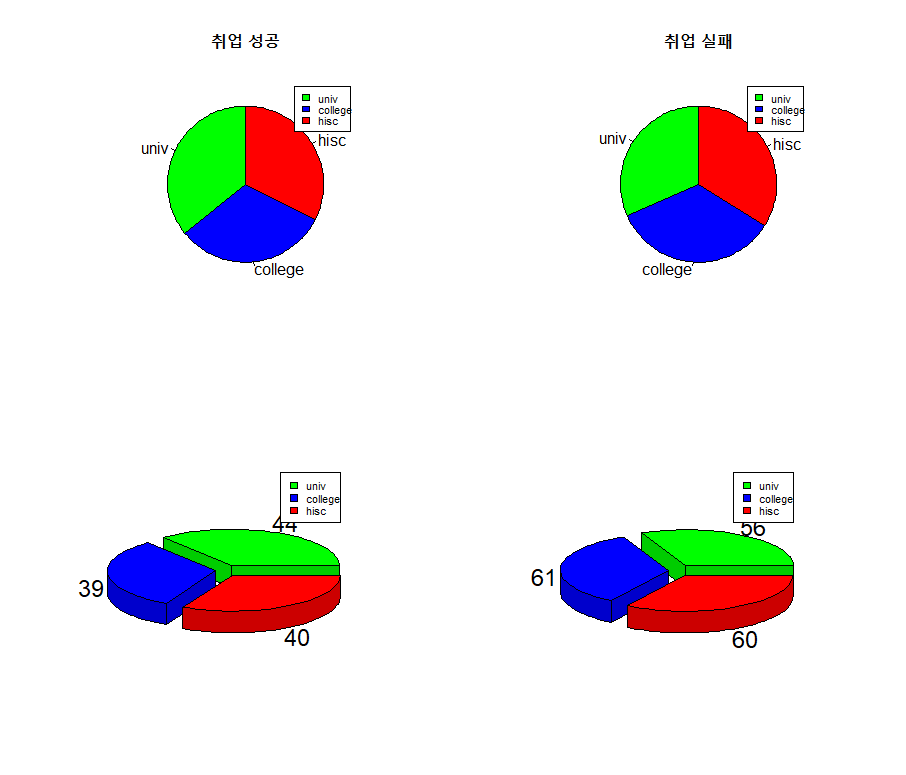


\*누적막대그래프, x축: 취업성공여부, y축: 사람 수(단위: 명)

\*누적막대그래프, x축: 취업성공여부, y축: 사람 수(단위: 명)

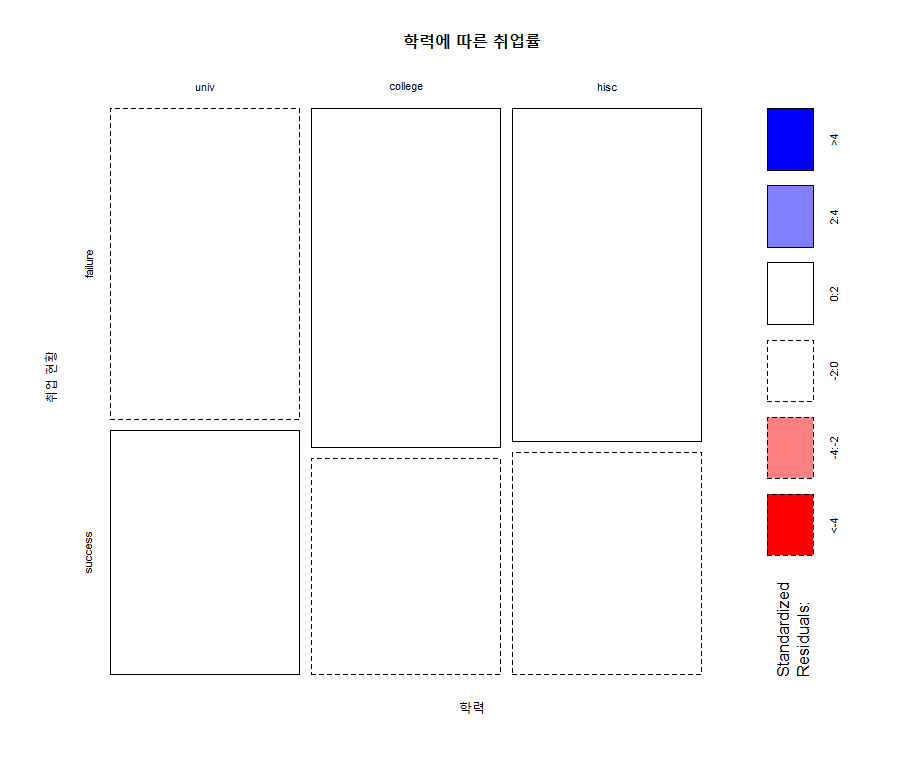


\*수평막대그래프: y축: 취업성공여부, x축: 사람 수(단위: 명)



원그래프&3D원그래프 (스크린4개)

데이터: 학력별 인원수(취업 여부에 따라 각각)



모자이크 그래프

데이터: 학력, 취업 여부

#자료 정보: https://vincentarelbundock.github.io/Rdatasets/doc/Ecdat/Diamond.html

#carat: 무게

#colour: 색깔 등급

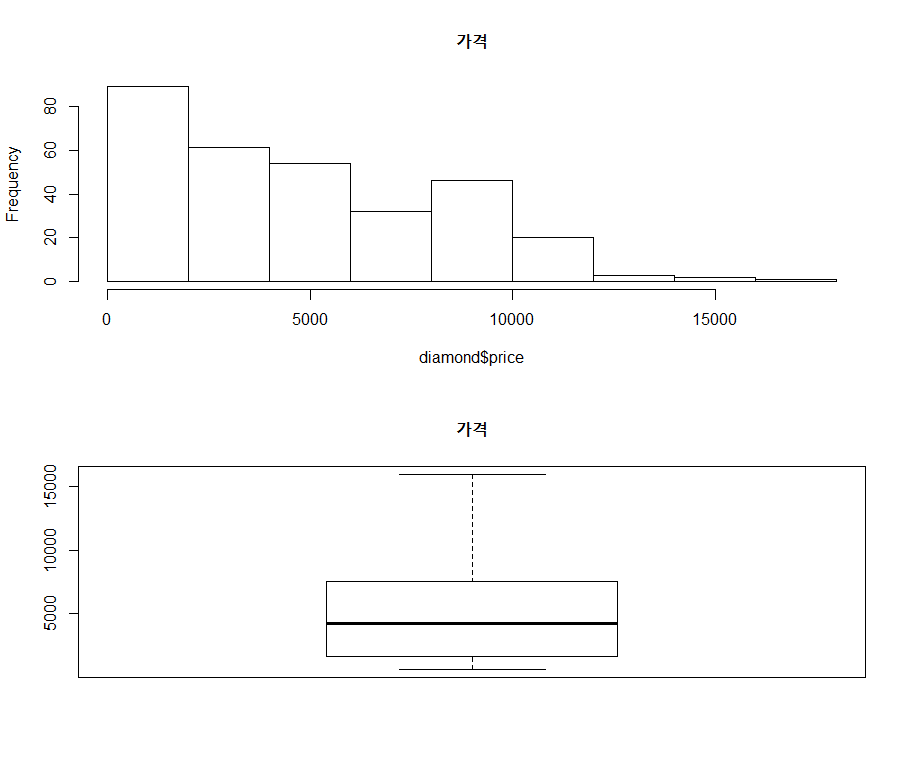
#clarity: 투명도

#certification: 인증기관

#price: 가격(싱가포르$)

# url = <https://vincentarelbundock.github.io/Rdatasets/csv/Ecdat/Diamond.csv>

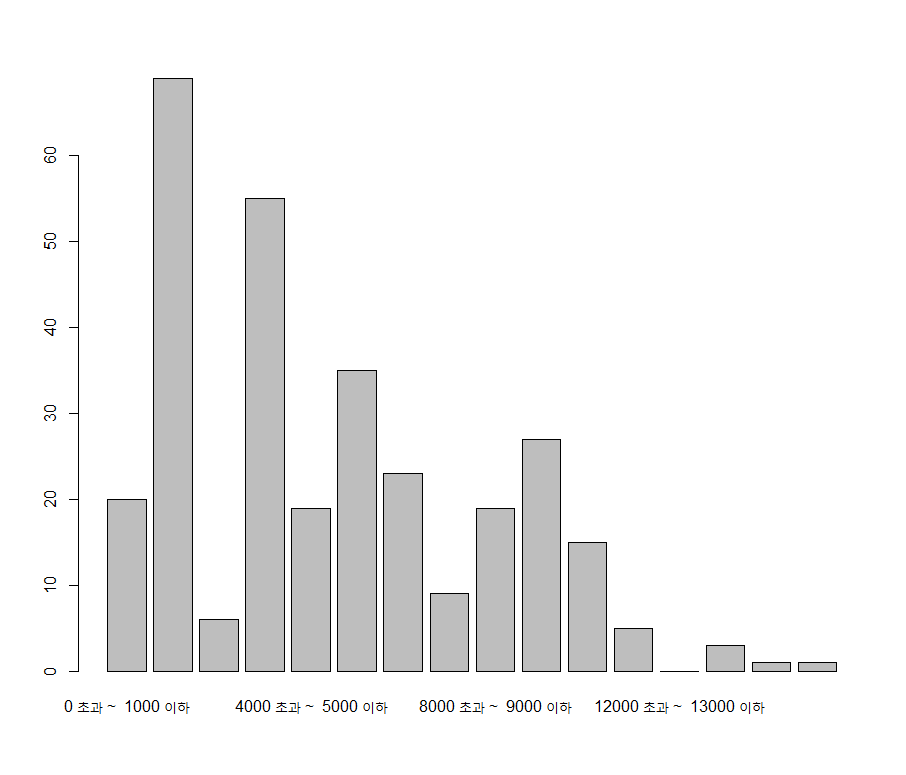
# 위의 주소를 통해 read.csv()로 읽어 사용해보세요.



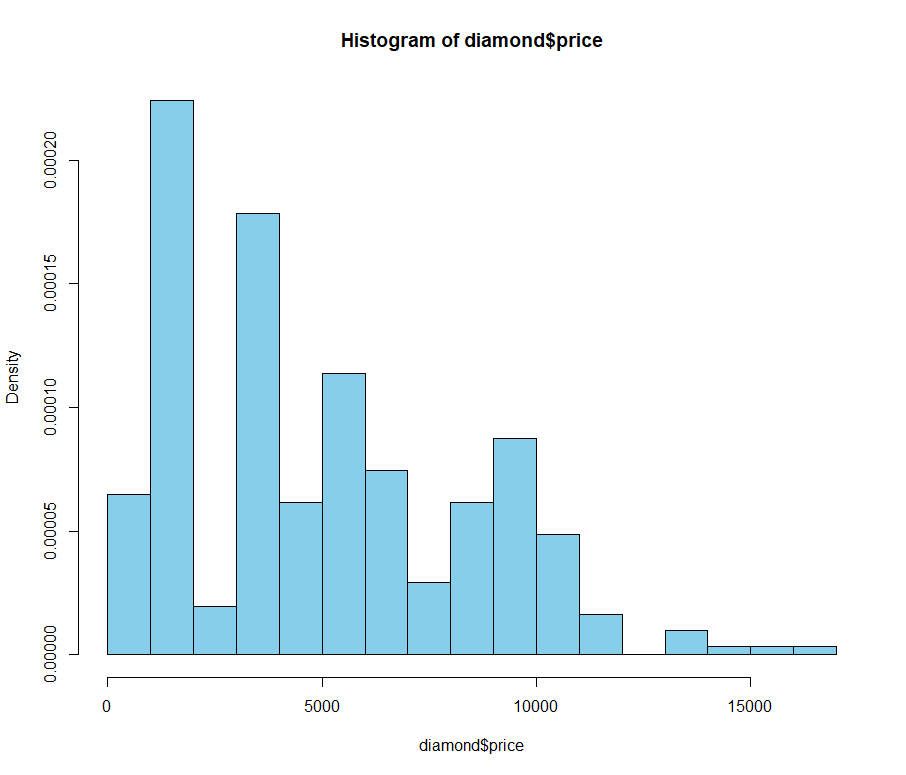
기본 hist(), boxplot(), 스크린 2개

데이터: 가격

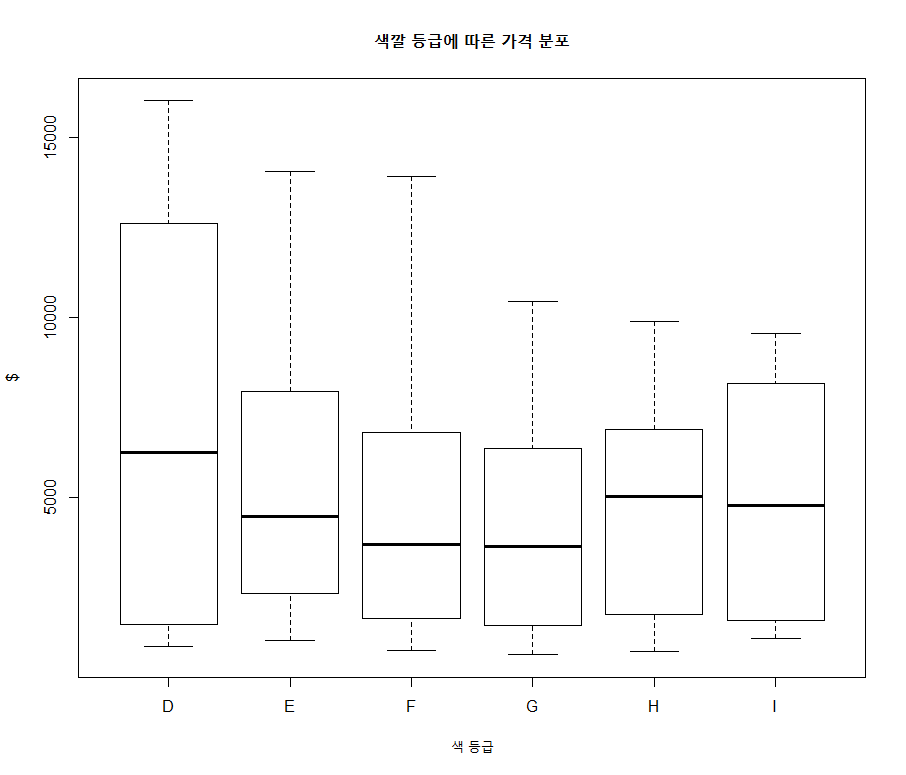
\*추가로 stem()도 사용해보세요



가격 데이터에 관한 범주형 변수 하나 생성한 뒤, 1000단위로 끊고, 히스토그램으로 만들어 보세요



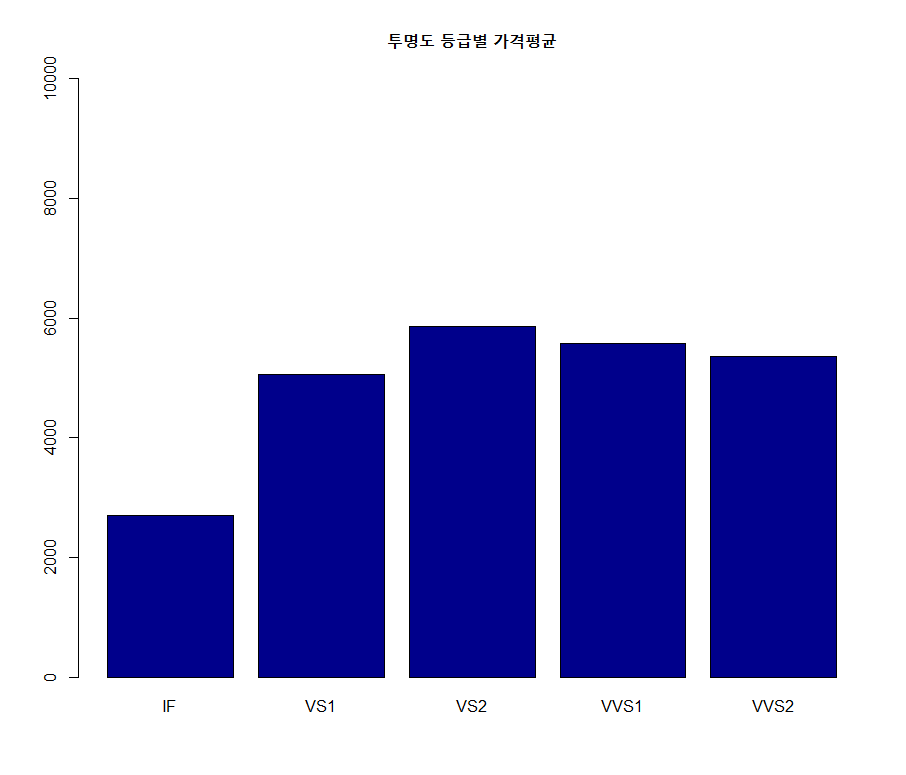
가격 변수를 히스토그램의 breaks 속성으로 1000단위씩 끊어보세요.



그룹간 비교

X축: 색 등급(colour)

y축: 가격

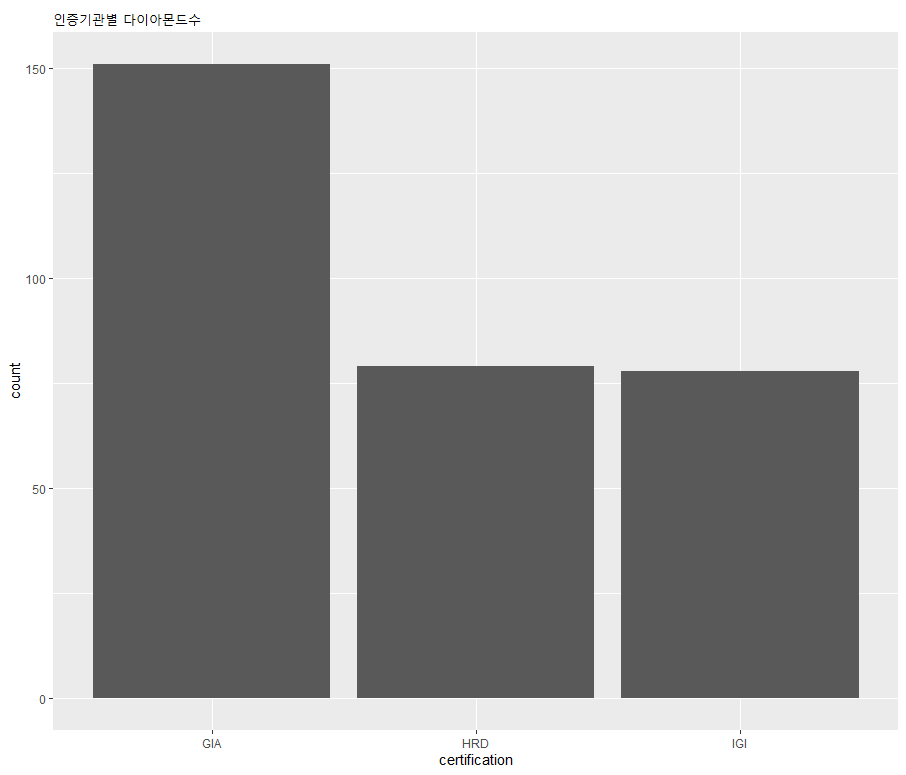


여러 연속자료 비교

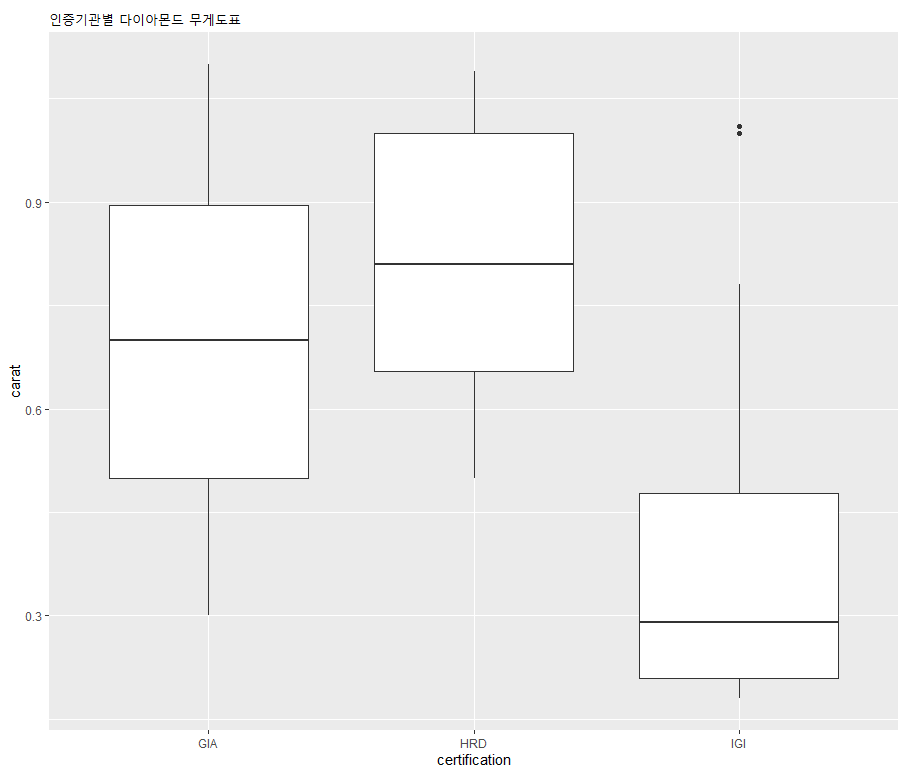
X 축: 투명도(clarity)

Y 축: 가격 평균\*

Ggplot2



데이터: 인증기관



X축: 인증기관

Y축: 무게(carat)

이 이외에도 자유롭게 데이터를 분석해보시면 도움이 많이 될 겁니다. 수고하세요!!