**2019 - 2 데이터 분석 기초**

**파이널 프로젝트 보고서**

**미디어학부**

**201221090 배문규**

**1. 프로젝트 개요**

**1) 주제선정배경**

체질량지수(Body Mass Index)란 것이 있습니다. 체중(kg)을 키(m)의 제곱으로 나눈 수치입니다. 비만 여부를 가늠하는데 널리 사용합니다. 체질량지수가 18.5 이하면 저체중입니다. 18.5-23이 정상입니다. 23-25면 과체중이며 25를 넘어가면 비만입니다. 30 이상이면 특별히 고도비만이라고 말합니다. 조금 복잡합니다만 25를 기억하면 쉽습니다. 25를 넘어가면 뚱뚱하고 25 아래면 괜찮다고 보는 것이지요. 그런데 최근 이러한 기준에 문제가 있다는 지적들이 꾸준히 제기되고 있습니다.

미국 CNN방송은 미국 일리노이 주에 사는 한 건강한 10대 배구 선수 소녀가 매일 운동해 탄탄한 몸을 가졌는데도, 의사는 살이 조금이라도 더 붙으면 BM1 지수가 ‘정상’ 범위를 벗어난다며 경고한다는 사연을 소개했습니다. 이 소녀는 “나는 과체중이나 비만이 아닌데도, 숫자가 ‘다른 얘기’를 하니 당황스럽다”며 BMI의 신뢰성을 의심했다.  
  
그런데 사실 BMI 계산법은 1800년대 사람들의 전반적인 비만 정도를 측정하기 위해 처음 고안됐습니다. 이후 1972년 한 연구에서, 7400명 이상을 대상으로 비만율을 계산할 때 BMI 계산법이 가장 유용하다는 사실이 밝혀졌습니다. BMI란 용어도 이 연구에서 처음 등장했습니다. 연구를 진행한 안셀 키즈 교수는 BMI는 대규모 집단의 비만 정도를 측정하는 연구에는 적합하지만, 개인의 비만 정도는 알 수 없다고 적시했습니다.

그래서 1980년대까지 개인의 불어난 살이 건강에 얼마나 위협적인지를 판단하는 여러 방법들이 고안됐습니다. 그러나 방법이 너무 다양해 하나로 통일할 수 없었습니다. 반면에, BMI는 앞서 일리노이주 배구 소녀의 경우에서도 드러났듯이 많은 오류를 담고 있지만, 계산법이 워낙 간단해 앞서 많은 의료 전문가가 개별 환자의 비만 척도로 이용하기 시작했습니다. 또 당시 사람들의 몸집이 점점 커지면서 자신의 건강 상태를 알고 싶은 사람들에게 간편한 BMI 계산법이 유행하기 시작했다고 합니다.  
  
하지만 BMI는 집단 연구에만 적합할 뿐, 개인의 건강과 비만 정도에 대해선 ‘빙산의 일각’만 알려줍니다. 심지어 일리노이주 소녀의 경우처럼 그릇된 진단을 내릴 수도 있습니다. 운동 선수들은 근육량이 많아, 체지방량이나 근육량을 구분하지 않는 BMI 계산법은 틀릴 수 밖에 없습니다. 예를 들어, 축구 선수들의 평균 BMI 지수는 31.35로 ‘비만’에 속합니다. 이들의 몸이 ‘비만’을 결정짓는 체지방보다, 근육 위주라는 것은 자명한 사실이죠.  
  
또 외견상 말랐지만 체지방이 많아 ‘마른 비만’에 속하는 사람들은 BMI로는 ‘비만’ 진단을 받을 수 없습니다. 팔 다리에는 살이 없지만 복부에 체지방이 많아 배가 불룩 나온 사람들이 ‘마른 비만’ 체형인데, 이 복부 지방이 다른 어떤 병보다 건강에 치명적이라는 사실은 이미 많은 연구에서 입증됐습니다.  
의학 저널인 ‘공공보건의 프런티어(Frontiers in Public Health)’에 게재된 한 논문에 따르면, BMI로만 비만 정도를 측정하면 실제 위험한 수치의 체지방량을 가진 환자들의 50%가 ‘정상’으로 나왔다고 합니다.

전문의들은 세계보건기구(WHO)가 정한 세계 비만기준보다 엄격한 국내 비만지수를 재조정할 필요가 있다는 데 공감합니다. 우리나라는 세계 비만기준이 아닌 아시아태평양 지역 비만기준을 채택했는데 이에 따르면 BMI 25부터 비만입니다. 세계 비만기준으로는 BMI 30이 비만입니다. 과체중 기준도 야박합니다. 과체중은 세계 비만기준으로는 BMI 25~29.9이지만 우리는 BMI 23~24.9입니다.

예컨대 키가 170cm라면 외국에선 87kg 이상부터 비만이지만 우리나라에선 73kg만 되어도 비만입니다. 키가 160cm라면 외국에서 77kg이 넘어야 비만이지만 우리나라에서 64kg만 넘어도 비만입니다.

우리가 아ㆍ태 비만기준을 채택한 것은 동ㆍ서양의 체형 차이에 주목했기 때문입니다. 서양인과 달리 동양인은 체중이 급증하면 당뇨병 등 대사질환과 암에 걸릴 가능성이 높아 비만지수를 엄격히 적용한 아ㆍ태 비만지수를 따랐습니다. 여기에 동양인은 서양인보다 육류 섭취가 적어 고기를 더 많이 먹게 되면 과잉지방에 취약해 비만기준을 BMI 25로 정했습니다. 오상우 동국대일산병원 가정의학과 교수는 “1998년 당시 국내에서 BMI 25 인구가 빠르게 증가해 아ㆍ태 비만기준을 채택한 것”이라면서 “하지만 우리도 미국처럼 BMI 25 인구가 정체되고, BMI 30 이상 비만인구가 급증해 비만지수를 개선할 다각적 연구가 필요하다”고 했습니다. 이 기준대로라면 우리나라 성인 3명중 1명이 체질량지수 25 이상의 비만에 해당되기 때문입니다.

대표적인 사례가 최근 사회적 파장을 일으킨 육군 간부 BMI조사입니다. 계명ㆍ대구대 간호대학 연구팀이 육군 간부 1,026명의 BMI를 분석한 결과, 34.9%가 비만, 25.9%가 과체중이었습니다. 하지만 전문의들은 “군인들의 비만 여부를 BMI 잣대만 가지고 판정 내린 것 자체가 무리”라고 지적합니다. BMI는 비만을 측정하는 다양한 지표의 하나일 뿐 절대적 지표가 아니라는 것입니다. 전문의들은 “현재 BMI 23부터 과체중으로 돼 있는데 현실적 적당하지 않다”면서 “세계 비만기준으로 조사했으면 결과가 달리 나타났을 것”이라고 했습니다. 유순집 부천성모병원 내분비내과 교수(대한비만학회 이사장)는 “BMI만으로 비만 여부를 가리려는 사회풍토가 문제”라면서 “BMI에 전적으로 의존하고 있는 비만 관련 지표를 개선할 필요가 있다”고 말했습니다. 박수경 서울대 의대 예방의학과 교수는 “BMI가 높으면 당뇨병, 고혈압 등 대사질환과 암에 걸릴 가능성이 높지만 상대 지표이기 때문에 맹신해서는 곤란하다”고 했습니다. 조정진 한림대동탄성심병원 가정의학과 교수는 “BMI는 체지방을 완전히 반영하지 못하고, 체지방의 체내분포도 반영하지 못하는 한계가 있다”고 말했습니다. 키와 몸무게를 이용해 비만여부를 판정하는 ‘겉보기 비만지수’이라는 것입니다.

서울대의대 예방의학과 유근영 교수팀이 한국인 1만6000명을 포함해 아시아 7개국 114만명을 대상으로 조사한 결과, 아시아인은 BMI가 22.6~27.5인 사람이 건강 악화로 인한 사망위험도가 가장 낮은 것으로 나타났습니다. 이 수치는 과체중에 근접한 사람부터 비만이 상당히 진행된 사람에 해당하는 범위입니다다. 즉, 아시아인의 경우 BMI 기준으로 약간 뚱뚱한 편에 속하는 사람이 더 오래 산다는 결과입니다.

게다가 체질량지수는 계산이 복잡합니다. 저는 체질량지수보다 줄자를 준비해달라는 말씀을 드립니다. "허리둘레 < 허벅지 제일 굵은 둘레+종아리 제일 굵은 둘레"를 목표로 하면 됩니다. 쉽게 말해 허리는 날씬하게 다리는 굵게 체형을 유지하자는 것입니다.

복잡하게 체질량지수를 계산할 이유가 없습니다. 몸에 나쁜 비만은 팔다리나 몸통의 피하지방보다 내장 사이에 낀 복부비만입니다. 그리고 근육의 대부분이 하체에 몰려있습니다. 그래서 "건강하게 오래 살려면 당신의 다리가 허리를 이기도록 하라"는 의학격언이 있습니다. 아무쪼록 체질량지수보다 날씬한 허리와 굵은 다리에 주목하시기 바랍니다.

◇ 건강 위협하는 내장지방 잡아내는 ‘허리둘레’

줄자로 허리둘레 측정하기

최근 건강검진 항목에 ‘허리둘레’가 추가 반영된 것은 복부비만 정도를 알아보기 위해서입니다. 성인을 기준으로 허리둘레가 여성은 85cm(33.46인치) 이상일 때, 남성은 90cm(35.43인치) 이상일 때 복부비만으로 진단합니다.

허리둘레를 측정할 때는 양발 간격을 25~30cm 정도 벌리고 서서 체중을 두 다리에 균등하게 분배한 후 숨을 편안히 내쉰 상태에서 진행합니다. 줄자는 갈비뼈의 가장 아랫부분과 골반의 가장 높은 부분의 중간부위를 측정하며 줄자가 압력을 주지 않을 정도로 느슨하게 하여 소수점 한자리까지 측정합니다

저는 올해 1학기 교내교양수업 운동과 건강관리를 수강하면서 과제로 국민체력100을 화성에 국민건강센터에 가서 체력검정을 받고왔습니다. 그런데 모든 체력평가 항목에서 1-2등급을 받았지만 BMI는 정상수치를 벗어나 경도비만을 받았습니다. 그렇기때문에 체중감량을 진단받았는데 인바디를 보니 체지방량은 정상이고 근육량이 과다여서 체중이 많이나간것이였습니다.

이처럼 BMI는 체지방과 제지방량을 구분하지않기 때문에 근육량으로 체중이나가거나 지방으로 체중이나가는지 알지못합니다. 그렇기 때문에 한개인의 신체평가지수로 적당하지 않다고 생각합니다. 그렇다면 2017년 국민체력실태조사 데이터에서 이 BMI가 체력과는 어느정도 상관관계가 있는지 분석해보고, 그리고 어떠한 요인들이 BMI보다 개인의 체력평가와 밀접한 영향이있는지 또한 분석 해보려고합니다.

먼저 데이터를 살펴보고 데이터를 정제한뒤 그래프를 이용해 시각적인 데이터간의 상관관계도 살펴보고 가설을 세워 교차분석, 상관분석, 연관분석을 거쳐서 가설의 결론을 도출하려고합니다.

1. [출처: 국제-조선일보] '비만 측정' 헛점 많은 1800년대 BMI 지수가 아직도 쓰이는 까닭은, <https://m.chosun.com/svc/article.html?sname=news&contid=2017081702861>

2. [출처] : BMI (비만도) 계산법, BMI는 당신을 속이고 있다?  
<https://www.hidoc.co.kr/healthstory/news/C0000479924> | 하이닥



**주요단어 텍스트분석 워드클라우드 결과**

**2) 데이터 분석 목적**

건강한 10대 운동선수 소녀가 매일 운동해 탄탄한 몸을 가졌는데도, 체중이 조금이라도 더 붙으면 BMI 지수가 ‘정상’ 범위를 벗어나 ‘경도비만’이 될수도 있습니다. 이러한 예시처럼 BMI가 가지고있는 허점을 2017년 국민체력실태조사 데이터를 바탕으로하여 분석해보고, 전문의들이 BMI를 재조정할 필요가 있다는 주장을 데이터를 통해 확인해보는것이 목적입니다.

그리고 BMI 뿐만아니라 함께 측정된 체지방률과 허리둘레데이터, 체력평가데이터를 이용하여 BMI와 체지방률, 허리둘레중 BMI가 다른 항목과 비교하여 체력과 연관성이 어느정도 있는지 알아보는것이 데이터 분석 목적입니다.

**3) 데이터 수집 전략**

별다른 데이터 수집 전략 없이 통계청 마이크로데이터 사이트(<https://mdis.kostat.go.kr/>)에서 공개한 2017년 국민체력실태조사 데이터를 수집했습니다.

**2. 데이터 분석 수행**

**1) 데이터 준비 및 데이터 탐색**

통계청 마이크로데이터 사이트(<https://mdis.kostat.go.kr/>)에서 공개한 2017년 국민체력실태조사 데이터를 수집하여 준비했습니다.

데이터는 총 4292명이고 남성 2146명, 여성 2146명입니다. 그리고 성별별로 19세부터 64세까지 연령이 골고루 분포되어있습니다.

체력조사데이터 중 분석목적과 관계없는 '지역' 을 제외한 모든 항목 추출했고 추출된 항목은 15가지로, 신체적항목 8가지와 체력적항목 7가지로 나눠져있습니다.

신체적 항목 8가지: 성별, 연령, 체중, 신장, BMI, 허리둘레, 체지방량, 연령집단

체력적 항목 7가지 : 윗몸일으키기, 악력, 상대악력, 제자리멀리뛰기, 20m왕복오래달리기, 10m왕복빨리달리기, 좌전굴

- 연령집단 19-24살,25-29살...60-64살까지 각 연령별로 적용되는 체력기준이 다른 연령집단을 9집단으로 나눈뒤 각 기준 적용하기위해 생성된 데이터입니다.

- 윗몸일으키기는 1분간 시행한횟수이고, 제자리멀리뛰기는 cm단위입니다.

- 악력은 순수 악력의힘을 측정한것이고, 상대악력은 순수악력을 체중과의 관계로 도출해낸 수치입니다.

- 좌전굴은 다리뻗고 앉아서 상체를 숙여 유연성을 체크하는것이고 숫자가 클수록 유연한것입니다.

- 10m왕복 빨리달리기는 10m를 빨리 달리는시간을 잰것이라 시간이 적을수록 체력이 좋은것입니다.

- 20m왕복 오래달리기는 20m 거리를 일정 시간간격마다 계속 달리는것이며, 만약 30이면 20m를 총 30번 뛴것이기때문에 숫자가 클수록 체력이 좋은것입니다.

데이터 탐색 목적은 BMI, 체지방량, 허리둘레가 성별별로 상관관계 및 분포를 알아보고 각 성별별로 어떤 차이가있는지 분석해보고 각 성별별로 BMI와 체지방률, 허리둘레중 어떤 항목이 다른 항목들과 더 밀접한 상관관계가 있는지 알아보고 적절한 가설을 세우고 분석를 진행하는것입니다.

*#csv파일불러와서 데이터생성*  
mdis <- **read.table**("국민체력실태조사\_성인(제공)\_2017\_20191112\_39075.csv",  
 header=FALSE, sep=",", colClasses = **c**("character", "numeric", "character", "numeric",   
 "numeric", "numeric", "numeric", "numeric", "numeric",   
 "numeric", "numeric", "numeric", "numeric", "numeric",   
 "numeric", "character"))  
*### 데이터 전처리* ###  
mdis**$**V3<- **recode\_factor**(mdis**$**V3, '1'="남자")  
mdis**$**V3<- **recode\_factor**(mdis**$**V3, '2'="여자")  
**colnames**(mdis) = **c**("지역", "연령", "성별", "신장", "체중", "BMI", "체지방률", "허리둘레", "윗몸일으키기",   
 "상대악력", "악력", "제자리멀리뛰기", "왕복오래달리기\_20m", "좌전굴", "왕복빨리달리기\_10m", "연령집단")  
  
*##필요한데이터만 추출(지역데이터 제외)*  
health <- **subset**(mdis, select=**c**(BMI, 체지방률, 성별, 연령, 체중, 신장, 허리둘레, 윗몸일으키기, 상대악력, 악력, 좌전굴, 제자리멀리뛰기, 왕복오래달리기\_20m, 왕복빨리달리기\_10m, 연령집단))  
   
*##성별별로 데이터분리*  
*#남성데이터 추출*  
health\_male <- health **%>%**  
 **filter**(성별 **==** "남자")  
*#남성만 추출한 데이터에서 성별항목제거 - 모든 데이터를 numeric타입으로 만들기 위해*  
health\_male <- **subset**(health\_male, select = **-**성별)  
*#남성데이터에서 BMI, 체지방률, 허리둘레를 대한비만학회의 기준에따라 등급분류*  
male\_eval <- health\_male **%>%**  
 **mutate**(분류\_BMI = **ifelse**(BMI **<**18.5, 1,  
 **ifelse**(BMI **<** 23, 2,  
 **ifelse**(BMI **<** 25, 3,  
 **ifelse**(BMI **<** 30, 4,5)))),*#대한비만학회 기준*  
 분류\_체지방률 = **ifelse**(체지방률 **<**15, 1,  
 **ifelse**(체지방률 **<** 20, 2,  
 **ifelse**(체지방률 **<** 25, 3,4))),*#대한비만학회 기준*  
 분류\_허리둘레 = **ifelse**(허리둘레 **>=** 95,3,  
 **ifelse**(허리둘레 **>=** 90,2,1)))*#WHO 아시아태평양 지역기준 - 디테일한 등급위해 95이상(고도복부비만) 추가*  
*#연령집단 19-24살,25-29살...60-64살까지 각 연령별로 적용되는 기준이 다른 연령집단을 9집단으로 나눈뒤 각 기준 적용*  
*#1=상(상위 33%),2=중(상위 33%-66%), 3=하(하위 33%)*  
male\_1 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 1) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**43,3,  
 **ifelse**(윗몸일으키기**<**55,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**57.4,3,  
 **ifelse**(상대악력**<**65.8,2,1)),  
 등급\_악력 = **ifelse**(악력**<**40.6,3,  
 **ifelse**(악력**<**46.2,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**7.1,3,  
 **ifelse**(좌전굴**<**14.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**209.1,3,  
 **ifelse**(제자리멀리뛰기**<**234.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**43,3,  
 **ifelse**(왕복오래달리기\_20m**<**61,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**10,1,  
 **ifelse**(왕복빨리달리기\_10m**<**11,2,3)))  
male\_2 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 2) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**41,3,  
 **ifelse**(윗몸일으키기**<**52,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**55.4,3,  
 **ifelse**(상대악력**<**65.8,2,1)),  
 등급\_악력 = **ifelse**(악력**<**41.2,3,  
 **ifelse**(악력**<**47.7,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**6.0,3,  
 **ifelse**(좌전굴**<**15.0,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**201.1,3,  
 **ifelse**(제자리멀리뛰기**<**226.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**35,3,  
 **ifelse**(왕복오래달리기\_20m**<**54,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**10.3,1,  
 **ifelse**(왕복빨리달리기\_10m**<**11.3,2,3)))  
male\_3 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 3) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**39,3,  
 **ifelse**(윗몸일으키기**<**50,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**56.0,3,  
 **ifelse**(상대악력**<**65.2,2,1)),  
 등급\_악력 = **ifelse**(악력**<**42.9,3,  
 **ifelse**(악력**<**49.9,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**7.3,3,  
 **ifelse**(좌전굴**<**14.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**199.1,3,  
 **ifelse**(제자리멀리뛰기**<**225.5,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**34,3,  
 **ifelse**(왕복오래달리기\_20m**<**51,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**10.4,1,  
 **ifelse**(왕복빨리달리기\_10m**<**11.5,2,3)))  
  
male\_4 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 4) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**35,3,  
 **ifelse**(윗몸일으키기**<**47,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**56.7,3,  
 **ifelse**(상대악력**<**64.7,2,1)),  
 등급\_악력 = **ifelse**(악력**<**42.0,3,  
 **ifelse**(악력**<**47.8,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**4.6,3,  
 **ifelse**(좌전굴**<**14.7,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**196.5,3,  
 **ifelse**(제자리멀리뛰기**<**215.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**32,3,  
 **ifelse**(왕복오래달리기\_20m**<**45,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**10.7,1,  
 **ifelse**(왕복빨리달리기\_10m**<**11.8,2,3)))  
male\_5 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 5) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**33,3,  
 **ifelse**(윗몸일으키기**<**44,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**54.4,3,  
 **ifelse**(상대악력**<**63.6,2,1)),  
 등급\_악력 = **ifelse**(악력**<**41.6,3,  
 **ifelse**(악력**<**48.4,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**3.8,3,  
 **ifelse**(좌전굴**<**14.0,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**191.1,3,  
 **ifelse**(제자리멀리뛰기**<**212.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**30,3,  
 **ifelse**(왕복오래달리기\_20m**<**42,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**11,1,  
 **ifelse**(왕복빨리달리기\_10m**<**12,2,3)))  
male\_6 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 6) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**32,3,  
 **ifelse**(윗몸일으키기**<**42,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**55.8,3,  
 **ifelse**(상대악력**<**63.9,2,1)),  
 등급\_악력 = **ifelse**(악력**<**41.5,3,  
 **ifelse**(악력**<**46.9,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**5.1,3,  
 **ifelse**(좌전굴**<**13.1,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**185.1,3,  
 **ifelse**(제자리멀리뛰기**<**203.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**26,3,  
 **ifelse**(왕복오래달리기\_20m**<**39,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**11.4,1,  
 **ifelse**(왕복빨리달리기\_10m**<**12.4,2,3)))  
male\_7 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 7) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**30,3,  
 **ifelse**(윗몸일으키기**<**39,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**55.2,3,  
 **ifelse**(상대악력**<**64.1,2,1)),  
 등급\_악력 = **ifelse**(악력**<**39.4,3,  
 **ifelse**(악력**<**44.8,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**5.9,3,  
 **ifelse**(좌전굴**<**13.4,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**177.1,3,  
 **ifelse**(제자리멀리뛰기**<**195.5,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**26,3,  
 **ifelse**(왕복오래달리기\_20m**<**36,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**11.6,1,  
 **ifelse**(왕복빨리달리기\_10m**<**12.7,2,3)))  
male\_8 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 8) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**28,3,  
 **ifelse**(윗몸일으키기**<**37,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**53.5,3,  
 **ifelse**(상대악력**<**62.0,2,1)),  
 등급\_악력 = **ifelse**(악력**<**38.0,3,  
 **ifelse**(악력**<**43.0,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**7.1,3,  
 **ifelse**(좌전굴**<**14.7,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**168.1,3,  
 **ifelse**(제자리멀리뛰기**<**186.5,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**21,3,  
 **ifelse**(왕복오래달리기\_20m**<**33,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**12.0,1,  
 **ifelse**(왕복빨리달리기\_10m**<**13.2,2,3)))  
male\_9 <- male\_eval **%>%**  
 **filter**(연령집단 **==** 9) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**23,3,  
 **ifelse**(윗몸일으키기**<**35,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**53.1,3,  
 **ifelse**(상대악력**<**60.3,2,1)),  
 등급\_악력 = **ifelse**(악력**<**36.2,3,  
 **ifelse**(악력**<**41.3,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**7.1,3,  
 **ifelse**(좌전굴**<**13.1,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**161.1,3,  
 **ifelse**(제자리멀리뛰기**<**182.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**17,3,  
 **ifelse**(왕복오래달리기\_20m**<**28,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**12.6,1,  
 **ifelse**(왕복빨리달리기\_10m**<**13.8,2,3)))  
*#각집단으로 나눠진 데이터를 하나의 데이터로 통합*  
male\_eval <- **rbind**(male\_1, male\_2)  
male\_eval <- **rbind**(male\_eval, male\_3)  
male\_eval <- **rbind**(male\_eval, male\_4)  
male\_eval <- **rbind**(male\_eval, male\_5)  
male\_eval <- **rbind**(male\_eval, male\_6)  
male\_eval <- **rbind**(male\_eval, male\_7)  
male\_eval <- **rbind**(male\_eval, male\_8)  
male\_eval <- **rbind**(male\_eval, male\_9)  
*#등급과 평가만 남겨두고 나머지 numeric 기록데이터들은 제거*  
male\_eval <- **subset**(male\_eval,select = **c**(분류\_BMI,분류\_체지방률,분류\_허리둘레,등급\_윗몸일으키기,등급\_좌전굴,등급\_제자리멀리뛰기,등급\_10m왕복빨리달리기,등급\_20m왕복오래달리기))  
  
  
*##여성도 남성과 똑같이 데이터 처리*  
health\_female <- health **%>%**  
 **filter**(성별 **==** "여자")  
health\_female <- **subset**(health\_female, select = **-**성별)  
female\_eval <- health\_female **%>%**  
 **mutate**(분류\_BMI = **ifelse**(BMI **<**18.5, 1,  
 **ifelse**(BMI **<** 23, 2,  
 **ifelse**(BMI **<** 25, 3,  
 **ifelse**(BMI **<** 30, 4,5)))),*#대한비만학회 기준*  
 분류\_체지방률 = **ifelse**(체지방률 **<**18, 1,  
 **ifelse**(체지방률 **<** 27, 2,  
 **ifelse**(체지방률 **<** 33, 3,4))),*#대한비만학회 기준*  
 분류\_허리둘레 = **ifelse**(허리둘레 **>=** 90,3,  
 **ifelse**(허리둘레 **>=** 85,2,1)))*#대한비만학회 기준*  
female\_1 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 1) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**26,3,  
 **ifelse**(윗몸일으키기**<**37,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**42.2,3,  
 **ifelse**(상대악력**<**50.8,2,1)),  
 등급\_악력 = **ifelse**(악력**<**23.4,3,  
 **ifelse**(악력**<**28.5,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**12.6,3,  
 **ifelse**(좌전굴**<**20.6,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**139.7,3,  
 **ifelse**(제자리멀리뛰기**<**165.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**19,3,  
 **ifelse**(왕복오래달리기\_20m**<**30,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**12.2,1,  
 **ifelse**(왕복빨리달리기\_10m**<**13.7,2,3)))  
female\_2 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 2) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**24,3,  
 **ifelse**(윗몸일으키기**<**36,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**41.4,3,  
 **ifelse**(상대악력**<**51.1,2,1)),  
 등급\_악력 = **ifelse**(악력**<**23.6,3,  
 **ifelse**(악력**<**27.8,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**11.5,3,  
 **ifelse**(좌전굴**<**21.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**140.1,3,  
 **ifelse**(제자리멀리뛰기**<**161.7,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**18,3,  
 **ifelse**(왕복오래달리기\_20m**<**29,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**12.3,1,  
 **ifelse**(왕복빨리달리기\_10m**<**13.8,2,3)))  
female\_3 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 3) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**24,3,  
 **ifelse**(윗몸일으키기**<**34,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**42.0,3,  
 **ifelse**(상대악력**<**49.0,2,1)),  
 등급\_악력 = **ifelse**(악력**<**24.2,3,  
 **ifelse**(악력**<**28.4,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**13.1,3,  
 **ifelse**(좌전굴**<**19.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**139.1,3,  
 **ifelse**(제자리멀리뛰기**<**158.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**18,3,  
 **ifelse**(왕복오래달리기\_20m**<**25,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**12.8,1,  
 **ifelse**(왕복빨리달리기\_10m**<**14.0,2,3)))  
female\_4 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 4) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**18,3,  
 **ifelse**(윗몸일으키기**<**30,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**41.6,3,  
 **ifelse**(상대악력**<**49.4,2,1)),  
 등급\_악력 = **ifelse**(악력**<**23.7,3,  
 **ifelse**(악력**<**28.0,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**10.0,3,  
 **ifelse**(좌전굴**<**17.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**134.1,3,  
 **ifelse**(제자리멀리뛰기**<**155.3,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**16,3,  
 **ifelse**(왕복오래달리기\_20m**<**24,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**13.2,1,  
 **ifelse**(왕복빨리달리기\_10m**<**14.9,2,3)))  
female\_5 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 5) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**21,3,  
 **ifelse**(윗몸일으키기**<**31,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**41.0,3,  
 **ifelse**(상대악력**<**50.2,2,1)),  
 등급\_악력 = **ifelse**(악력**<**24.4,3,  
 **ifelse**(악력**<**28.5,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**11.1,3,  
 **ifelse**(좌전굴**<**17.8,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**135.1,3,  
 **ifelse**(제자리멀리뛰기**<**155.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**16,3,  
 **ifelse**(왕복오래달리기\_20m**<**23,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**13.0,1,  
 **ifelse**(왕복빨리달리기\_10m**<**14.0,2,3)))  
female\_6 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 6) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**20,3,  
 **ifelse**(윗몸일으키기**<**30,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**41.3,3,  
 **ifelse**(상대악력**<**48.7,2,1)),  
 등급\_악력 = **ifelse**(악력**<**23.8,3,  
 **ifelse**(악력**<**28.0,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**10.4,3,  
 **ifelse**(좌전굴**<**19.4,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**130.1,3,  
 **ifelse**(제자리멀리뛰기**<**150.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**14,3,  
 **ifelse**(왕복오래달리기\_20m**<**21,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**13.5,1,  
 **ifelse**(왕복빨리달리기\_10m**<**14.4,2,3)))  
female\_7 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 7) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**16,3,  
 **ifelse**(윗몸일으키기**<**26,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**39.5,3,  
 **ifelse**(상대악력**<**47.3,2,1)),  
 등급\_악력 = **ifelse**(악력**<**23.2,3,  
 **ifelse**(악력**<**26.6,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**13.0,3,  
 **ifelse**(좌전굴**<**19.4,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**122.1,3,  
 **ifelse**(제자리멀리뛰기**<**142.5,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**13,3,  
 **ifelse**(왕복오래달리기\_20m**<**20,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**13.7,1,  
 **ifelse**(왕복빨리달리기\_10m**<**14.9,2,3)))  
female\_8 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 8) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**12,3,  
 **ifelse**(윗몸일으키기**<**21,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**37.9,3,  
 **ifelse**(상대악력**<**46.0,2,1)),  
 등급\_악력 = **ifelse**(악력**<**21.7,3,  
 **ifelse**(악력**<**25.9,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**13.4,3,  
 **ifelse**(좌전굴**<**19.9,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**113.1,3,  
 **ifelse**(제자리멀리뛰기**<**133.5,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**12,3,  
 **ifelse**(왕복오래달리기\_20m**<**17,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**14.1,1,  
 **ifelse**(왕복빨리달리기\_10m**<**15.4,2,3)))  
female\_9 <- female\_eval **%>%**  
 **filter**(연령집단 **==** 9) **%>%**  
 **mutate**(등급\_윗몸일으키기 = **ifelse**(윗몸일으키기**<**8,3,  
 **ifelse**(윗몸일으키기**<**17,2,1)),  
 등급\_상대악력 = **ifelse**(상대악력**<**38.4,3,  
 **ifelse**(상대악력**<**45.2,2,1)),  
 등급\_악력 = **ifelse**(악력**<**22.3,3,  
 **ifelse**(악력**<**25.7,2,1)),  
 등급\_좌전굴 = **ifelse**(좌전굴**<**13.6,3,  
 **ifelse**(좌전굴**<**20.4,2,1)),  
 등급\_제자리멀리뛰기 = **ifelse**(제자리멀리뛰기**<**101.1,3,  
 **ifelse**(제자리멀리뛰기**<**126.1,2,1)),  
 등급\_20m왕복오래달리기 = **ifelse**(왕복오래달리기\_20m**<**10,3,  
 **ifelse**(왕복오래달리기\_20m**<**15,2,1)),  
 등급\_10m왕복빨리달리기 = **ifelse**(왕복빨리달리기\_10m**<**15.0,1,  
 **ifelse**(왕복빨리달리기\_10m**<**16.2,2,3)))  
  
female\_eval <- **rbind**(female\_1, female\_2)  
female\_eval <- **rbind**(female\_eval, female\_3)  
female\_eval <- **rbind**(female\_eval, female\_4)  
female\_eval <- **rbind**(female\_eval, female\_5)  
female\_eval <- **rbind**(female\_eval, female\_6)  
female\_eval <- **rbind**(female\_eval, female\_7)  
female\_eval <- **rbind**(female\_eval, female\_8)  
female\_eval <- **rbind**(female\_eval, female\_9)  
female\_eval <- **subset**(female\_eval,select = **c**(분류\_BMI,분류\_체지방률,분류\_허리둘레,등급\_윗몸일으키기,등급\_좌전굴,등급\_제자리멀리뛰기,등급\_10m왕복빨리달리기,등급\_20m왕복오래달리기))  
  
  
*#소수점데이터 소수점한자리까지 반올림*  
health**$**BMI <- **round**(health**$**BMI, digits = 0)  
health**$**체지방률 <- **round**(health**$**체지방률, digits = 0)  
health**$**체중 <- **round**(health**$**체중, digits = 0)  
health**$**신장 <- **round**(health**$**신장, digits = 0)  
health**$**허리둘레 <- **round**(health**$**허리둘레, digits = 0)  
health**$**상대악력 <- **round**(health**$**상대악력, digits = 0)  
health**$**악력 <- **round**(health**$**악력, digits = 0)  
health**$**좌전굴 <- **round**(health**$**좌전굴, digits = 0)  
health**$**제자리멀리뛰기 <- **round**(health**$**제자리멀리뛰기, digits = 0)  
health**$**왕복빨리달리기\_10m <- **round**(health**$**왕복빨리달리기\_10m, digits = 0)  
  
*#수치가 반올림된 남성데이터 재추출*  
health\_male <- health **%>%**  
 **filter**(성별 **==** "남자")  
*#남성만 추출한 데이터에서 성별항목제거 - 모든 데이터를 numeric타입으로 만들기 위해*  
health\_male <- **subset**(health\_male, select = **-c**(성별, 연령집단))  
  
*#수치가 반올린된 여성데이터 재추출*  
health\_female <- health **%>%**  
 **filter**(성별 **==** "여자")  
*#남성만 추출한 데이터에서 성별항목제거 - 모든 데이터를 numeric타입으로 만들기 위해*  
health\_female <- **subset**(health\_female, select = **-c**(성별, 연령집단))  
  
*#분석할 데이터파악, 결측치가 없음은 데이터추출전 미리 파악완료*  
**summary**(health**$**BMI)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 16.00 22.00 23.00 23.78 26.00 44.00

**summary**(health**$**체지방률)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 6.00 21.00 26.00 26.24 32.00 52.00

**summary**(health**$**상대악력)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 10.00 25.00 33.00 34.07 43.00 70.00

**summary**(health**$**악력)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 9.00 23.00 30.00 32.05 40.00 68.00

**summary**(health**$**좌전굴) *#좌전굴이란, 다리를 쭉뻗고 앉아서 앞으로 팔을 내미는 유연성테스트임*

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -20.00 7.00 13.00 12.19 19.00 39.00

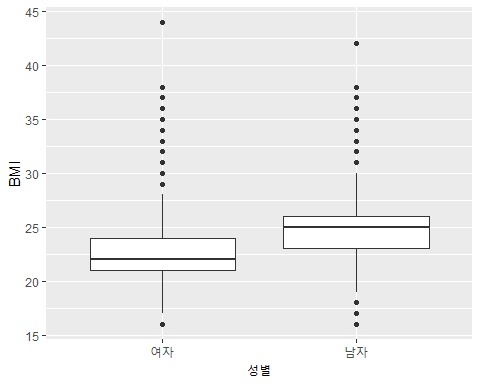
**summary**(health**$**왕복빨리달리기\_10m) *#왕복달리기를 수행한 초(second)*

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 8.00 11.00 13.00 12.89 14.00 30.00

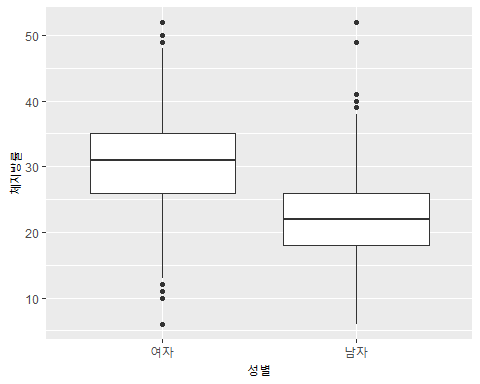
**summary**(health**$**왕복오래달리기\_20m) *#왕복오래달리기를 수행한 횟수(count)*

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.00 15.00 24.00 28.53 38.00 117.00

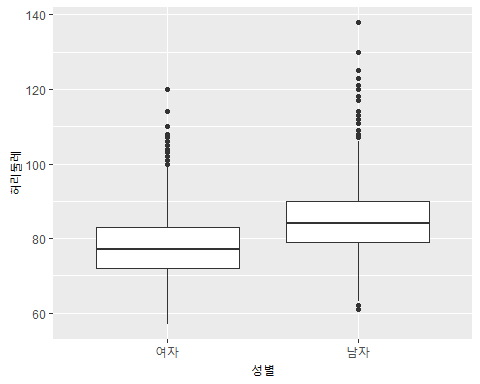
*##### geom\_boxplot() - 박스플롯으로 데이터(성별)분석 #####*  
*##성별별 BMI, 체지방률,허리둘레 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = 성별, y = BMI)) **+**  
 **geom\_boxplot**() *# 여자의 중앙값은 22.5정도, 남자는 25정도로 남성이 전체적으로 BMI가 높은 경향이있다.*



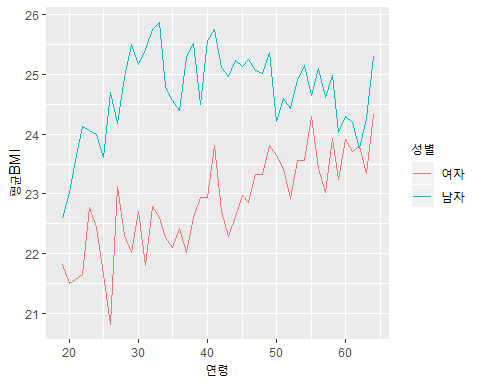
**ggplot**(data = health, **aes**(x = 성별, y = 체지방률)) **+**  
 **geom\_boxplot**() *# 여자의 중앙값은 31정도, 남자는 22정도로 여성이 전체적으로 체지방률이 높은 경향이있다.*



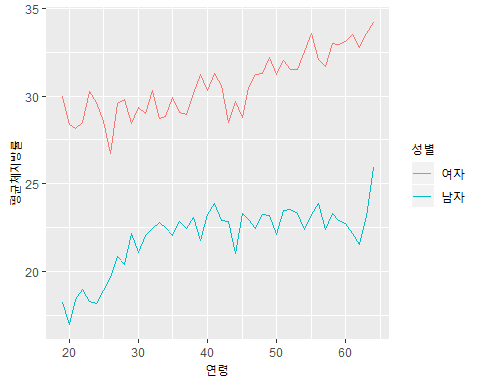
**ggplot**(data = health, **aes**(x = 성별, y = 허리둘레)) **+**  
 **geom\_boxplot**() *# 여자의 중앙값은 78정도, 남자는 82정도로 남성이 전체적으로 허리둘레가 두꺼운 경향이있다.*



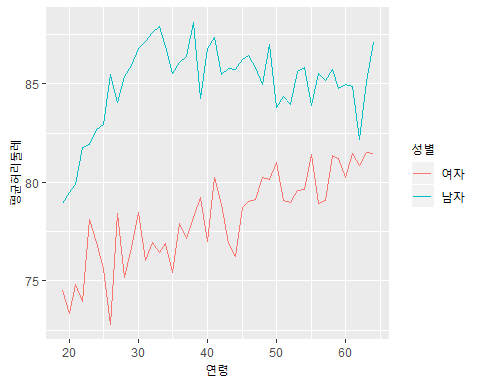
*## 체지방률과 BMI,허리둘레는 상관관계가 있지만, 남성이 여성과 비교해 체지방률은 명확히 낮고 BMI,허리둘레는 크다.*  
*## 그렇기 때문에 데이터를 분석할때 남성과 여성을 반드시 구분해주어야한다.*  
  
  
  
  
*##### geom\_line() - 선그래프로 데이터(연령)분석 #####*  
*##geom\_line()그래프에 적합하게 연령별 평균체지방률, 평균BMI, 평균허리둘레데이터 생성*  
연령별평균BMI <- health **%>%**  
 **group\_by**(연령, 성별) **%>%**  
 **summarise**(평균BMI = **mean**(BMI))  
연령별평균체지방률 <- health **%>%**  
 **group\_by**(연령, 성별) **%>%**  
 **summarise**(평균체지방률 = **mean**(체지방률))  
연령별평균허리둘레 <- health **%>%**  
 **group\_by**(연령, 성별) **%>%**  
 **summarise**(평균허리둘레 = **mean**(허리둘레))  
  
*##연령별 평균체지방률,평균BMI,평균허리둘레 그래프생성*  
**ggplot**(data = 연령별평균BMI , **aes**(x = 연령, y = 평균BMI, col = 성별)) **+**  
 **geom\_line**() *# 등락이 있지만 여성은 대체적으로 나이가 들어감에따라 평균BMI가 상승하는 경향을보인다.*



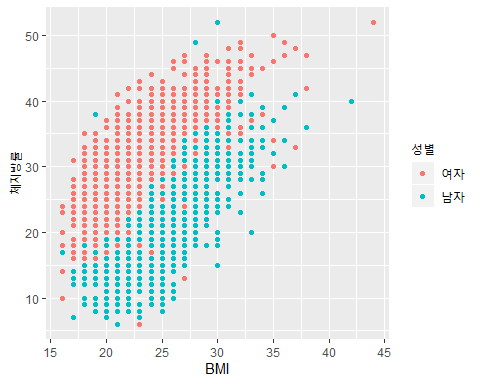
*# 남성은 30대 초반에 평균 BMI가 최대치를 찍고 그후 미세하게 감소하는 경향을 보인다.*  
**ggplot**(data = 연령별평균체지방률, **aes**(x = 연령, y = 평균체지방률, col = 성별)) **+**  
 **geom\_line**() *# 등락이 있지만 남성과 여성 모두 대체적으로 나이가 들어감에따라 평균체지방률이 상승하는 경향을보인다.*



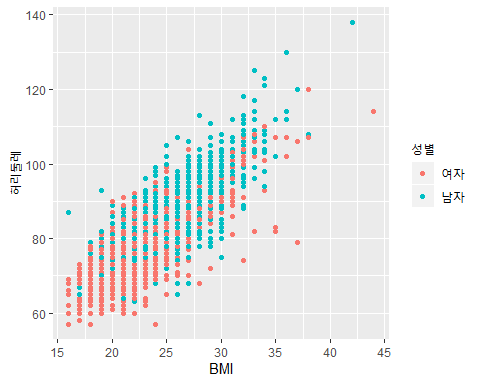
**ggplot**(data = 연령별평균허리둘레, **aes**(x = 연령, y = 평균허리둘레, col = 성별)) **+**  
 **geom\_line**() *# 등락이 있지만 BMI와 유사하게 여성은 대체적으로 나이가 들어감에따라 평균허리둘레가 상승하는 경향을보인다.*



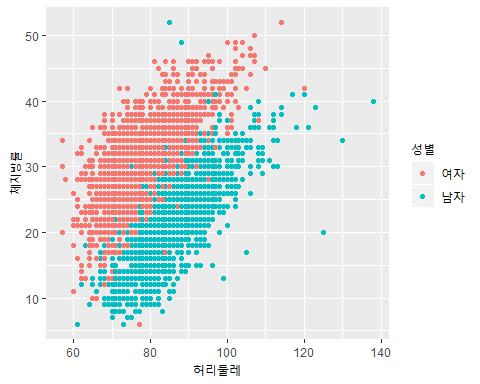
*# 남성은 30대 초반에 평균 허리둘레가 최대치를 찍고 그후 미세하게 감소하는 경향을 보인다.*  
  
  
  
  
*##### geom\_point() - 산점도로 데이터분석 #####*  
*##BMI와 체지방률,허리둘레의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 체지방률, col = 성별)) **+**  
 **geom\_point**() *# 각 성별별로 BMI와 체지방률은 강한 양의 상관관계를 보인다.*



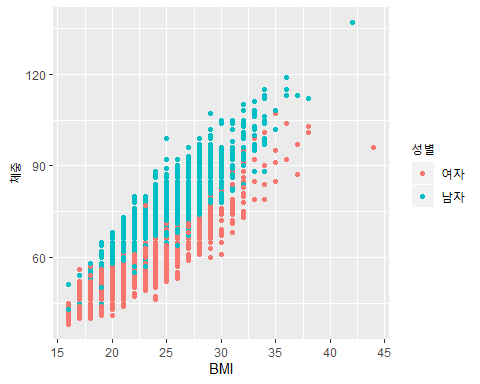
**ggplot**(data = health, **aes**(x = BMI, y = 허리둘레, col = 성별)) **+**  
 **geom\_point**() *# 각 성별별로 BMI와 허리둘레도 강한 양의 상관관계를 보인다.*



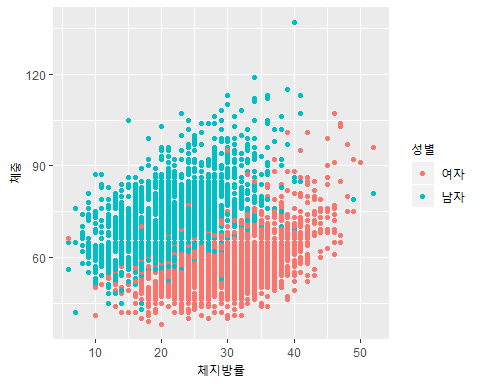
**ggplot**(data = health, **aes**(x = 허리둘레, y = 체지방률, col = 성별)) **+**  
 **geom\_point**() *# 각 성별별로 허리둘레와 체지방률은 강한 양의 상관관계를 보인다.*



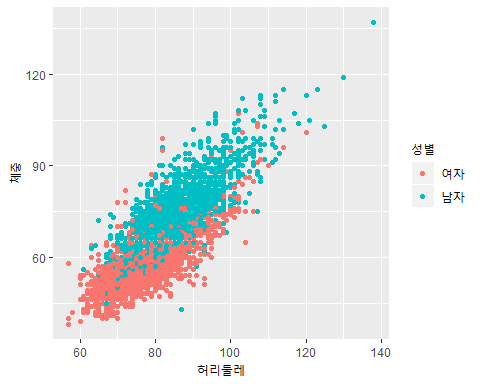
*## BMI,체지방률, 허리둘레 셋다 서로 강한 양의 상관관계를 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 체중의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 체중, col = 성별)) **+**  
 **geom\_point**()*# BMI와 체중은 강한 양의 상관관계를 보인다.*



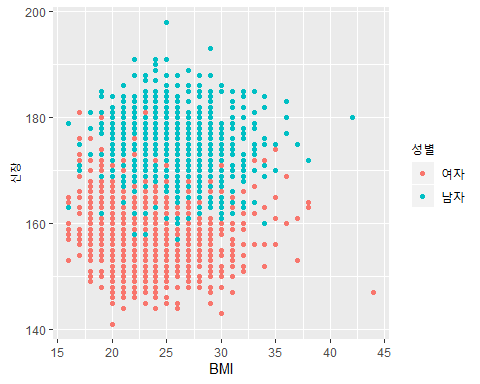
**ggplot**(data = health, **aes**(x = 체지방률, y = 체중, col = 성별)) **+**  
 **geom\_point**()*# 체지방률과 체중은 약한 양의 상관관계를 보인다.*



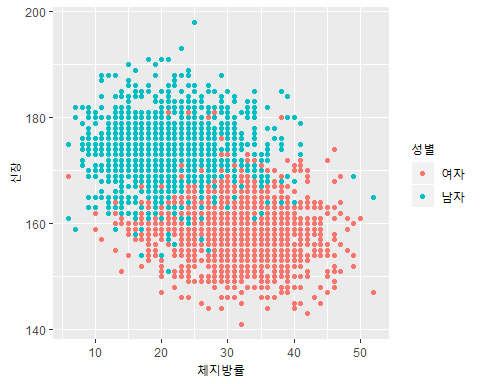
**ggplot**(data = health, **aes**(x = 허리둘레, y = 체중, col = 성별)) **+**  
 **geom\_point**()*# 허리둘레와 체중은 강한 양의 상관관계를 보인다.*



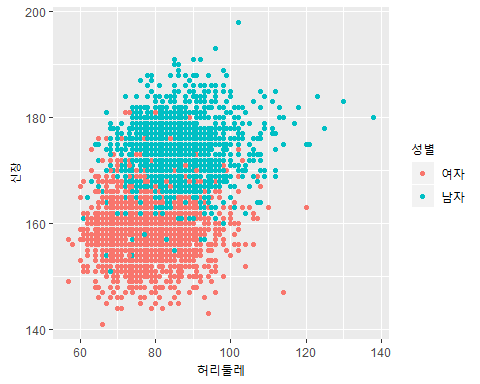
*## 체중은 체지방률보다 BMI, 허리둘레와 더 강한 양의 상관관계를보인다.*  
*## 체중을 신장의 제곱으로 나눠서 구한 BMI이기 때문에 당연한 결과라고할수있다.*  
*## 체중과 더욱관련있는것은 체지방보다 허리둘레 라는 것이 의미있다.*  
  
  
*##BMI와 체지방률,허리둘레 - 신장의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 신장, col = 성별)) **+**  
 **geom\_point**() *# BMI와 신장은 상관관계가 없다*



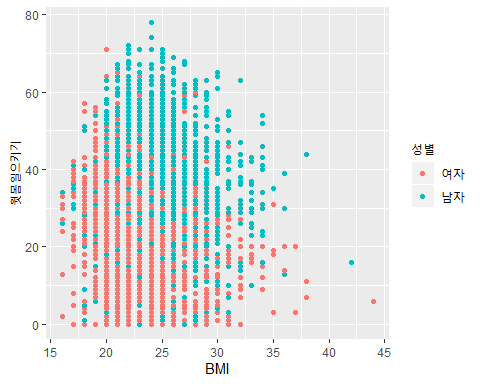
**ggplot**(data = health, **aes**(x = 체지방률, y = 신장, col = 성별)) **+**  
 **geom\_point**() *# 체지방률과 신장은 상관관계가 없다*



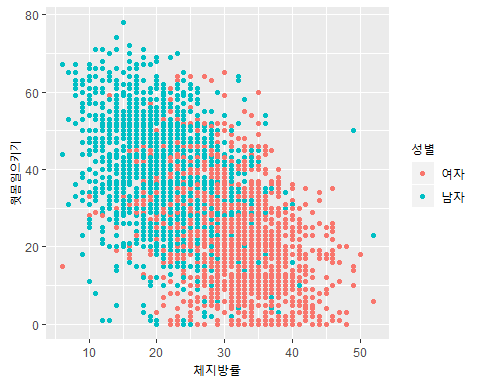
**ggplot**(data = health, **aes**(x = 허리둘레, y = 신장, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 신장은 상관관계가 없다*



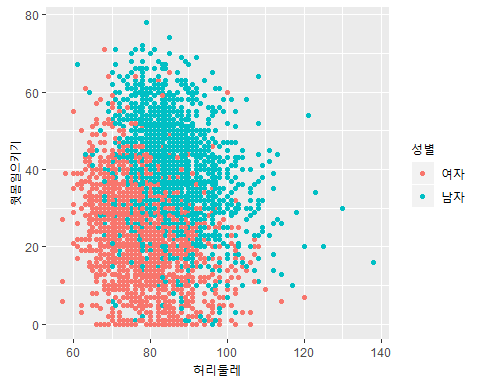
*## 체중에 신장의 제곱을 나눠서 BMI를 구한 BMI와 체지방률, 허리둘레 모두 신장과 상관관계가 없음을 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 윗몸일으키기의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 윗몸일으키기, col = 성별)) **+**  
 **geom\_point**() *# BMI와 윗몸일으키기는 상관관계가 거의 없다.*



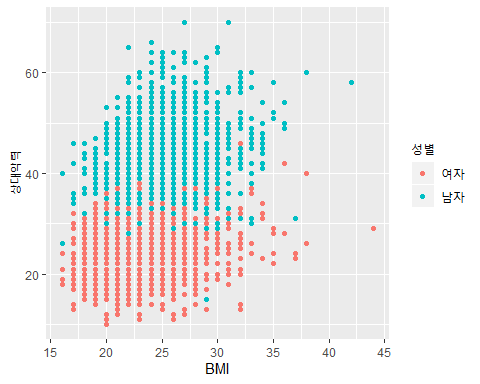
**ggplot**(data = health, **aes**(x = 체지방률, y = 윗몸일으키기, col = 성별)) **+**  
 **geom\_point**() *# 체지방률이 증가할수록 윗몸일으키기하는 수가 적어지는 약한 음의 상관관계를 보인다.*



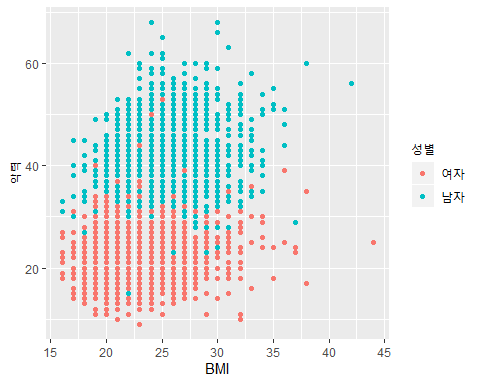
**ggplot**(data = health, **aes**(x = 허리둘레, y = 윗몸일으키기, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 윗몸일으키기는 상관관계가 거의 없다.*



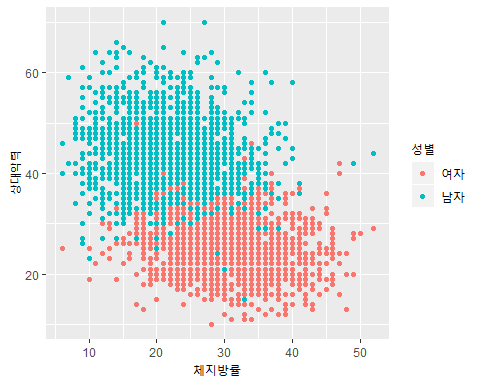
*## BMI, 허리둘레보다 체지방률이 윗몸일으키기에 더 큰 상관관계를 보인다.*  
*## 허리둘레와 윗몸일으키기가 상관관계가 거의 없다는것이 의미있다.*  
*## 여자보다 남자가 대체적으로 윗몸일으키기를 더 많이함을 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 악력의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 상대악력, col = 성별)) **+**  
 **geom\_point**()



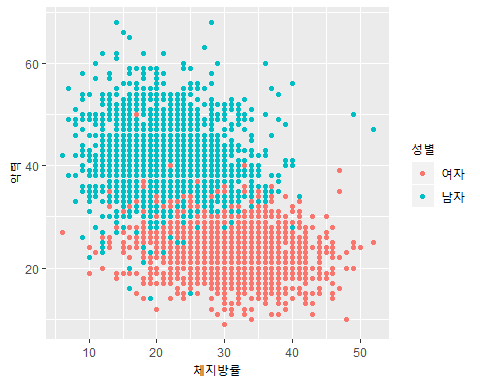
**ggplot**(data = health, **aes**(x = BMI, y = 악력, col = 성별)) **+**  
 **geom\_point**() *# BMI와 악력은 상관관계가 없다*



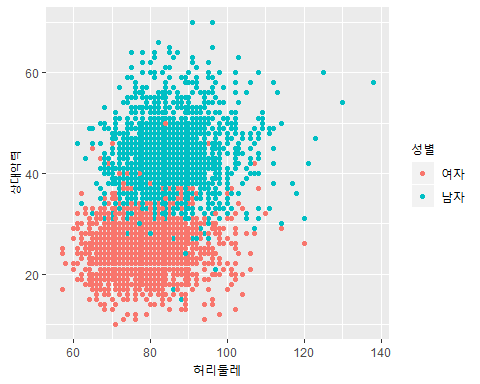
**ggplot**(data = health, **aes**(x = 체지방률, y = 상대악력, col = 성별)) **+**  
 **geom\_point**()



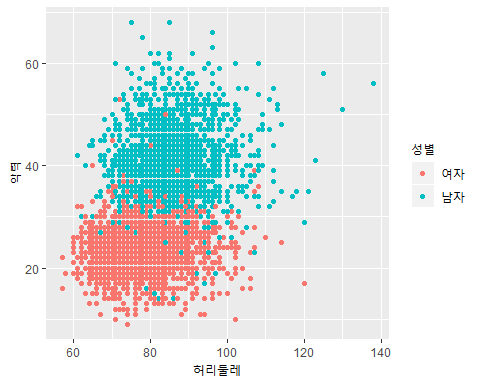
**ggplot**(data = health, **aes**(x = 체지방률, y = 악력, col = 성별)) **+**  
 **geom\_point**() *# 체지방률과 악력은 상관관계가 없다.*



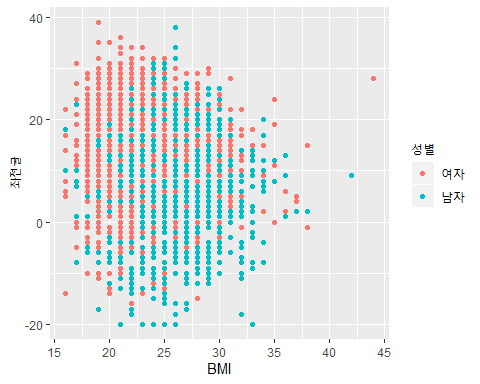
**ggplot**(data = health, **aes**(x = 허리둘레, y = 상대악력, col = 성별)) **+**  
 **geom\_point**()



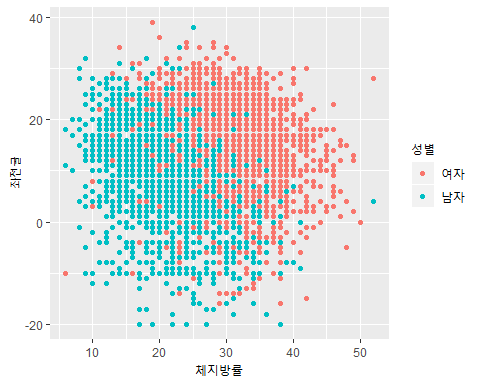
**ggplot**(data = health, **aes**(x = 허리둘레, y = 악력, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 악력은 상관관계가 없다*



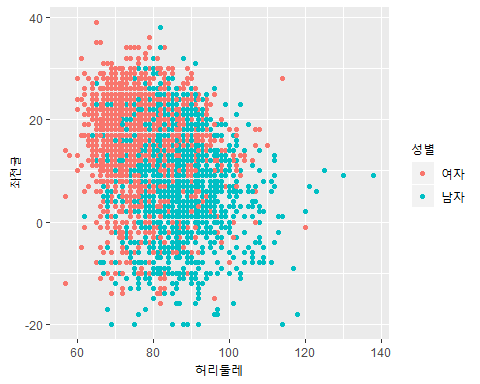
*## BMI와 체지방률,허리둘레 모두 악력과 상관관계가 없음을 보인다.*  
*## 남자가 여자보다 명확하게 악력이 강함을 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 좌전굴의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 좌전굴, col = 성별)) **+**  
 **geom\_point**() *# BMI와 좌전굴은 상관관계가 없다.*



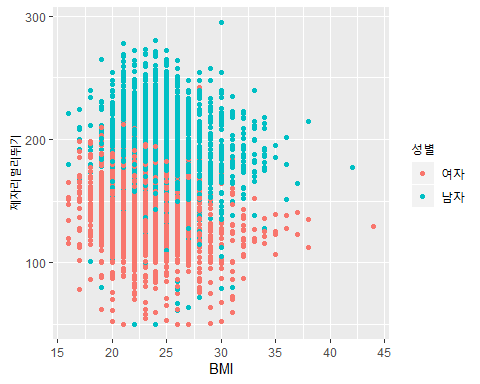
**ggplot**(data = health, **aes**(x = 체지방률, y = 좌전굴, col = 성별)) **+**  
 **geom\_point**() *# 체지방률과 좌전굴은 약한 음의 상관관계를 보인다.*



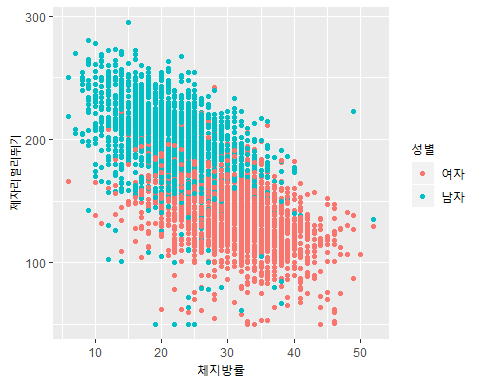
**ggplot**(data = health, **aes**(x = 허리둘레, y = 좌전굴, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 좌전굴은 상관관계가 거의 없다.*



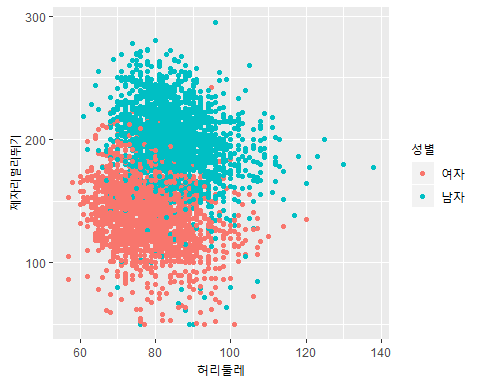
*## 유연성을 측정하는 좌전굴은 성별, BMI와 상관관계가 없고, 체지방률과 약한 상관관계를 보인다.*  
*## 앉아서 앞으로 숙이는 좌전굴 특성상 허리둘레가 상관이 있을것같았지만, 상관관계가 거의 없었다.*  
  
  
*##BMI와 체지방률,허리둘레 - 제자리멀리뛰기의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 제자리멀리뛰기, col = 성별)) **+**  
 **geom\_point**() *# BMI와 제자리멀리뛰기는 상관관계가 거의 없다.*



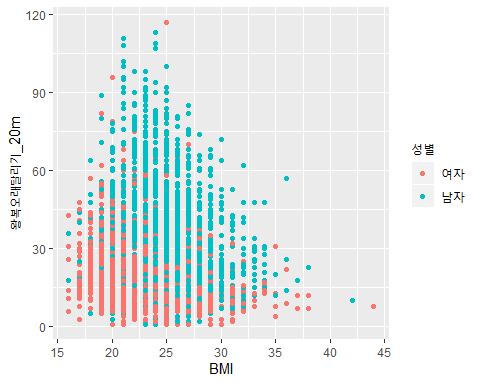
**ggplot**(data = health, **aes**(x = 체지방률, y = 제자리멀리뛰기, col = 성별)) **+**  
 **geom\_point**() *# 체지방룰은 제자리멀리뛰기와 약한 음의 상관관계를 보인다.*



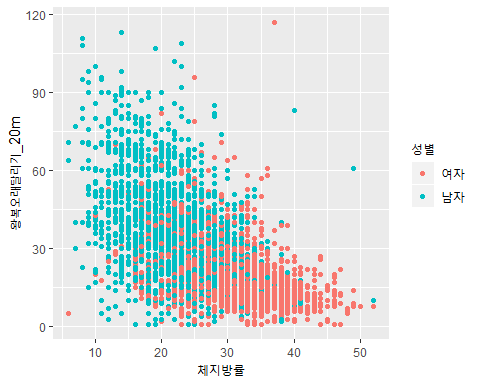
**ggplot**(data = health, **aes**(x = 허리둘레, y = 제자리멀리뛰기, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 제자리멀리뛰기는 상관관계가 거의 없다.*



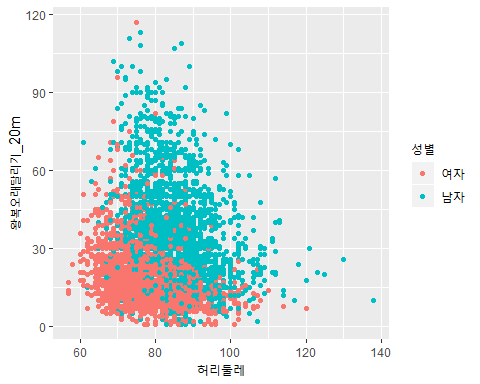
*## BMI,허리둘레보다 체지방률이 제자리멀리뛰기에 더 큰 상관관계를 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 20m왕복오래달리기의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 왕복오래달리기\_20m, col = 성별)) **+**  
 **geom\_point**() *# BMI와 20m왕복오래달리기는 상관관계가 거의 없다.*



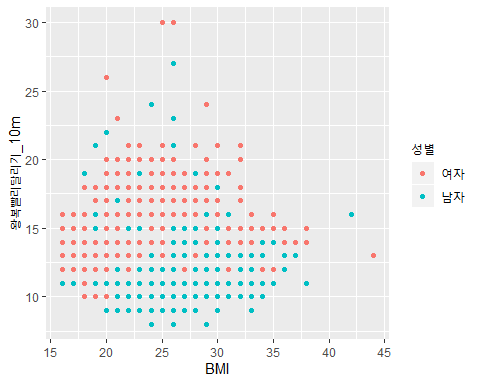
**ggplot**(data = health, **aes**(x = 체지방률, y = 왕복오래달리기\_20m, col = 성별)) **+**  
 **geom\_point**() *# 체지방률과 20m왕복오래달리기는 아주 약한 음의 상관관계를 보인다.*



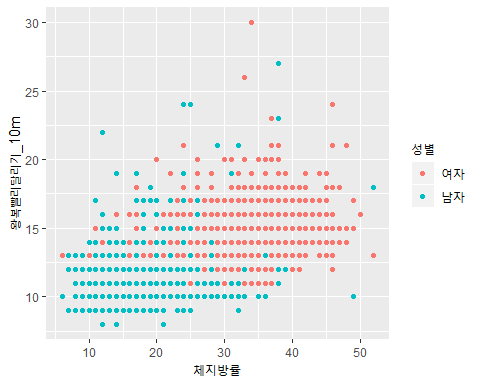
**ggplot**(data = health, **aes**(x = 허리둘레, y = 왕복오래달리기\_20m, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 20m왕복오래달리기는 상관관계가 거의 없다.*



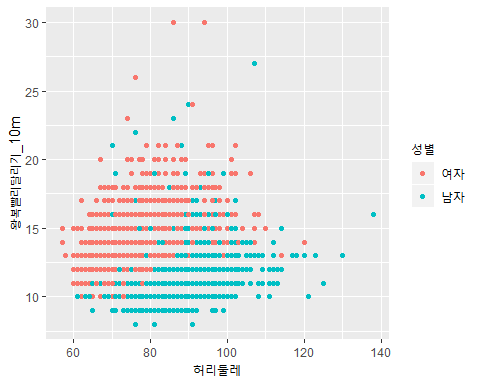
*## BMI, 허리둘레보다 체지방률이 20m왕복오래달리기에 그나마 더 큰 상관관계를 보인다.*  
*## 대체적으로 남성이 여성보다 더 오래달리는 경향을 보인다.*  
  
  
*##BMI와 체지방률,허리둘레 - 10m왕복빨리달리기의 상관관계그래프 생성*  
**ggplot**(data = health, **aes**(x = BMI, y = 왕복빨리달리기\_10m, col = 성별)) **+**  
 **geom\_point**() *# BMI와 10m왕복빨리달리기는 상관관계가 거의 없다.*



**ggplot**(data = health, **aes**(x = 체지방률, y = 왕복빨리달리기\_10m, col = 성별)) **+**  
 **geom\_point**() *# 체지방률과 10m왕복빨리달리기는 아주약한 양의 상관관계를 보인다.*



**ggplot**(data = health, **aes**(x = 허리둘레, y = 왕복빨리달리기\_10m, col = 성별)) **+**  
 **geom\_point**() *# 허리둘레와 10m왕복빨리달리기는 양의 상관관계를 보인다.*



*## 허리둘레, 체지방률, BMI순으로 10m왕복빨리달리기와 상관관계를 보인다.*

**2) 데이터 분석 전략**

먼저 가설을 세운다. 그리고 데이터를 성별별로 나누고, BMI, 체지방률, 허리둘레과 체력의 상관관계를 분석하기위해 교차분석, 상관분석, 연관분석 세가지 분석을 사용하여 가설을 검정하려고 한다.

*###### 가설 : BMI가 다른수치들보다 체력수준을 파악하는데 용이하지않다. #####*  
*###### 흔히 본인 신체에대한 평가로 BMI를 가장 흔하게 사용하는데, BMI가 다른 수치들보다 신체의 체력과 상관관계가 밀접한지 알기위해 가설을 세움.*  
*###### 검정방법 : BMI를 체지방률, 허리둘레와 비교하여 체력과의 상관관계를 교차분석과 상관분석, 연관분석까지 총 세가지 분석을 사용하여 검정 #####*  
  
  
**3) 데이터 분석 결과**

분석을 실시하고 분석결과를 살펴보고 설명한다.

*##### 교차분석실시 #####*  
*##BMI와 체지방률,허리둘레는 좌전굴과 수치적으로 유의미한 관계를 갖는지 확인*  
*#남성*  
**chisq.test**(health\_male**$**BMI,health\_male**$**좌전굴) *#귀무가설 : 남성의 BMI와 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_male$BMI, health\_male$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$BMI and health\_male$좌전굴  
## X-squared = 1105, df = 1219, p-value = 0.9911

*#p = 0.99로 유의수준 0.05보다 크므로 귀무가설승인, 남성의 BMI과 좌전굴을 서로 관계가 없다.*  
**chisq.test**(health\_male**$**체지방률,health\_male**$**좌전굴) *#귀무가설 : 남성의 체지방률과 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_male$체지방률, health\_male$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$체지방률 and health\_male$좌전굴  
## X-squared = 2336, df = 1961, p-value = 7.845e-09

*#p = 7.8e-09로 유의수준 0.05보다 작으므로 귀무가설기각, 남성의 BMI와 좌전굴은 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_male**$**허리둘레,health\_male**$**좌전굴) *#귀무가설 : 남성의 BMI와 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_male$허리둘레, health\_male$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$허리둘레 and health\_male$좌전굴  
## X-squared = 3905.3, df = 3180, p-value < 2.2e-16

*#p < 2.2e-16로 유의수준 0.05보다 작으므로 귀무가설기각, 남성의 허리둘레와 좌전굴은 서로 유의미한 관계를 갖는다.*  
*#여성*  
**chisq.test**(health\_female**$**BMI,health\_female**$**좌전굴) *#귀무가설 : 여성의 BMI와 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_female$BMI, health\_female$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$BMI and health\_female$좌전굴  
## X-squared = 1247.3, df = 1196, p-value = 0.1473

*#p = 0.14로 유의수준 0.05보다 크므로 귀무가설승인, 여성의 BMI과 좌전굴을 서로 관계가 없다.*  
**chisq.test**(health\_female**$**체지방률,health\_female**$**좌전굴) *#귀무가설 : 여성의 체지방률과 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_female$체지방률, health\_female$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$체지방률 and health\_female$좌전굴  
## X-squared = 2489.6, df = 2184, p-value = 4.625e-06

*#p = 4.6e-06로 유의수준 0.05보다 작으므로 귀무가설기각, 여성의 BMI와 좌전굴은 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_female**$**허리둘레,health\_female**$**좌전굴) *#귀무가설 : 여성의 체지방률과 좌전굴은 서로 관계없다.*

## Warning in chisq.test(health\_female$허리둘레, health\_female$좌전굴): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$허리둘레 and health\_female$좌전굴  
## X-squared = 3660, df = 2756, p-value < 2.2e-16

*#p < 2.2e-16로 유의수준 0.05보다 작으므로 귀무가설기각, 여성의 허리둘레와 좌전굴은 서로 유의미한 관계를 갖는다.*  
  
*#그래프에서 분석한것처럼 남녀 구분없이 BMI은 좌전굴과 서로 관계가 없다고나오고,*   
*#체지방률은 좌전굴과 서로 유의미한 관계가있다고 나왔다.*  
*#하지만 허리둘레는 그래프에선 상관관계가 거의없는것처럼 보였지만, 수치적으로는 서로 유의미한 관계가있다고 나왔다.*  
*#다만, Warning message: Chi-squared approximation may be incorrect메시지가 떠서 정확하지않을수있다.*  
  
  
*##BMI와 체지방률, 허리둘레는 10m왕복빨리달리기와 수치적으로 유의미한 관계를 갖는지 확인*  
*#남성*  
**chisq.test**(health\_male**$**BMI,health\_male**$**왕복빨리달리기\_10m) *#귀무가설 : BMI와 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$BMI, health\_male$왕복빨리달리기\_10m): Chi-  
## squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$BMI and health\_male$왕복빨리달리기\_10m  
## X-squared = 553.58, df = 368, p-value = 1.212e-09

*#p = 1.2e-09으로 유의수준 0.05보다 작으므로 귀무가설기각,남성의 BMI와 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_male**$**체지방률,health\_male**$**왕복빨리달리기\_10m) *#귀무가설 : 체지방률과 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$체지방률, health\_male$왕복빨리달리기\_10m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$체지방률 and health\_male$왕복빨리달리기\_10m  
## X-squared = 2225.2, df = 592, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 작으므로 귀무가설기각,남성의 체지방률과 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_male**$**허리둘레,health\_male**$**왕복빨리달리기\_10m) *#귀무가설 : 체지방률과 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$허리둘레, health\_male$왕복빨리달리기\_10m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$허리둘레 and health\_male$왕복빨리달리기\_10m  
## X-squared = 1511.9, df = 960, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 작으므로 귀무가설기각,남성의 허리둘레와 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
*#여성*  
**chisq.test**(health\_female**$**BMI,health\_female**$**왕복빨리달리기\_10m) *#귀무가설 : BMI와 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_female$BMI, health\_female$왕복빨리달리기\_10m): Chi-  
## squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$BMI and health\_female$왕복빨리달리기\_10m  
## X-squared = 465.23, df = 345, p-value = 1.61e-05

*#p = 1.6e-05으로 유의수준 0.05보다 작으므로 귀무가설기각,여성의 BMI와 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_female**$**체지방률,health\_female**$**왕복빨리달리기\_10m) *#귀무가설 : 체지방률과 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_female$체지방률, health\_female$왕복빨리달리기\_10m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$체지방률 and health\_female$왕복빨리달리기\_10m  
## X-squared = 1267.6, df = 630, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 작으므로 귀무가설기각,여성의 체지방률과 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_female**$**체지방률,health\_female**$**왕복빨리달리기\_10m) *#귀무가설 : 체지방률과 10m왕복빨리달리기는 서로 관계없다.*

## Warning in chisq.test(health\_female$체지방률, health\_female$왕복빨리달리기\_10m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$체지방률 and health\_female$왕복빨리달리기\_10m  
## X-squared = 1267.6, df = 630, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 작으므로 귀무가설기각,여성의 허리둘레와 10m왕복빨리달리기는 서로 유의미한 관계를 갖는다.*  
  
*#그래프에서 분석한것과 다르게 BMI와 체지방률 모두 10m왕복빨리달리기와 서로 유의미한 관계를 갖는다고 나왔다.*  
*#그래프에서도 10m빨리달리기와 상관관계가 보였던 허리둘레도 역시 서로 유의미한 관계를 갖는다고 나왔다.*  
*#다만, Warning message: Chi-squared approximation may be incorrect메시지가 떠서 정확하지않을수있다.*  
  
  
*##BMI와 체지방률, 허리둘레는 20m왕복오래달리기와 수치적으로 유의미한 관계를 갖는지 확인*  
*#남성*  
**chisq.test**(health\_male**$**BMI,health\_male**$**왕복오래달리기\_20m) *#귀무가설 : BMI와 20m왕복오래달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$BMI, health\_male$왕복오래달리기\_20m): Chi-  
## squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$BMI and health\_male$왕복오래달리기\_20m  
## X-squared = 2289.5, df = 2346, p-value = 0.7942

*#p = 0.79로 유의수준 0.05보다 크므로 귀무가설승인, 남성의 BMI와 10m왕복빨리달리기는 서로 관계가 없다고 볼수있다.*  
**chisq.test**(health\_male**$**체지방률,health\_male**$**왕복오래달리기\_20m) *#귀무가설 : 체지방률과 20m왕복오래달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$체지방률, health\_male$왕복오래달리기\_20m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$체지방률 and health\_male$왕복오래달리기\_20m  
## X-squared = 5719.2, df = 3774, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 낮으므로 귀무가설기각, 남성의 체지방률과 20m왕복오래달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_male**$**허리둘레,health\_male**$**왕복오래달리기\_20m) *#귀무가설 : 체지방률과 20m왕복오래달리기는 서로 관계없다.*

## Warning in chisq.test(health\_male$허리둘레, health\_male$왕복오래달리기\_20m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_male$허리둘레 and health\_male$왕복오래달리기\_20m  
## X-squared = 6483.9, df = 6120, p-value = 0.0006148

*#p = 0.0006으로 유의수준 0.05보다 낮으므로 귀무가설기각, 남성의 허리둘레와 20m왕복오래달리기는 서로 유의미한 관계를 갖는다.*  
*#여성*  
**chisq.test**(health\_female**$**BMI,health\_female**$**왕복오래달리기\_20m) *#귀무가설 : BMI와 20m왕복오래달리기는 서로 관계없다.*

## Warning in chisq.test(health\_female$BMI, health\_female$왕복오래달리기\_20m): Chi-  
## squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$BMI and health\_female$왕복오래달리기\_20m  
## X-squared = 1492.3, df = 1656, p-value = 0.9983

*#p = 0.99로 유의수준 0.05보다 크므로 귀무가설승인, 여성의 BMI와 10m왕복빨리달리기는 서로 관계가 없다고 볼수있다.*  
**chisq.test**(health\_female**$**체지방률,health\_female**$**왕복오래달리기\_20m) *#귀무가설 : 체지방률과 20m왕복오래달리기는 서로 관계없다.*

## Warning in chisq.test(health\_female$체지방률, health\_female$왕복오래달리기\_20m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$체지방률 and health\_female$왕복오래달리기\_20m  
## X-squared = 4918, df = 3024, p-value < 2.2e-16

*#p < 2.2e-16으로 유의수준 0.05보다 낮으므로 귀무가설기각, 여성의 체지방률과 20m왕복오래달리기는 서로 유의미한 관계를 갖는다.*  
**chisq.test**(health\_female**$**허리둘레,health\_female**$**왕복오래달리기\_20m) *#귀무가설 : 체지방률과 20m왕복오래달리기는 서로 관계없다.*

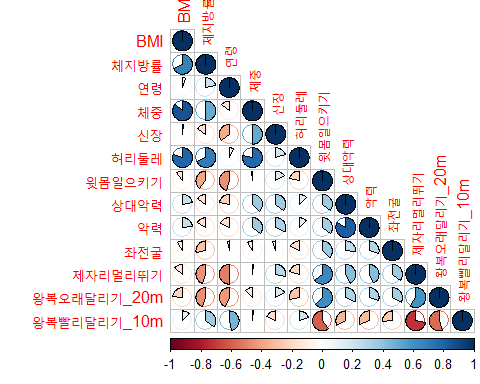
## Warning in chisq.test(health\_female$허리둘레, health\_female$왕복오래달리기\_20m):  
## Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: health\_female$허리둘레 and health\_female$왕복오래달리기\_20m  
## X-squared = 3275.6, df = 3816, p-value = 1

*#p = 1으로 유의수준 0.05보다 크므로 귀무가설승인, 여성의 허리둘레와 10m왕복빨리달리기는 서로 관계가 없다고 볼수있다.*  
  
*#그래프에서 분석한것처럼 BMI,허리둘레보다 체지방률이 20m왕복오래달리기와 서로 유의미한 관계를 갖는다.*  
*#다만, Warning message: Chi-squared approximation may be incorrect메시지가 떠서 정확하지않을수있다.*  
  
  
  
*##### 상관분석실시 #####*  
*#남성*  
**cor**(health\_male)

## BMI 체지방률 연령 체중 신장  
## BMI 1.00000000 0.6718890 0.05209121 0.86583427 0.02090379  
## 체지방률 0.67188899 1.0000000 0.21850847 0.50401614 -0.15332652  
## 연령 0.05209121 0.2185085 1.00000000 -0.13431960 -0.36853726  
## 체중 0.86583427 0.5040161 -0.13431960 1.00000000 0.50612089  
## 신장 0.02090379 -0.1533265 -0.36853726 0.50612089 1.00000000  
## 허리둘레 0.79823311 0.6923275 0.07195485 0.78925010 0.19925247  
## 윗몸일으키기 -0.10103794 -0.3965191 -0.45640993 -0.01010980 0.16623670  
## 상대악력 0.22958306 -0.1542541 -0.22223169 0.36966554 0.35084830  
## 악력 0.23272687 -0.1579498 -0.17732940 0.35656738 0.31916669  
## 좌전굴 -0.08123980 -0.2840626 -0.02707541 -0.09406164 -0.05049090  
## 제자리멀리뛰기 -0.13830146 -0.4493711 -0.50819984 0.01834145 0.28875671  
## 왕복오래달리기\_20m -0.22733433 -0.4509577 -0.43728421 -0.11700583 0.16472869  
## 왕복빨리달리기\_10m 0.11766274 0.3512964 0.46939650 0.01481657 -0.18440563  
## 허리둘레 윗몸일으키기 상대악력 악력 좌전굴  
## BMI 0.79823311 -0.1010379 0.2295831 0.2327269 -0.08123980  
## 체지방률 0.69232750 -0.3965191 -0.1542541 -0.1579498 -0.28406258  
## 연령 0.07195485 -0.4564099 -0.2222317 -0.1773294 -0.02707541  
## 체중 0.78925010 -0.0101098 0.3696655 0.3565674 -0.09406164  
## 신장 0.19925247 0.1662367 0.3508483 0.3191667 -0.05049090  
## 허리둘레 1.00000000 -0.2104997 0.1349733 0.1293835 -0.17860851  
## 윗몸일으키기 -0.21049973 1.0000000 0.3579722 0.3452623 0.36731498  
## 상대악력 0.13497333 0.3579722 1.0000000 0.8297819 0.25635935  
## 악력 0.12938347 0.3452623 0.8297819 1.0000000 0.28602225  
## 좌전굴 -0.17860851 0.3673150 0.2563594 0.2860222 1.00000000  
## 제자리멀리뛰기 -0.20435937 0.6444296 0.4256653 0.4222635 0.35087265  
## 왕복오래달리기\_20m -0.28898772 0.6189662 0.3170936 0.3014675 0.33021849  
## 왕복빨리달리기\_10m 0.20218016 -0.6033399 -0.3200720 -0.3104401 -0.26438547  
## 제자리멀리뛰기 왕복오래달리기\_20m 왕복빨리달리기\_10m  
## BMI -0.13830146 -0.2273343 0.11766274  
## 체지방률 -0.44937108 -0.4509577 0.35129636  
## 연령 -0.50819984 -0.4372842 0.46939650  
## 체중 0.01834145 -0.1170058 0.01481657  
## 신장 0.28875671 0.1647287 -0.18440563  
## 허리둘레 -0.20435937 -0.2889877 0.20218016  
## 윗몸일으키기 0.64442959 0.6189662 -0.60333989  
## 상대악력 0.42566529 0.3170936 -0.32007199  
## 악력 0.42226350 0.3014675 -0.31044011  
## 좌전굴 0.35087265 0.3302185 -0.26438547  
## 제자리멀리뛰기 1.00000000 0.5959346 -0.72543733  
## 왕복오래달리기\_20m 0.59593459 1.0000000 -0.56243484  
## 왕복빨리달리기\_10m -0.72543733 -0.5624348 1.00000000

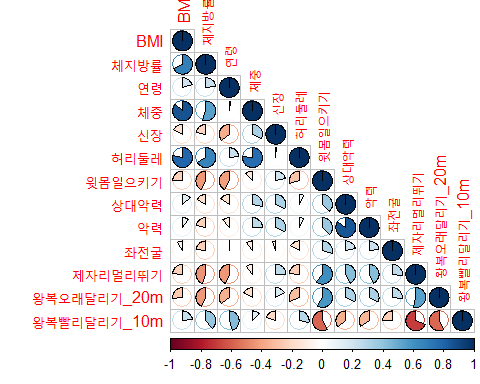
**corrplot**(**cor**(health\_male), method="pie", type="lower")



*#여성*  
**cor**(health\_female)

## BMI 체지방률 연령 체중 신장  
## BMI 1.00000000 0.6858744 0.206138412 0.86633675 -0.16882328  
## 체지방률 0.68587436 1.0000000 0.224673422 0.54857332 -0.21691137  
## 연령 0.20613841 0.2246734 1.000000000 0.01975218 -0.36115743  
## 체중 0.86633675 0.5485733 0.019752184 1.00000000 0.32869908  
## 신장 -0.16882328 -0.2169114 -0.361157428 0.32869908 1.00000000  
## 허리둘레 0.79088792 0.6608233 0.224536749 0.76589802 0.01423909  
## 윗몸일으키기 -0.23548194 -0.4333738 -0.414988773 -0.12118496 0.20602577  
## 상대악력 0.13928018 -0.1744315 -0.160900107 0.29421695 0.31620449  
## 악력 0.10165770 -0.2232561 -0.109516714 0.25555245 0.31560459  
## 좌전굴 -0.07684778 -0.2323389 0.007009447 -0.10270527 -0.05807370  
## 제자리멀리뛰기 -0.23300760 -0.4483363 -0.436367040 -0.09773468 0.25011161  
## 왕복오래달리기\_20m -0.26474387 -0.4321952 -0.359069052 -0.17285955 0.15950829  
## 왕복빨리달리기\_10m 0.23210130 0.3897675 0.439870752 0.12182820 -0.19739011  
## 허리둘레 윗몸일으키기 상대악력 악력  
## BMI 0.79088792 -0.2354819 0.13928018 0.10165770  
## 체지방률 0.66082329 -0.4333738 -0.17443148 -0.22325615  
## 연령 0.22453675 -0.4149888 -0.16090011 -0.10951671  
## 체중 0.76589802 -0.1211850 0.29421695 0.25555245  
## 신장 0.01423909 0.2060258 0.31620449 0.31560459  
## 허리둘레 1.00000000 -0.2899022 0.07988528 0.07222015  
## 윗몸일으키기 -0.28990219 1.0000000 0.37691374 0.37432076  
## 상대악력 0.07988528 0.3769137 1.00000000 0.85255784  
## 악력 0.07222015 0.3743208 0.85255784 1.00000000  
## 좌전굴 -0.15593718 0.2940510 0.20851768 0.21959178  
## 제자리멀리뛰기 -0.24518021 0.6135950 0.42817480 0.42690575  
## 왕복오래달리기\_20m -0.30732912 0.5892482 0.31589938 0.32411636  
## 왕복빨리달리기\_10m 0.27514110 -0.5844102 -0.35567486 -0.33781537  
## 좌전굴 제자리멀리뛰기 왕복오래달리기\_20m  
## BMI -0.076847779 -0.23300760 -0.2647439  
## 체지방률 -0.232338881 -0.44833628 -0.4321952  
## 연령 0.007009447 -0.43636704 -0.3590691  
## 체중 -0.102705267 -0.09773468 -0.1728596  
## 신장 -0.058073697 0.25011161 0.1595083  
## 허리둘레 -0.155937184 -0.24518021 -0.3073291  
## 윗몸일으키기 0.294051019 0.61359502 0.5892482  
## 상대악력 0.208517680 0.42817480 0.3158994  
## 악력 0.219591780 0.42690575 0.3241164  
## 좌전굴 1.000000000 0.26847071 0.2671578  
## 제자리멀리뛰기 0.268470707 1.00000000 0.5438635  
## 왕복오래달리기\_20m 0.267157828 0.54386351 1.0000000  
## 왕복빨리달리기\_10m -0.241850555 -0.69164504 -0.5762520  
## 왕복빨리달리기\_10m  
## BMI 0.2321013  
## 체지방률 0.3897675  
## 연령 0.4398708  
## 체중 0.1218282  
## 신장 -0.1973901  
## 허리둘레 0.2751411  
## 윗몸일으키기 -0.5844102  
## 상대악력 -0.3556749  
## 악력 -0.3378154  
## 좌전굴 -0.2418506  
## 제자리멀리뛰기 -0.6916450  
## 왕복오래달리기\_20m -0.5762520  
## 왕복빨리달리기\_10m 1.0000000

**corrplot**(**cor**(health\_female), method="pie", type="lower")



*##성별 관계없이 BMI는 체지방률보다 허리둘레에 더 밀접한 상관관계를 가지고 있다.*  
*##남자는 체지방률이 BMI보다 허리둘레와 미세하게나마 더 밀접한 상관관계를 가지고 있다.*  
*##반면에 여자는 체지방률이 허리둘레보다 BMI와 미세하게나마 더 밀접한 상관관계를 가지고 있다.*  
*##성별 관계없이 체지방률, 허리둘레, BMI순으로 악력을 제외한 체력과 더 밀접한 상관관계를 가지고 있다.*  
  
  
  
*##### 연관분석실시 #####*  
*###연관분석위해 factor형 변환*  
*###남성데이터 변환*  
male\_eval**$**분류\_BMI <- **factor**(male\_eval**$**분류\_BMI, levels=**c**(1**:**5), labels=**c**('낮음', '정상', '경도비만', '비만', '고도비만'))  
male\_eval**$**분류\_체지방률 <- **factor**(male\_eval**$**분류\_체지방률, levels=**c**(1**:**4), labels=**c**('낮음', '정상', '경도비만', '비만'))  
male\_eval**$**분류\_허리둘레 <- **factor**(male\_eval**$**분류\_허리둘레, levels=**c**(1**:**3), labels=**c**('정상', '복부비만', '고도복부비만'))  
male\_eval**$**등급\_윗몸일으키기 <- **factor**(male\_eval**$**등급\_윗몸일으키기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
male\_eval**$**등급\_좌전굴 <- **factor**(male\_eval**$**등급\_좌전굴, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
male\_eval**$**등급\_제자리멀리뛰기 <- **factor**(male\_eval**$**등급\_제자리멀리뛰기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
male\_eval**$**등급\_10m왕복빨리달리기 <- **factor**(male\_eval**$**등급\_10m왕복빨리달리기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
male\_eval**$**등급\_20m왕복오래달리기 <- **factor**(male\_eval**$**등급\_20m왕복오래달리기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
*#male\_eval$등급\_상대악력 <- factor(male\_eval$등급\_상대악력, levels=c(1:3), labels=c('상', '중', '하'))*  
*#male\_eval$등급\_악력 <- factor(male\_eval$등급\_악력, levels=c(1:3), labels=c('상', '중', '하'))*  
  
*## 조건 (support=0.04, confidence=0.01, minlen=1, maxlen=2)으로 연관분석 시작*  
*#arules 및 tm 라이브러리에 inspect () 메소드가 있음*  
*#로드 순서에 따라이 메소드가 구현되는 방식에 영향을 미치는 오류가있어서 tm을 detach하고 다시 arules라이브러리를 로드.*  
**detach**(package**:**tm, unload=TRUE)  
**library**(arules)  
rule\_male <- **apriori**(male\_eval, parameter=**list**(support=0.04, confidence=0.3, minlen=1, maxlen=2))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.3 0.1 1 none FALSE TRUE 5 0.04 1  
## maxlen target ext  
## 2 rules FALSE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 85   
##   
## set item appearances ...[0 item(s)] done [0.00s].  
## set transactions ...[27 item(s), 2146 transaction(s)] done [0.00s].  
## sorting and recoding items ... [26 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2

## Warning in apriori(male\_eval, parameter = list(support = 0.04, confidence  
## = 0.3, : Mining stopped (maxlen reached). Only patterns up to a length of 2  
## returned!

## done [0.00s].  
## writing ... [322 rule(s)] done [0.00s].  
## creating S4 object ... done [0.00s].

*## 규칙 요약 보기*  
**summary**(rule\_male)

## set of 322 rules  
##   
## rule length distribution (lhs + rhs):sizes  
## 1 2   
## 17 305   
##   
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 2.000 2.000 1.947 2.000 2.000   
##   
## summary of quality measures:  
## support confidence lift count   
## Min. :0.04007 Min. :0.3001 Min. :0.5738 Min. : 86.0   
## 1st Qu.:0.10158 1st Qu.:0.3328 1st Qu.:0.9803 1st Qu.: 218.0   
## Median :0.12022 Median :0.3682 Median :1.0647 Median : 258.0   
## Mean :0.14512 Mean :0.4246 Mean :1.2320 Mean : 311.4   
## 3rd Qu.:0.16356 3rd Qu.:0.4729 3rd Qu.:1.3735 3rd Qu.: 351.0   
## Max. :0.75163 Max. :0.9932 Max. :7.3929 Max. :1613.0   
##   
## mining info:  
## data ntransactions support confidence  
## male\_eval 2146 0.04 0.3

*## 규칙 상세 보기*  
rule\_male.df <- **as**(rule\_male, "data.frame")  
*#남성 - BMI에 따른 운동능력의 등급 연관분석*  
rule\_male.BMI<- **inspect**(**subset**(rule\_male, subset=lhs **%pin%** "분류\_BMI" **&** rhs **%pin%** "등급"))

## lhs rhs support confidence  
## [1] {분류\_BMI=정상} => {등급\_20m왕복오래달리기=상} 0.10158434 0.3682432   
## [2] {분류\_BMI=정상} => {등급\_윗몸일으키기=상} 0.08620690 0.3125000   
## [3] {분류\_BMI=정상} => {등급\_제자리멀리뛰기=상} 0.09599254 0.3479730   
## [4] {분류\_BMI=정상} => {등급\_좌전굴=상} 0.09086673 0.3293919   
## [5] {분류\_BMI=정상} => {등급\_좌전굴=하} 0.08993476 0.3260135   
## [6] {분류\_BMI=정상} => {등급\_제자리멀리뛰기=하} 0.08341100 0.3023649   
## [7] {분류\_BMI=정상} => {등급\_윗몸일으키기=하} 0.09179870 0.3327703   
## [8] {분류\_BMI=정상} => {등급\_10m왕복빨리달리기=상} 0.09878844 0.3581081   
## [9] {분류\_BMI=정상} => {등급\_10m왕복빨리달리기=중} 0.10577819 0.3834459   
## [10] {분류\_BMI=정상} => {등급\_좌전굴=중} 0.09506058 0.3445946   
## [11] {분류\_BMI=정상} => {등급\_20m왕복오래달리기=중} 0.09878844 0.3581081   
## [12] {분류\_BMI=정상} => {등급\_제자리멀리뛰기=중} 0.09645853 0.3496622   
## [13] {분류\_BMI=정상} => {등급\_윗몸일으키기=중} 0.09785648 0.3547297   
## [14] {분류\_BMI=경도비만} => {등급\_20m왕복오래달리기=상} 0.10205033 0.3584288   
## [15] {분류\_BMI=경도비만} => {등급\_윗몸일으키기=상} 0.10158434 0.3567921   
## [16] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=상} 0.10344828 0.3633388   
## [17] {분류\_BMI=경도비만} => {등급\_좌전굴=상} 0.10764212 0.3780687   
## [18] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=상} 0.11183597 0.3927987   
## [19] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=중} 0.10671016 0.3747954   
## [20] {분류\_BMI=경도비만} => {등급\_좌전굴=중} 0.09878844 0.3469722   
## [21] {분류\_BMI=경도비만} => {등급\_20m왕복오래달리기=중} 0.11369991 0.3993453   
## [22] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=중} 0.10391426 0.3649755   
## [23] {분류\_BMI=경도비만} => {등급\_윗몸일으키기=중} 0.10764212 0.3780687   
## [24] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=하} 0.12907735 0.3445274   
## [25] {분류\_BMI=비만} => {등급\_좌전굴=상} 0.11276794 0.3009950   
## [26] {분류\_BMI=비만} => {등급\_좌전굴=하} 0.12255359 0.3271144   
## [27] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=하} 0.13280522 0.3544776   
## [28] {분류\_BMI=비만} => {등급\_윗몸일으키기=하} 0.12301957 0.3283582   
## [29] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=하} 0.15097856 0.4029851   
## [30] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=상} 0.12022367 0.3208955   
## [31] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=중} 0.12534949 0.3345771   
## [32] {분류\_BMI=비만} => {등급\_좌전굴=중} 0.13932898 0.3718905   
## [33] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=중} 0.13000932 0.3470149   
## [34] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=중} 0.14212488 0.3793532   
## [35] {분류\_BMI=비만} => {등급\_윗몸일으키기=중} 0.14165890 0.3781095   
## lift count  
## [1] 1.2083333 218   
## [2] 1.0238550 185   
## [3] 1.1246235 206   
## [4] 1.0334430 195   
## [5] 1.0037661 193   
## [6] 0.9243234 179   
## [7] 1.0143821 197   
## [8] 1.0385135 212   
## [9] 1.0784731 227   
## [10] 0.9666667 204   
## [11] 1.0032637 212   
## [12] 0.9620192 207   
## [13] 0.9672808 210   
## [14] 1.1761288 219   
## [15] 1.1689709 218   
## [16] 1.1742847 222   
## [17] 1.1861630 231   
## [18] 1.1391162 240   
## [19] 1.0541428 229   
## [20] 0.9733363 212   
## [21] 1.1187925 244   
## [22] 1.0041504 223   
## [23] 1.0309219 231   
## [24] 1.1498534 277   
## [25] 0.9443499 242   
## [26] 1.0071558 263   
## [27] 1.0836310 285   
## [28] 1.0009328 264   
## [29] 1.1911928 324   
## [30] 0.9305970 258   
## [31] 0.9410255 269   
## [32] 1.0432381 299   
## [33] 0.9721854 279   
## [34] 1.0437077 305   
## [35] 1.0310329 304

*#남성 - 체지방률에 따른 운동능력의 등급 연관분석*  
rule\_male.FAT<- **inspect**(**subset**(rule\_male, subset=lhs **%pin%** "분류\_체지방률=" **&** rhs **%pin%** "등급"))

## lhs rhs support   
## [1] {분류\_체지방률=낮음} => {등급\_20m왕복오래달리기=상} 0.07082945  
## [2] {분류\_체지방률=낮음} => {등급\_윗몸일으키기=상} 0.06383970  
## [3] {분류\_체지방률=낮음} => {등급\_제자리멀리뛰기=상} 0.07502330  
## [4] {분류\_체지방률=낮음} => {등급\_좌전굴=상} 0.06943150  
## [5] {분류\_체지방률=낮음} => {등급\_10m왕복빨리달리기=상} 0.07129543  
## [6] {분류\_체지방률=낮음} => {등급\_20m왕복오래달리기=중} 0.04426841  
## [7] {분류\_체지방률=낮음} => {등급\_윗몸일으키기=중} 0.04287046  
## [8] {분류\_체지방률=정상} => {등급\_20m왕복오래달리기=상} 0.09739049  
## [9] {분류\_체지방률=정상} => {등급\_윗몸일으키기=상} 0.09785648  
## [10] {분류\_체지방률=정상} => {등급\_제자리멀리뛰기=상} 0.10065238  
## [11] {분류\_체지방률=정상} => {등급\_좌전굴=상} 0.08527493  
## [12] {분류\_체지방률=정상} => {등급\_10m왕복빨리달리기=상} 0.10531221  
## [13] {분류\_체지방률=정상} => {등급\_10m왕복빨리달리기=중} 0.08014911  
## [14] {분류\_체지방률=정상} => {등급\_좌전굴=중} 0.09273066  
## [15] {분류\_체지방률=정상} => {등급\_20m왕복오래달리기=중} 0.08527493  
## [16] {분류\_체지방률=정상} => {등급\_제자리멀리뛰기=중} 0.08527493  
## [17] {분류\_체지방률=정상} => {등급\_윗몸일으키기=중} 0.07921715  
## [18] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=하} 0.13653308  
## [19] {분류\_체지방률=비만} => {등급\_좌전굴=하} 0.13793103  
## [20] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=하} 0.15377446  
## [21] {분류\_체지방률=비만} => {등급\_윗몸일으키기=하} 0.14631873  
## [22] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=하} 0.16356011  
## [23] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=중} 0.09925443  
## [24] {분류\_체지방률=비만} => {등급\_좌전굴=중} 0.10531221  
## [25] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=중} 0.09599254  
## [26] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=중} 0.10624418  
## [27] {분류\_체지방률=비만} => {등급\_윗몸일으키기=중} 0.10997204  
## [28] {분류\_체지방률=경도비만} => {등급\_윗몸일으키기=상} 0.10158434  
## [29] {분류\_체지방률=경도비만} => {등급\_좌전굴=상} 0.10904007  
## [30] {분류\_체지방률=경도비만} => {등급\_좌전굴=하} 0.10018639  
## [31] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=하} 0.10484623  
## [32] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=하} 0.10391426  
## [33] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=상} 0.10577819  
## [34] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=중} 0.13513514  
## [35] {분류\_체지방률=경도비만} => {등급\_좌전굴=중} 0.12395154  
## [36] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=중} 0.13140727  
## [37] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=중} 0.13280522  
## [38] {분류\_체지방률=경도비만} => {등급\_윗몸일으키기=중} 0.13466915  
## confidence lift count  
## [1] 0.5152542 1.6907272 152   
## [2] 0.4644068 1.5215526 137   
## [3] 0.5457627 1.7638656 161   
## [4] 0.5050847 1.5846665 149   
## [5] 0.5186441 1.5040678 153   
## [6] 0.3220339 0.9021994 95   
## [7] 0.3118644 0.8503952 92   
## [8] 0.4213710 1.3826638 209   
## [9] 0.4233871 1.3871583 210   
## [10] 0.4354839 1.4074524 216   
## [11] 0.3689516 1.1575587 183   
## [12] 0.4556452 1.3213710 226   
## [13] 0.3467742 0.9753308 172   
## [14] 0.4012097 1.1254849 199   
## [15] 0.3689516 1.0336425 183   
## [16] 0.3689516 1.0150900 183   
## [17] 0.3427419 0.9345924 170   
## [18] 0.4578125 1.5279403 293   
## [19] 0.4625000 1.4239957 296   
## [20] 0.5156250 1.5762553 330   
## [21] 0.4906250 1.4955700 314   
## [22] 0.5484375 1.6211389 351   
## [23] 0.3328125 0.9360624 213   
## [24] 0.3531250 0.9905964 226   
## [25] 0.3218750 0.9017542 206   
## [26] 0.3562500 0.9801442 228   
## [27] 0.3687500 1.0055114 236   
## [28] 0.3048951 0.9989388 218   
## [29] 0.3272727 1.0267943 234   
## [30] 0.3006993 0.9258260 215   
## [31] 0.3146853 0.9619867 225   
## [32] 0.3118881 0.9219172 223   
## [33] 0.3174825 0.9206993 227   
## [34] 0.4055944 1.1407675 290   
## [35] 0.3720280 1.0436236 266   
## [36] 0.3944056 1.1049535 282   
## [37] 0.3986014 1.0966649 285   
## [38] 0.4041958 1.1021654 289

*#남성 - 허리둘레에 따른 운동능력의 등급 연관분석*  
rule\_male.WRIST<- **inspect**(**subset**(rule\_male, subset=lhs **%pin%** "분류\_허리둘레=" **&** rhs **%pin%** "등급"))

## lhs rhs support   
## [1] {분류\_허리둘레=고도복부비만} => {등급\_10m왕복빨리달리기=하} 0.06197577  
## [2] {분류\_허리둘레=고도복부비만} => {등급\_좌전굴=하} 0.06104380  
## [3] {분류\_허리둘레=고도복부비만} => {등급\_제자리멀리뛰기=하} 0.06011184  
## [4] {분류\_허리둘레=고도복부비만} => {등급\_윗몸일으키기=하} 0.05871389  
## [5] {분류\_허리둘레=고도복부비만} => {등급\_20m왕복오래달리기=하} 0.07082945  
## [6] {분류\_허리둘레=고도복부비만} => {등급\_좌전굴=중} 0.04007456  
## [7] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=하} 0.05079217  
## [8] {분류\_허리둘레=복부비만} => {등급\_좌전굴=하} 0.04799627  
## [9] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=하} 0.04986021  
## [10] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=하} 0.05032619  
## [11] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=하} 0.06337372  
## [12] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=상} 0.04054054  
## [13] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=중} 0.04240447  
## [14] {분류\_허리둘레=복부비만} => {등급\_좌전굴=중} 0.04799627  
## [15] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=중} 0.04240447  
## [16] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=중} 0.04846226  
## [17] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=중} 0.05405405  
## [18] {분류\_허리둘레=정상} => {등급\_20m왕복오래달리기=상} 0.26188257  
## [19] {분류\_허리둘레=정상} => {등급\_윗몸일으키기=상} 0.25908667  
## [20] {분류\_허리둘레=정상} => {등급\_제자리멀리뛰기=상} 0.25768872  
## [21] {분류\_허리둘레=정상} => {등급\_좌전굴=상} 0.26747437  
## [22] {분류\_허리둘레=정상} => {등급\_10m왕복빨리달리기=상} 0.28471575  
## [23] {분류\_허리둘레=정상} => {등급\_10m왕복빨리달리기=중} 0.28005592  
## [24] {분류\_허리둘레=정상} => {등급\_좌전굴=중} 0.26840634  
## [25] {분류\_허리둘레=정상} => {등급\_20m왕복오래달리기=중} 0.28564772  
## [26] {분류\_허리둘레=정상} => {등급\_제자리멀리뛰기=중} 0.27679404  
## [27] {분류\_허리둘레=정상} => {등급\_윗몸일으키기=중} 0.27353215  
## confidence lift count  
## [1] 0.5406504 1.8044102 133   
## [2] 0.5325203 1.6395819 131   
## [3] 0.5243902 1.6030505 129   
## [4] 0.5121951 1.5613221 126   
## [5] 0.6178862 1.8264239 152   
## [6] 0.3495935 0.9806897 86   
## [7] 0.3797909 1.2675449 109   
## [8] 0.3588850 1.1049745 103   
## [9] 0.3728223 1.1397103 107   
## [10] 0.3763066 1.1470938 108   
## [11] 0.4738676 1.4007161 136   
## [12] 0.3031359 0.8790941 87   
## [13] 0.3170732 0.8917943 91   
## [14] 0.3588850 1.0067546 103   
## [15] 0.3170732 0.8883016 91   
## [16] 0.3623693 0.9969803 104   
## [17] 0.4041812 1.1021256 116   
## [18] 0.3484191 1.1432835 562   
## [19] 0.3446993 1.1293507 556   
## [20] 0.3428394 1.1080323 553   
## [21] 0.3558586 1.1164805 574   
## [22] 0.3787973 1.0985121 611   
## [23] 0.3725976 1.0479614 601   
## [24] 0.3570986 1.0017432 576   
## [25] 0.3800372 1.0646995 613   
## [26] 0.3682579 1.0131814 594   
## [27] 0.3639182 0.9923359 587

*#남성 - 분류가 비만일때 따른 운동능력 '하'등급 연관분석*  
rule\_male.LOW<- **inspect**(**subset**(rule\_male, subset=lhs **%pin%** "비만" **&** rhs **%pin%** "하"))

## lhs rhs support   
## [1] {분류\_허리둘레=고도복부비만} => {등급\_10m왕복빨리달리기=하} 0.06197577  
## [2] {분류\_허리둘레=고도복부비만} => {등급\_좌전굴=하} 0.06104380  
## [3] {분류\_허리둘레=고도복부비만} => {등급\_제자리멀리뛰기=하} 0.06011184  
## [4] {분류\_허리둘레=고도복부비만} => {등급\_윗몸일으키기=하} 0.05871389  
## [5] {분류\_허리둘레=고도복부비만} => {등급\_20m왕복오래달리기=하} 0.07082945  
## [6] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=하} 0.05079217  
## [7] {분류\_허리둘레=복부비만} => {등급\_좌전굴=하} 0.04799627  
## [8] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=하} 0.04986021  
## [9] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=하} 0.05032619  
## [10] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=하} 0.06337372  
## [11] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=하} 0.13653308  
## [12] {분류\_체지방률=비만} => {등급\_좌전굴=하} 0.13793103  
## [13] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=하} 0.15377446  
## [14] {분류\_체지방률=비만} => {등급\_윗몸일으키기=하} 0.14631873  
## [15] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=하} 0.16356011  
## [16] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=하} 0.12907735  
## [17] {분류\_체지방률=경도비만} => {등급\_좌전굴=하} 0.10018639  
## [18] {분류\_BMI=비만} => {등급\_좌전굴=하} 0.12255359  
## [19] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=하} 0.10484623  
## [20] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=하} 0.13280522  
## [21] {분류\_BMI=비만} => {등급\_윗몸일으키기=하} 0.12301957  
## [22] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=하} 0.10391426  
## [23] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=하} 0.15097856  
## confidence lift count  
## [1] 0.5406504 1.8044102 133   
## [2] 0.5325203 1.6395819 131   
## [3] 0.5243902 1.6030505 129   
## [4] 0.5121951 1.5613221 126   
## [5] 0.6178862 1.8264239 152   
## [6] 0.3797909 1.2675449 109   
## [7] 0.3588850 1.1049745 103   
## [8] 0.3728223 1.1397103 107   
## [9] 0.3763066 1.1470938 108   
## [10] 0.4738676 1.4007161 136   
## [11] 0.4578125 1.5279403 293   
## [12] 0.4625000 1.4239957 296   
## [13] 0.5156250 1.5762553 330   
## [14] 0.4906250 1.4955700 314   
## [15] 0.5484375 1.6211389 351   
## [16] 0.3445274 1.1498534 277   
## [17] 0.3006993 0.9258260 215   
## [18] 0.3271144 1.0071558 263   
## [19] 0.3146853 0.9619867 225   
## [20] 0.3544776 1.0836310 285   
## [21] 0.3283582 1.0009328 264   
## [22] 0.3118881 0.9219172 223   
## [23] 0.4029851 1.1911928 324

*##기준을 충족한 규칙들중 연관성을 표현할수있는 lift를 기준으로 평가해보면(lift > 1만 다룸),*  
*##BMI가 정상으로 분류된 집단은 9가지 룰을 가지고있는데, 그 중 "하" 2개, "중" 2개, "상" 5개였다.*  
*## 경도비만으로 분류된 집단은 9가지 룰을 가지고있는데, 그 중 "중" 4개, "상" 5개 "하"등급은 1개도 없다.*  
*## 비만으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "하" 5개, "중" 3개로 "상"등급은 1개도 없다.*  
*## 비만등급은 "상"등급이 1개도 없는 걸로보아 연관성이 있어보이기는 하나,*   
*## 경도비만으로 분류된 집단이 정상보다도 더 좋은성적을 거두었다.*  
*## BMI는 제지방량과 지방량을 구분하지 못해서 근육량이 많은 사람들이 경도비만으로 분류되었을 가능성이 있어서 이런경향이 나타나는것 같다.*  
*##체지방률이 낮음으로 분류된 집단은 5가지 룰을 가지고있는데, 모두 "상"이었다.*  
*## 정상으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "중" 3개, 상" 5개로, "하"등급은 1개도 없다.*  
*## 경도비만으로 분류된 집단은 6가지룰을 가지고있는데, 그 중 "중" 5개, "상" 1개 "하"등급은 1개도 없다.*  
*## 비만으로 분류된 집단은 6가지 룰을 가지고있는데, 그 중 "하" 5개, "중" 1개로 "상"등급은 1개도 없다.*  
*## 낮음에서 비만으로 체지방률이 증가함에따라 "상"등급은 줄어들고, "하"등급이 늘어나는 패턴이 보인다.*  
*## 정상과 경도비만의 경계가 BMI는 모호한데에 비해 체지방률은 정확하다고 볼수있다.*  
*##허리둘레가 정상으로 분류된 집단은 9가지 룰을 가지고있는데, 그 중 "중" 4개, "상" 5개으로, "하"등급은 1개도 없다.*  
*## 복부비만으로 분류된 집단은 7가지룰을 가지고있는데, 그 중 "하" 5개, "중" 2개로, "상"등급은 1개도 없다.*  
*## 고도복부비만으로 분류된 집단은 5가지 룰을 가지고있는데, 모두 "하" 등급이다.*  
*##체지방률과 허리둘레는 신체능력등급에서 서로 비슷한 수준의 경향을보이며 BMI보다 더 신체능력과 밀접한 연관성이 있다고 보인다.*  
  
  
  
*###여성데이터 변환*  
female\_eval**$**분류\_BMI <- **factor**(female\_eval**$**분류\_BMI, levels=**c**(1**:**5), labels=**c**('낮음', '정상', '경도비만', '비만', '고도비만'))  
female\_eval**$**분류\_체지방률 <- **factor**(female\_eval**$**분류\_체지방률, levels=**c**(1**:**4), labels=**c**('낮음', '정상', '경도비만', '비만'))  
female\_eval**$**분류\_허리둘레 <- **factor**(female\_eval**$**분류\_허리둘레, levels=**c**(1**:**3), labels=**c**('정상', '복부비만','고도복부비만'))  
female\_eval**$**등급\_윗몸일으키기 <- **factor**(female\_eval**$**등급\_윗몸일으키기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
female\_eval**$**등급\_좌전굴 <- **factor**(female\_eval**$**등급\_좌전굴, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
female\_eval**$**등급\_제자리멀리뛰기 <- **factor**(female\_eval**$**등급\_제자리멀리뛰기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
female\_eval**$**등급\_10m왕복빨리달리기 <- **factor**(female\_eval**$**등급\_10m왕복빨리달리기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
female\_eval**$**등급\_20m왕복오래달리기 <- **factor**(female\_eval**$**등급\_20m왕복오래달리기, levels=**c**(1**:**3), labels=**c**('상', '중', '하'))  
*#female\_eval$등급\_상대악력 <- factor(female\_eval$등급\_상대악력, levels=c(1:3), labels=c('상', '중', '하'))*  
*#female\_eval$등급\_악력 <- factor(female\_eval$등급\_악력, levels=c(1:3), labels=c('상', '중', '하'))*  
  
*## 조건 (support=0.04, confidence=0.3, minlen=1, maxlen=2)으로 연관분석 시작*  
rule\_female <- **apriori**(female\_eval, parameter=**list**(support=0.04, confidence=0.3, minlen=1,maxlen=2))

## Apriori  
##   
## Parameter specification:  
## confidence minval smax arem aval originalSupport maxtime support minlen  
## 0.3 0.1 1 none FALSE TRUE 5 0.04 1  
## maxlen target ext  
## 2 rules FALSE  
##   
## Algorithmic control:  
## filter tree heap memopt load sort verbose  
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE  
##   
## Absolute minimum support count: 85   
##   
## set item appearances ...[0 item(s)] done [0.00s].  
## set transactions ...[27 item(s), 2146 transaction(s)] done [0.00s].  
## sorting and recoding items ... [25 item(s)] done [0.00s].  
## creating transaction tree ... done [0.00s].  
## checking subsets of size 1 2

## Warning in apriori(female\_eval, parameter = list(support = 0.04, confidence  
## = 0.3, : Mining stopped (maxlen reached). Only patterns up to a length of 2  
## returned!

## done [0.00s].  
## writing ... [305 rule(s)] done [0.00s].  
## creating S4 object ... done [0.00s].

*## 규칙 요약 보기*  
**summary**(rule\_female)

## set of 305 rules  
##   
## rule length distribution (lhs + rhs):sizes  
## 1 2   
## 18 287   
##   
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 2.000 2.000 1.941 2.000 2.000   
##   
## summary of quality measures:  
## support confidence lift count   
## Min. :0.04054 Min. :0.3008 Min. :0.3942 Min. : 87.0   
## 1st Qu.:0.10904 1st Qu.:0.3391 1st Qu.:0.9837 1st Qu.: 234.0   
## Median :0.13467 Median :0.3720 Median :1.0426 Median : 289.0   
## Mean :0.15956 Mean :0.4403 Mean :1.1953 Mean : 342.4   
## 3rd Qu.:0.18919 3rd Qu.:0.4975 3rd Qu.:1.3048 3rd Qu.: 406.0   
## Max. :0.79357 Max. :1.0000 Max. :3.4952 Max. :1703.0   
##   
## mining info:  
## data ntransactions support confidence  
## female\_eval 2146 0.04 0.3

*## 규칙 상세 보기*  
rule\_female.df <- **as**(rule\_female, "data.frame")  
*#여성 - BMI에 따른 운동능력의 등급 연관분석*  
rule\_female.BMI<- **inspect**(**subset**(rule\_female, subset=lhs **%pin%** "분류\_BMI=" **&** rhs **%pin%** "등급"))

## lhs rhs support confidence  
## [1] {분류\_BMI=비만} => {등급\_좌전굴=하} 0.06383970 0.3663102   
## [2] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=하} 0.07502330 0.4304813   
## [3] {분류\_BMI=비만} => {등급\_윗몸일으키기=하} 0.06989748 0.4010695   
## [4] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=하} 0.07362535 0.4224599   
## [5] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=중} 0.06150979 0.3529412   
## [6] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=중} 0.05871389 0.3368984   
## [7] {분류\_BMI=비만} => {등급\_좌전굴=중} 0.06710158 0.3850267   
## [8] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=중} 0.05917987 0.3395722   
## [9] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=하} 0.08667288 0.4973262   
## [10] {분류\_BMI=비만} => {등급\_윗몸일으키기=중} 0.06570363 0.3770053   
## [11] {분류\_BMI=경도비만} => {등급\_윗몸일으키기=상} 0.06849953 0.3426573   
## [12] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=상} 0.06943150 0.3473193   
## [13] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=상} 0.06570363 0.3286713   
## [14] {분류\_BMI=경도비만} => {등급\_좌전굴=상} 0.06710158 0.3356643   
## [15] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=하} 0.06430568 0.3216783   
## [16] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=하} 0.06383970 0.3193473   
## [17] {분류\_BMI=경도비만} => {등급\_20m왕복오래달리기=중} 0.07269338 0.3636364   
## [18] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=중} 0.06989748 0.3496503   
## [19] {분류\_BMI=경도비만} => {등급\_좌전굴=중} 0.07595527 0.3799534   
## [20] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=중} 0.06663560 0.3333333   
## [21] {분류\_BMI=경도비만} => {등급\_20m왕복오래달리기=하} 0.07082945 0.3543124   
## [22] {분류\_BMI=경도비만} => {등급\_윗몸일으키기=중} 0.07362535 0.3682984   
## [23] {분류\_BMI=정상} => {등급\_20m왕복오래달리기=상} 0.19291705 0.3529412   
## [24] {분류\_BMI=정상} => {등급\_윗몸일으키기=상} 0.18499534 0.3384484   
## [25] {분류\_BMI=정상} => {등급\_제자리멀리뛰기=상} 0.18313141 0.3350384   
## [26] {분류\_BMI=정상} => {등급\_10m왕복빨리달리기=상} 0.18918919 0.3461211   
## [27] {분류\_BMI=정상} => {등급\_좌전굴=하} 0.17008388 0.3111679   
## [28] {분류\_BMI=정상} => {등급\_좌전굴=상} 0.19617894 0.3589088   
## [29] {분류\_BMI=정상} => {등급\_윗몸일으키기=하} 0.16635601 0.3043478   
## [30] {분류\_BMI=정상} => {등급\_20m왕복오래달리기=중} 0.19105312 0.3495311   
## [31] {분류\_BMI=정상} => {등급\_10m왕복빨리달리기=중} 0.20270270 0.3708440   
## [32] {분류\_BMI=정상} => {등급\_좌전굴=중} 0.18033551 0.3299233   
## [33] {분류\_BMI=정상} => {등급\_제자리멀리뛰기=중} 0.20083877 0.3674339   
## [34] {분류\_BMI=정상} => {등급\_윗몸일으키기=중} 0.19524697 0.3572038   
## lift count  
## [1] 1.1376289 137   
## [2] 1.3048204 161   
## [3] 1.2156712 150   
## [4] 1.2679705 158   
## [5] 1.0125826 132   
## [6] 0.9538047 126   
## [7] 1.0900625 144   
## [8] 0.9563280 127   
## [9] 1.3969398 186   
## [10] 1.0425947 141   
## [11] 1.1107895 147   
## [12] 1.1141216 149   
## [13] 1.0372480 141   
## [14] 1.0334801 144   
## [15] 0.9750306 138   
## [16] 0.9584886 137   
## [17] 1.0432669 156   
## [18] 0.9899072 150   
## [19] 1.0756991 163   
## [20] 0.9387577 143   
## [21] 0.9952282 152   
## [22] 1.0185158 158   
## [23] 1.1946558 414   
## [24] 1.0971455 397   
## [25] 1.0747269 393   
## [26] 1.0923173 406   
## [27] 0.9663769 365   
## [28] 1.1050477 421   
## [29] 0.9225006 357   
## [30] 1.0027992 410   
## [31] 1.0499092 435   
## [32] 0.9340572 387   
## [33] 1.0347942 431   
## [34] 0.9878341 419

*#여성 - 체지방률에 따른 운동능력의 등급 연관분석*  
rule\_female.FAT<- **inspect**(**subset**(rule\_female, subset=lhs **%pin%** "분류\_체지방률=" **&** rhs **%pin%** "등급"))

## lhs rhs support   
## [1] {분류\_체지방률=정상} => {등급\_20m왕복오래달리기=상} 0.12861137  
## [2] {분류\_체지방률=정상} => {등급\_윗몸일으키기=상} 0.12068966  
## [3] {분류\_체지방률=정상} => {등급\_제자리멀리뛰기=상} 0.12907735  
## [4] {분류\_체지방률=정상} => {등급\_10m왕복빨리달리기=상} 0.12954334  
## [5] {분류\_체지방률=정상} => {등급\_좌전굴=상} 0.11742777  
## [6] {분류\_체지방률=정상} => {등급\_20m왕복오래달리기=중} 0.08294501  
## [7] {분류\_체지방률=정상} => {등급\_10m왕복빨리달리기=중} 0.08387698  
## [8] {분류\_체지방률=정상} => {등급\_좌전굴=중} 0.08853681  
## [9] {분류\_체지방률=정상} => {등급\_제자리멀리뛰기=중} 0.08247903  
## [10] {분류\_체지방률=정상} => {등급\_윗몸일으키기=중} 0.08900280  
## [11] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=상} 0.10810811  
## [12] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=상} 0.11090401  
## [13] {분류\_체지방률=경도비만} => {등급\_좌전굴=하} 0.11509786  
## [14] {분류\_체지방률=비만} => {등급\_좌전굴=하} 0.14771668  
## [15] {분류\_체지방률=경도비만} => {등급\_좌전굴=상} 0.11463187  
## [16] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=하} 0.10810811  
## [17] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=하} 0.17054986  
## [18] {분류\_체지방률=경도비만} => {등급\_윗몸일으키기=하} 0.11463187  
## [19] {분류\_체지방률=비만} => {등급\_윗몸일으키기=하} 0.16309413  
## [20] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=하} 0.10904007  
## [21] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=하} 0.17101584  
## [22] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=중} 0.13466915  
## [23] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=중} 0.12395154  
## [24] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=중} 0.13420317  
## [25] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=중} 0.12721342  
## [26] {분류\_체지방률=경도비만} => {등급\_좌전굴=중} 0.12348555  
## [27] {분류\_체지방률=비만} => {등급\_좌전굴=중} 0.13140727  
## [28] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=중} 0.13606710  
## [29] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=하} 0.11835974  
## [30] {분류\_체지방률=경도비만} => {등급\_윗몸일으키기=중} 0.13280522  
## [31] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=중} 0.12628145  
## [32] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=하} 0.18592731  
## [33] {분류\_체지방률=비만} => {등급\_윗몸일으키기=중} 0.13140727  
## confidence lift count  
## [1] 0.4955117 1.6772367 276   
## [2] 0.4649910 1.5073576 259   
## [3] 0.4973070 1.5952479 277   
## [4] 0.4991023 1.5751082 278   
## [5] 0.4524237 1.3929717 252   
## [6] 0.3195691 0.9168387 178   
## [7] 0.3231598 0.9149088 180   
## [8] 0.3411131 0.9657371 190   
## [9] 0.3177738 0.8949377 177   
## [10] 0.3429084 0.9483009 191   
## [11] 0.3060686 0.9817985 232   
## [12] 0.3139842 0.9908971 238   
## [13] 0.3258575 1.0119974 247   
## [14] 0.4122237 1.2802200 317   
## [15] 0.3245383 0.9992240 246   
## [16] 0.3060686 0.9277164 232   
## [17] 0.4759428 1.4426175 366   
## [18] 0.3245383 0.9836993 246   
## [19] 0.4551365 1.3795523 350   
## [20] 0.3087071 0.9265531 234   
## [21] 0.4772432 1.4323970 367   
## [22] 0.3812665 1.0938474 289   
## [23] 0.3459038 0.9923924 266   
## [24] 0.3799472 1.0756817 288   
## [25] 0.3550065 1.0050712 273   
## [26] 0.3496042 0.9897766 265   
## [27] 0.3667100 1.0382054 282   
## [28] 0.3852243 1.0848967 292   
## [29] 0.3350923 0.9412411 254   
## [30] 0.3759894 1.0397852 285   
## [31] 0.3524057 0.9924707 271   
## [32] 0.5188557 1.4574139 399   
## [33] 0.3667100 1.0141233 282

*#여성 - 허리둘레에 따른 운동능력의 등급 연관분석*  
rule\_female.WRIST<- **inspect**(**subset**(rule\_female, subset=lhs **%pin%** "분류\_허리둘레=" **&** rhs **%pin%** "등급"))

## lhs rhs support   
## [1] {분류\_허리둘레=고도복부비만} => {등급\_좌전굴=하} 0.04566636  
## [2] {분류\_허리둘레=고도복부비만} => {등급\_10m왕복빨리달리기=하} 0.05312209  
## [3] {분류\_허리둘레=고도복부비만} => {등급\_윗몸일으키기=하} 0.05125815  
## [4] {분류\_허리둘레=고도복부비만} => {등급\_제자리멀리뛰기=하} 0.05312209  
## [5] {분류\_허리둘레=고도복부비만} => {등급\_20m왕복오래달리기=하} 0.05731594  
## [6] {분류\_허리둘레=복부비만} => {등급\_좌전굴=하} 0.04473439  
## [7] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=하} 0.04659832  
## [8] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=하} 0.04193849  
## [9] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=하} 0.04054054  
## [10] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=중} 0.04380242  
## [11] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=중} 0.04287046  
## [12] {분류\_허리둘레=복부비만} => {등급\_좌전굴=중} 0.04380242  
## [13] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=중} 0.04287046  
## [14] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=하} 0.05265610  
## [15] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=중} 0.04846226  
## [16] {분류\_허리둘레=정상} => {등급\_20m왕복오래달리기=상} 0.26980429  
## [17] {분류\_허리둘레=정상} => {등급\_윗몸일으키기=상} 0.27027027  
## [18] {분류\_허리둘레=정상} => {등급\_제자리멀리뛰기=상} 0.26654240  
## [19] {분류\_허리둘레=정상} => {등급\_10m왕복빨리달리기=상} 0.27632805  
## [20] {분류\_허리둘레=정상} => {등급\_좌전굴=상} 0.28471575  
## [21] {분류\_허리둘레=정상} => {등급\_제자리멀리뛰기=하} 0.23951538  
## [22] {분류\_허리둘레=정상} => {등급\_20m왕복오래달리기=중} 0.27772600  
## [23] {분류\_허리둘레=정상} => {등급\_10m왕복빨리달리기=중} 0.28704567  
## [24] {분류\_허리둘레=정상} => {등급\_좌전굴=중} 0.27726002  
## [25] {분류\_허리둘레=정상} => {등급\_제자리멀리뛰기=중} 0.28751165  
## [26] {분류\_허리둘레=정상} => {등급\_20m왕복오래달리기=하} 0.24603914  
## [27] {분류\_허리둘레=정상} => {등급\_윗몸일으키기=중} 0.28657968  
## confidence lift count  
## [1] 0.4974619 1.5449397 98   
## [2] 0.5786802 1.7540222 114   
## [3] 0.5583756 1.6924776 110   
## [4] 0.5786802 1.7368500 114   
## [5] 0.6243655 1.7537805 123   
## [6] 0.3902439 1.2119586 96   
## [7] 0.4065041 1.2321437 100   
## [8] 0.3658537 1.1089293 90   
## [9] 0.3536585 1.0614702 87   
## [10] 0.3821138 1.0962784 94   
## [11] 0.3739837 1.0587983 92   
## [12] 0.3821138 1.0818156 94   
## [13] 0.3739837 1.0532403 92   
## [14] 0.4593496 1.2902673 113   
## [15] 0.4227642 1.1691392 104   
## [16] 0.3399883 1.1508120 579   
## [17] 0.3405755 1.1040407 580   
## [18] 0.3358779 1.0774199 572   
## [19] 0.3482090 1.0989068 593   
## [20] 0.3587786 1.1046470 611   
## [21] 0.3018203 0.9058831 514   
## [22] 0.3499706 1.0040602 596   
## [23] 0.3617146 1.0240628 616   
## [24] 0.3493834 0.9891515 595   
## [25] 0.3623018 1.0203408 617   
## [26] 0.3100411 0.8708746 528   
## [27] 0.3611274 0.9986849 615

*#여성 - 분류가 비만일때 따른 운동능력 '하'등급 연관분석*  
rule\_female.LOW<- **inspect**(**subset**(rule\_female, subset=lhs **%pin%** "비만" **&** rhs **%pin%** "하"))

## lhs rhs support   
## [1] {분류\_허리둘레=고도복부비만} => {등급\_좌전굴=하} 0.04566636  
## [2] {분류\_허리둘레=고도복부비만} => {등급\_10m왕복빨리달리기=하} 0.05312209  
## [3] {분류\_허리둘레=고도복부비만} => {등급\_윗몸일으키기=하} 0.05125815  
## [4] {분류\_허리둘레=고도복부비만} => {등급\_제자리멀리뛰기=하} 0.05312209  
## [5] {분류\_허리둘레=고도복부비만} => {등급\_20m왕복오래달리기=하} 0.05731594  
## [6] {분류\_허리둘레=복부비만} => {등급\_좌전굴=하} 0.04473439  
## [7] {분류\_허리둘레=복부비만} => {등급\_10m왕복빨리달리기=하} 0.04659832  
## [8] {분류\_허리둘레=복부비만} => {등급\_윗몸일으키기=하} 0.04193849  
## [9] {분류\_허리둘레=복부비만} => {등급\_제자리멀리뛰기=하} 0.04054054  
## [10] {분류\_허리둘레=복부비만} => {등급\_20m왕복오래달리기=하} 0.05265610  
## [11] {분류\_BMI=비만} => {등급\_좌전굴=하} 0.06383970  
## [12] {분류\_BMI=비만} => {등급\_10m왕복빨리달리기=하} 0.07502330  
## [13] {분류\_BMI=비만} => {등급\_윗몸일으키기=하} 0.06989748  
## [14] {분류\_BMI=비만} => {등급\_제자리멀리뛰기=하} 0.07362535  
## [15] {분류\_BMI=비만} => {등급\_20m왕복오래달리기=하} 0.08667288  
## [16] {분류\_BMI=경도비만} => {등급\_10m왕복빨리달리기=하} 0.06430568  
## [17] {분류\_BMI=경도비만} => {등급\_제자리멀리뛰기=하} 0.06383970  
## [18] {분류\_BMI=경도비만} => {등급\_20m왕복오래달리기=하} 0.07082945  
## [19] {분류\_체지방률=경도비만} => {등급\_좌전굴=하} 0.11509786  
## [20] {분류\_체지방률=비만} => {등급\_좌전굴=하} 0.14771668  
## [21] {분류\_체지방률=경도비만} => {등급\_10m왕복빨리달리기=하} 0.10810811  
## [22] {분류\_체지방률=비만} => {등급\_10m왕복빨리달리기=하} 0.17054986  
## [23] {분류\_체지방률=경도비만} => {등급\_윗몸일으키기=하} 0.11463187  
## [24] {분류\_체지방률=비만} => {등급\_윗몸일으키기=하} 0.16309413  
## [25] {분류\_체지방률=경도비만} => {등급\_제자리멀리뛰기=하} 0.10904007  
## [26] {분류\_체지방률=비만} => {등급\_제자리멀리뛰기=하} 0.17101584  
## [27] {분류\_체지방률=경도비만} => {등급\_20m왕복오래달리기=하} 0.11835974  
## [28] {분류\_체지방률=비만} => {등급\_20m왕복오래달리기=하} 0.18592731  
## confidence lift count  
## [1] 0.4974619 1.5449397 98   
## [2] 0.5786802 1.7540222 114   
## [3] 0.5583756 1.6924776 110   
## [4] 0.5786802 1.7368500 114   
## [5] 0.6243655 1.7537805 123   
## [6] 0.3902439 1.2119586 96   
## [7] 0.4065041 1.2321437 100   
## [8] 0.3658537 1.1089293 90   
## [9] 0.3536585 1.0614702 87   
## [10] 0.4593496 1.2902673 113   
## [11] 0.3663102 1.1376289 137   
## [12] 0.4304813 1.3048204 161   
## [13] 0.4010695 1.2156712 150   
## [14] 0.4224599 1.2679705 158   
## [15] 0.4973262 1.3969398 186   
## [16] 0.3216783 0.9750306 138   
## [17] 0.3193473 0.9584886 137   
## [18] 0.3543124 0.9952282 152   
## [19] 0.3258575 1.0119974 247   
## [20] 0.4122237 1.2802200 317   
## [21] 0.3060686 0.9277164 232   
## [22] 0.4759428 1.4426175 366   
## [23] 0.3245383 0.9836993 246   
## [24] 0.4551365 1.3795523 350   
## [25] 0.3087071 0.9265531 234   
## [26] 0.4772432 1.4323970 367   
## [27] 0.3350923 0.9412411 254   
## [28] 0.5188557 1.4574139 399

*##기준을 충족한 규칙들중 연관성을 표현할수있는 lift를 기준으로 평가해보면(lift > 1만 다룸),*  
*##BMI가 정상으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "중" 3개, "상" 5개로 "하"등급은 1개도 없다.*  
*## 경도비만으로 분류된 집단은 7가지 룰을 가지고있는데, 그 중 "중" 3개, "상" 4개 "하"등급은 1개도 없다.*  
*## 비만으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "하" 5개, "중" 3개로 "상"등급은 1개도 없다.*  
*## 비만등급은 "상"등급이 1개도 없는 걸로보아 연관성이 있어보이기는 하나,*   
*## 경도비만으로 분류된 집단과 정상으로 분류된 집단 사이에 별다른 차이가 없다.*  
*## 여성 또한 BMI는 제지방량과 지방량을 구분하지 못해서 근육량이 많은 사람들이 경도비만으로 분류되었을 가능성이 있어서 이런경향이 나타나는것 같다.*  
*## 남성과 여성 모두 경도비만으로 분류된 집단은 좋은 신체등급을 가지고있고,*  
*## 비만으로 분류된 집단부터 좋지못한 신체등급을 받고있다.*  
*## BMI 지수의 경도비만의 경계를 좀더 높여야될 필요성이 보인다.*  
*##체지방률이 정상으로 분류된 집단은 5가지 룰을 가지고있는데, 모두 "상" 등급이다.*   
*## 경도비만으로 분류된 집단은 5가지룰을 가지고있는데, 그 중 "하" 1개, "중" 4개 "상"등급은 1개도 없다.*  
*## 비만으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "하" 5개, "중" 3개로 "상"등급은 1개도 없다.*  
*## 여성 또한 정상에서 비만으로 체지방률이 증가함에따라 눈에띄게 "상"등급은 줄어들고, "하"등급이 늘어나는 패턴이 보인다.*  
*## 남성과 여성 모두 정상과 경도비만의 경계가 BMI는 모호한데에 비해 체지방률은 정확하다고 볼수있다.*  
*##허리둘레가 정상으로 분류된 집단은 8가지 룰을 가지고있는데, 그 중 "중" 3개, "상" 5개으로, "하"등급은 1개도 없다.*  
*## 복부비만으로 분류된 집단은 10가지룰을 가지고있는데, 그 중 "하" 5개, "중" 5개로, "상"등급은 1개도 없다.*  
*## 고도복부비만으로 분류된 집단은 5가지 룰을 가지고있는데, 모두 "하" 등급이다.*  
*##남성과 여성 모두 체지방률과 허리둘레는 신체능력등급에서 서로 비슷한 수준의 경향을보이며 BMI보다 더 신체능력과 밀접한 연관성이 있다고 보인다.*  
  
  
  
  
  
*### 최종 분석결과 : 체지방률이 가장 정확하고 BMI보다 허리둘레가 체력수준을 더 잘 파악할수있다.*   
  
*#체지량지수인 BMI는 단순히 몸무게에 키의제곱을 나누는것으로 구할수있기 때문에 몸무게에 비례한다.*  
*#체지방률은 쉽게 알기힘들기때문에 비교적 쉽게 알수있는 BMI를 이용하지만,*   
*#근육질인 운동선수들도 BMI지수로 바라본다면 비만일수있고 팔다리마르고 배만나온마른비만들도 BMI에서는 정상으로 나올수있다.*   
*#그렇기 때문에 BMI는 우리 신체건강상태를 파악할수있는 적절한 기준이 되지못한다.*   
*#우리는 질병을 제외하고서 신체적으로 건강하다는것은 신체적능력이 뛰어난것이라고 말할수있다.*   
*#이러한 신체적 퍼포먼스가 BMI에 영향을 많이 받는지 혹은 체지방률이나 허리둘레가 더 영향을 많이 받는것인지 데이터를 통해 알아보니,*  
*#만약 우리가 운동을하거나 다이어트를 하려고한다면 근육이나 수분같은 제지방량이 빠져도 줄어들수있는 체중이나 BMI를 줄이는것보다*   
*#체지방이 집중되있는 복부지방을 줄이는 방향으로 허리둘레줄이기를 목표로하는것이 더욱 바람직 하다.*  
*#그리고 다른 분류기준들에 비해 BMI는 정상과 경도비만의 구분이 모호하기때문에 BMI수치을 좀더 적절하게 이용하기 위해서는 기준을 현대체형에 맞게 조절해야할 필요가 있어보인다.*

**3. 프로젝트 수행 성과**

**1) 프로젝트 성과 및 기대효과**

이 프로젝트에서 신체능력과 밀접한 관련이 있는것은 BMI가 아니라 허리둘레라는것을 데이터를 통해 확인 하였습니다. 그렇기 때문에 앞으로 만약 다이어트를 하고 건강해지기 위해서는 체중을 줄이기보다는 허리둘레를 줄이는 것을 기대해봅니다.

만약 가족들이나 친구들이 다이어트를 한다고 체중계를 사거나, 사려고 고민한다면 줄자를 하나 선물해주면서 체중보다는 허리둘레를 줄이는게 더 바람직하다고 이야기해줄것입니다.

그리고 이번 프로젝트를 통해 대한비만학회에서 정한 BMI기준이 체지방률과 허리둘레와 달리 정상과 경도비만이 구분이 모호하다는 것을 보았고, 미약하지만 오히려 정상보다도 경도비만이 더뛰어난 체력평가를 받았는것을 확인하였습니다. 그렇기 때문에 대한비만학회에서는 BMI가 서구화된 한국인의 체형에 맞게 적절하게 적용되려면 기준을 좀 더 상향시키기를 기대합니다.

**2) Lessons Learned**

미디어학부지만 C언어, 자바등 일반 프로그래밍언어만 쓸줄 알았지 데이터를 다룰줄을 몰랐습니다.

그래서 이번 프로젝트를 통해서 R이 데이터를 분석하기 정말 강력한 언어라고 느꼈습니다. 그리고 통계청 마이크로데이터 사이트(<https://mdis.kostat.go.kr/>)라는 양질의 우리 사회데이터를 제공하는 사이트를 알았고,

만약 나의 데이터가 포함 되어있을 수도 있는 우리사회 데이터를 가지고 이리저리 탐색하고 분석해보니 통계와 데이터가 훨씬 친근하고 가깝게 느껴졌습니다.

그리고 2017년이후 데이터를 선택해야되서 정말로 분석해보고싶은 데이터를 분석하지 못해서 아쉬웠지만 시대가 워낙 빠르게 변화하다보니 2년보다 더 지난 데이터를 분석해서 현시대에 그 패턴이나 분석결과를 적용시키는것이 적절하지 못할수도 있기 때문이라 생각하니 이해가 되었습니다.

이번 프로젝트로 R을 어느정도 능숙하게 다룰수있게되면서 덩달아 데이터를 읽을줄알고, 다룰수있게 되어서 도움이 많이 되었습니다.

**3) 프로젝트 산출물 목록**

데이터파일 및 가이드는 2차과제제출하여 파이널과제는

R프로젝트 폴더(R파일, csv파일, text파일), 프로젝트 보고서, 프로젝트 포스터 3가지만 제출합니다.