

What make us fatty?

Team#12

김성연 심성호 황세현



한국인, 얼마나 건강할까...남성 10명 중 4명은 비만, 흡연 늘고 여성 폭음도 증가

홍진수 기자 soo43@kyunghyang.com

33



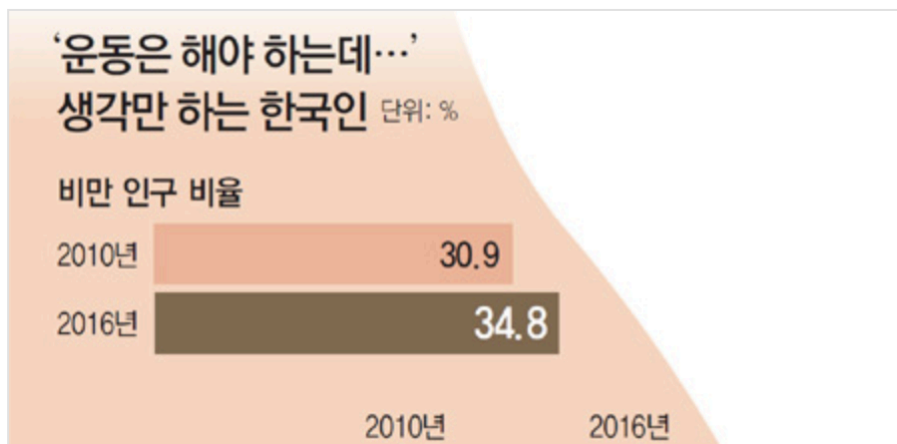
치맥 즐기고 운동 안 하고,비만·당뇨 건강지표 사상 최악

안건고 폭식-폭음, 40대男 절반 비만

조건희기자 | 입력 2017-11-07 03:00:00 | 수정 2017-11-07 03:00:00

댓글보기 0 | 폰트 + - | 뉴스듣기 여성 남성 |

질병관리본부 '국민건강영양조사'



BMI analysis on “Young People Survey”

- In **2013**, FSEV (University in **Slovakia**) surveyed young people **aged between 15-30**.
- **1010 rows** and **150 columns** (139 integer and 11 categorical)
 - Music preferences (19 items)
 - Movie preferences (12 items)
 - Hobbies & interests (32 items)
 - Phobias (10 items)
 - Health habits (3 items)
 - Personality traits, views on life, & opinions (57 items)
 - Spending habits (7 items)
 - Demographics (10 items) → **BMI**
- Reference : <https://www.kaggle.com/miroslavsabo/young-people-survey>



Analysis method

- 데이터의 타상성을 높이기 위해 **outlier**를 제거한다.
- BMI 지수와 다른 integer value를 갖는 column과의 **correlation** 분석
: BMI에 따라 집단을 나눈 뒤 (저체중, 정상, 과체중, 비만 등) EDA를 통해 집단별 특징을 추출한다.
- Machine learning (logistic regression, random forest, etc.) 을 이용하여 **BMI** 지수를 예측한다.



Expectancy effects

- **Correlation analysis**

- BMI가 높은 집단과 연관성이 높은 항목들 찾아보기
- Ex) 스스로 건강하다고 생각하는 정도
- Ex) 질 높은 음식에 대한 투자 유무

- **Statistical hypothesis testing**

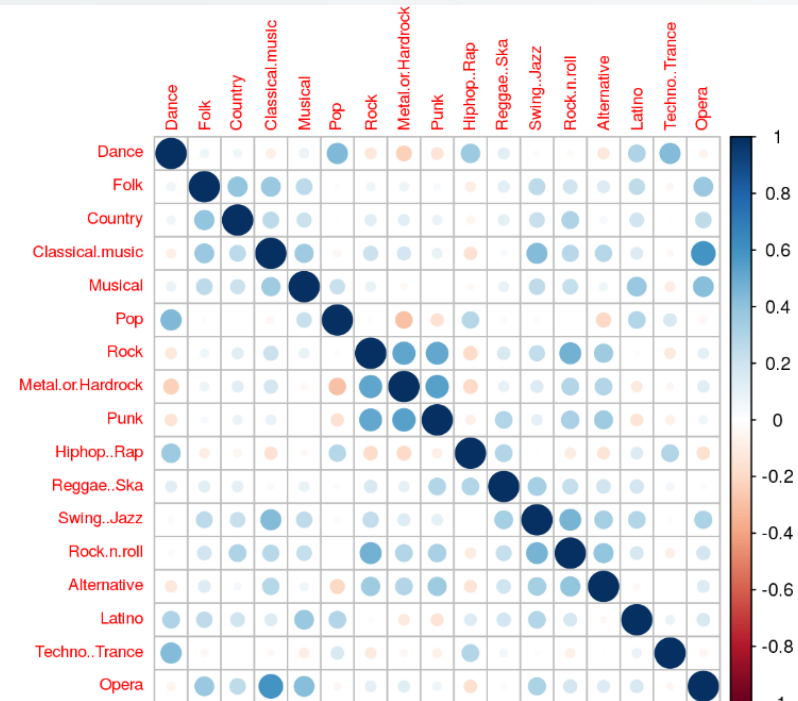
- 흡연과 비만
- 음주와 비만

- **Prediction**

- 체중 집단 예측(classification)
- BMI 지수 예측(regression)

- **Visualization**

- Correlation 시각화



Thank you

