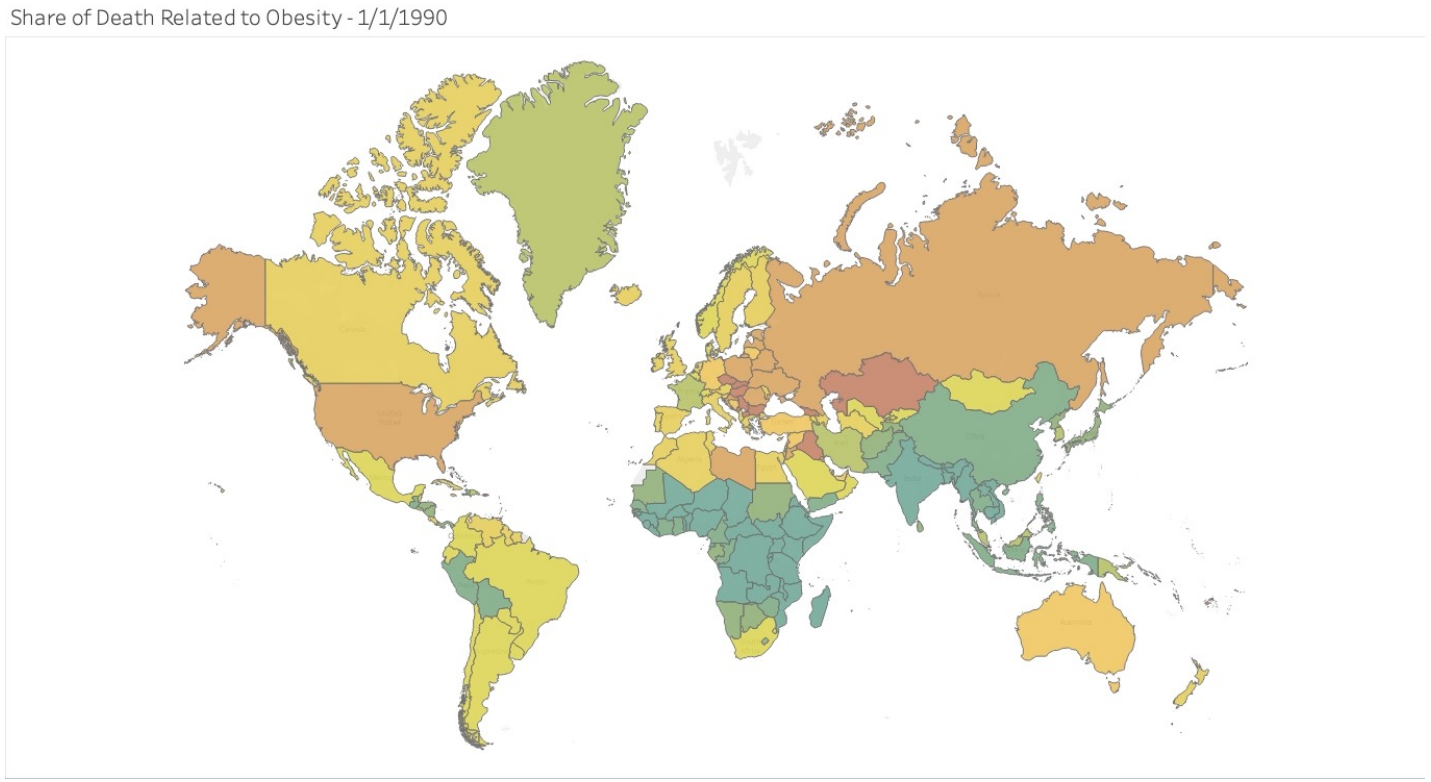
Project:
OBESITY TREND & INFLUENCING FACTORS

Li, Peifeng / Wang, Conghao

Dec 10th , 2021

OBESITY DEATH RATE IN THE WORLD



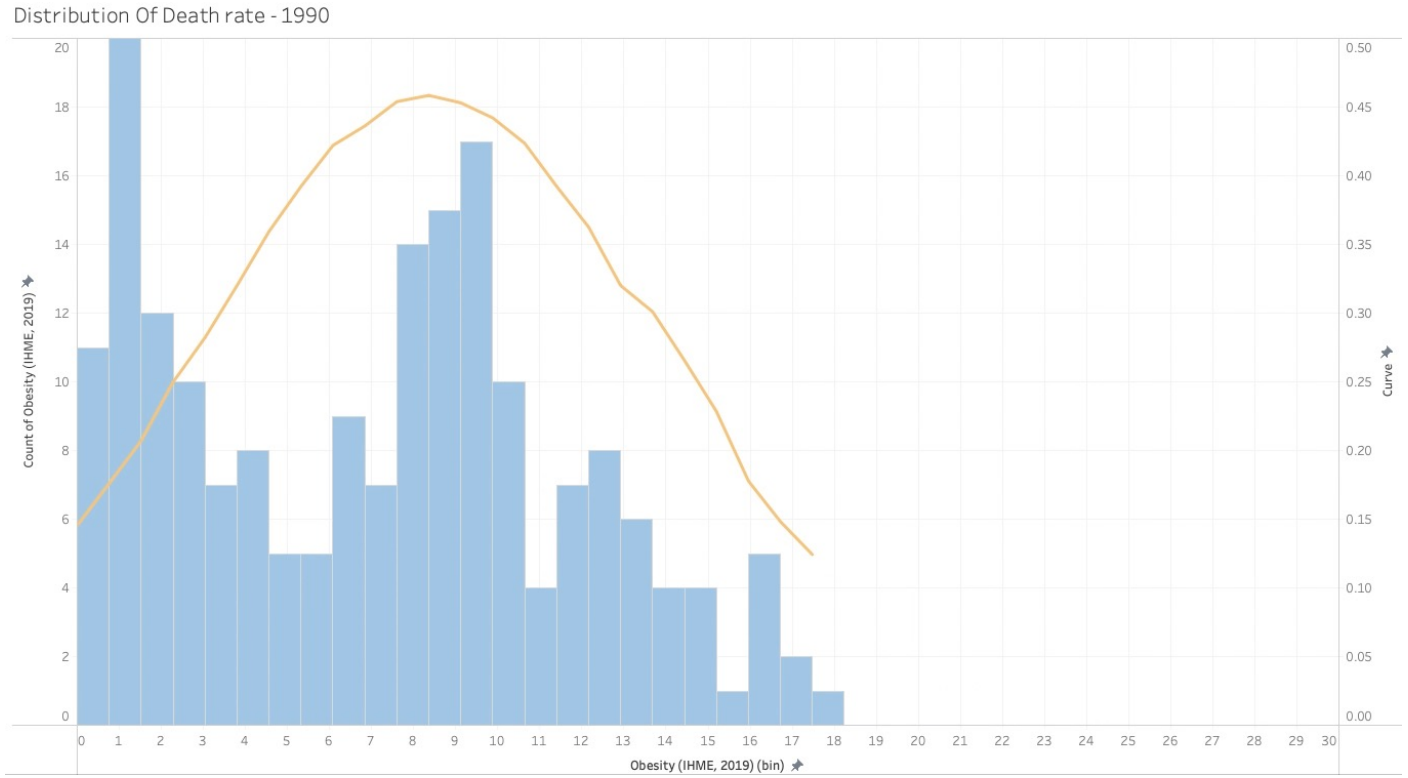
In recent 30 years, the obesity death rate keep increasing in the world.

- High: Developed & Poor[1] countries
- Medium: Developing countries
- Low: Poor countries

As in developed societies, the risk for obesity in developing countries is also strongly influenced by diet and lifestyle[2], which are changing dramatically as a result of the economic and nutrition transition[3]

References:
[1] *Obesity prevalence and trends in Latin-American countries, C.Filozof et al., 2008*
[2] *Andrew M Prentice, The emerging epidemic of obesity in developing countries, International Journal of Epidemiology, Volume 35, Issue 1, February 2006, Pages 93-99*
[3] *Caballero B. Introduction. Symposium: Obesity in developing countries: biological and ecological factors. The Journal of Nutrition. 2001 Mar;131(3):866S-870S. DOI: 10.1093/jn/131.3.866s. PMID: 11238776.*

OBESITY DEATH RATE IN THE WORLD



The distribution of obesity death rate in the world from 1990 to 2017. There are about 13% of the world's adult population were obese in 2017[1].

- Long tail in high death rate ~ 30%
- Majority around 10%
- Fewer low death rate countries

As in economy growth, the obesity death rate increase gradually[2]. The distribution of obesity death rate change from a tall and thin into a short and fat one with a long tail.

References:
[1] *https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight*
[2] *Andrew M Prentice, The emerging epidemic of obesity in developing countries, International Journal of Epidemiology, Volume 35, Issue 1, February 2006, Pages 93-99*

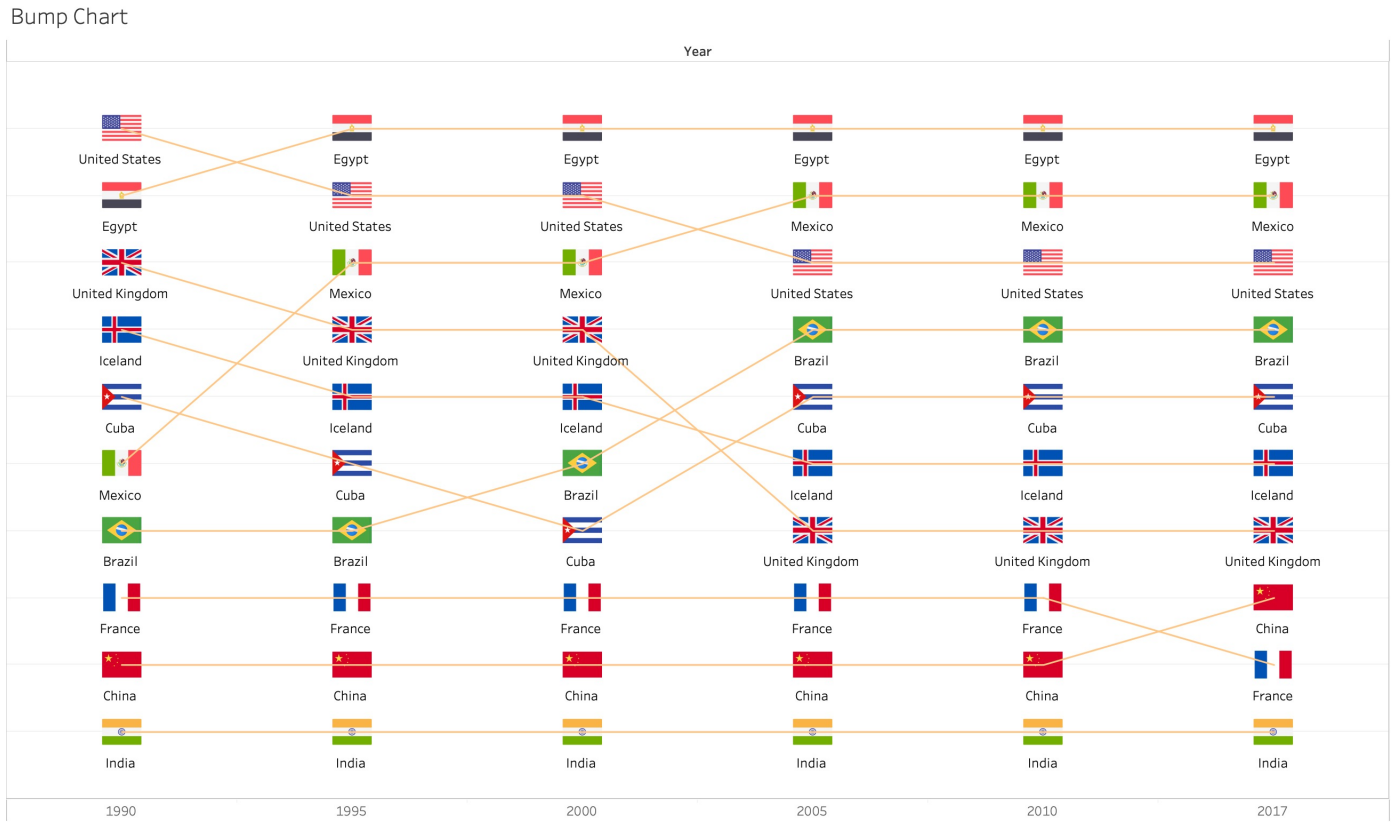
OBESITY DEATH RATE IN THE WORLD

With people's attention to scientific diet and the popularity of fitness, the degree of obesity in the United States, United Kingdom and Iceland is gradually decreasing[1], but it is still at a high level.

The obesity rate in Egypt is mainly due to eating a lot of staple diet and sugar[2], while Mexico and Brazil are based on high-fat barbecue.

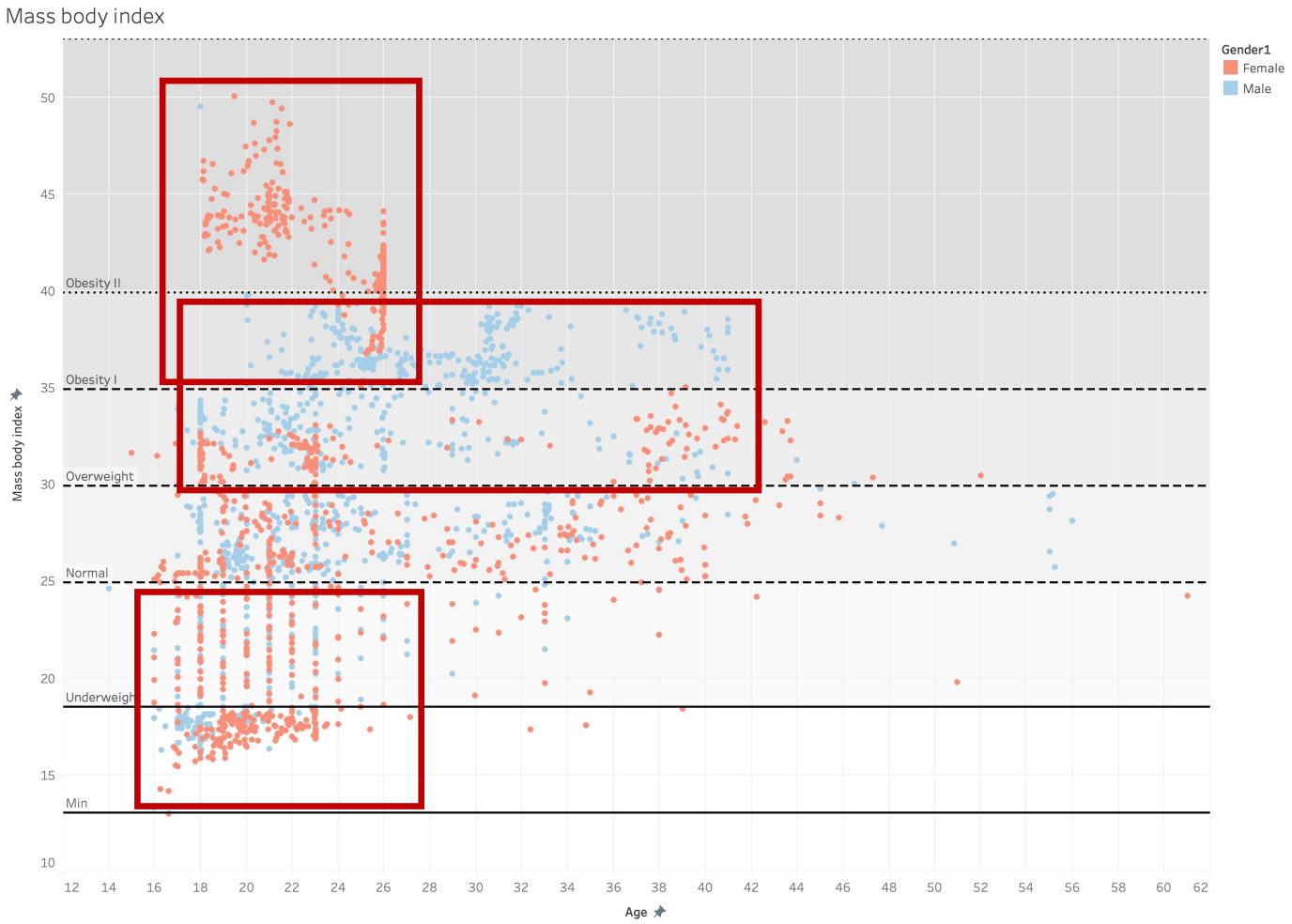
The French diet is relatively healthy, so the obesity rate has been maintained at a low level[3].

China[4] and India[5] have a large population base, their economies are in the developing stage, and a large part of the population is on the verge of food scarcity, so the overall obesity rate is low.



References:
[1] *https://www.zhihu.com/question/27564352/answer/2166490622*
[2] *https://www.zhihu.com/question/287955997*
[3] *Carballo-Casla, A., Ortolá, R., García-Esquinas, E. et al. The Southern European Atlantic Diet and all-cause mortality in older adults. BMC Med* ***19,*** *36 (2021). https://doi.org/10.1186/s12916-021-01911-y*
[4] *https://www.freedieting.com/china-study-diet*
[5] *Sharma, M., Kishore, A., Roy, D. et al. A comparison of the Indian diet with the EAT-Lancet reference diet. BMC Public Health* ***20,*** *812 (2020). https://doi.org/10.1186/s12889-020-08951-8*

INFLUENCING FACTORS OF OBESITY



The scatter plot shows the mass body index distribution of male and female in the obesity data set[1].

Key findings

- Most severe obesity basically happens in female under 25[2]
- Numbers of Male in age 18 – 40 have mild obesity higher than female
- Majority people do not have a obesity problem
- There are a promising number of people are under a health weight, especially young girl in age 18 -25

References:

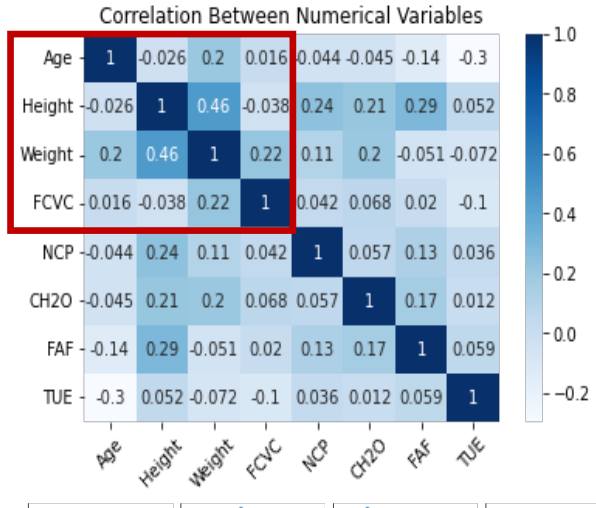
[1] *https://archive.ics.uci.edu/ml/datasets/Estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition+*

[2] *https://www.cdc.gov/obesity/data/childhood.html*

INFLUENCING FACTORS OF OBESITY

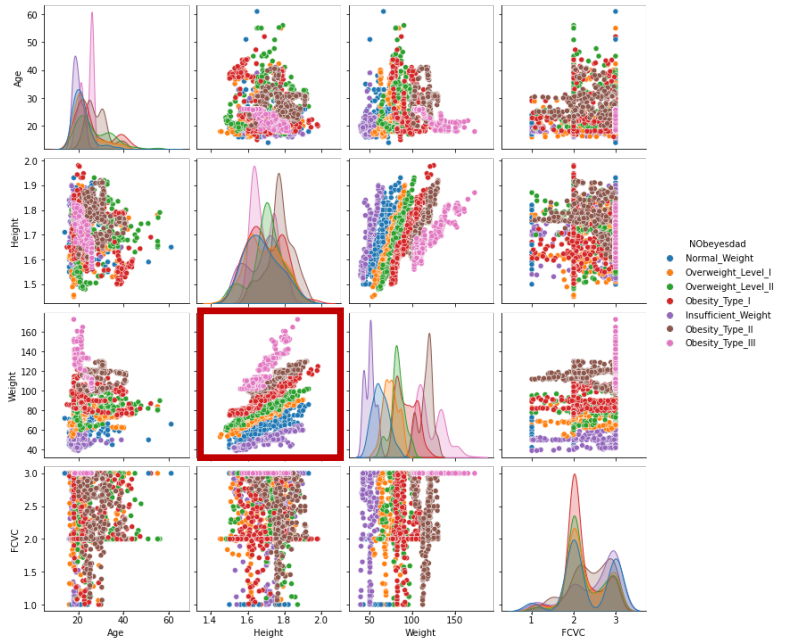
The heatmap shows the correlation between numerical variables:

- The Height and Weight have the highest correlation as 0.46
- The Weight have weak correlation with Age, FCVC (Vegetable intake), CH2O(Water intake), NCP(Number of meals)
- There are almost little correlations between different lifestyles



The pair plot shows the relationship between each two variables:

- There is a quite clear linear relationship between Height and Weight in each weight condition: a obese person tend to have larger weight than normal person if they have the same height (Slop between Height and Weight increases as the level of obesity increases)

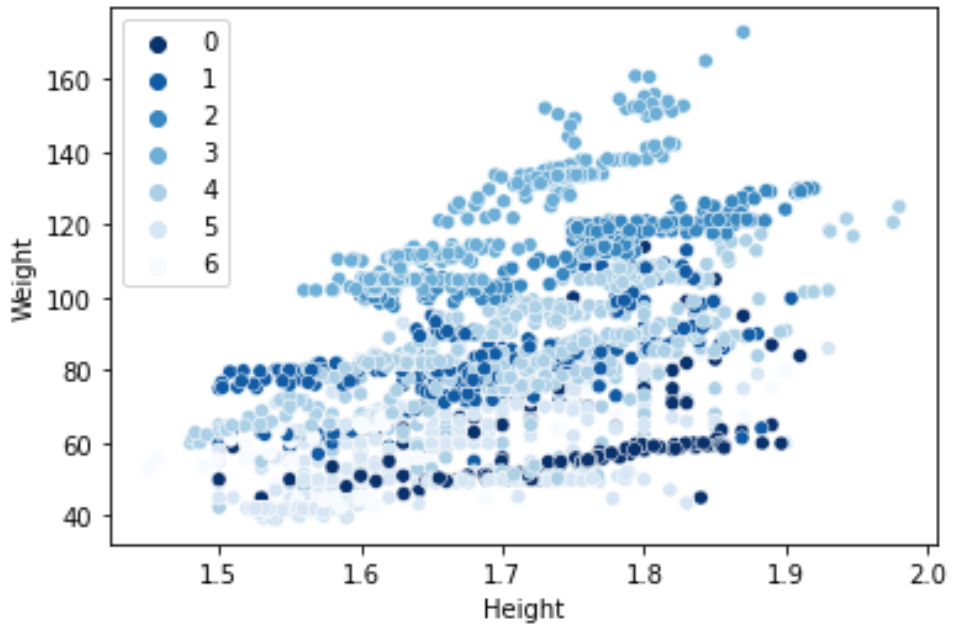


INFLUENCING FACTORS OF OBESITY

Deploy Random Forest Classifier and Decision Tree Classifier to do classification projection, using Naïve bayes as baseline model (prediction accuracy of **83.7%**)

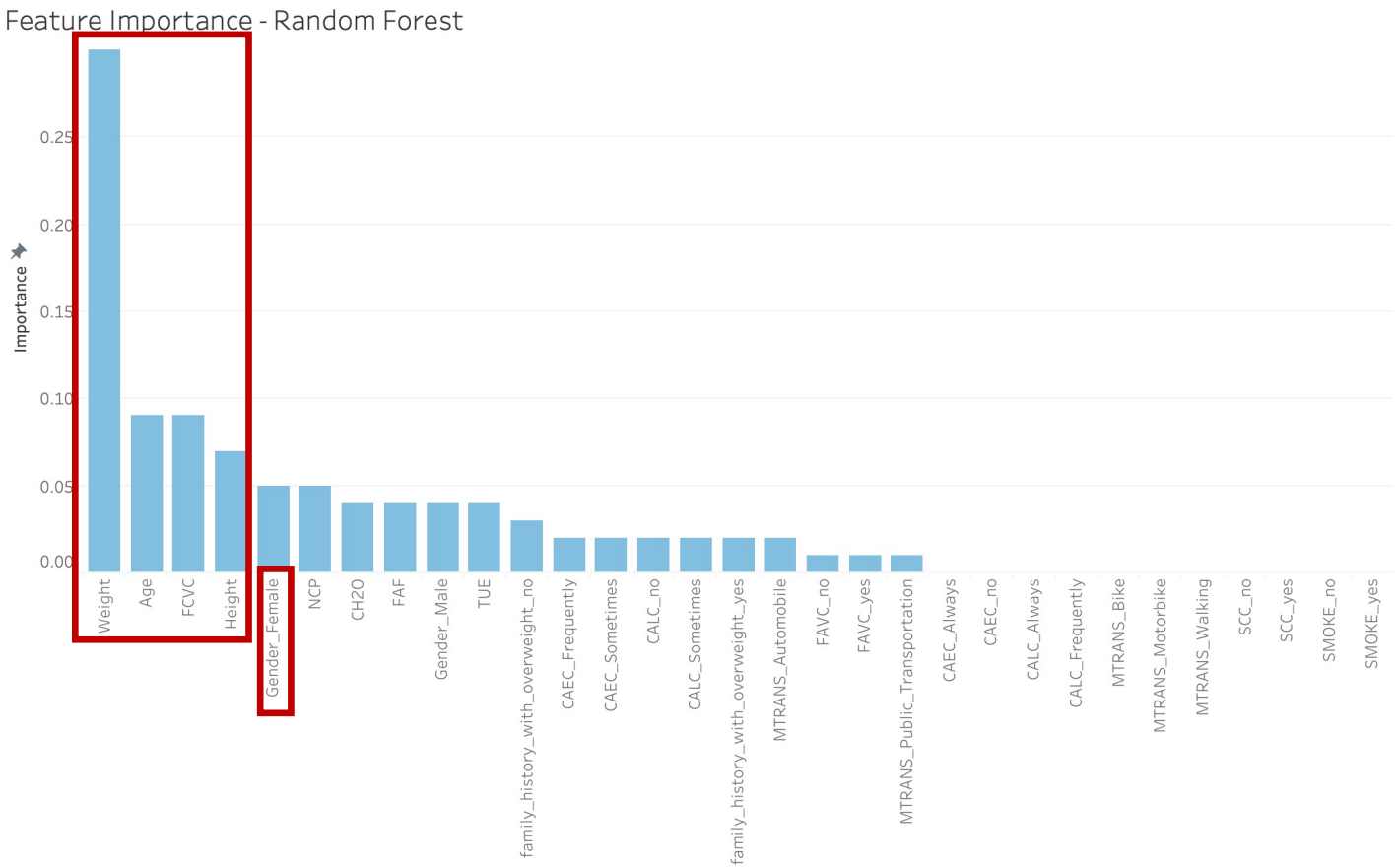
The best Random Forest Classifier receive a quite well accuracy of 93.75%.

Apply KMeans to do clustering, but the result is not so good as we wish.




```
Random Forest:
----------------------------------------------------------------
Accuracy: 0.9375
Accuracy w/Scaled Data (ss): 0.9375

Classification Report (ss):
                     precision    recall  f1-score   support

Insufficient_Weight       1.00      0.92      0.96        78
      Normal_Weight       0.73      0.99      0.84        69
     Obesity_Type_I       0.99      0.94      0.96        86
    Obesity_Type_II       1.00      1.00      1.00        73
   Obesity_Type_III       1.00      1.00      1.00        88
 Overweight_Level_I       0.93      0.81      0.87        68
Overweight_Level_II       0.95      0.88      0.91        66

           accuracy                           0.94       528
          macro avg       0.94      0.93      0.93       528
       weighted avg       0.95      0.94      0.94       528

Decision Tree:
----------------------------------------------------------------
Accuracy: 0.92235
Accuracy w/Scaled Data (ss): 0.92803

Classification Report (ss):
                     precision    recall  f1-score   support

Insufficient_Weight       0.97      0.94      0.95        78
      Normal_Weight       0.84      0.84      0.84        69
     Obesity_Type_I       0.95      0.97      0.96        86
    Obesity_Type_II       1.00      0.99      0.99        73
   Obesity_Type_III       0.98      1.00      0.99        88
 Overweight_Level_I       0.85      0.82      0.84        68
Overweight_Level_II       0.87      0.91      0.89        66

           accuracy                           0.93       528
          macro avg       0.92      0.92      0.92       528
       weighted avg       0.93      0.93      0.93       528
```

INFLUENCING FACTORS OF OBESITY



Extract the feature importance from Random Forest classifier:

Key Findings:

- Weight is the most important issue lead to obesity – Of course!
- Age and FCVC is another two important features close to overweight[1]
- Female tend to be overweight than Male[2]
- FAF(Physical Activity Frequency), CH2O(Water in take), TUE(Time using electronics) also have a impact which can not be ignored on body weight[3]

References:

[1] *K He, et al. "Changes in intake of fruits and vegetables in relation to risk of obesity and weight gain among middle-aged women." International Journal of Obesity (2004) 28:1569-74. Accessed November 30, 2014, doi:10.1038/sj.ijo.0802795*

[2] *https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight*

[3] *World Health Organization. Notes for the media: New physical activity guidance can help reduce risk of breast, colon cancers; 2011. Accessed January 28, 2012.*