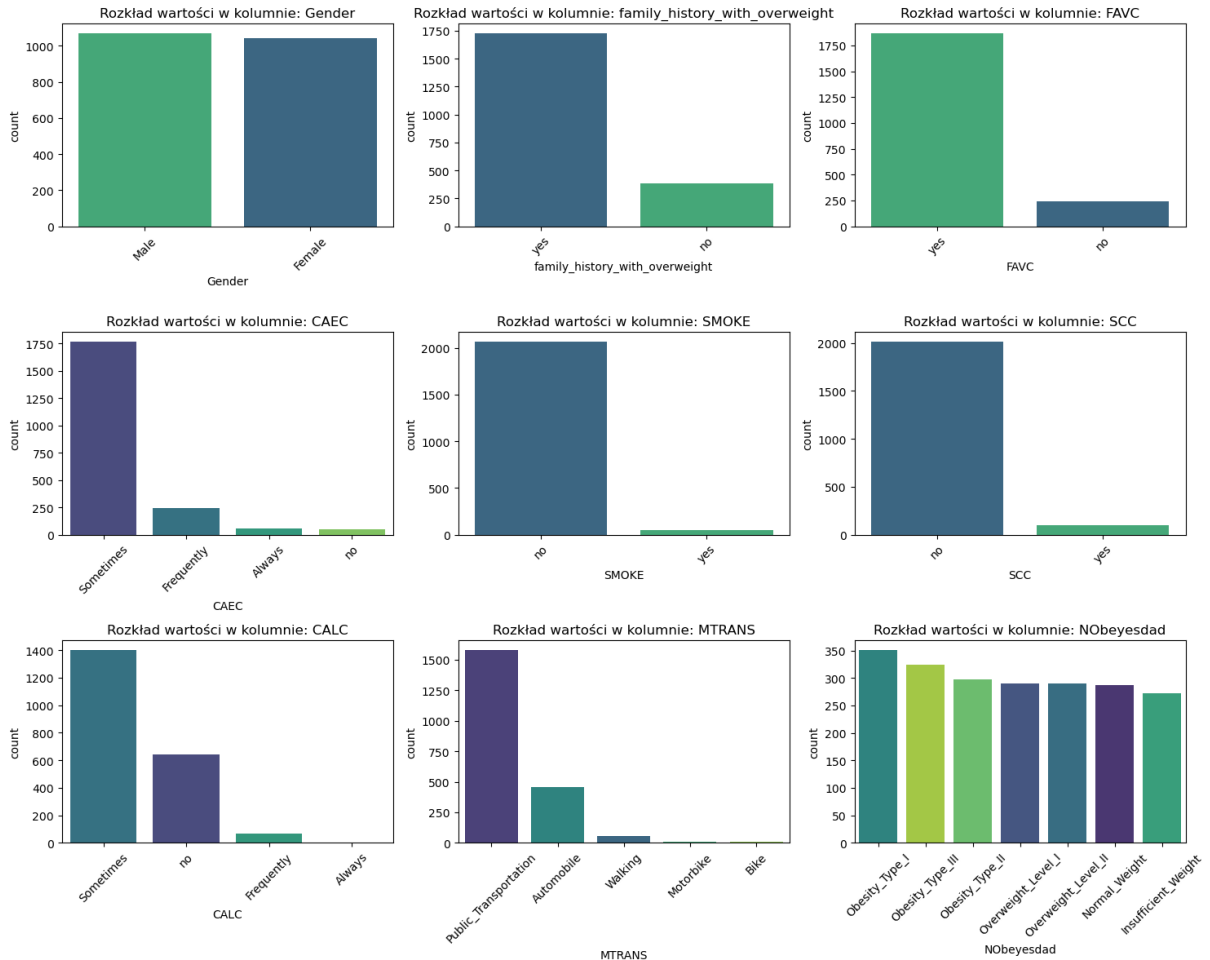


## 1. Wprowadzenie

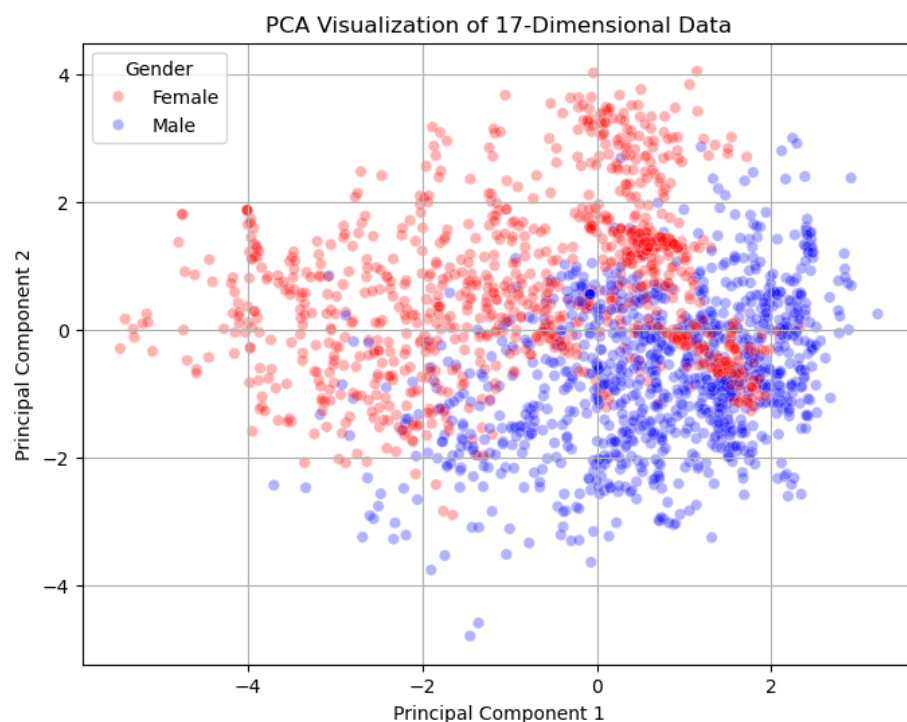
Celem analizy jest zrozumienie zależności pomiędzy cechami zawartymi w zbiorze danych oraz identyfikacja zmiennych o najwyższej wartości informacyjnej. Analiza uwzględnia wizualizacje i statystyki opisowe.

## 2. Obserwacje

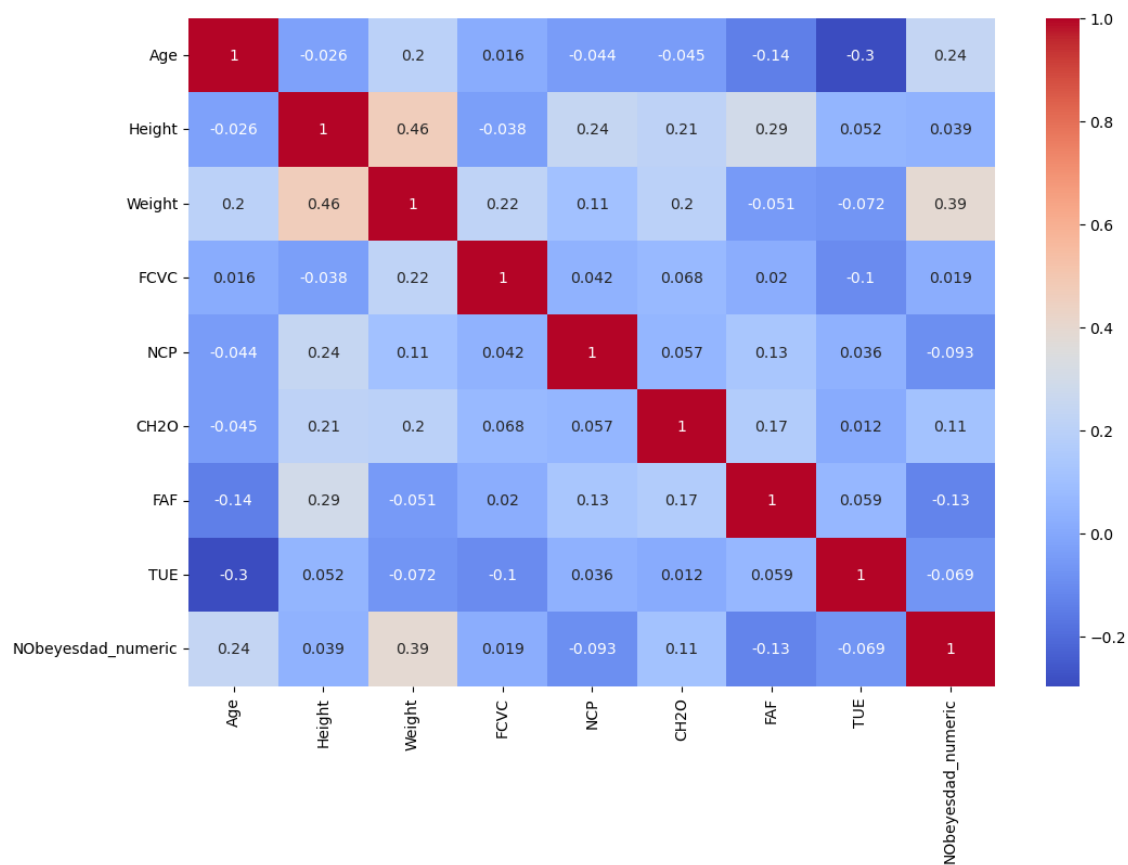
- Na cechy numeryczne skrajne wartości bardzo nie wpływają, ponieważ średnia i mediana wszędzie mają podobne wartości
- Cechy płeć i stopień otyłości są zbalansowane w przeciwieństwie do reszty cech.



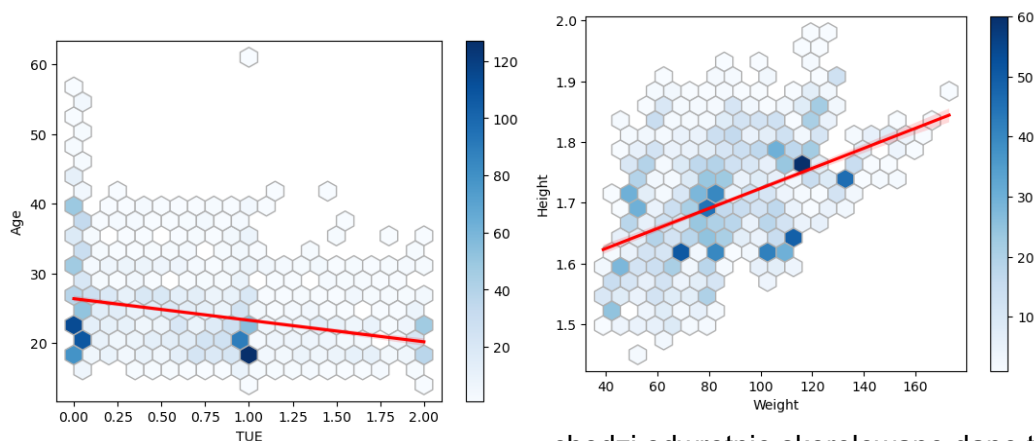
- W cechach numerycznych które się odnoszą do częstości wykonywania czynności jako wartości przeważają liczby całkowite



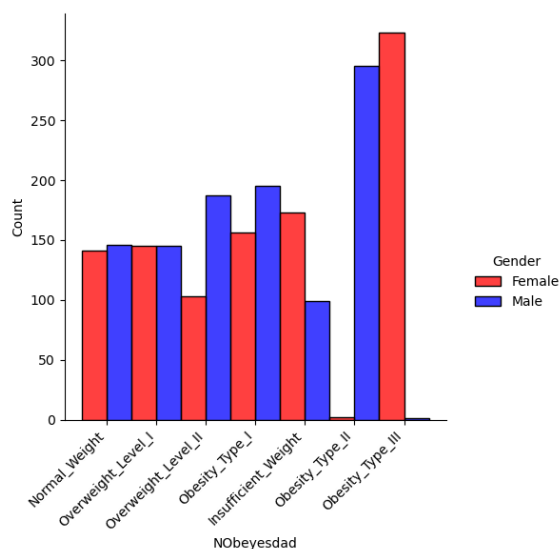
Dwie płcie wykazują pewien stopień separacji w dwuwymiarowej przestrzeni PCA. W środkowej części występuje znaczące nakładanie się danych. Punkty danych są szeroko rozproszone wzdłuż obu głównych składowych, co wskazuje, że te dwie składowe przechwytują znaczną wariancję oryginalnych 17-wymiarowych danych.



W zbiorze danych można zauważyć silnie skorelowane dane takie jak wzrost i waga. Ponadto poziom otyłości jest skorelowany z wagą i ilością spożywanej wody i negatywnie skorelowany z częstotliwością uprawianego sportu.



A jeśli chodzi odwrotnie skorelowane dane to najlepiej wypada częstotliwość używania urządzeń elektronicznych z wiekiem.



Wiele kobiet ma otyłość typu 3, a prawie żadnych mężczyzn. Dokładnie odwrotna sytuacja ma miejsce w przypadku typu 2.

### 3. Wnioski

Na podstawie wykresu o typie otyłości rozdzielającego płcie można wywnioskować, że dane mogą nie być prawidłowe co sugeruje ilość kobiet w obesity type II i mężczyzn obesity type III.

Ciekawą kwestią jest to, że osoby otyłe piją dużo wody co się kłóci z powszechnymi przekonaniem. W zbiorze danych występują korelacje takie jak wzrost z wagą, poziom otyłości z wagą itd.

