

1. Gender column has two categories M and F so encoding it
2. The age is in the format of xx-xx, so the outcomes can be
   1. avg
   2. Segment into categories (teen, working, adults etc) and then encode it
3. Ocupation column is from 0-20 making it an array of 21
   1. Group them and encode
   2. Depending on the