

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
4. Drop userid and prodcut id
5. Gender ko ohe male(1) female (0)
6. Age class ko make separate columns and drop first
7. Categies ko ohe and drop first
8. Stay use regex and only the first digit