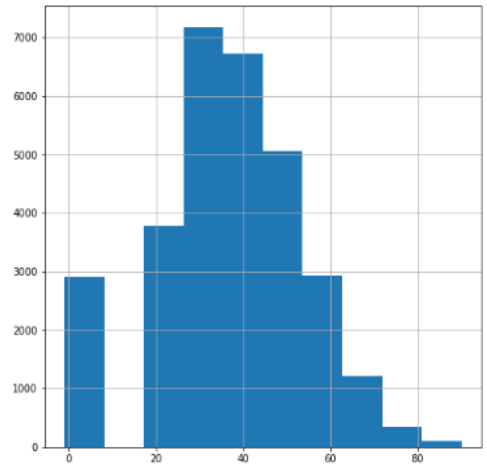
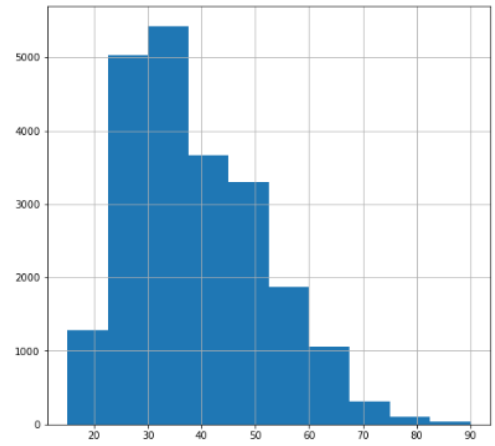
AGE

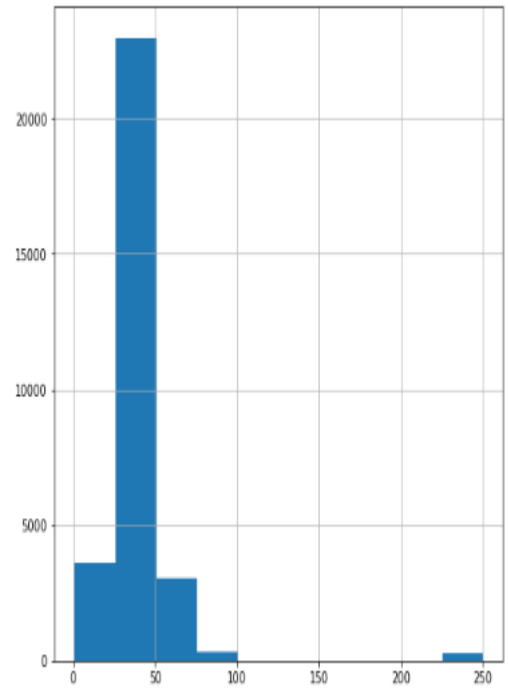
```
data['age'].hist(figsize=(8,8))
plt.show()
```

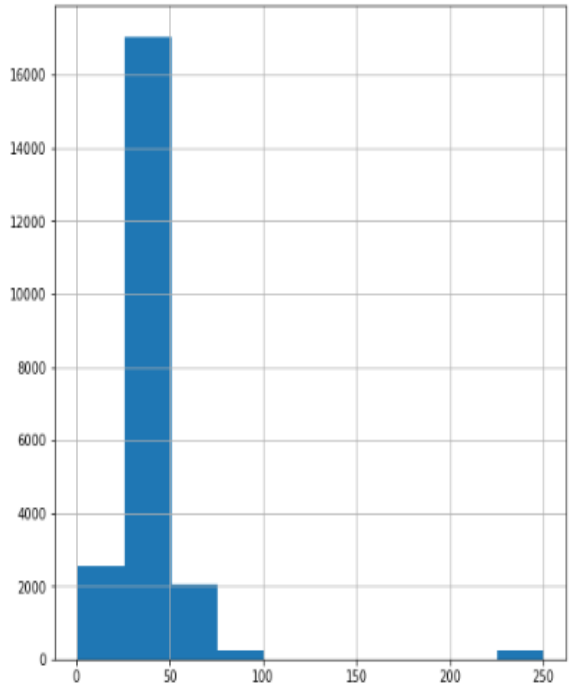




Hour per week

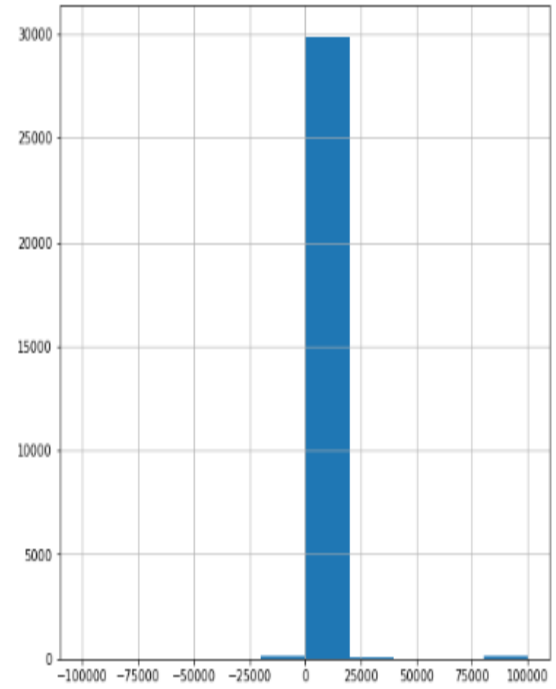
```
data['hours_per_week'].hist(figsize=(8,8))
plt.show()
```

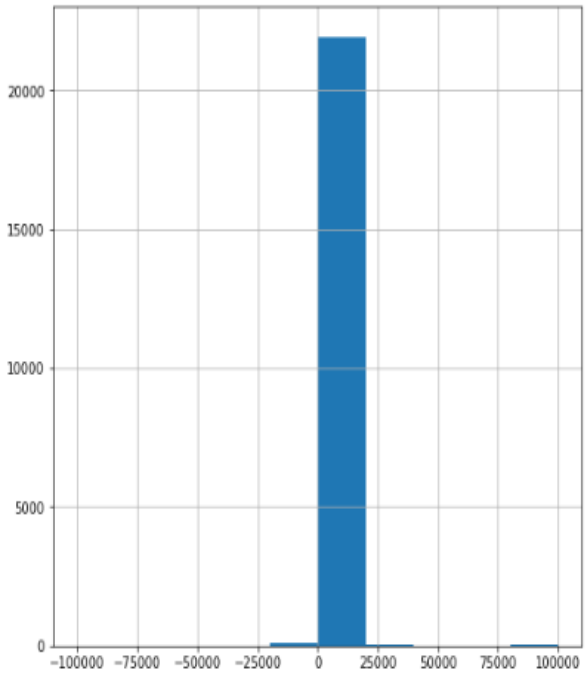




Capitan gain

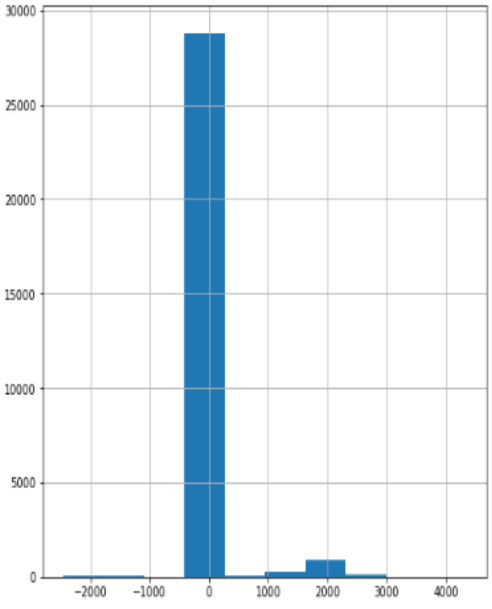
```
1 data["capital_gain"].hist(figsize=(8,8))
2 plt.show()
```

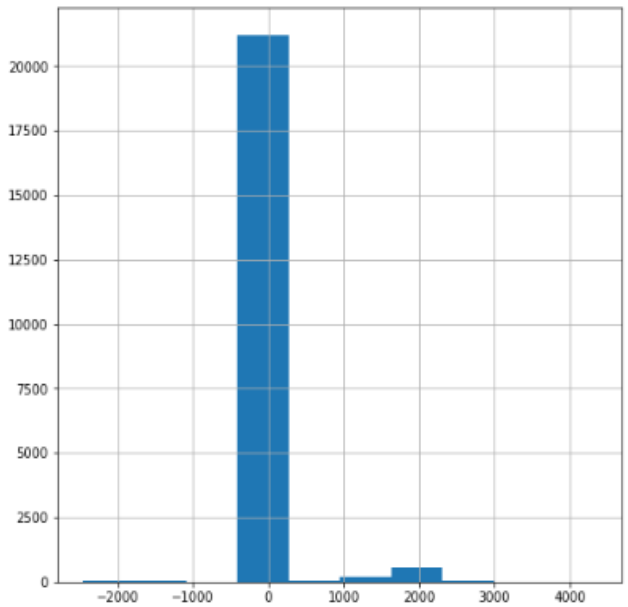




Capital loss

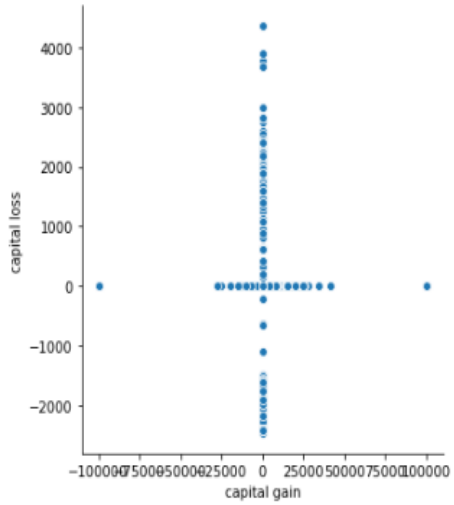
```
1 data["capital_loss"].hist(figsize=(8,8))
2 plt.show()
```

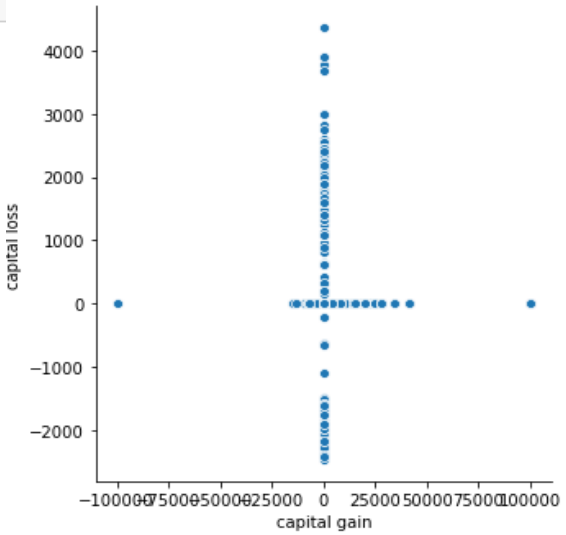




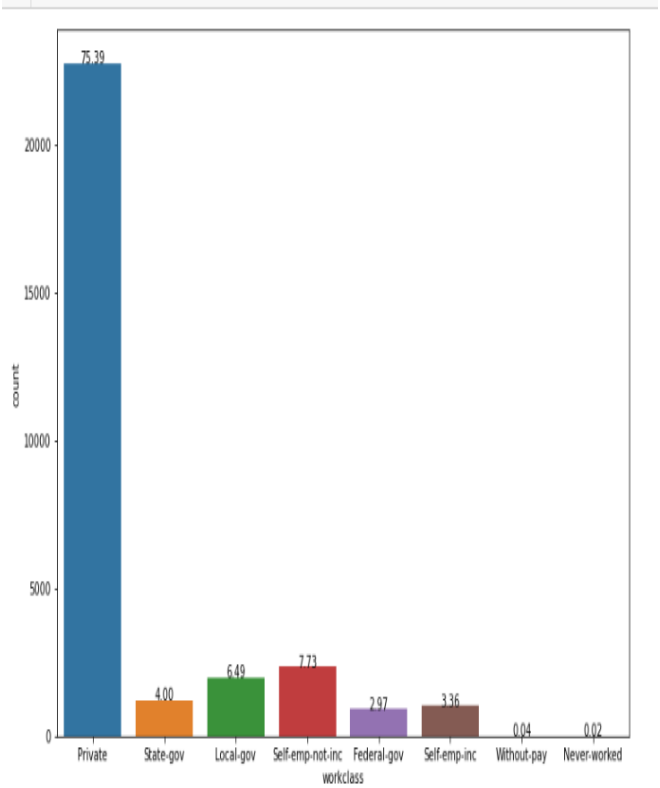
Capital loss vs Capital gain

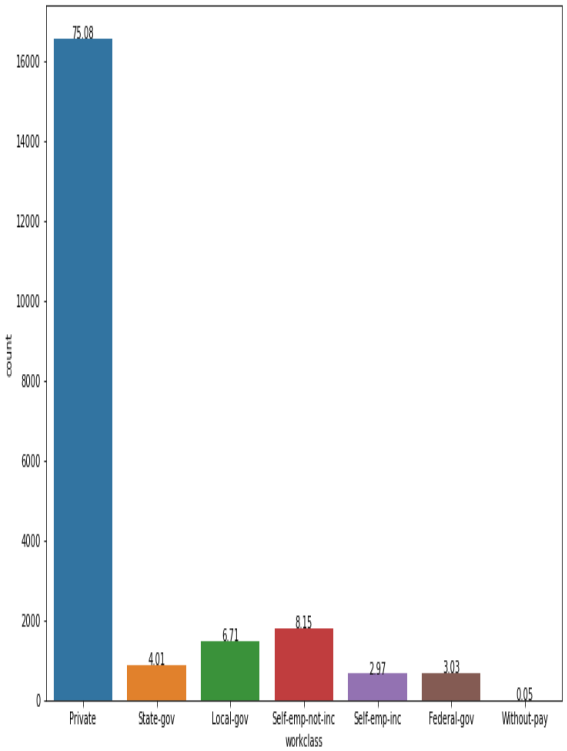
```
:  1 sns.relplot('capital_gain','capital_loss', data= data)
   2 plt.xlabel("capital gain")
   3 plt.ylabel("capital loss")
   4 plt.show()
```



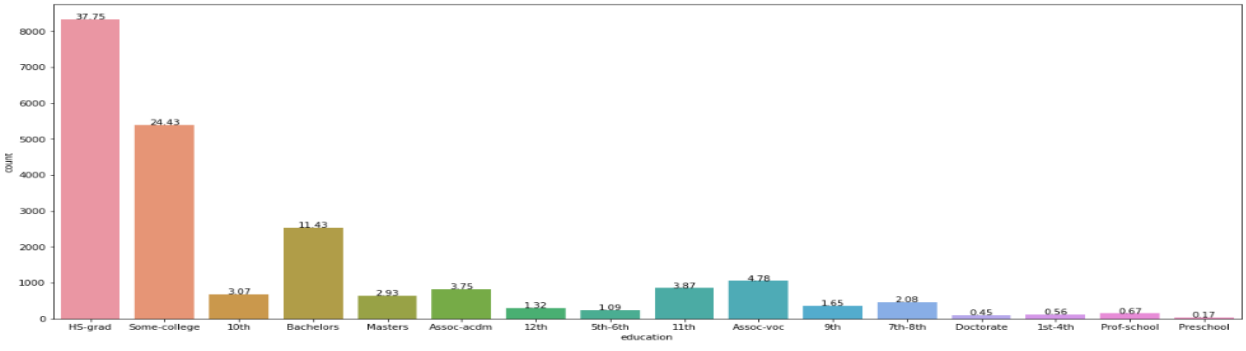


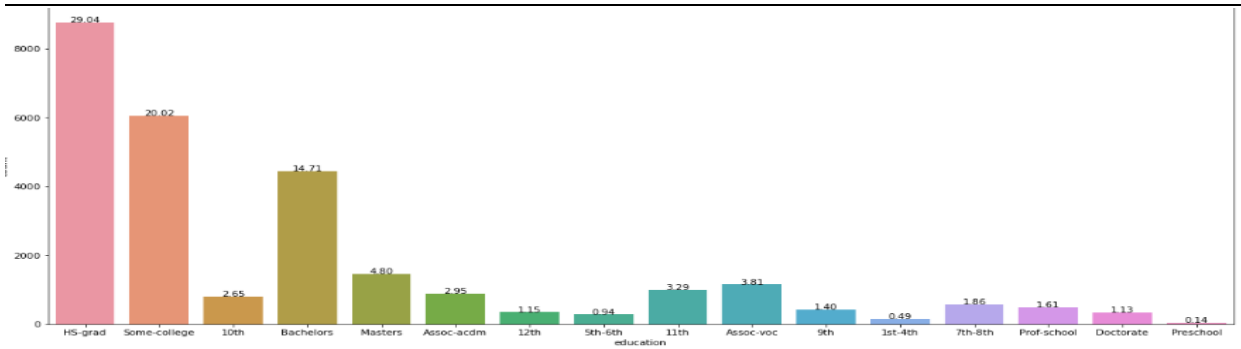
Workclass



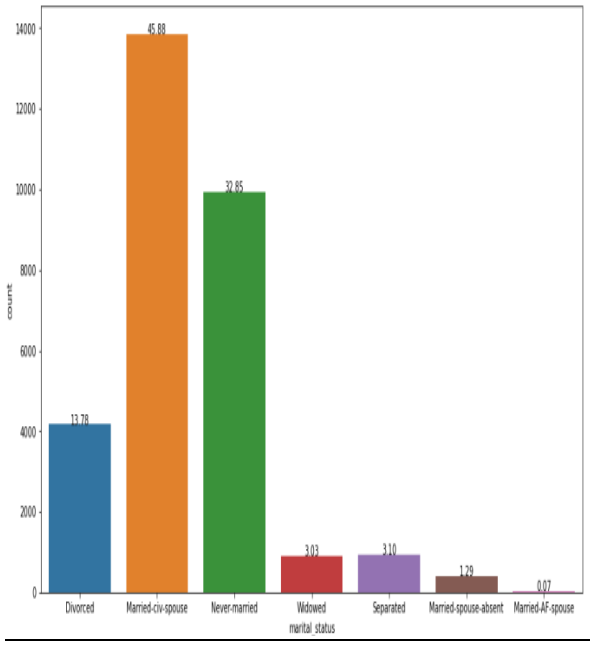


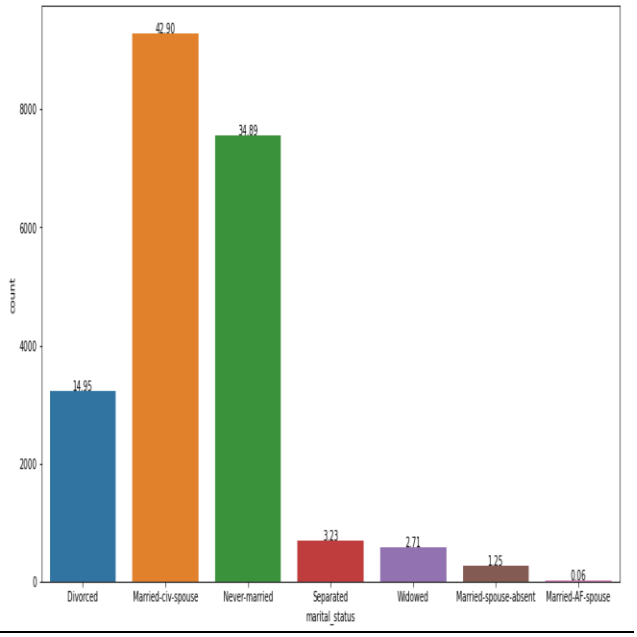
Education



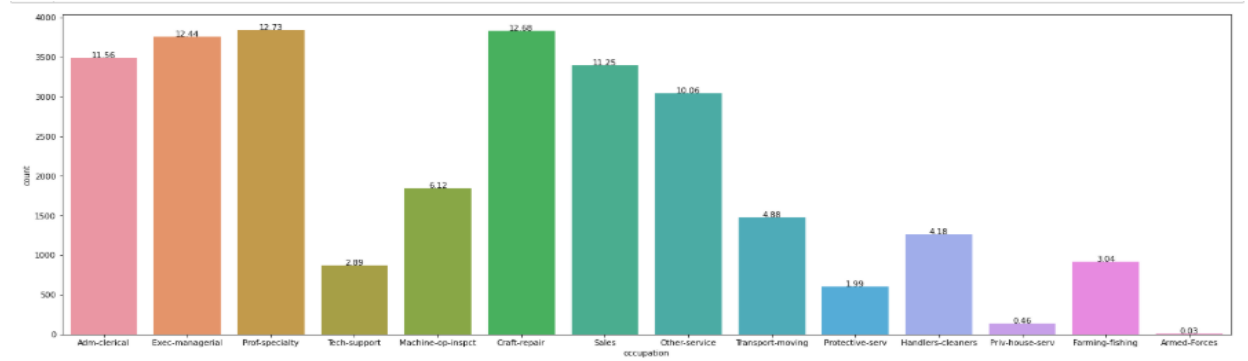


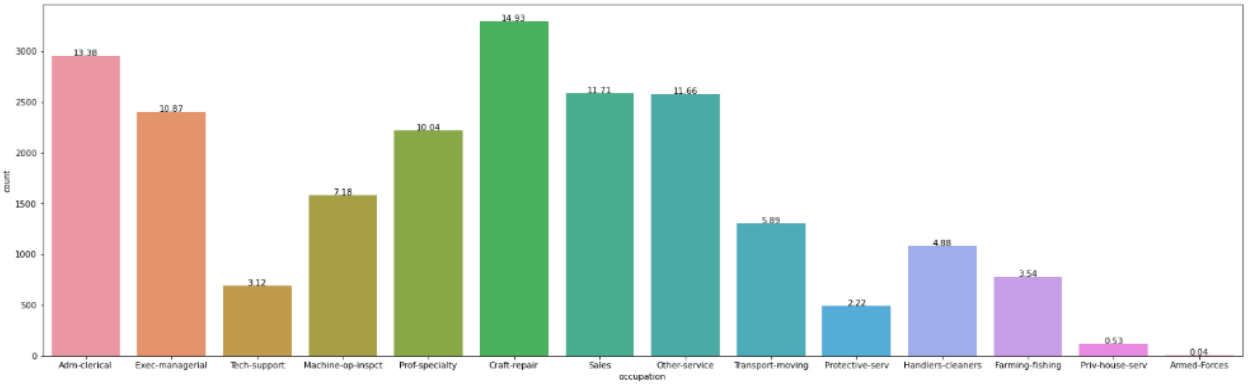
Maritial status



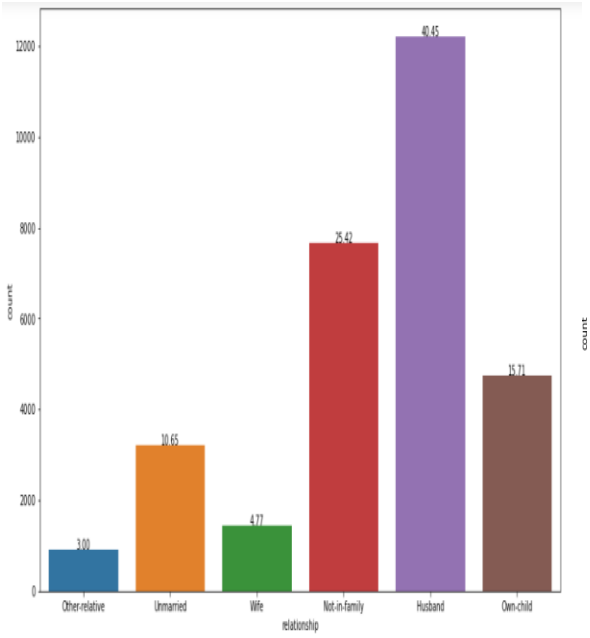


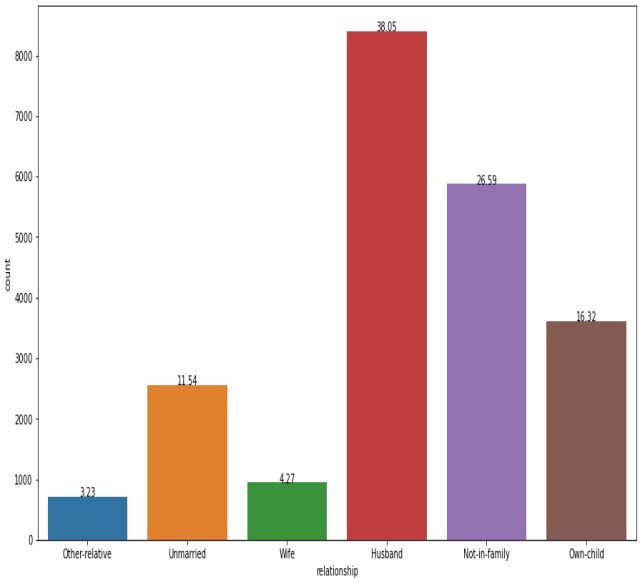
Occupation-



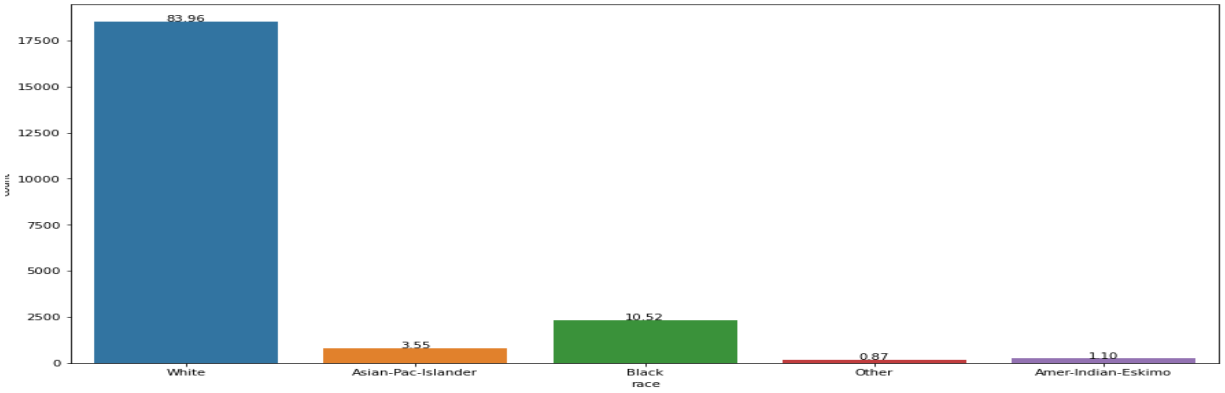


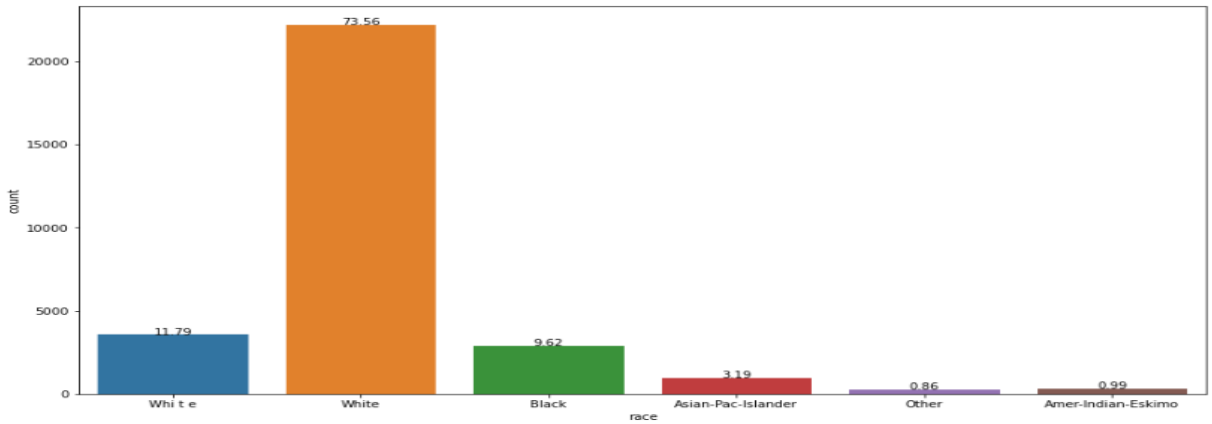
Relationship-



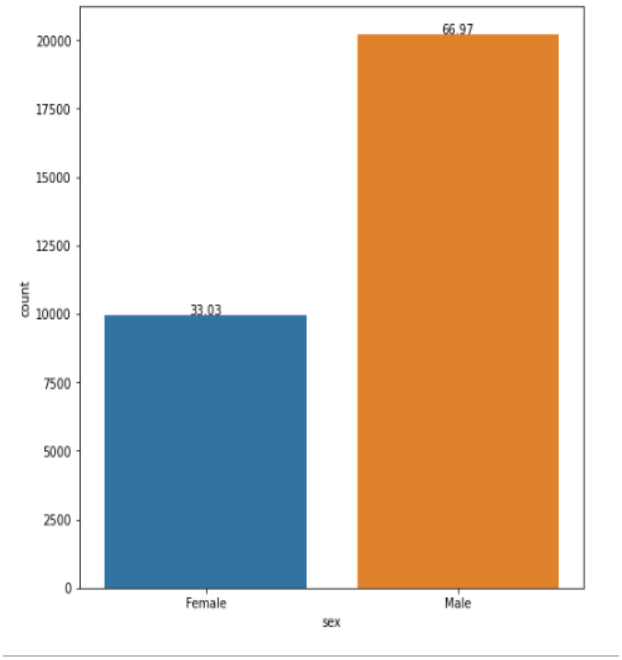


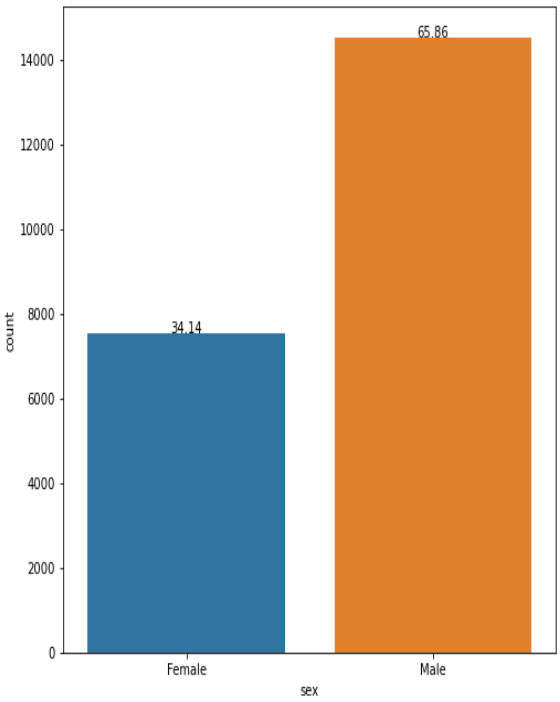
Race-


83.96
3.55
10.52
0.87
1.10
White
Asian-Pac-Islander
Black
Other
Amer-Indian-Eskimo
race

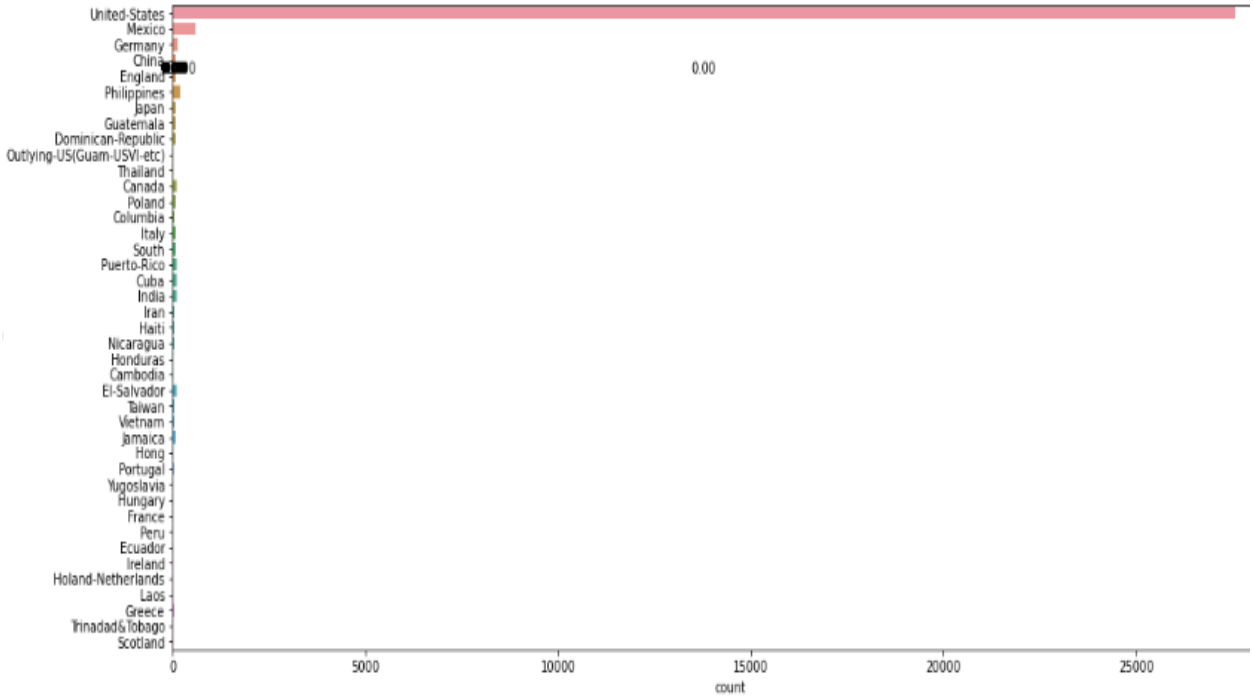

73.56
11.79
9.62
3.19
0.86
0.99
Whi t e
White
Black
Asian-Pac-Islander
Other
Amer-Indian-Eskimo
count
race

Sex-

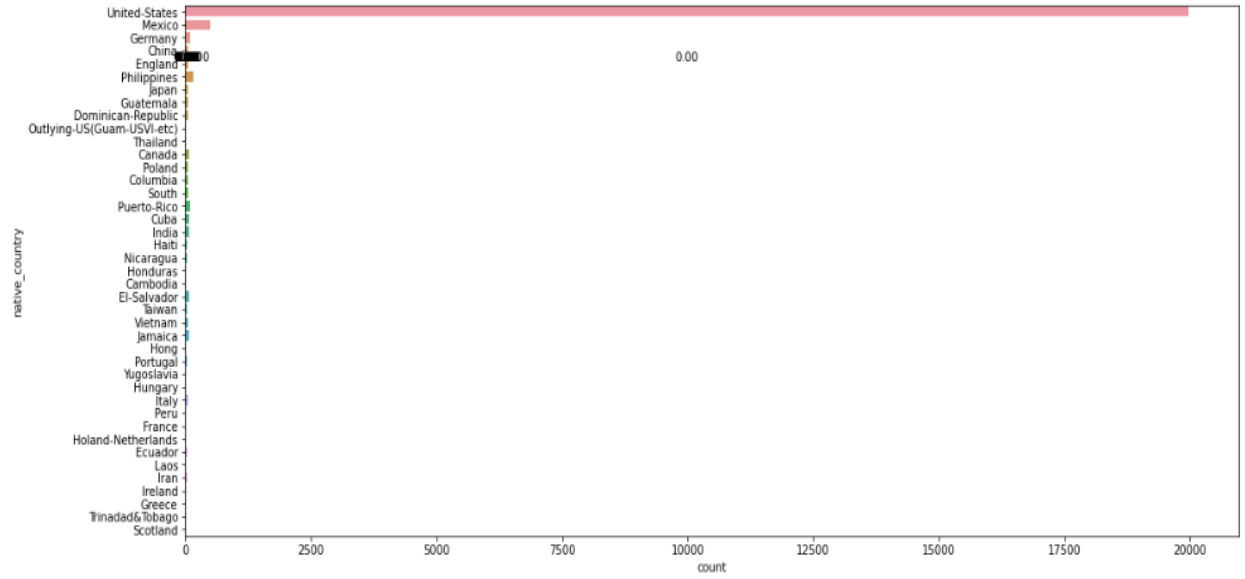

66.97
33.03
Female
Male
sex
count


65.86
34.14
Female
Male
sex
count

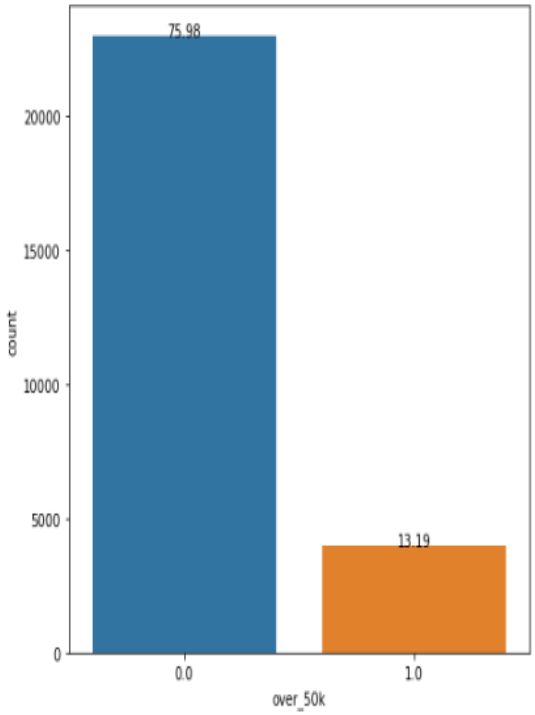
Native country-

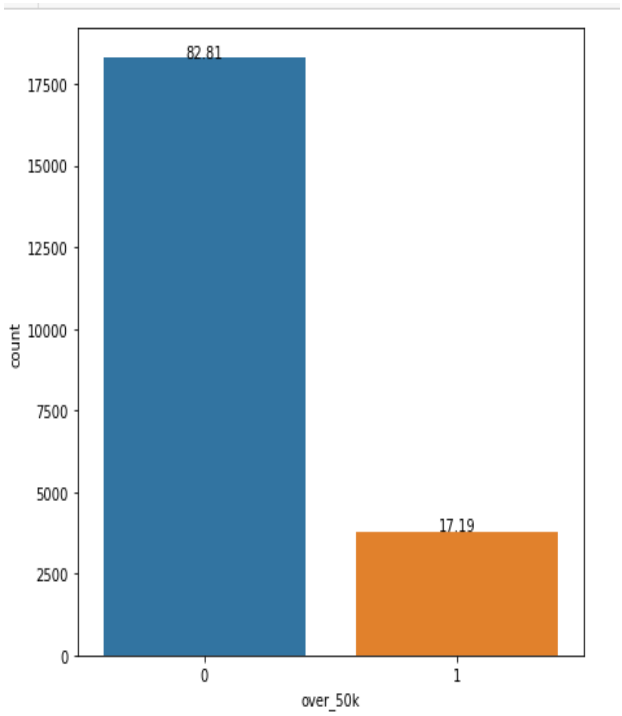


Over50k-

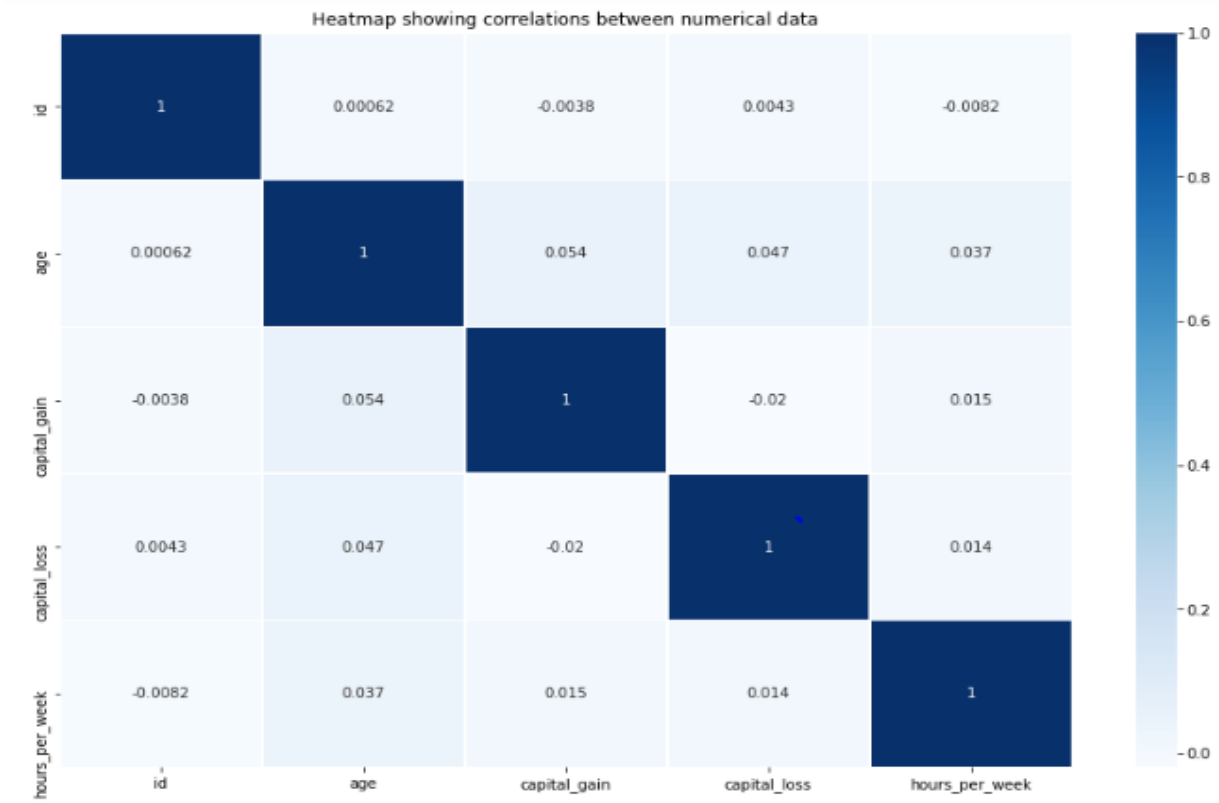


Over_50


75.98
13.19
20000
15000
10000
5000
0
count
0.0
1.0
over_50k


82.81
17.19
17500
15000
12500
10000
7500
5000
2500
0
count
0
1
over_50k


Heatmap showing correlations between numerical data
id
age
capital_gain
capital_loss
hours_per_week
1
-0.00039
0.001
0.00077
-0.0056
-0.00039
1
0.068
0.051
0.085
0.001
0.068
1
-0.025
0.031
0.00077
0.051
-0.025
1
0.025
-0.0056
0.085
0.031
0.025
1
1.0
0.8
0.6
0.4
0.2
0.0


Heatmap showing correlations between numerical data
id
age
capital_gain
capital_loss
hours_per_week
1
0.00062
-0.0038
0.0043
-0.0082
0.00062
1
0.054
0.047
0.037
-0.0038
0.054
1
-0.02
0.015
0.0043
0.047
-0.02
1
0.014
-0.0082
0.037
0.015
0.014
1
1.0
0.8
0.6
0.4
0.2
0.0