analysis_3_3

Kevin Ros

3/2/2020

```
orig_data = data.frame(read.csv("ACSdata_org.csv",header=TRUE))
syn_data = data.frame(read.csv("ACSdata_syn.csv",header=TRUE))
```

```
m = nrow(syn_data)
n = nrow(orig_data)

expected_match_risk = 0
true_match_rate = 0

for(s in 1:m){
  c_i = 0
  records = orig_data[orig_data$SEX == syn_data$SEX[s] & orig_data$RACE == syn_data$RACE[s] & orig_data$

  if(nrow(records) == 1){
    true_match_rate = true_match_rate + 1/m
  }

  if (nrow(records) != 0) {
    expected_match_risk = expected_match_risk + (1/nrow(records))
  }

}
print(expected_match_risk)
```

```
## [1] 57
```

```
print(true_match_rate)
```

```
## [1] 4e-04
```

I think the false match rate is 0 because neither SEX, RACE, nor MAR are synthesized.