

Xiao Zhou

Mohammad Soleymani

CS535

17 April 2024

### HW3

#### 1. The Means group by race 1,2,3

	Confident	Egotistic	Intelligent	Kind	Responsible	Trustworthy	Aggressive	Caring	Emotional	Friendly	Sociable
1	5.925109	4.057195	5.863676	5.715496	5.905479	5.645065	3.636961	5.580169	4.869951	5.867516	5.851906
2	5.967909	4.322545	5.528636	5.404500	5.514227	5.346682	4.128409	5.305455	4.732318	5.498636	5.647455
3	5.836190	3.604444	6.317937	5.857302	6.146032	5.866508	3.217619	5.751429	4.845238	6.043333	5.830635

#### The Medians group by race 1, 2, 3

	Confident	Egotistic	Intelligent	Kind	Responsible	Trustworthy	Aggressive	Caring	Emotional	Friendly	Sociable
1	6.03	4.00	6.00	5.930	6.07	5.80	3.40	5.80	4.87	6.19	6.07
2	6.13	4.20	5.70	5.835	5.80	5.67	3.73	5.73	4.73	5.93	6.13
3	5.87	3.33	6.47	6.190	6.27	6.00	2.93	6.00	4.93	6.33	6.07

### *The results of One-way-ANOVA across race 1,2,3*

```
[('rIntelligent',
  F_onewayResult(statistic=34.578276662469264, pvalue=1.6711926732092035e-15)),
 ('rAggressive',
  F_onewayResult(statistic=24.614244796632537, pvalue=2.7078370672104353e-11)),
 ('rResponsible',
  F_onewayResult(statistic=23.50567211273655, pvalue=8.005057161350754e-11)),
 ('rEgotistic',
  F_onewayResult(statistic=15.870459650352897, pvalue=1.4412211501758323e-07)),
 ('rTrustworthy',
  F_onewayResult(statistic=12.588919683611826, pvalue=3.672594877204227e-06)),
 ('rFriendly',
  F_onewayResult(statistic=9.95454713118648, pvalue=4.9774464052824096e-05)),
 ('rKind',
  F_onewayResult(statistic=8.995422410126517, pvalue=0.00012878104123732718)),
 ('rCaring',
  F_onewayResult(statistic=7.806449234445434, pvalue=0.00041893692225684995)),
 ('rEmotional',
  F_onewayResult(statistic=3.7705266249058407, pvalue=0.02319487807768113)),
 ('rSociable',
  F_onewayResult(statistic=3.3722702134696423, pvalue=0.03449616164486741)),
 ('rConfident',
  F_onewayResult(statistic=0.6685383548264834, pvalue=0.5125652685487156))]
```

If the p-value is less than the significance level (0.05), it indicates that there are statistically significant differences among the groups. Among all the features, Intelligent, Aggressive, Responsible, Egotistic, Trustworthy, Friendly, Kind, Caring, Emotional, and Sociable have a  $p < 0.05$ , which indicates significant bias exists in these features between races.

### *2. The Means group by gender*

	Confident	Egotistic	Intelligent	Kind	Responsible	Trustworthy	Aggressive	Caring	Emotional	Friendly	Sociable
0	5.994449	3.601511	5.947671	6.172770	6.118059	6.008562	3.127545	6.081123	5.289098	6.314092	6.257156
1	5.865760	4.432829	5.752246	5.303822	5.667762	5.309094	4.110788	5.139165	4.528345	5.450024	5.488266

### *The Medians group by gender*

	Confident	Egotistic	Intelligent	Kind	Responsible	Trustworthy	Aggressive	Caring	Emotional	Friendly	Sociable
0	6.13	3.5	6.00	6.33	6.20	6.13	3.00	6.27	5.33	6.53	6.40
1	5.93	4.4	5.87	5.53	5.87	5.47	3.88	5.33	4.53	5.73	5.67

### *The results of t-test across gender*

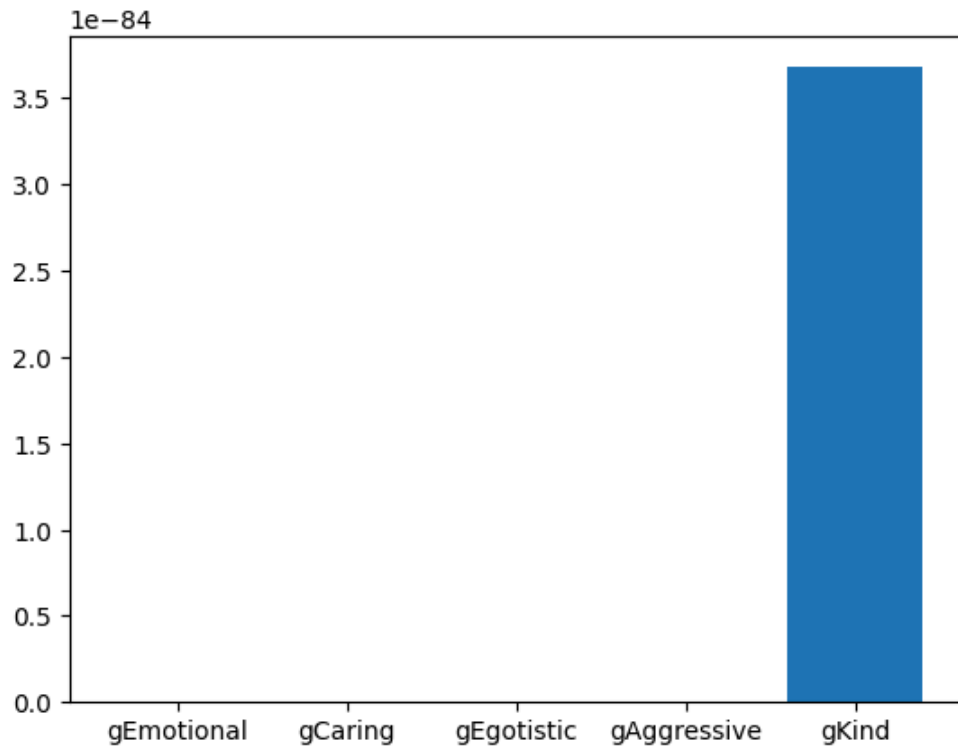
```
[('gEmotional',
  TtestResult(statistic=29.83017935114266, pvalue=1.029377924430242e-164, df=2220.0)),
 ('gCaring',
  TtestResult(statistic=22.965908793068422, pvalue=6.745136923994714e-105, df=2220.0)),
 ('gEgotistic',
  TtestResult(statistic=-22.7262686753705, pvalue=5.724535735483598e-103, df=2220.0)),
 ('gAggressive',
  TtestResult(statistic=-22.47095116212666, pvalue=6.29311355462135e-101, df=2220.0)),
 ('gKind',
  TtestResult(statistic=20.297689926853188, pvalue=3.6772754632964984e-84, df=2220.0)),
 ('gTrustworthy',
  TtestResult(statistic=18.755142226724622, pvalue=5.706506511591532e-73, df=2220.0)),
 ('gFriendly',
  TtestResult(statistic=17.47298898451281, pvalue=3.7108186188491455e-64, df=2220.0)),
 ('gSociable',
  TtestResult(statistic=17.18204476074865, pvalue=3.1946748840509533e-62, df=2220.0)),
 ('gResponsible',
  TtestResult(statistic=12.36280978838484, pvalue=5.314276336798082e-34, df=2220.0)),
 ('gIntelligent',
  TtestResult(statistic=6.202285041707829, pvalue=6.617541566954897e-10, df=2220.0)),
 ('gConfident',
  TtestResult(statistic=3.66917416692047, pvalue=0.0002490671279198179, df=2220.0))]
```

If the p-value is less than the significance level (0.05), it indicates that there are statistically significant differences among the groups. Among all the features, all features have a  $p < 0.05$ , which indicates significant bias exists in all these features between genders.

### *3. Rank the top 5 most significant biases you have found and p-values*

```
{'gEmotional': 1.029377924430242e-164,
 'gCaring': 6.745136923994714e-105,
 'gEgotistic': 5.724535735483598e-103,
 'gAggressive': 6.29311355462135e-101,
 'gKind': 3.6772754632964984e-84}
```

4. *Plot the histograms for the most significant biases you found from the previous task and interpret the results.*



The top 5 significant biases are Emotional, Caring, Egotistic, Aggressive, and Kind across gender, so the bias toward genders might be more significant compared to bias toward race. All 5 p-values reported are far less than 0.05, then that result is statistically significant, indicating there is a severe bias between genders. The features except for Kind are too small to be displayed.

5. *Identify any possible adversities toward certain groups*

The maximum acceptance rate is 0.183 by race 1. According to the four-fifth rule, the minority groups need to have an acceptance rate of  $0.183 \times 0.8 = 0.146$ . The acceptance rate for all groups is: [0.18300654, 0.12272727, 0.17460317]. So only

group 3 from the minority group fulfill the four-fifth rule. There is possible adversity toward group 2.

The maximum acceptance rate is 0.209 by race 1. According to the four-fifth rule, the minority groups need to have an acceptance rate of  $0.209 \times 0.8 = 0.167$ . The acceptance rate for all groups is:  $[0.11857293, 0.20882585]$ . Group 0 from the minority group doesn't fulfill the four-fifth rule, there is possible adversity toward group 0.