The first analysis is to check if our annotated dataset is biased for the gender or ethnicity. To do so we will consider all the documents in annotation collection (October 31st Analysis).

Gender bias analysis:

On this first plot, in the left top corner there is the gender distribution, we can see that we have more female than male (60% to 40%). Then all the other plots are here to compare the average value for each ocean variable:
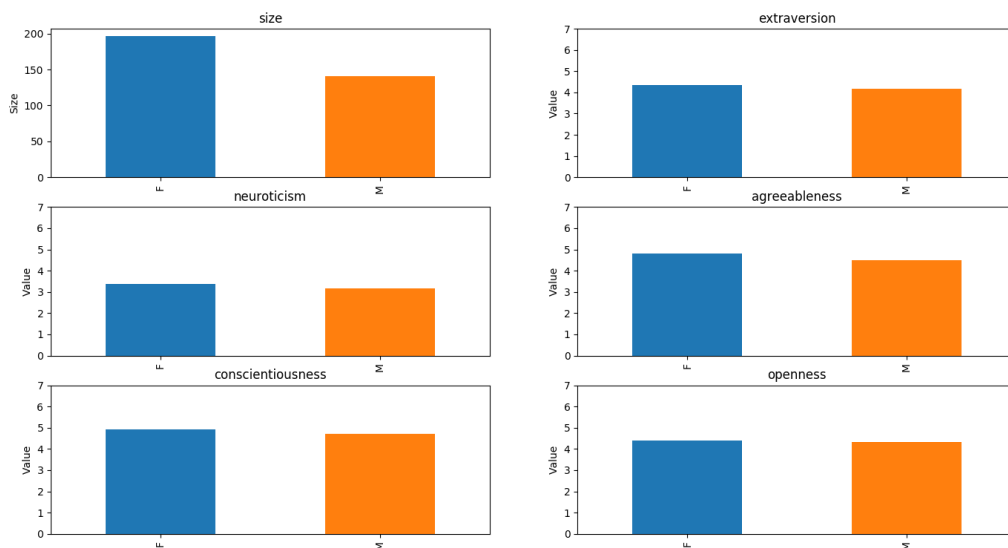

By looking at those plots I don't see any variable which has a big difference between genders. To be sure, let's run a student test and examine the p-value for each ocean variable:

The only ocean variable for which there is a bias is agreeableness (p_value < 0.1%) and there might be one for conscientiousness and neuroticism (p_value < 5%):

ocean_item	average_M	average_F	p_value
extraversion	4.170213	4.347927	0.171133
neuroticism	3.162234	3.373942	0.045781
agreeableness	4.481678	4.821489	0.000918
conscientiousness	4.71513	4.93824	0.026657
openness	4.324764	4.39192	0.52489

Ethnicity bias analysis

By looking at the ethnicity distribution, we can see that we have more than 80% of white candidates, 11% of black and less than 10% for all the other ethnicities combined. Because this dataset is really imbalanced, none of the statical test doesn't reach significant values, and so comparing the average values between ethnicities doesn't prove anything.
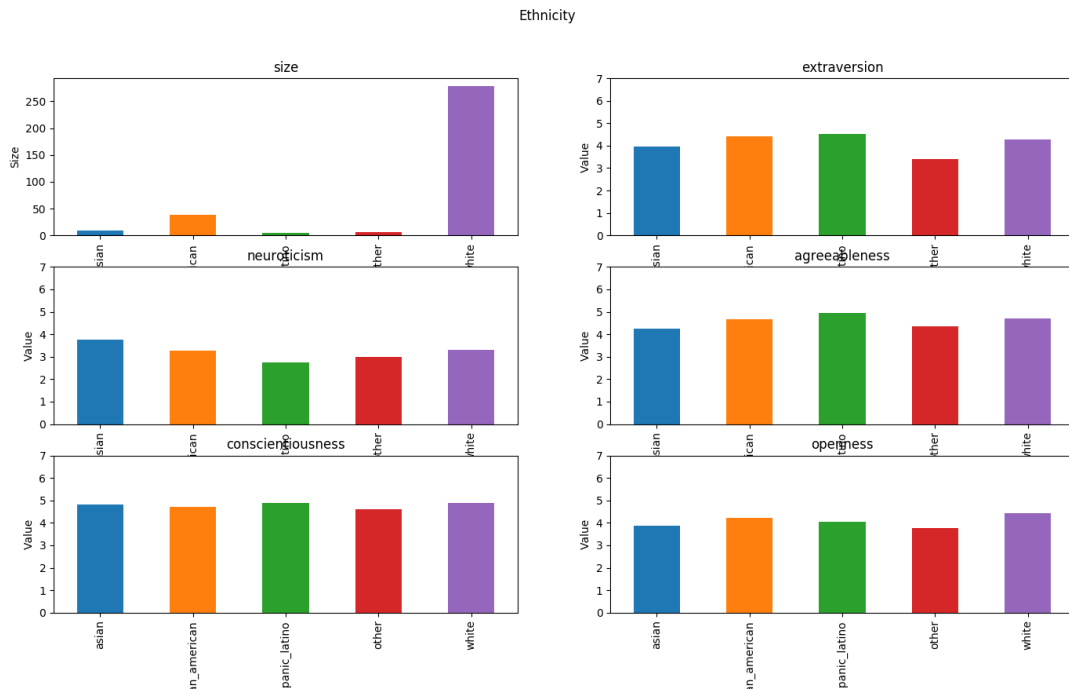
I run the statical student test only to compare the white population and black population scores and even if some average are pretty different, we can't conclude to any significant bias:

ocean_item	average_White	average_Black	p_value
extraversion	4.278076	4.417735	0.479109
neuroticism	3.290472	3.255342	0.83066
agreeableness	4.69773	4.667735	0.852886
conscientiousness	4.869624	4.707265	0.301723
openness	4.419654	4.208333	0.190438