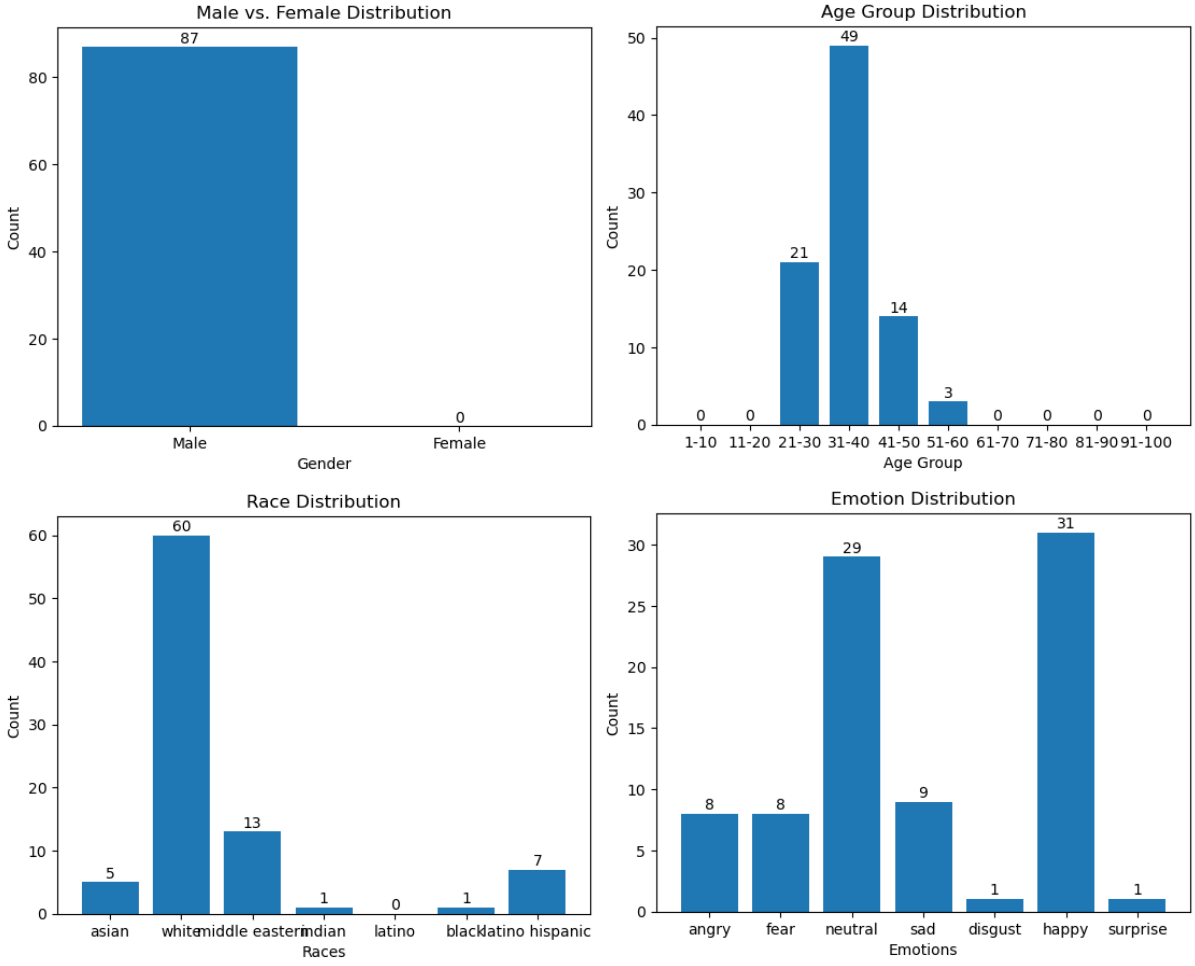
FYP – Implemented Metrics

This document outlines the metrics considered thus, far in relation to determining the presence of bias in visual datasets. The examples provided below are sample outputs retrieved from a set of generative images using the **Stable Diffusion** model using the prompt “Doctor”. (Last Updated 16/02/24)

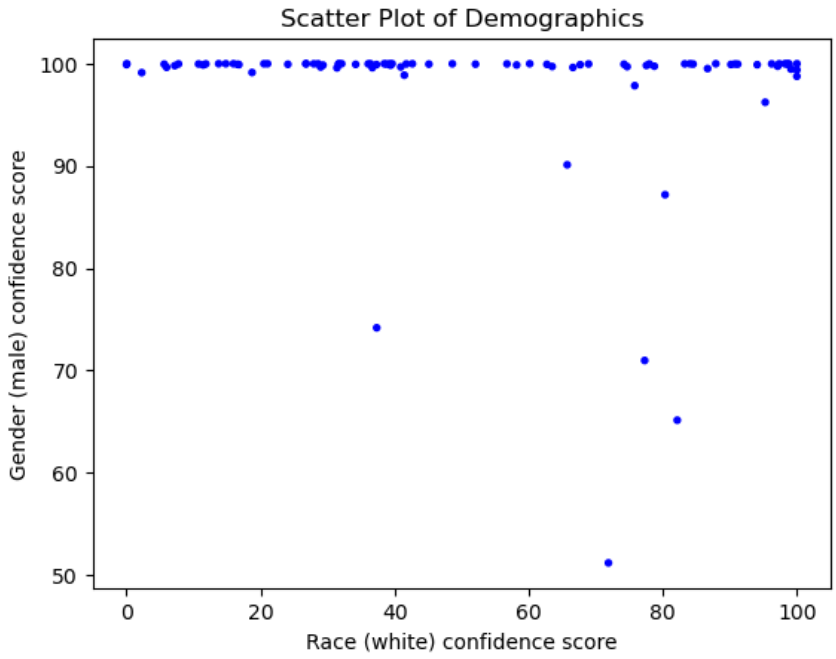
Count

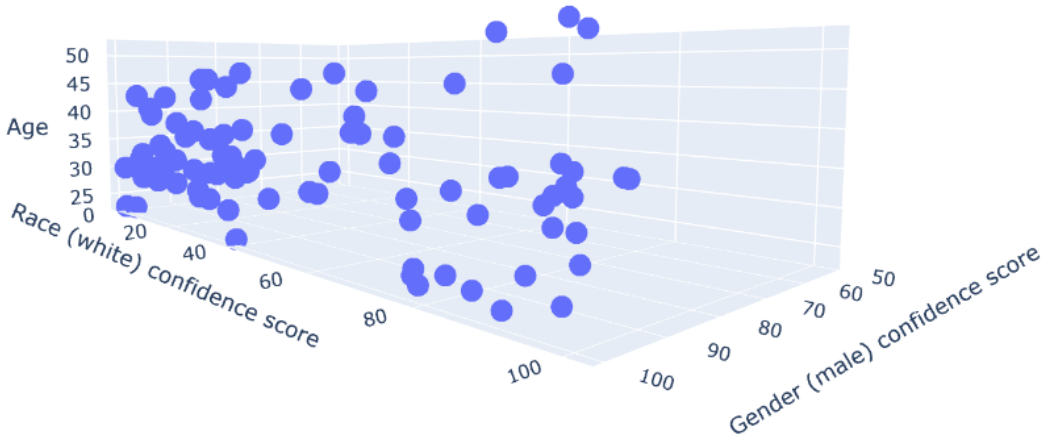
The first metric considered is a simple count, once the image parameters (Gender, Race, Age, Emotion) are extracted via the DeepFace model these are depicted as bar graphs the figures below showcase said graphs. These graphs are also modular meaning that the data can be altered to fit a specific subset of individuals such as depicting the Male vs Female distribution for individuals of white ethnicity which are 30 years old.



Scatter Plot

This metric uses the percentages assigned to images denoting their belonging to each class relevant to Gender/Age/Race/Emotion. It visualises the data in a scatter plot format 2D & 3D. This is going to be used to determine if any cluster/groups appear in relation to the data which might provide further insight into the bias present. For instance, the 2D scatter plot below plots the Gender (male) confidence score vs Race (white) confidence score, thus the higher the percentage value the greater that datapoints association with the label in parenthesis (in this case male and white). This scatter plot can thus lead to conclusion such as males seeming to have a varied distribution of races whereas there are barely any woman in this considered dataset and the few that there are seem to be white. The metrics in parenthesis are dynamic meaning they can be changed to provide differing insight. This also applies to the 3D variation seen below.





Person Prominence

This is a metric based on the area which the person in the image takes up as well as their distance from the centre of the image which is taken as proxy for importance. This metric was used in the REVISE paper (Table 2) link. As of now this metric is simply being depicted as a line graph as showcased below. Whereas the Y-axis are limited to Person Area & Distance, the x-axis can be changed to denote the data in relation to Gender, Age, Race, Emotion.