CS 6603 - Final Project
Spring 2022

Group Members: Elissa Rupley, Ben Mighall, Wesley Addo, Kofi Neizer-Ashun

Step 1

Dataset Selected: Personal Key Indicators of Heart Disease (Kaggle)
https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease

Regulated Domain: Healthcare (Health Insurance Portability and Accountability Act of 1996, Title VI Civil Rights Act of 1964)

Observations in Dataset: 319,795
Variables in Dataset: 18
Dependent Variables: HeartDisease, Stroke

Protected Class Variables:
Sex - Equal Pay Act of 1963; Civil Rights Act of 1964, 1991
Age - Age Discrimination in Employment Act of 1967
Race - Civil Rights Act of 1964, 1991

Step 2

1.) Relationship between members and membership categories for each protected class variable

Protected Class Variables	Members
Sex	Male, Female
Age	18-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 or older
Race	White, Black, Asian, American Indian/Alaskan, Native, Hispanic, Other

2.) Values associated with dependent variables (Heart Disease, Stroke)

Protected Class - Sex	Heart Disease	Stroke
Male	3309	1242
Female	2411	1421

Protected Class - Age	Heart Disease	Stroke
18-24	30	13
25-29	22	13
30-34	48	31
35-39	62	44
40-44	100	53
45-49	149	97
50-54	281	166
55-59	437	241
60-64	668	309
65-69	850	374
70-74	1015	418
75-79	880	402
80 or older	1178	502

Protected Class - Race	Heart Disease	Stroke
White	4368	1771
Black	524	398
Asian	104	77
Hispanic	337	159
American Indian/Alaskan Native	119	94
Other	268	164

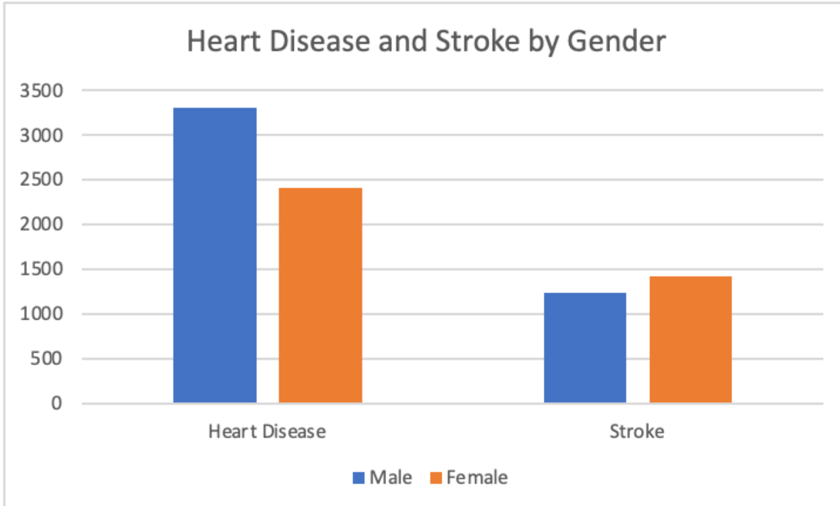
3.) Frequency of each protected class

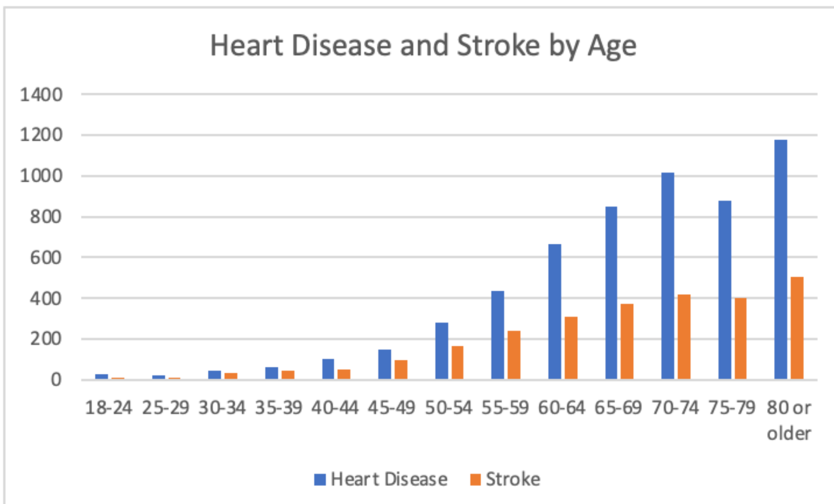
Sex	Frequency
Male	30793
Female	34742

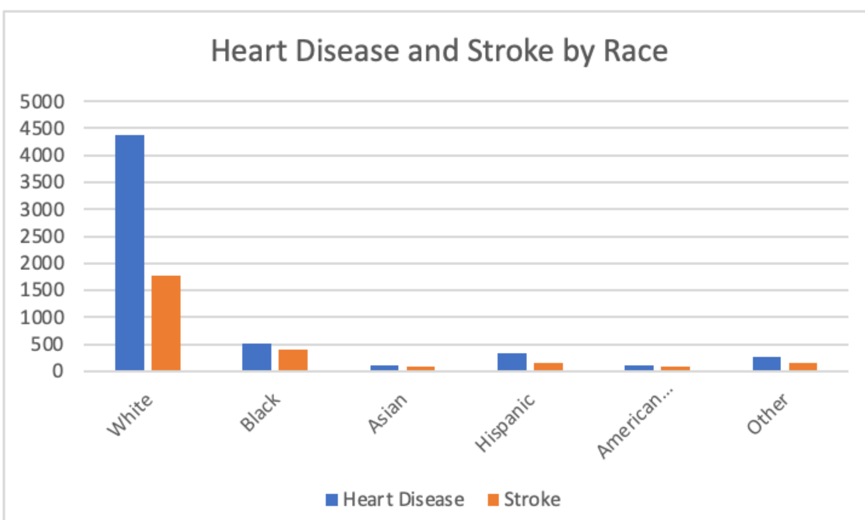
Age	Frequency
18-24	3969
25-29	3481
30-34	3941
35-39	4146
40-44	4116
45-49	4448
50-54	5142
55-59	5893
60-64	6695
65-69	6899
70-74	6670
75-79	4880
80 or older	5255

Race	Frequency
White	44866
Black	6674
Asian	2599
Hispanic	6934
American Indian/Alaskan Native	1303
Other	3159

4.) Histograms







Step 3

	Privileged	**Unprivileged**
Sex	Male	Female
Age	18-24 25-29 30-34 35-39	40-44 45-49 50-54 55-59 60-64 65-69 70-74 75-79 80 or older
Race	White	Black Asian American Indian/Alaskan Native Hispanic Other

Reweighting was the pre-processing bias mitigation algorithm selected. Reweighting was done by looking at the GenHealth variable. If GenHealth is Excellent, Very Good, Good, or Fair, an unfavorable outcome becomes favorable if a person is unprivileged.

	Disparate Impact	Disparate Impact (WEIGHTED)	Statistical Parity	Statistical Parity (Weighted)
Sex vs. Heart Disease	1.0439	1.1066	0.0392	0.0953
Sex vs. Stroke	0.9988	1.0321	-0.0012	-0.0054
Age vs. Heart Disease	0.8995	0.9945	-0.0995	0.0803
Age vs. Stroke	0.9581	0.9984	-0.0417	0.0309
Race vs. Heart Disease	1.0292	1.0884	0.0265	-0.0016
Race vs. Stroke	1.0005	1.0321	0.0005	0.0309

<u>Step 4</u>

1) The privileged and unprivileged groups were determined by age: individuals younger than 40 were considered privileged, while individuals 40 and older were considered unprivileged.

2) The dependent variable used was HeartDisease.

3) **Fairness metrics results on classified output:**

	Disparate Impact	Statistical Parity
Original dataset	0.960721768	-0.039224352
Transformed Dataset	0.904977264	-0.094029618

4)

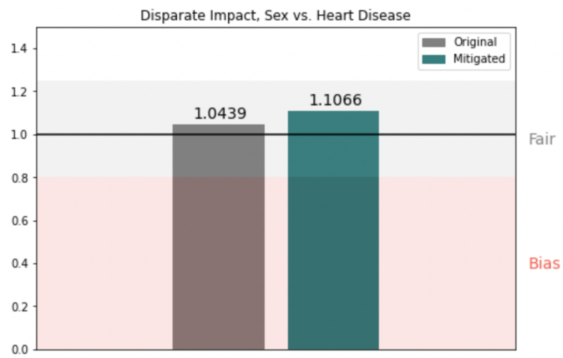
		Positive	Negative	No Change
Transformed Dataset	**Disparate Impact**	YES		
Trained Original Dataset			YES	
Trained Transformed Dataset			YES	
Transformed Dataset	**Statistical Parity**	YES		
Trained Original Dataset			YES	
Trained Transformed Dataset			YES	

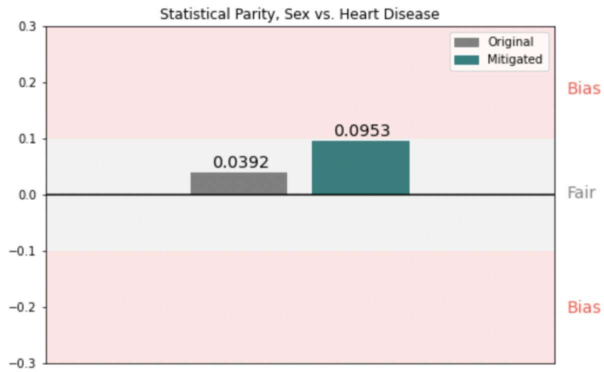
<u>Step 5</u>

Project Team Members:
Elissa Rupley
Ben Mighall
Wesley Addo
Kofi Neizer-Ashun

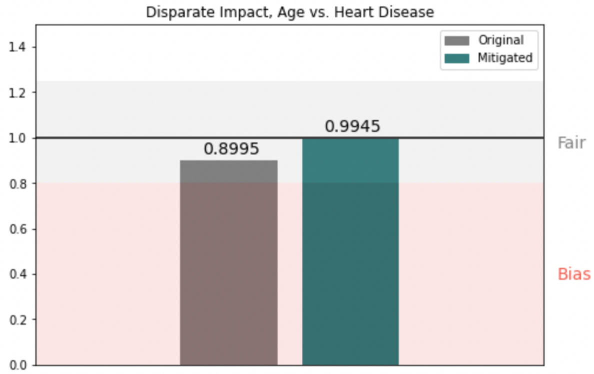
<u>3.2 and 3.4 Graphs (Original & Mitigated combined):</u>

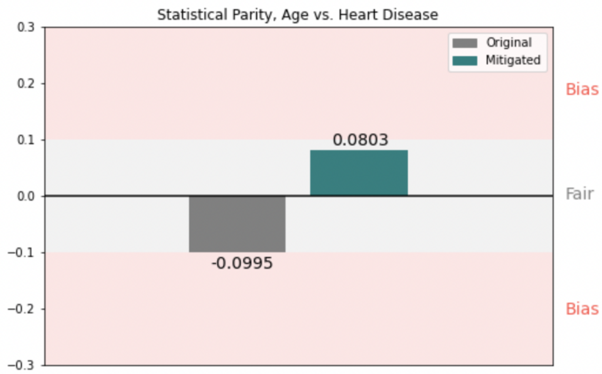
Sex vs. Heart Disease



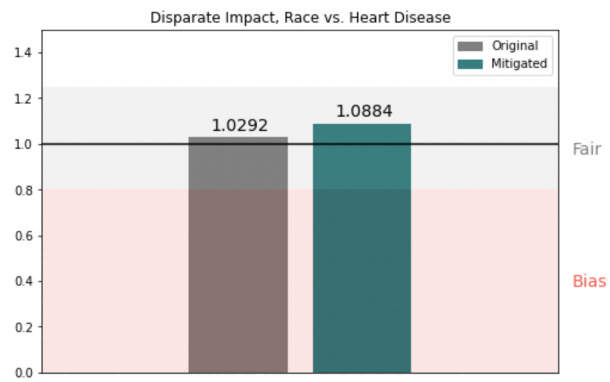


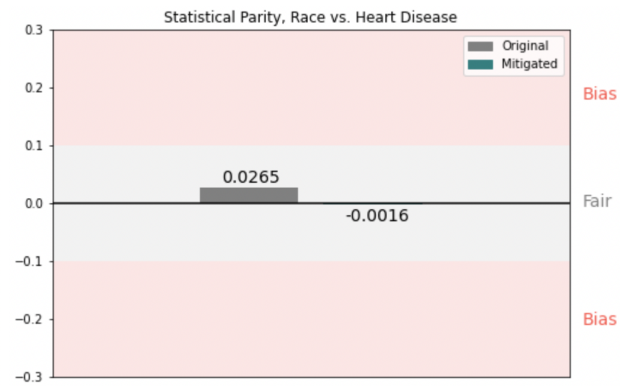
Age vs. Heart Disease



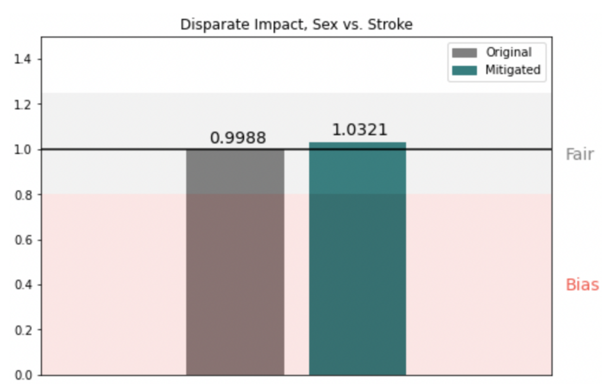


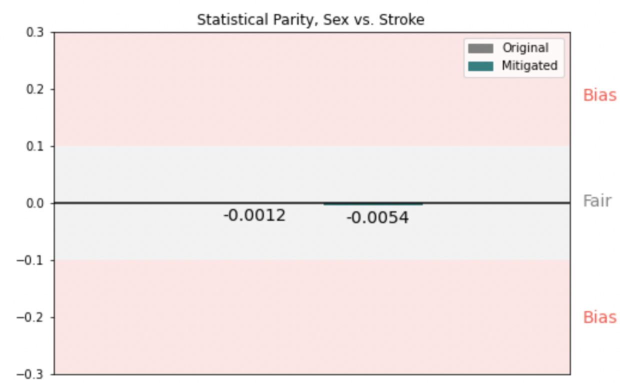
Race vs. Heart Disease



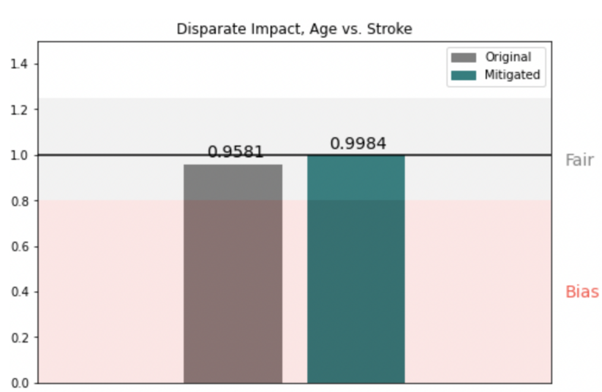


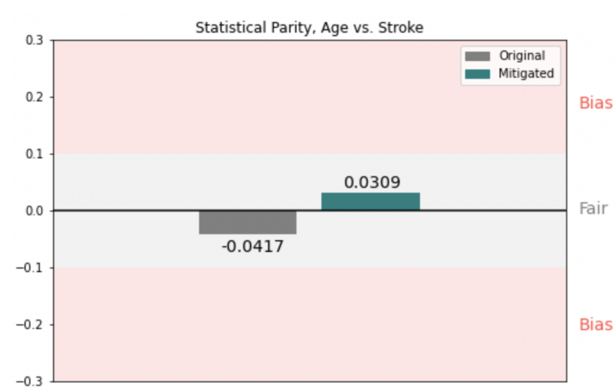
Sex vs. Stroke



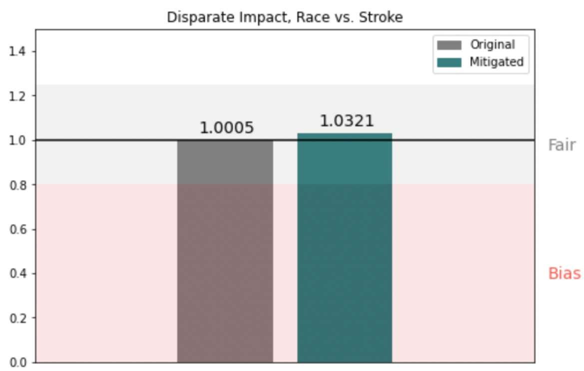


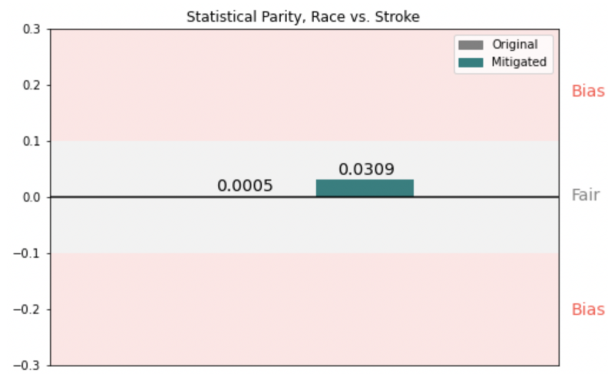
Age vs. Stroke



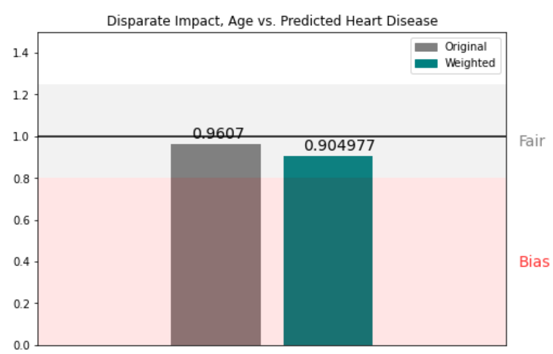


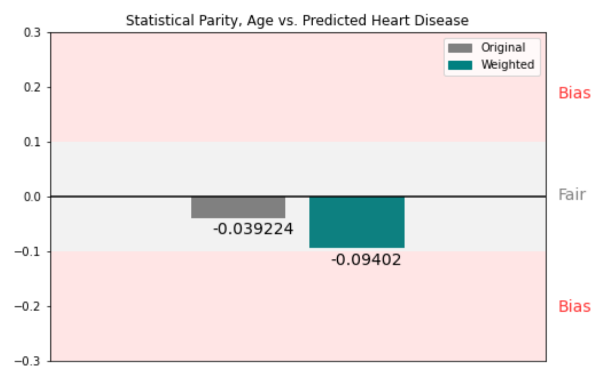
Race vs. Stroke





4.5





Did any of these approaches seem to work to mitigate bias (or increase fairness)? Explain your reasoning. Did any group receive a positive advantage? Was any group disadvantaged by these approaches? What issues would arise if you used these methods to mitigate bias?

<u>Wesley Addo:</u>
Reweighting mostly seemed to mitigate bias where it mattered. Protected classes which previously had a more marked difference in outcome had those differences mitigated after reweighting was done. The unprivileged group in the age protected class seemed to benefit the most with both dependent variables.They had the most fair outcome after the bias mitigation, with the difference reaching near 0 bias. All the other protected classes, however, had more bias towards the privileged group after the reweighting process. Thankfully, the increment was negligible. Using this method to mitigate bias could result in unbalanced outcomes for different groups. Since the process is supposed to be a one-size-fits-all solution for all groups, it is going to mitigate bias for different groups to different degrees, and in some cases increase bias (albeit slightly).

Kofi Neizer-Ashun:
Training a classifier on the original and transformed data set actually increased fairness among privileged and unprivileged groups, especially age. Using the heart disease variable as a dependent variable, the unprivileged age group (>=40) had the positive advantage of benefiting more than the privileged group (< 40). Since the original and the transformed data had a fair outcome for age, there might be a result of biasness in other protected class if this approach is used to mitigate bias.

Elissa Rupley:
The Disparate Impact (weighted) approach worked to mitigate bias for some of the variables including Age vs. Heart Disease and Age vs. Stroke. This helped the data become fairer for the unprivileged group (over the age of 40). This group received a positive advantage on both the Hearth Disease and Stroke data. Most of the original data is quite close to fair. The Disparate Impact and the Statistical Parity for didn't help a few of the variables include Sex vs. Heart and Sex vs. Stroke groups. A few groups were disadvantaged by implementing these approaches including race and sex for both the Heart Disease and Stroke data. If this method was used to mitigate bias, these groups could be at a disadvantage for things that this data might be used for. If doctors or pharmacies were using this data to do things like help diagnose, treat, or prevent heart disease or stroke, the unprivileged people in the race and sex category including women and people who are black, Asian, American Indian/Alaskan Native, Hispanic, and others could be missed diagnoses or treated more frequently than the privileged groups like men and people who are white.

Ben Mighall:
The reweighting approach was a mixed bag; according to the two fairness metrics, it did successfully mitigate bias in some circumstances, though actually increased it in others. For example, race versus heart disease had its bias almost completely mitigated according to Statistical Parity (SP), but became less fair according to Disparate Impact (DI). In other cases such as age versus heart disease, the DI showed the bias as mitigated but the SP was only slightly mitigated. Others, such as sex versus stroke, had both SP and DI show that bias actually increased. Overall, it did seem to mitigate bias about half the time, but only decreased it according to both metrics in a few cases. As for whether groups received a positive or negative advantage, SP did not appear to show a major difference, though DI did show that the mitigated values showed a bias against the "privileged" groups more disparately in some cases, such as sex and race in heart disease. The difference appeared to be negligible, however. Using this approach to mitigate bias did not improve the fairness clearly by both metrics and increased the bias in many cases; therefore, using this strategy could decrease accuracy and increase bias for this specific dataset.