

# Bias in AI Report

Date: 30/04/2022

## 1 INTRODUCTION

IN this report, I will simply be outlining the method in which I used to solve Tasks 1 and 2, related to an interview and job offer system which determines whether a job applicant should be given an offer.

## 2 TASK 1 (DATASET ANALYSIS) SOLUTIONS

### 2.1 Question 1

In order to determine the sensitive attributes within the dataset, I simply analysed its columns, which can be seen below.

ApplicantCode	Gender	BAMEyn	ShortlistedNY	Interviewed	FemaleONpanel	OfferNY	AcceptNY	JoinYN
0	1	1	2	1	1	1.0	1.0	1.0
1	2	1	2	1	1	1.0	1.0	1.0
2	3	1	2	1	1	1.0	1.0	1.0
3	4	1	2	1	1	2.0	1.0	1.0
4	5	1	2	1	1	2.0	0.0	NaN

**Fig 1.** A table showing the columns withing the dataset, along with its first few rows.

The attributes which can be discriminated against include the gender of job applicants ("Gender") and whether applicants are Black, Asian or of ethnic minorities ("BAMEyn"). This was determined through ethnicity and gender both being attributes belonging to others. This may provide reason for such people of different groups to be treated differently, such as by an interview panel member. The privileged groups are the groups that receive a greater proportion of job offers than another group, which is unrepresentative to the size of a group since discrimination has occurred.

In particular, the privileged group for "Gender" are males and the unprivileged group are females, since 4.95% of female applicants receive offers compared to 23.1% of male applicants. Also, the unprivileged group for "BAMEyn" are those that are BAME, since 6.61% of BAME applicants receive offers compared to 12.6% of BAME applicants. See Fig 3 for full details of these statistics.

"FemaleONpanel" would not be a sensitive attribute, since we are solely concerned with those attributes related to members that may potentially receive a job.

### 2.2 Question 1

The key information for the privileged and unprivileged groups can be seen below.

	Gender	BAMEyn
Minimum Value	1	1
Maximum Value	2	2
Range	1	1
Mean	1.721	1.568
Standard Deviation	0.4491	0.4963
Variance	0.2017	0.2463

**Fig 2.** A table detailing the spread of the "Gender" and "BAMEyn" groups.

	Gender	
	Male	Female
Total	78	202
Shortlisted	38	50
Interviewed	27	28
Offered	18	10
Accepted	12	6
Joined	12	6

**Fig 3.** A table showing the number of applicants, for the privileged and unprivileged sub-groups of "Gender", at each stage of the application process.

	BAMEyn	
	BAME	Non-BAME
Total	121	159
Shortlisted	19	69
Interviewed	13	42
Offered	8	20
Accepted	4	14
Joined	4	14

**Fig 4.** A table showing the number of applicants, for the privileged and unprivileged sub-groups of "BAMEyn", at each stage of the application process.

### 2.3 Question 3

The information for the disparity between privileged and unprivileged groups, when invited to an interview and offered a job, can be seen below. This includes the percentages of being invited to an interview or being offered a job, given the group that an applicant belongs to.

	Invited to an Interview	Offered a Job
Male	48.72%	23.08%
Female	24.75%	4.950%

**Fig 5.** A table showing the statistical disparity between the privileged “Male” group and unprivileged “Female” group, at two stages of the job application process.

$P(\text{Invited to an interview} | \text{Male}) = 38/78$   
 $0.48717948717948717$   
 $P(\text{Offered a Job} | \text{Male}) = 18/78 = 0.23076923076923078$   
 $P(\text{Invited to an interview} | \text{Female}) = 50/202$   
 $0.24752475247524752$   
 $P(\text{Offered a Job} | \text{Female}) = 10/20 = 0.04950495049504951$

	Invited to an Interview	Offered a Job
BAME	15.70%	6.612%
Non-BAME	43.40%	12.58%

**Fig 6.** A table showing the statistical disparity between the unprivileged “BAME” group and privileged “Non-BAME” group, at two stages of the job application process.

$P(\text{Invited to an interview} | \text{BAME}) = 19/121$   
 $0.15702479338842976$   
 $P(\text{Offered a Job} | \text{BAME}) = 8/121 = 0.06611570247933884$   
 $P(\text{Invited to an interview} | \text{Non-BAME}) = 69/159$   
 $0.4339622641509434$   
 $P(\text{Offered a Job} | \text{Non-BAME}) = 20/159 = 0.1257861635220125$

### 2.4 Question 4

To prove that the dataset is biased towards the privileged groups in the shortlisting process, I carried out a Chi-squared tests between the number of shortlisted applicants and the variable containing a privileged group.

A Chi-squared test tells us how likely a certain distribution is due to chance. This procedure includes defining a null hypothesis, which we plan to reject or accept. We then carry out the Chi-square formula on normalized values, within a table containing information between our two variables. This results in a Chi-squared value. Along with the degree of freedom value (the number of independent variables), we can look up the p-value which represents the probability that our null hypothesis occurs. We can compare this to our significance value to either accept or reject  $h_0$ .

To prove that the dataset is biased in the shortlisting process towards the non-BAME group:

$h_0$ : There is no relationship between an applicant that is BAME / not BAME and whether they were shortlisted or not.

Significance value: 0.05

	Shortlisted	Not Shortlisted	Total
BAME	19 (38.03)	102 (82.97)	121
Non-BAME	69 (49.97)	90 (109.0)	159
Total	88	192	280

Chi-squared value =  $\frac{(19-38.03)^2}{38.03} + \frac{(102-82.97)^2}{82.97} + \frac{(69-49.97)^2}{49.97} + \frac{(90-109.0)^2}{109.0} = 24.45$

Degrees of freedom =  $(2-1)*(2-1) = 1$

p-value =  $7.626 \times 10^{-7} = 0.0000763\%$

$7.626 \times 10^{-7} < 0.05$ , hence we can say that there is insignificant evidence to accept the null hypothesis. Therefore, being shortlisted is dependent on whether you belong to a BAME group.

To prove that the dataset is biased in the shortlisting process towards the non-BAME group:

$h_0$ : There is no relationship between an applicant’s gender and whether they were shortlisted or were not.

Significance value = 0.05

	Shortlisted	Not Shortlisted	Total
Male	38 (24.51)	40 (53.49)	78
Female	50 (63.49)	152 (138.5)	202
Total	88	192	280

Chi-squared value =  $\frac{(38-24.51)^2}{24.51} + \frac{(40-53.49)^2}{53.49} + \frac{(50-63.49)^2}{63.49} + \frac{(152-138.5)^2}{138.5} = 15.01$

Degrees of freedom =  $(2-1)*(2-1) = 1$

p-value =  $1.069 \times 10^{-4} = 0.0107\%$

$1.069 \times 10^{-4} < 0.05$ , hence we can say that there is insignificant evidence to accept the null hypothesis. Therefore, being shortlisted is dependent on an applicant’s gender.

### 2.4 Question 5

Please see the source code I have attached to see my implementation of Question 4. In order to assist this process, I made use of “scipy” which is a scientific computing library. This contains “chi2\_contingency()”, which automatically calculates the Chi-squared value, p-value, degrees of freedom and the expected values (used in the calculation for the Chi-squared value). The result of the p-values was then used to evaluate the acceptance or rejection of the null hypothesis.