

Study of Ageism in England Based on ELSA Data

yt1369¹

¹New York University (NYU)

December 15, 2017

PUI2017 Extra Credit Project Report

<Yixuan Tang, yixuantang, yt1369>

Introduction

The word “Ageism” first appeared in 1969, soon after the Federal Discrimination in Employment Act set 40 as a lower age bound that workers can complain. With population aging, the problem of ageism to older people is increasingly severe. In this project, I want to explore whether older people experienced ageism and what factors leads to more serious ageism problem. I plan to use the data from the wave 7 of the English Longitudinal Study of Ageing (ELSA) in the England and perform z test and multilinear regression analysis to answer the problems.

Data

- Data acquisition

I used the data from *English Longitudinal Study of Aging, Wave5(2010 - 2011)*, this data is collected by UK [NatCen Social Research](#), ELSA follows the lives of adults in England ages 50 and over. This dataset is valuable and suitable in this project, first because it has sufficient samples: 10,274 cases, secondly, in Wave5 survey, ELSA asked respondents about the frequency of [five discriminatory situations](#) as follows,

1. You are treated with less courtesy or respect than other people
2. You receive poorer service than other people at restaurants or stores
3. People act as if they think you are not clever
4. You are threatened or harassed
5. You receive poorer service or treatment than other people from doctors or hospitals

These perceived ageism frequency information combined with the other variables in the data set(age, marital status, income, education level) can be used to find the answer of whether older people experienced ageism and what factors leads to more serious ageism problem.

- Data cleaning

UK Data Service provide ELSA data in sas, sav and dta format, I transformed the sav file into csv file by R, and used pandas in python to extract and manipulate the parameters I want, for the purposes of this research, here I only select ‘Age’, ‘gender’, ‘financial status’, ‘work status’, ‘perceived discrimination frequency’ and ‘self reported reason for discrimination’ in the final dataset.

The raw data contains 5 column for ‘frequency of discrimination’, to convert the column into a format that can be used in computational models, the 5 conditions(‘Never’, ‘Less than once a year’, ‘A few times a year’, ‘At least once a week’, ‘Almost every day’, ‘A few times a month’) each was assigned a score from 0 - 5,

and sum up into a new column named ‘DisScore’, the higher the value is, the severer the discrimination the person perceived. After data cleaning, the data frame has 3545 observations.

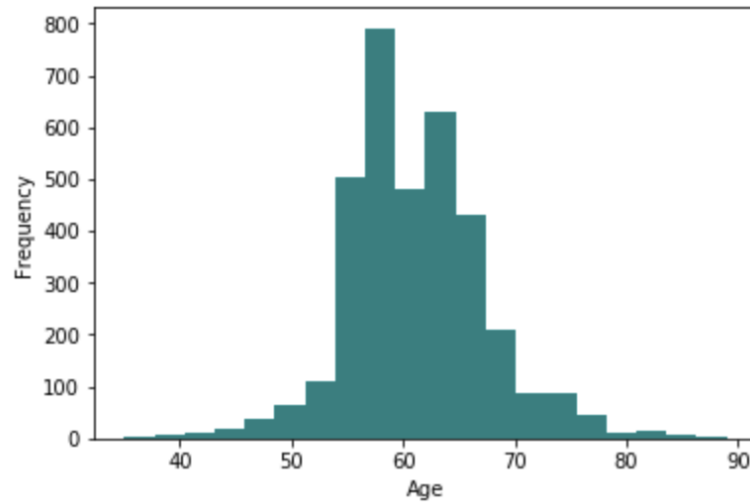


Figure 1:
distribution in dataset

Age

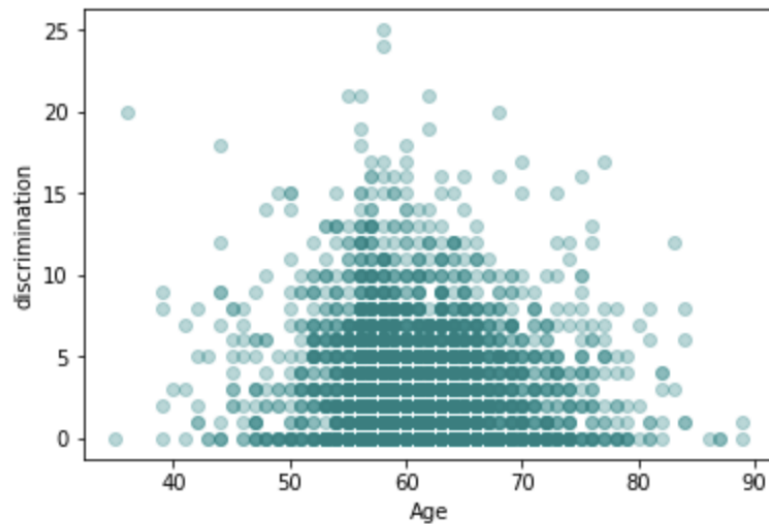


Figure 2:
discrimination score vs age

Perceived

Analysis

I used statsmodels in python to perform linear regression on respondents’ age, household income and ‘DisScore’ to explore how well does discrimination relate to age and financial status. I also set my null hypothesis that *the ratio of female experienced discrimination is the same or lower than male in this survey*, I then performed z tests at significance level $p=0.05$ to test it.

Results

In the descriptive analysis, among 3545 valid respondents, 72% of them think they perceived discrimination in life, and most people perceived discrimination think the cause is their age, and the second and third causation is gender and financial status respectively. In the linear regression model, the relationship between age and perceived discrimination or household income and perceived discrimination is not significant indicated by the R^2 (Fig3-4).

OLS Regression Results

Dep. Variable:	DisScore	R-squared:	0.008			
Model:	OLS	Adj. R-squared:	0.008			
Method:	Least Squares	F-statistic:	29.92			
Date:	Thu, 14 Dec 2017	Prob (F-statistic):	4.80e-08			
Time:	18:00:54	Log-Likelihood:	-9406.9			
No. Observations:	3545	AIC:	1.882e+04			
Df Residuals:	3543	BIC:	1.883e+04			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t 	[0.025	0.975]
const	6.4337	0.574	11.217	0.000	5.309	7.558
indoby	-0.0511	0.009	-5.470	0.000	-0.069	-0.033
Omnibus:	900.659	Durbin-Watson:	1.951			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2206.748			
Skew:	1.390	Prob(JB):	0.00			
Kurtosis:	5.685	Cond. No.	610.			

Figure 3:
perceived discrimination linear regression

Age and

The primary outcome was the z test of whether female perceived more discrimination than male, under significance level $p=0.05$, the null hypothesis that *the ratio of female experienced discrimination is the same or lower than male in this survey* can be rejected.

Limitations and Future work

The result from regression model can not explain the ageism phenomenon, to find out the underlying relationship between ageism and causation, I need to look into other aspects, such as race, sexual orientation and health status, however, many types of these variables have been deleted in the dataset in order to reduce the chance to identify individuals. To fix this, I plan to involve more dataset in the future work, such as regional distribution data and employment data. Also, it could be meaningful to look into a time range and different areas to see whether social development has effect on ageism phenomenon.

As for the method part, if more variables involved, using decision tree to see the discrimination probability would be useful.

References:

1. [Why Ageism Never Gets Old](#) This The New Yorker article gives inspiration to me to explore the ageism problem, it also briefly explained the history of ageism and emphasized the seriousness in professional and

OLS Regression Results

Dep. Variable:	DisScore	R-squared:	0.010
Model:	OLS	Adj. R-squared:	0.009
Method:	Least Squares	F-statistic:	17.11
Date:	Thu, 14 Dec 2017	Prob (F-statistic):	4.01e-08
Time:	18:01:50	Log-Likelihood:	-9404.7
No. Observations:	3545	AIC:	1.882e+04
Df Residuals:	3542	BIC:	1.883e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	6.9679	0.629	11.081	0.000	5.735	8.201
indobyr	-0.0578	0.010	-5.848	0.000	-0.077	-0.038
iasinc	-4.538e-06	2.2e-06	-2.067	0.039	-8.84e-06	-2.35e-07

Omnibus:	896.467	Durbin-Watson:	1.950
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2189.966
Skew:	1.385	Prob(JB):	0.00
Kurtosis:	5.674	Cond. No.	4.33e+05

Figure 4: Age + household income and perceived discrimination multi-variables linear regression

technology industry.

2. [Age Discrimination and Hiring of Older Workers](#) This is a recent work from David Neumark, a professor at University of California, it gives the evidence that age discrimination makes it harder for older individuals.
3. [NatCen](#) Britain's largest independent social research agency, the ELSA provider.

Appendix:

The raw dataset mistakenly labeled the answer of 'frequency of five discriminatory situations' in 4 columns, after I confirmed with the data provider *UK Data Service*, I correct the label by assign them the new label(Table1).

0 Disagree	0 Never
1 Slightly disagree	1 Less than once a year
3 Slightly agree	2 A few times a year
2 Neither agree nor disagree	4 At least once a week
4 Agree	3 A few times a month
5 Strongly agree	5 Almost every day

Table 1: left column in the wrong label and right column is the correct label