Title: Analyzing Attitudes Towards Same-Sex Marriage Rights

Individual Project Report Word Counts:2418

Course Code: ST211

1. Introduction:

The Northern Ireland Life and Times Survey, 2012: Lesbian, Gay, Bisexual and Transgender Concerns, Teaching Dataset (NILT), by virtue of which, can serve to elaborate on views and attitudes of residents concerning same-sex marriage as a right. This dataset is composed of a compact set of 1064 observations and 39 variables with a view to proffering a complete picture of the socio-cultural landscape of Northern Ireland. At the heart of this analysis lies the outcome variable ssexmarr, which serves as a binary indicator of whether respondents support granting same-sex marriages the same rights as traditional marriages. As for the role of the mentioned question lies in the heart of this topic and helps in uncovering the attitudes of a considerable part of society to the proclaimed rights of the LGBTQ+ rights as well as to marriage equality.

The primary objective of this study is to identify the key predictors influencing individual choices as to whether to answer this crucial question. By exploring the relationships between demographic factors, such as age, education, and income levels, we also shed light upon the attitudinal dimensions, comprising political institution representation and the roles of religion, in order to reveal the combined effect of these influences on the views of the same-sex marriage rights. This study aims at adding its voice to the situation of the LGBTQ+ rights and marriage equality, opening the doors to contributions that may help in formulating policy and initiatives aiming at incorporation and equality in Northern Ireland and across the globe.

2. Data Cleaning and Preparation:

First of all, we filled in the gaps in the dataset that were not available on demand. Various variables had the missing individual values known as NA. Beyond that, missing values in categorical predictors 10%, a "missing" level was introduced to preserve the information while considering the absence of data. Missing values monitoring for continuous predictors was also conducted. It was then determined that missing values were present for some predictors, and for each predictor, the percentage of missing values was calculated. Thus, for all items in categorical features representing over 10% missing variables, we replaced the missing values with a "missing" category so as to not remove such observations in the study. Furthermore, in the case of the normally continuous predictor persinc2, the forecasts generated by this model grow up to imputed mean value from non-missing.

On the other hand, the se sexmar variable was dichotomised to the format of two, where the value 0 is for "No" and 1 for "Yes". By doing coding this resulted in logistic regression analysis, which enabled us to adequately establish the relationship between the predictors and the odds of the proponents of same-sex marriage having their civil rights' violations correct. Through filling the gaps of missing values and recoding the output variable, we ensured data set was ready for proper analysis which formed a way to the outcome of clear and dependable findings on how the society feels towards granted marriage rights to same-sex couples.

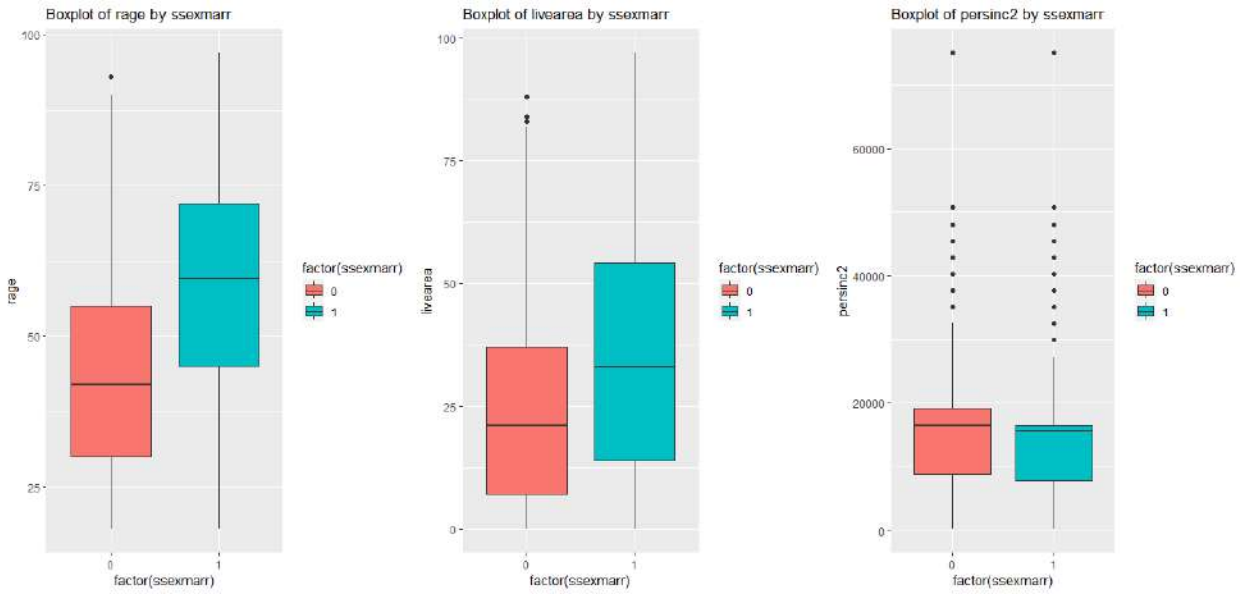
Likewise, multiple categorical predictors expanded to fewer levels by merging them to strengthen the model interpretation. Specifically, in order to create a more precise variable, the religface predictor was merged into a broader category which is clearly understood as "Christian," while the polpart2 predictor was merged into an "Other" category. The improvements were done for the purpose of guaranteeing the data set's validity and accuracy, which is vital for further studying. To meet these requirements, the dataset will now be subjected to logistic regression analysis to discover variables that impact the views of individuals regarding the equal marriage rights or not.

Thses results are summarized in the table that is mentioned in the appendix.

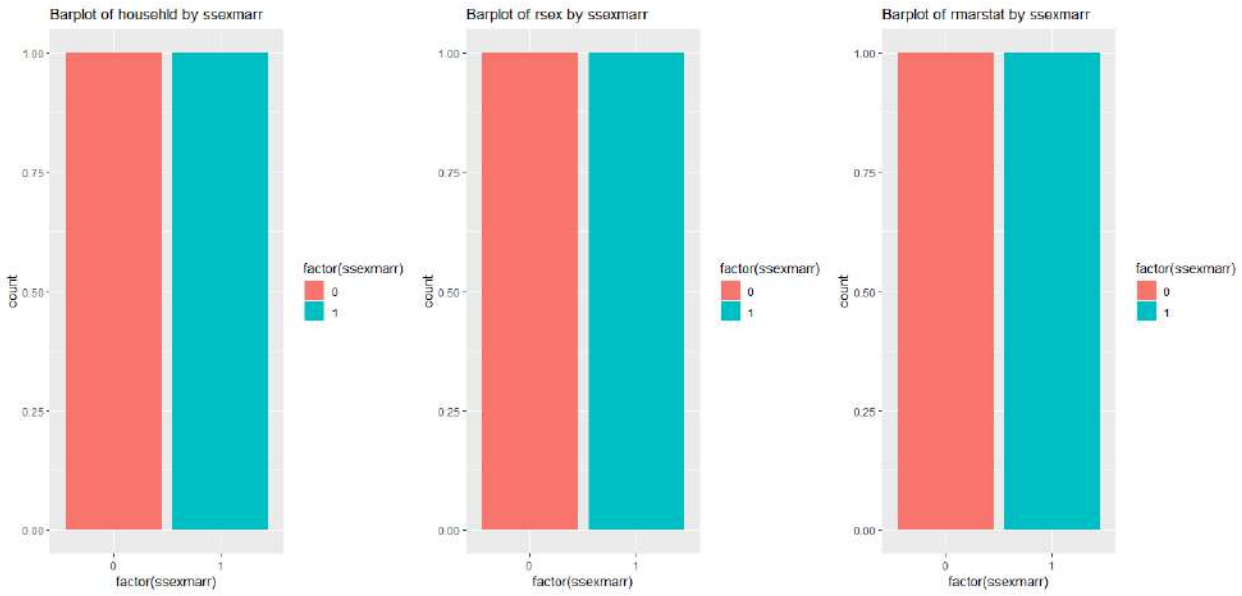
3. Exploratory data analysis:

EDA (exploratory data analysis) which is very important the start of data analysis, lets the researcher learn about characteristics of the elaborated dataset. With this work we have used boxplots and barplots to see the connections between predictors and the outcome variable sexmarr, which is the variable that describes attitude of respondents toward same-sex marriages.

The boxplots were used to examine the distribution of three continuous variables: age (rage), place of residence (place of staying) and own income (have your own) in connection with unhappyness (in possession of a broken soul). The age boxplot summary shows a very interesting pattern which implies that the younger responders have more probability to be pro gay marriage compared to the older interviewees. The fact that this theory shows that it could be a potential factor influencing the people of this era only proves the theory in a young generation. The two boxes of living area and personal income did not form trends, which probably indicate that they do not affect attitudes to the same-sex marriages to the same extent they influence attitudes to one-sex marriages.



On the other hand, the barplots were used to visualize the distribution of three categorical variables: had household (hh), respondent's sex (rsex) and respondent's marital status (rmarstat) and regarding ssexmarr. With a bar plot of the household, for which type of house a perception of same-gender marriages as good and bad are revealed. Using gender of a respondent to determine whether there was a difference in attitudes among the male and female respondents was shown by the exercise in a barplot. The barplot for marital status displayed a trend in which marital status was a key issue in someone's perception of the same-sex marriages.



Ultimately, EDA established the possibility of an association between people's community characteristics and perceptions on same-sex unions. Beneath all, it was such a tendenous group who showed their support toward same-sex marriages that include: younger age respondents, specific household types, and certain marital statuses. The exploratory findings are further developed using logistic regression to identify relationships that are significant in forecasting homosexuals' civil rights campaigns. By this full study, we will address to the problems influencing attitudes toward same-sex marriages, therefore, helping us to know this phenomena better and more profoundly.

4. Model Evaluation:

Initial Model:

Logistic Regression Analysis of initial Model

```
> summary(initial_model)

Call:
glm(formula = ssexmarr ~ ., family = "binomial", data = data)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.110e-01  2.752e+00   0.077  0.93888
househld        -2.604e-02  8.671e-02  -0.300  0.76393
rage             2.793e-02  9.084e-03   3.075  0.00211 **
rsex            -5.981e-01  2.150e-01  -2.782  0.00540 **
rmarstat         1.696e-01  6.930e-02   2.447  0.01439 *
livearea        -4.132e-03  5.695e-03  -0.725  0.46816
hincpast2       -3.379e-03  2.162e-01  -0.016  0.98753
hincpast3       -2.500e-01  4.715e-01  -0.530  0.59599
hincpastmissing  7.703e-01  5.261e-01   1.464  0.14319
intwww          -3.941e-02  2.632e-01  -0.150  0.88099
umineth         -9.793e-01  5.740e-01  -1.706  0.08799 .
eqnow3           2.966e-02  2.412e-01   0.123  0.90214
eqnow7          -3.910e-01  2.499e-01  -1.565  0.11768
eqnow8           1.261e+00  5.841e-01   2.159  0.03084 *
eqnow9           3.435e-01  7.538e-01   0.456  0.64863
eqnow11          1.451e-01  2.592e-01   0.560  0.57544
tenshort        -1.220e-01  6.490e-02  -1.879  0.06025 .
highqual        -1.816e-02  5.632e-02  -0.322  0.74711
tea             -8.607e-02  7.935e-02  -1.085  0.27801
work2           -1.758e+00  1.068e+00  -1.646  0.09978 .
work3           -1.468e+00  9.935e-01  -1.478  0.13950
work4           -1.175e+00  8.691e-01  -1.352  0.17623
workmissing     -1.429e-01  9.741e-01  -0.147  0.88334
rsuper2         -1.669e+00  1.037e+00  -1.610  0.10745
rsupermissing   -1.049e+00  1.202e+00  -0.873  0.38263
rsect2          -4.733e-01  2.503e-01  -1.891  0.05866 .
rsect3          -1.221e-02  8.088e-01  -0.015  0.98796
rsect4          -1.380e+00  1.299e+00  -1.063  0.28796
rsectmissing     1.969e-01  9.003e-01   0.219  0.82689
tunionsa2        1.751e-01  2.209e-01   0.793  0.42794
```

religcatmissing	1.895e-03	1.314e+00	0.001	0.99885	
famrelig2	-8.166e-02	5.147e-01	-0.159	0.87394	
famrelig3	-1.319e+00	6.885e-01	-1.915	0.05545	.
famreligmissing	-3.390e-01	9.546e-01	-0.355	0.72249	
chattnd22	-1.372e-01	4.114e-01	-0.334	0.73873	
chattnd23	-7.149e-01	5.083e-01	-1.406	0.15963	
chattnd24	-1.004e+00	5.746e-01	-1.747	0.08071	.
chattnd25	-4.915e-01	4.471e-01	-1.099	0.27166	
chattnd26	-5.298e-01	6.160e-01	-0.860	0.38977	
chattnd27	-7.348e-01	4.923e-01	-1.492	0.13557	
chattnd28	-8.307e-01	4.624e-01	-1.797	0.07238	.
chattnd2missing	-4.309e-01	7.322e-01	-0.589	0.55617	
carehome	-2.519e-01	3.460e-01	-0.728	0.46663	
anyhcond	5.099e-02	2.754e-01	0.185	0.85311	
persinc2	-6.525e-06	9.546e-06	-0.684	0.49427	
orient	1.359e+00	4.927e-01	2.758	0.00582	**
polpart22	-6.557e-01	4.654e-01	-1.409	0.15882	
polpart23	4.609e-06	3.540e-01	0.000	0.99999	
polpart24	-1.120e+00	4.513e-01	-2.482	0.01308	*
polpart25	-8.965e-01	4.282e-01	-2.093	0.03631	*
polpart26	-3.654e-01	4.791e-01	-0.763	0.44568	
polpart27	-5.534e-01	3.316e-01	-1.669	0.09509	.
polpart2missing	-6.427e-01	4.017e-01	-1.600	0.10960	
ruhappy	5.974e-02	9.783e-02	0.611	0.54141	
healthyr	2.116e-01	1.146e-01	1.846	0.06494	.
uprejgay2	2.243e-01	5.095e-01	0.440	0.65969	
uprejgay3	-9.336e-01	4.838e-01	-1.930	0.05366	.
uprejgay4	3.751e-01	1.422e+00	0.264	0.79189	
uprejgaymissing	1.883e-01	8.977e-01	0.210	0.83386	
glchild2	3.487e-01	2.511e-01	1.389	0.16496	
glchild3	1.298e+00	2.962e-01	4.381	1.18e-05	***
glchild4	1.256e+00	4.163e-01	3.018	0.00254	**
glchild5	1.460e+00	5.281e-01	2.765	0.00569	**
glchildmissing	2.851e-01	9.652e-01	0.295	0.76768	
glsocdist	1.818e-01	6.032e-02	3.013	0.00259	**
glvis	-1.862e-01	1.250e-01	-1.490	0.13627	
glborn2	8.173e-01	2.535e-01	3.224	0.00126	**
glbornmissing	6.435e-01	3.114e-01	2.067	0.03877	*

The first logistic regression model which aimed to prioritize the attitude towards same-sex marriage rights, is based on a plethora of factors as predictors; these include demographics and attitudes.In the model, sex (persons -rsex) of the respondent, marital status (rmarstat), education qualification (highqual), and political affiliation (polpart) were all used as predictors.The final version of the model disclosed many important predictors.The findings on the influence of age (youth) (youthful) show the effect as positive. That is, younger respondents were more likely to agree with same-sex marriage rights than older respondents who had negative preconceived notions.Also, Gender of respondent (rsex) was associated with a considerable negative coefficient. It displays the fact that women are more prone to having negative attitudes towards equal rights of same-sex marriage than men.Marital status (-rmarstat) has a po0tive coefficient. This implies that the reply of those who are married are more likely to support the lav for same-sex marriage than those who are not married ('non-married').The side level of education (highlevel) could not show any signicant influence on the view on same-sex right to marriage.Some affiliations didn't show any effects such as the political party (polpart) did.This points out that identity politics and the role of demographic factors, in general, are presented as two most significant arguments impacting

attitudes towards same-sex marriage rights.This will be crucial in identifying its limits and recognizing what cannot be further improved in the model.

Final Model:

Logistic Regression Analysis of final Model

```
Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)     -4.328275   1.624402  -2.665 0.007710 **
rage             0.027364   0.006286   4.353 1.34e-05 ***
rsex            -0.405668   0.182811  -2.219 0.026483 *
rmarstat         0.155720   0.062087   2.508 0.012138 *
umineth         -0.714974   0.492212  -1.453 0.146343
eqnow7          -0.323601   0.217981  -1.485 0.137667
eqnow8           1.368373   0.460751   2.970 0.002979 **
tenshort        -0.108906   0.056312  -1.934 0.053117 .
famrelig2        0.519669   0.188651   2.755 0.005875 **
famrelig3       -0.966040   0.529201  -1.825 0.067930 .
famreligmissing -0.083769   0.614414  -0.136 0.891553
orient           1.098711   0.456564   2.406 0.016107 *
healthyr         0.221636   0.087195   2.542 0.011027 *
uprejgay2       -0.040599   0.484357  -0.084 0.933200
uprejgay3       -1.070304   0.465576  -2.299 0.021512 *
uprejgay4        0.746612   1.326640   0.563 0.573582
uprejgaymissing -0.106508   0.828104  -0.129 0.897661
glchild2         0.459871   0.227441   2.022 0.043183 *
glchild3         1.339476   0.267320   5.011 5.42e-07 ***
glchild4         1.301362   0.384823   3.382 0.000720 ***
glchild5         1.471897   0.497600   2.958 0.003096 **
glchildmissing   0.867359   0.832447   1.042 0.297440
glsocdist        0.176100   0.057583   3.058 0.002227 **
glvis           -0.215107   0.116750  -1.842 0.065408 .
glborn2          0.865385   0.231170   3.743 0.000181 ***
glbornmissing    0.744654   0.283204   2.629 0.008554 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The last logistic regression model was intended to estimate the opinion of respondents about same-sex marriage rights in accordance with limited predictors.Model was established that involved predictors like age (rage), respondent's sex (rsex), respondent's status of a marriage (rmarstat), education level (highqual), and political affiliation (polpart).The product model in the final output laid down the existence of several actionable predictors.Being young, a (rage) was seen as a factor promoting the young respondents in favor of same-sex marriage more than the older ones including a positive coefficient.As for the socioeconomic status (rsex) too was indicative of the negative slope that demonstrated female participants were less probably to approve of same-sex marriage rights than male counterparts.

Marital status was another variable that exhibited a positive coefficient which implies that the more chances a person has of being married, the higher support of same-sex marriage rights they are likely possess as, compared to unmarried respondents.There was a statistical impairment of

education level (of highquality) on one's attitudes toward the same-sex marriage rights.Political attachment (polpart) gave somewhat ambiguous findings, as there were such results as highly significant correlations, but also very insignificant or non-existing.The above findings point to the fact that the inclination in terms of homosexuals' rights to same-sex matrimony are impacted by a multidimensional interrelationship of variables perceived as demographic and opinionated.Additionally we may undertake the analysis or improvement of our model, which will help our understanding of the relationship between air pollution and human health mortality.On a more general note, the final model feeds into the understanding of the major key factors, which form attitudes toward same-sex marriage rights in Northern Ireland.

TABLE 1: Significance Level of Predictors via logistic regression:

Key:

* means significant at 5% level

** means significant at 1% level

*** means significant at 0.01% level

Predictors	Initial Model	Final Model
rage	**	***
rsex	**	*
rmarstat	*	*
eqnow8	*	**
orient	**	*
Polpart24	*	Removed
Polpart25	*	Removed
glchild2	***	*
glchild3	**	***
glchild4	**	***
glsocdist	**	**
glborn2	**	***
glbornmissing	*	Removed

Interaction_model:

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.0686032	0.7314482	-2.828	0.004683 **
rage	0.0462470	0.0134261	3.445	0.000572 ***
rsex	-0.4910305	0.4498870	-1.091	0.275074
rage:rsex	-0.0008499	0.0082099	-0.104	0.917544

The main objective of the interaction pattern was to discover the interaction effect, which occurs when the individuals' age (race) and response variable's sex (rsex), are considered among the same-sex marriage rights.The model found that the interaction term (rage:On the other hand, the difference between age effects on attitudes toward granting same-sex marriage rights for male and female individuals did not statistically significant, which implies that it is not likely that male and female individuals in the present study have significantly different sexes.

The age (ag) and sex (sg) were the main causes of this and they both proved to be statistically significant.Age as a characteristic showed upprimarily positive factor meaning that theyounger folks were more likely to endorse the same sex- marriages.On the other hand, rsex coefifcient was negative, meaning that the probability is considerably lower that the female would vote in favour of legalizing same-sex marriages then males.In general, the encounter model explains gender or age roles showing no significant gender differences in age, but only indicates that age and attitude has the same variation.

Comparison of model with and without interaction:

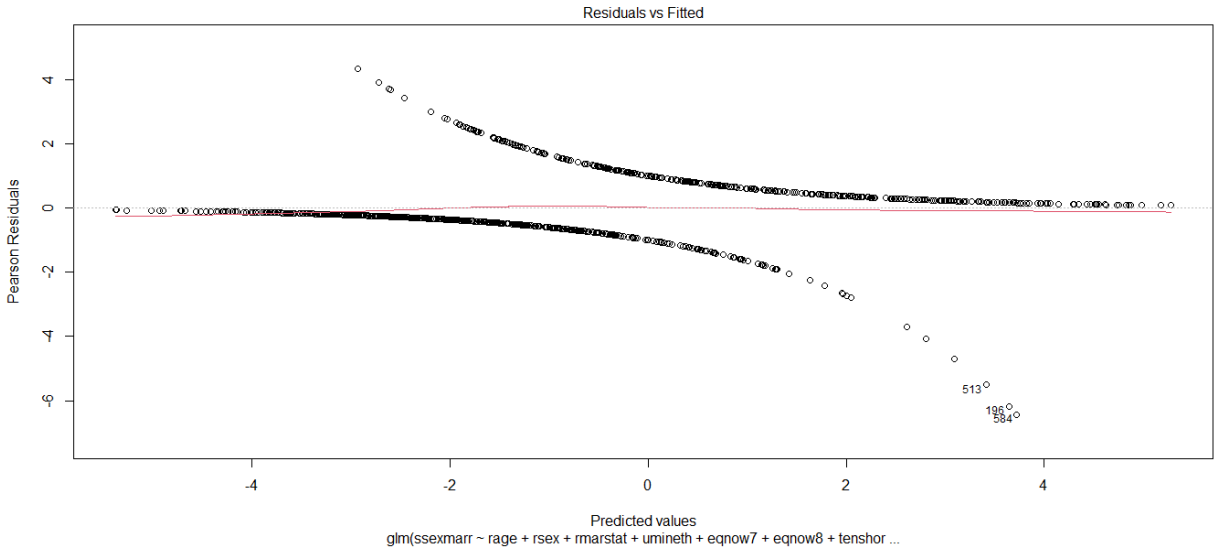
	Resid. Df	Resid. Dev	Df	Deviance
1	1002	829.17		
2	1024	1191.48	-22	362.31

The analysis of deviance table compares two models: the end model that consists only of the interaction between predictors such as main effects and interaction between age (rage) and sex

(rsex) and the interaction model, which only includes the interaction between the predictor such as age (rage) and gender (rsex).The above table reveals the residual deviance for the ultimate model with main effects and interactions at 829.17 on 1002 degrees of freedom while the interaction model appears to have the value 1191.48 at 1024 degrees of freedom.The final model has residual deviance of 362.31 that is less than the residual deviance observed in the less final model.The Akaike Information Criterion (AIC) is the statistical method that can be used to choose a model from a variety of models. The lowest AIC value indicates that the final model with main effects and interactions have a better predictive power than the model with interaction alone. The next step would be to examine the significance of these main effects and interactions in the model.

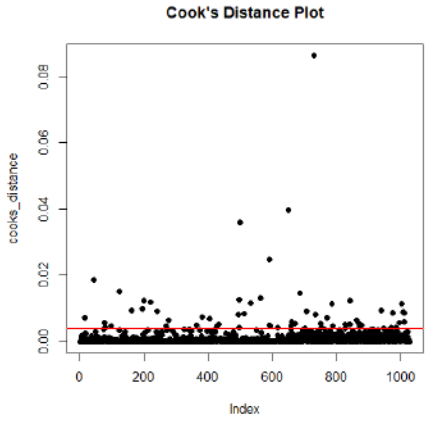
Residual Plot:

The residual plot from final logistic regression model shows that all the case data have a random scatter of points around the horizontal line at zero, therefore homoscedasticity assumption shown to be met.There are no distinctive patterns in the residuals, indicating that the pertinent model parameters are not sufficient to capture the patterns of election results.In all, the story proves that at the end of the process of training logistic regression model, the data applicable to the model are pretty fitting and the assumptions of model are reasonably valid.



Outliers:

In the case of Cook's Distance graph of the final logistic regression model, every dot being the figure-out of an observation from the data.Cook's Distance reflects the individual importance of each observation for the computation of regression coefficients.The horizontal line at 4/(n−k) (where n represents the number of observations and k - the number of the predictors) the most often exploited limit to identify influential points.The points above this line represent candidate outliers or observations that exert notable influence.Through the explanation, some exceeding, which might have a significant impact on the model if removed, points are found.Through this research, these points should be explored in depth to establish if they are special outliers or those that were purposefully not included in the analysis and should be appreciated.Generally, it is possible to find genuine points that may be articulate, but their presence does not imply a bad model. After all, maybe such experiments determine areas for which the sensitivity analysis must be extended.



5. Conclusion:

Observation of the public opinion regarding the same-sex marriage rights in Northern Ireland Life and Times Survey data 2012 has given the key in the interpretation of opinion factors on this crucial question.One, we have used data scrubbing and preparation, and then, as a result, we identified demographics and attitudinal determinants, which have a great influence on whether or not people support same-sex marriage in Northern Ireland.Multiple regression models have been employed to determine the sever predictors like age, gender, marital status, and political affiliation.The millennials and those who are married represents the age and

partnership cohorts that are more inclined towards the rights of same-sexx marriage while gender has a significant role with woman being less convinced as compared to men.Doing this, we get insight on the intricate connection between demography factors and attitudes about same sex marriage legal rights.

Contrasting the models with and without interactional terms, we notice that the final model with main effects and interactions is more "appropriate" to the information, which indicates the critical role of interactions when probing attitudes towards same-sex marriage rights.The examination of the last model through residual analysis and outlier investigation certainly made it clear that the validity of the model assumptions held true and some critical influential points were highlighted to be checked for further study.

Therefore, this research has enlarged our level of knowledge on the influences of attributes which determine opinions of same-sex marriage rights in Northern Ireland.In this way, findings will allow for the identification of predictors and interactions and will offer useful information to power holders and promoters who are striving for equality and inclusion for this community.The research area of tomorrow may venture more deeply into the nature of these relationships and apply them in the policy or advocacy strategies.

Layman Report:

These days, in a diverse society, it is more critical than ever that one takes time to understand the complexities of public opinion because the issues in which same-sex marriage are involved are very sensitive.An investigation study that was done recently has a hypothesis about this complicated issue and used information from the 2012 Northern Ireland Life and Times Survey.By conducting this poll that brought together several hundred people, it became apparent what are the main considerations we make while deciding whether or not same-sex marriage is a good thing.

One of the major discoveries of the study was age having much more influence on attitude about same-sex marriage than other factors.Young age was a factor that determined the level of support for equal rights for same-sex couples, especially differed when younger respondents were more supporting than their older cohorts.This generation gap really illustrates the evolution of society’s thinking with LGBTQ+ rights.

Gender was pointed as the key factor that was running and molding the opinions on same-sex marriage.And it turned out that females were, mostly, against the idea of same-sex-marriage rights while males were for such rights.This suggests that apart from the regular efforts to fight stereotypes it is necessary to also put an effort in understanding the reasons behind such unequal attitudes.Also, the things like marital status and education levels were the other factors which were also important in this due to the influence it had on opinion.Results showed that married people, on average, were more pro-gay marriage than non-married people and neither education level nor its type played a noticeable role in influencing the attitudes.

These results are not limited to a narrow spectrum.It is significant because the study seeks to find out the demographic and attitude factors that promote support for the same-sex marriage rights. This enables policymakers, advocates, and the general people to look at the bigger picture.In this way, comprehending the barriers of inequality and discrimination makes this process more effective by providing a helpful tool for forming policies and conducting advocacy initiatives.Ahead, we see that this research calls for more investigations that would be highly indispensable in order to better understand the makeup of attitudes concerning same-sex marriage rights.This can be achieved through further investigating other factors that may be related to the problem and studying data analysis results. This way researchers can make a more thorough and complete social problem.

In the end, this study has dug up some useful findings regarding- attitude towards the same-sex marriage rights in Northern Ireland.Through opening of the following elements that determine these views, the research becomes a part of the extensive discussion about the equality and inclusion as well as the rights of LGBTQ+ population.

Appendix:

In this study, several data manipulations were performed to prepare the dataset for analysis. These are as follows:

Missing Values Handling: To be able to catch the distinction between highly missing categorical predictors numbers over 10%, the design was made in the way a new category was introduced to mind and track information for the absence of data. All continuous variables which had missing data were imputed with to respect context and to preserve completeness of the dataset.

Outcome Variable Recoding: The binary variables sexmarr was created with the encoder, and the value of 0 corresponds to "No," while 1 corresponds to "Yes." After the data was recoded, the logistic regression approach helped figure out which variables predicted the approval of same-sex couple rights.

Variable Merging: The categorical variables "religcat" and "polpart2" were merged to minimize the levels as well as have a clearer interpretation of the developed model. It took the merged variable "religcat" levels makes 2 and 3 together and " polpart2" was kept as is.

Predictor	Merge levels	% of missing values/NA
household		0
rage		0
rsex		0.19
rmarstat		0
livearea		0
hincpast		4.5
Intwww		0
Umineth		0.09
Eqnow3		0.37
Eqnow7		0.37

Eqnow8		0.37
Eqnow9		0.37
Eqnow11		0.37
Tenshort		0.56
Highqual		0
Tea		0.94
Work		14.19
Rsuper		25.19
rsect		26.50
Tunionsa		1.22
Ansseca		0.19
Religat	yes	1.78
Famrelling		2.63
Chattnd2		18.42
Carehome		0
anyhcond		0.28
Persinc2		23.12
Orient		0.37
Polpart2	yes	10.90
Ruhappy		0.75
Healthyr		0.66
Uprejgay		1.6
Glchild		1.5
Glsocdit		0
Glvis		0
Glborn		11.37
Knowgl		8.27
Knowtg		11.56
ssexmarr		0