Shyama Prasad Mukherji College for Women
University of Delhi

Crime Against Women in India

Prepared By:

Isheta Tangri (21/1305)
Nishtha Narula (21/1323)
Nishtha Roy (21/1330)
Shruti Goyal (21/1370)

Introductory Econometrics

Ms. Gita Golani

March 31st, 2023

Shyama Prasad Mukherji College for Women
University of Delhi

Introductory Econometrics
Ms. Gita Golani
March 31st, 2023

Crime Against Women in India

Isheta Tangri (21/1305)
Nishtha Narula (21/1323)
Nishtha Roy (21/1330)
Shruti Goyal (21/1370)

INDEX

ABSTRACT

In India, crime against women is a severe problem that has long been present. Women still experience violence, discrimination, and harassment on a daily basis, including domestic abuse, sexual assault, acid assaults, honour killings, and human trafficking, in spite of several laws and programmes. This paper provides an overview of the extent and nature of crime against women in India, the socio-cultural factors that contribute to it, and the legal framework for addressing it. It also covers the difficulties with execution and the necessity of a comprehensive strategy that includes women's empowerment through education and awareness, as well as efficient law enforcement and accountability mechanisms. In the end, it makes the case that combating crime against women requires

INTRODUCTION

Crimes against women in India have been a significant issue for several years. Despite various measures taken by the government, the problem continues to persist. Women in India face various forms of violence and discrimination, including sexual harassment, domestic violence, rape, honour killings, female foeticide, trafficking, and others.

The patriarchal and conservative societal norms, coupled with inadequate law enforcement and judicial system, have contributed to the prevalence of such crimes. Women in India continue to face discrimination and are not given equal opportunities and rights as men. This has led to a power imbalance, which often results in violence and exploitation of women.

One of the most common forms of violence against women in India is sexual harassment. According to a report by the National Crime Records Bureau (NCRB), in 2019, there were 88,974 reported cases of sexual harassment in India. However, the actual number is likely to be much higher, as many cases go unreported due to fear of retaliation or shame.

Rape is another prevalent form of violence against women in India. The NCRB reported 32,033 cases of rape in 2019, with an average of 88 cases reported every day. The issue gained international attention in 2012 when a 23-year-old woman was gang-raped on a bus in Delhi, leading to widespread protests and demands for stricter laws.

Domestic violence is another major issue that women in India face. According to the National Family Health Survey (NFHS) 2019-20, one in three women in India has experienced physical violence since the age of 15. Women who are victims of domestic violence often suffer in silence, as they fear retaliation from their abusers or social stigma.

Female foeticide, the practice of selectively aborting female fetuses due to a preference for male children, is another form of violence against women in India. The practice is illegal, but it continues to be prevalent in many parts of the country. According to the NCRB, there were 2,848 reported cases of female foeticide in 2019.

Honour killings, a practice where family members kill women who are perceived to have brought shame to the family, is another form of violence against women in India. The victims are often accused of marrying outside their caste or religion or engaging in premarital sex. According to the NCRB, there were 70 reported cases of honour killings in 2019.

Trafficking of women and children is another significant issue in India. Women and girls are often trafficked for sexual exploitation or forced labor. According to the Global Slavery Index, India has the highest number of people living in modern slavery, including forced labor and forced marriage.

The Indian government has taken several measures to address the issue of crimes against women. In 2013, the government enacted the Criminal Law (Amendment) Act, which increased the punishment for rape and introduced new offenses, such as acid attacks and stalking. The government also set up special courts for fast-track trials of sexual offences.

However, these measures have not been enough to curb the problem. There are several factors that contribute to the prevalence of crimes against women in India, including social norms, lack of education, poverty, and weak law enforcement.

It requires a concerted effort from the government, civil society, and the people of India to work together to create a safe and equitable society for women. This includes providing education and awareness programs, strengthening law enforcement and judicial system, and creating a culture that promotes gender equality and respect for women.

LITERATURE REVIEW

Crime against women in India is a serious issue that has garnered significant attention in recent years. The literature on this topic is vast, and several studies have been conducted to understand the various aspects of this issue. Here is a brief literature review of crime against women in India.
Violence against women: A statistical overview, challenges and gaps in data collection and methodology in the EU. This report by the European Union Agency for Fundamental Rights (FRA) provides a statistical overview of violence against women in the EU. It highlights the challenges and gaps in data collection and methodology in the EU and provides recommendations to improve data collection and methodology.
"Understanding the nature and dynamics of violence against women in India: An overview". This article by Jyoti Grewal provides an overview of violence against women in India, including the various forms of violence, the socio-economic and cultural factors that contribute to violence, and the legal framework for addressing violence against women.
"Crimes against women in India: Analysis of official statistics". This article by NCRB (National Crime Records Bureau) provides an analysis of crimes against women in India based on official statistics. It covers various aspects of crimes against women, including the number and types of crimes, age and marital status of victims, and the location of crimes.
"Rape in India: A comprehensive analysis". This article by Dhananjay Kumar and G. Srinivasan provides a comprehensive analysis of rape in India. It covers various aspects of rape, including the number of reported cases, the age and marital status of victims, and the location and timing of the crime.
"Gender-based violence in India: A review of the literature". This article by Sabrina Juran provides a comprehensive review of the literature on gender-based violence in India. It covers various aspects of gender-based violence, including domestic violence, sexual harassment, and rape, and provides recommendations for addressing the issue.
"Sexual harassment in India: A review of the literature". This article by Saumya Pandey and Manju Mehta provides a review of the literature on sexual harassment in India. It covers various aspects of sexual harassment, including the prevalence and nature of the problem, the legal framework for addressing sexual harassment, and the challenges in implementing the law.
Overall, the literature on crime against women in India highlights the severity of the issue and the need for a comprehensive approach to address the problem. The studies provide insights into the various forms of violence against women, the socio-economic and cultural factors that contribute to violence, and the legal framework for addressing the issue.

ECONOMETRIC METHODOLOGY

- **MULTIPLE LINEAR REGRESSION**

WE HAVE USED MLR FOR THIS STATISTICS. A regression model with more than one explanatory variable is known as a **multiple regression model,** multiple because *multiple influences* (i.e., variables) may affect the dependent variable. the simplest of the multiple regression models, namely, the three-variable model in which the behaviour of the dependent variable *Y* is examined in relation to two explanatory variables, *X*2 and *X*3.

Yt =B1 +B2X2t +B3X3t +ut
where *Y* = the dependent variable
*X*2 and *X*3 = the explanatory variables
u = the stochastic disturbance term
t = the *t* th observation

- **Ordinary Least Squares (OLS) method**

It is widely used to estimate the parameter of a linear regression model. OLS estimators minimise the sum of the squared errors (a difference between observed values and predicted values). While OLS is computationally feasible and can be easily used while doing any econometrics test, it is important to know the underlying assumptions of OLS regression. This is because a lack of knowledge of OLS assumptions would result in its misuse and give incorrect results for the econometrics test completed. The importance of OLS assumptions cannot be overemphasised. Following are the assumptions of OLS Method:

1) The regression model is linear in the parameters and it is correctly specified.

2) X^2 and X^3 are uncorrelated with the disturbance term u therefore are non-stochastic in nature.

3) The error term u has a zero-mean value.

4) Homoscedasticity is constant.

5) No autocorrelation exists between the error terms.

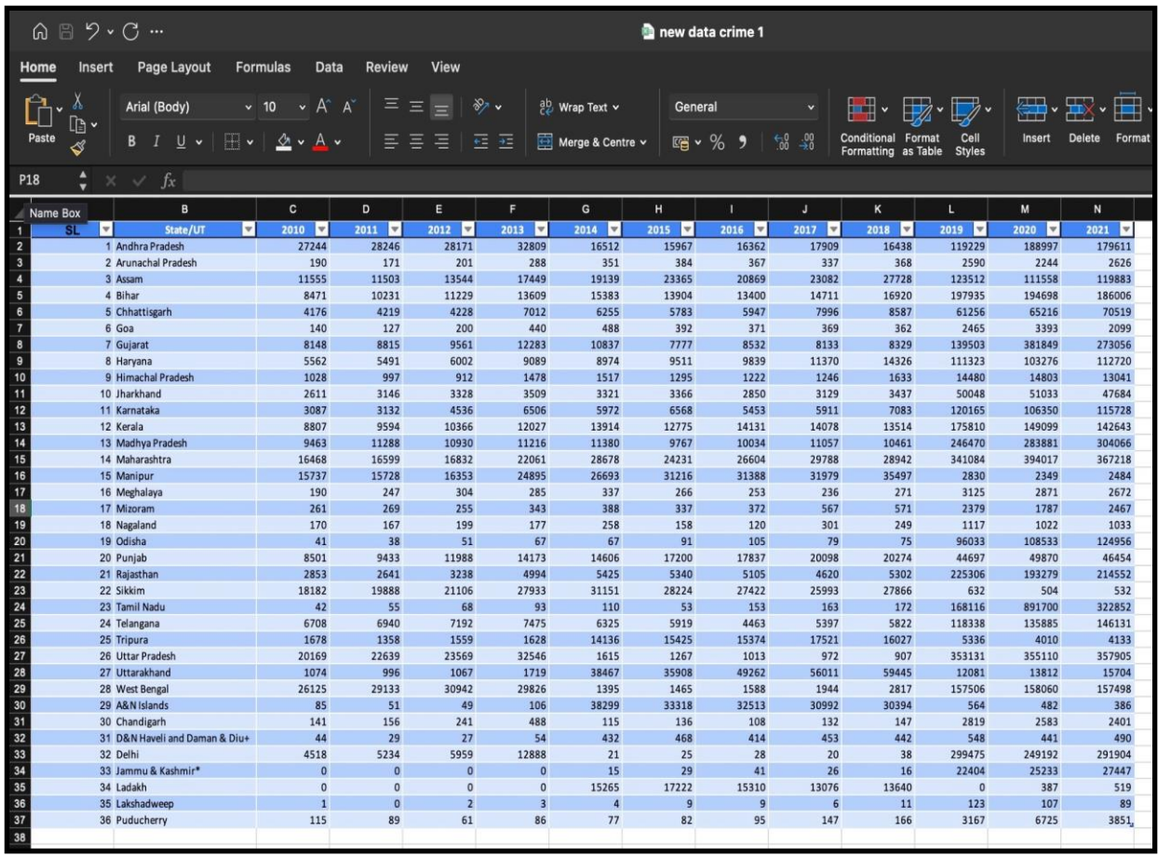
6) The error term u follows the normal distribution with mean zero and (homoscedastic) variance 2. $u_i \sim$ N(0, α^2)

7) There is no multicollinearity or we can say there is no exact linear relationship between the explanatory variables X^2 and X^3.

Regression

Regression model: Yt =B1 +B2X2t +B3X3t +ut

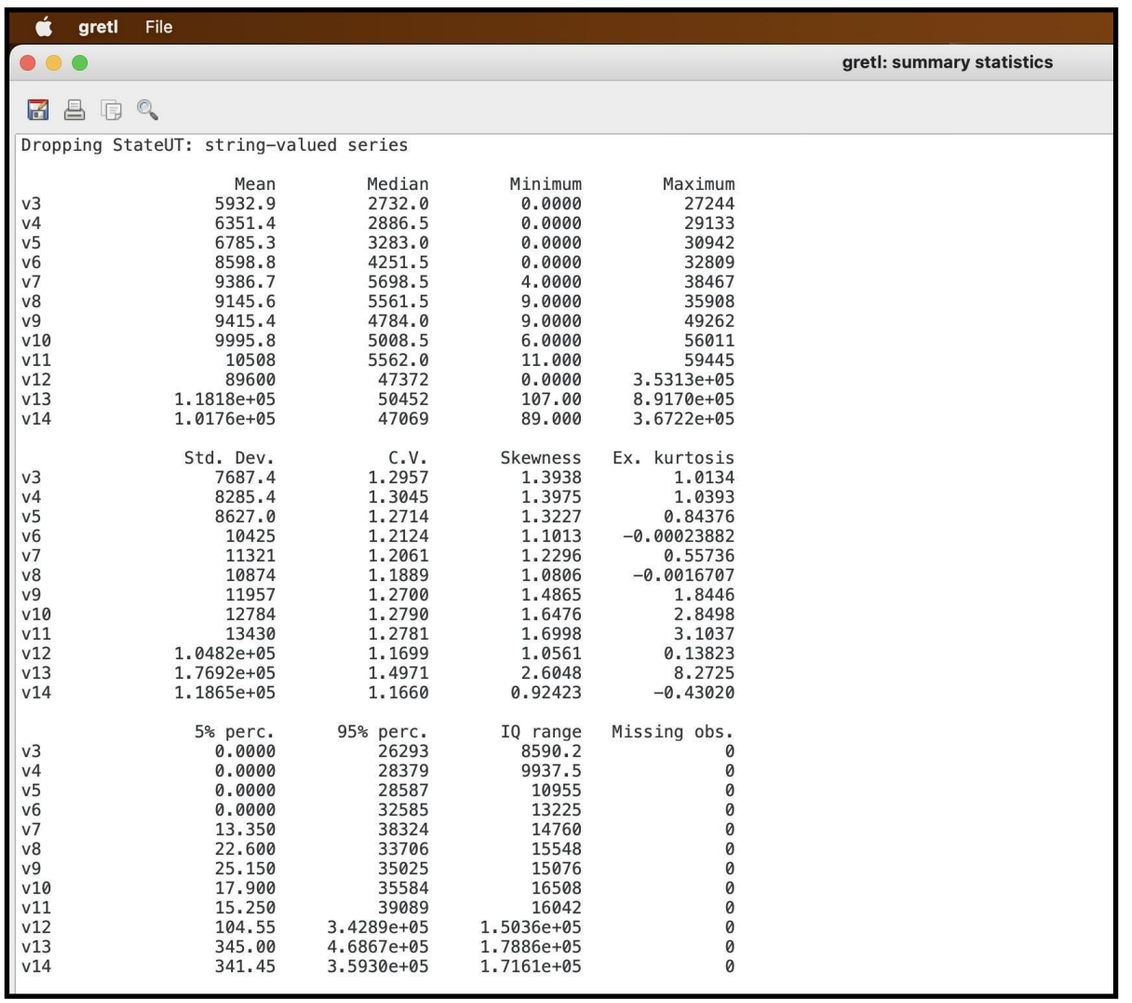
MAIN DATA



	SL	State/UT	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
2	1	Andhra Pradesh	27244	28246	28171	32809	16512	15967	16362	17909	16438	119229	188997	179611
3	2	Arunachal Pradesh	190	171	201	288	351	384	367	337	368	2590	2244	2626
4	3	Assam	11555	11503	13544	17449	19139	23365	20869	23082	27728	123512	111558	119883
5	4	Bihar	8471	10231	11229	13609	15383	13904	13400	14711	16920	197935	194698	186006
6	5	Chhattisgarh	4176	4219	4228	7012	6255	5783	5947	7996	8587	61256	65216	70519
7	6	Goa	140	127	200	440	488	392	371	369	362	2465	3393	2099
8	7	Gujarat	8148	8815	9561	12283	10837	7777	8532	8133	8329	139503	381849	273056
9	8	Haryana	5562	5491	6002	9089	8974	9511	9839	11370	14326	111323	103276	112720
10	9	Himachal Pradesh	1028	997	912	1478	1517	1295	1222	1246	1633	14480	14803	13041
11	10	Jharkhand	2611	3146	3328	3509	3321	3366	2850	3129	3437	50048	51033	47684
12	11	Karnataka	3087	3132	4536	6506	5972	6568	5453	5911	7083	120165	106350	115728
13	12	Kerala	8807	9594	10366	12027	13914	12775	14131	14078	13514	175810	149099	142643
14	13	Madhya Pradesh	9463	11288	10930	11216	11380	9767	10034	11057	10461	246470	283881	304066
15	14	Maharashtra	16468	16599	16832	22061	28678	24231	26604	29788	28942	341084	394017	367218
16	15	Manipur	15737	15728	16353	24895	26693	31216	31388	31979	35497	2830	2349	2484
17	16	Meghalaya	190	247	304	285	337	266	253	236	271	3125	2871	2672
18	17	Mizoram	261	269	255	343	388	337	372	567	571	2379	1787	2467
19	18	Nagaland	170	167	199	177	258	158	120	301	249	1117	1022	1033
20	19	Odisha	41	38	51	67	67	91	105	79	75	96033	108533	124956
21	20	Punjab	8501	9433	11988	14173	14606	17200	17837	20098	20274	44697	49870	46454
22	21	Rajasthan	2853	2641	3238	4994	5425	5340	5105	4620	5302	225306	193279	214552
23	22	Sikkim	18182	19888	21106	27933	31151	28224	27422	25993	27866	632	504	532
24	23	Tamil Nadu	42	55	68	93	110	53	153	163	172	168116	891700	322852
25	24	Telangana	6708	6940	7192	7475	6325	5919	4463	5397	5822	118338	135885	146131
26	25	Tripura	1678	1358	1559	1628	14136	15425	15374	17521	16027	5336	4010	4133
27	26	Uttar Pradesh	20169	22639	23569	32546	1615	1267	1013	972	907	353131	355110	357905
28	27	Uttarakhand	1074	996	1067	1719	38467	35908	49262	56011	59445	12081	13812	15704
29	28	West Bengal	26125	29133	30942	29826	1395	1465	1588	1944	2817	157506	158060	157498
30	29	A&N Islands	85	51	49	106	38299	33318	32513	30992	30394	564	482	386
31	30	Chandigarh	141	156	241	488	115	136	108	132	147	2819	2583	2401
32	31	D&N Haveli and Daman & Diu+	44	29	27	54	432	468	414	453	442	548	441	490
33	32	Delhi	4518	5234	5959	12888	21	25	28	20	38	299475	249192	291904
34	33	Jammu & Kashmir*	0	0	0	0	15	29	41	26	16	22404	25233	27447
35	34	Ladakh	0	0	0	0	15265	17222	15310	13076	13640	0	387	519
36	35	Lakshadweep	1	0	2	3	4	9	9	6	11	123	107	89
37	36	Puducherry	115	89	61	86	77	82	95	147	166	3167	6725	3851

SOURCE : NCRB ……IPC CRIMES (STATE/UT WISE)-2010-21

Summary Statistics:



Dropping StateUT: string-valued series

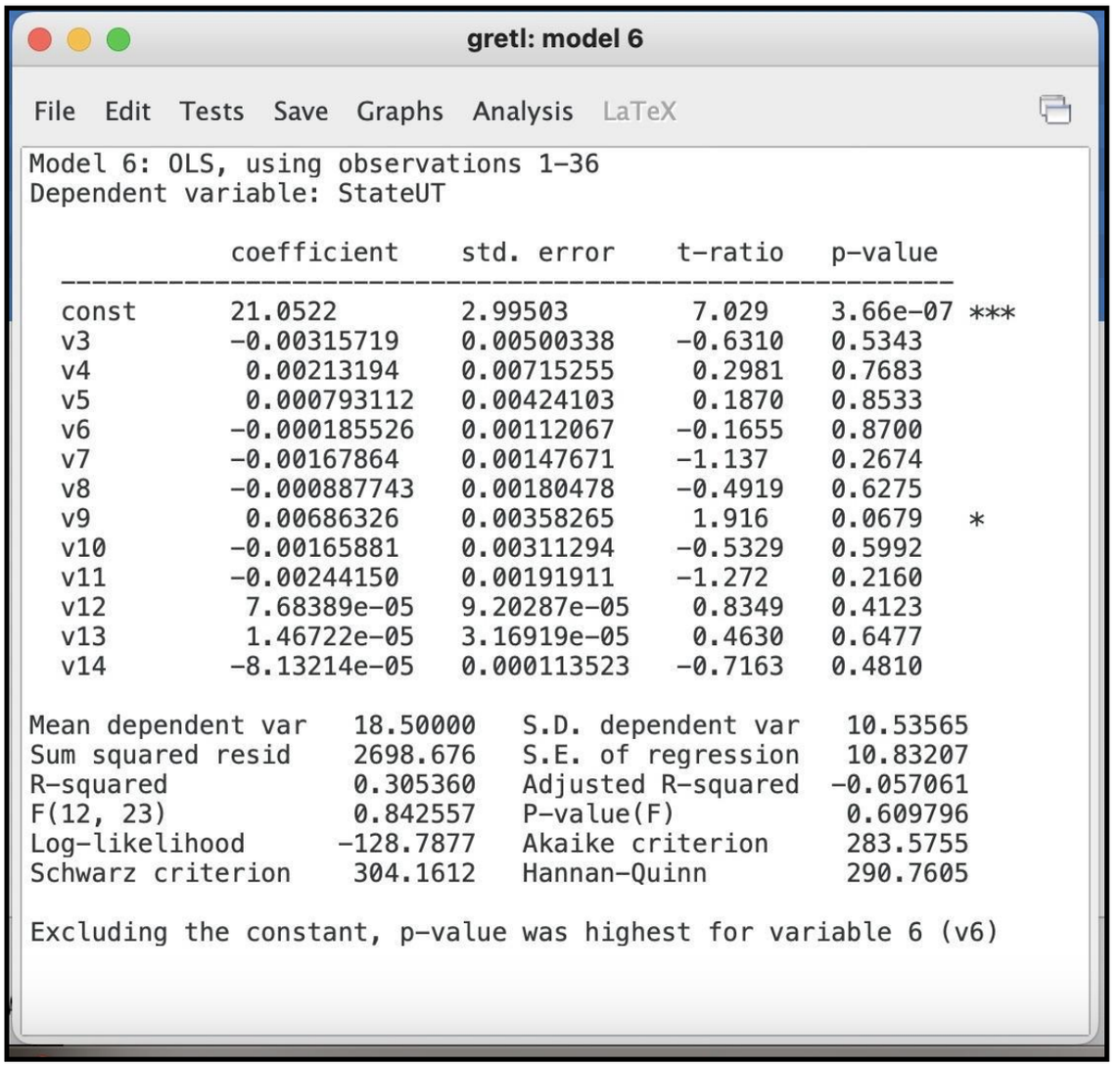
	Mean	Median	Minimum	Maximum
v3	5932.9	2732.0	0.0000	27244
v4	6351.4	2886.5	0.0000	29133
v5	6785.3	3283.0	0.0000	30942
v6	8598.8	4251.5	0.0000	32809
v7	9386.7	5698.5	4.0000	38467
v8	9145.6	5561.5	9.0000	35908
v9	9415.4	4784.0	9.0000	49262
v10	9995.8	5008.5	6.0000	56011
v11	10508	5562.0	11.000	59445
v12	89600	47372	0.0000	3.5313e+05
v13	1.1818e+05	50452	107.00	8.9170e+05
v14	1.0176e+05	47069	89.000	3.6722e+05

	Std. Dev.	C.V.	Skewness	Ex. kurtosis
v3	7687.4	1.2957	1.3938	1.0134
v4	8285.4	1.3045	1.3975	1.0393
v5	8627.0	1.2714	1.3227	0.84376
v6	10425	1.2124	1.1013	-0.00023882
v7	11321	1.2061	1.2296	0.55736
v8	10874	1.1889	1.0806	-0.0016707
v9	11957	1.2700	1.4865	1.8446
v10	12784	1.2790	1.6476	2.8498
v11	13430	1.2781	1.6998	3.1037
v12	1.0482e+05	1.1699	1.0561	0.13823
v13	1.7692e+05	1.4971	2.6048	8.2725
v14	1.1865e+05	1.1660	0.92423	-0.43020

	5% perc.	95% perc.	IQ range	Missing obs.
v3	0.0000	26293	8590.2	0
v4	0.0000	28379	9937.5	0
v5	0.0000	28587	10955	0
v6	0.0000	32585	13225	0
v7	13.350	38324	14760	0
v8	22.600	33706	15548	0
v9	25.150	35025	15076	0
v10	17.900	35584	16508	0
v11	15.250	39089	16042	0
v12	104.55	3.4289e+05	1.5036e+05	0
v13	345.00	4.6867e+05	1.7886e+05	0
v14	341.45	3.5930e+05	1.7161e+05	0

The summary statistics for the taken variables shows various results:

We then run the regression by using the method of OLS estimation in Gretl and the following results were obtained:



gretl: model 6

File Edit Tests Save Graphs Analysis LaTeX

Model 6: OLS, using observations 1-36
Dependent variable: StateUT

	coefficient	std. error	t-ratio	p-value	
const	21.0522	2.99503	7.029	3.66e-07	***
v3	-0.00315719	0.00500338	-0.6310	0.5343	
v4	0.00213194	0.00715255	0.2981	0.7683	
v5	0.000793112	0.00424103	0.1870	0.8533	
v6	-0.000185526	0.00112067	-0.1655	0.8700	
v7	-0.00167864	0.00147671	-1.137	0.2674	
v8	-0.000887743	0.00180478	-0.4919	0.6275	
v9	0.00686326	0.00358265	1.916	0.0679	*
v10	-0.00165881	0.00311294	-0.5329	0.5992	
v11	-0.00244150	0.00191911	-1.272	0.2160	
v12	7.68389e-05	9.20287e-05	0.8349	0.4123	
v13	1.46722e-05	3.16919e-05	0.4630	0.6477	
v14	-8.13214e-05	0.000113523	-0.7163	0.4810	

Mean dependent var	18.50000	S.D. dependent var	10.53565
Sum squared resid	2698.676	S.E. of regression	10.83207
R-squared	0.305360	Adjusted R-squared	-0.057061
F(12, 23)	0.842557	P-value(F)	0.609796
Log-likelihood	-128.7877	Akaike criterion	283.5755
Schwarz criterion	304.1612	Hannan-Quinn	290.7605

Excluding the constant, p-value was highest for variable 6 (v6)

From the result obtained we obtain the regression equation as follows:

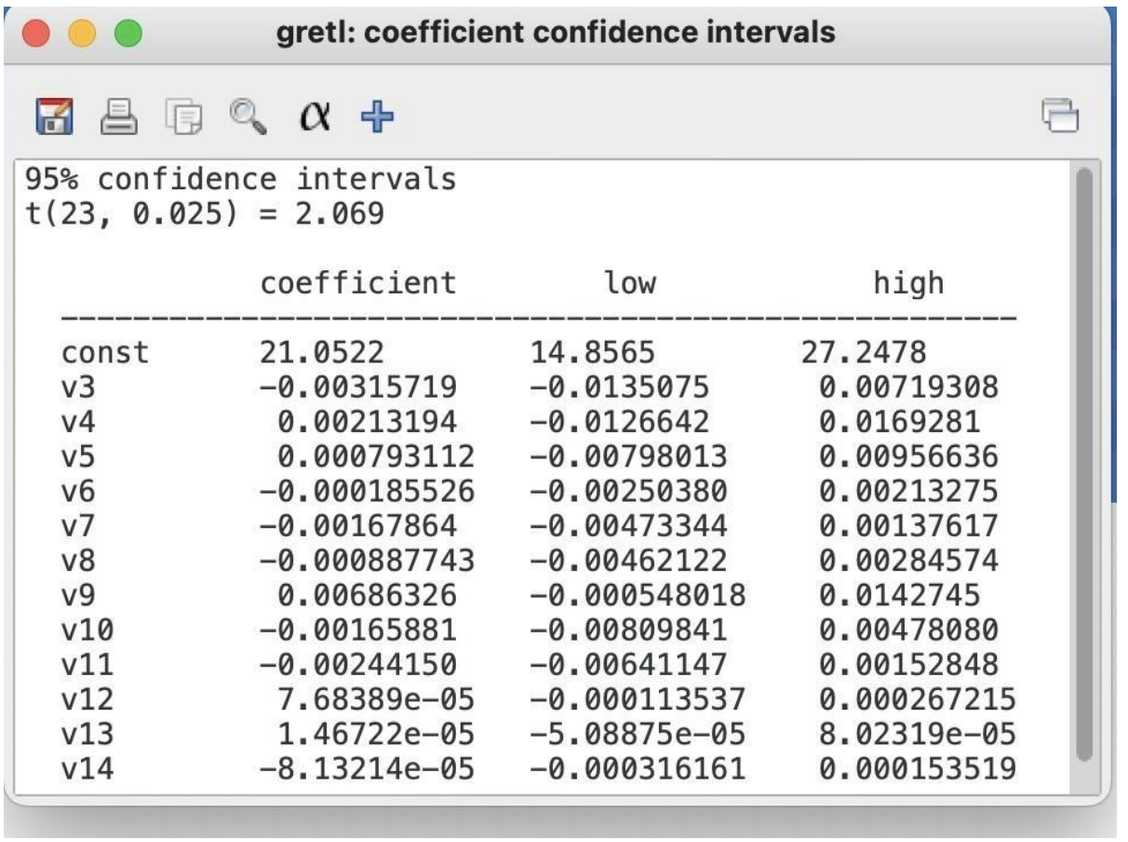
$state_t$ = 21.0522 - 0.00315719 v^3+0.00213194 v^4 + 0.000793112v^5 - 0.000185526 V^6 - 0.00167864 V^7 - 0.000887743 V^8 + 0.00686326 V^9 - 0.00165881 V^10 - 0.00244150 V^11 + 7.68389e-05 V^12 + 1.46722e-05 V^13 - 8.13214-05 V^14

- **R- Square Value** R-Squared is a statistical method that is used for calculating the data point closed with the regression line. It has another name called

coefficient determination and multiple determination for multiple regression. The below formula is used for calculating the R-Squared values.

HERE , k > 1, ADJUSTED R2 < R2; that is, as the number of explanatory variables increases in a model, the adjusted *R*2 becomes increasingly smaller than the unadjusted *R*2. There seems to be a "penalty" involved in adding more explanatory variables to a regression model. Although the unadjusted *R*2 is positive, the adjusted *R*2 turns out to be negative.

- **The Confidence Interval Approach to Hypothesis Testing**



The coefficients (βi's) signify the relation of the respective independent variable with the given dependent variable.

- **The estimated value of the constant is 21.0522 which has an economic significance. Statistically, the mean predicted value of state/ut is 21.0522.**
- **The partial slope coefficient β̂t of V4, V5, V9, V12, V13 are positive which means rates of crime against women in these particular year (state wise) increases.**

ANOVA Table:

- A study of these components of TSS is known as the analysis of variance (ANOVA) from the regression viewpoint.

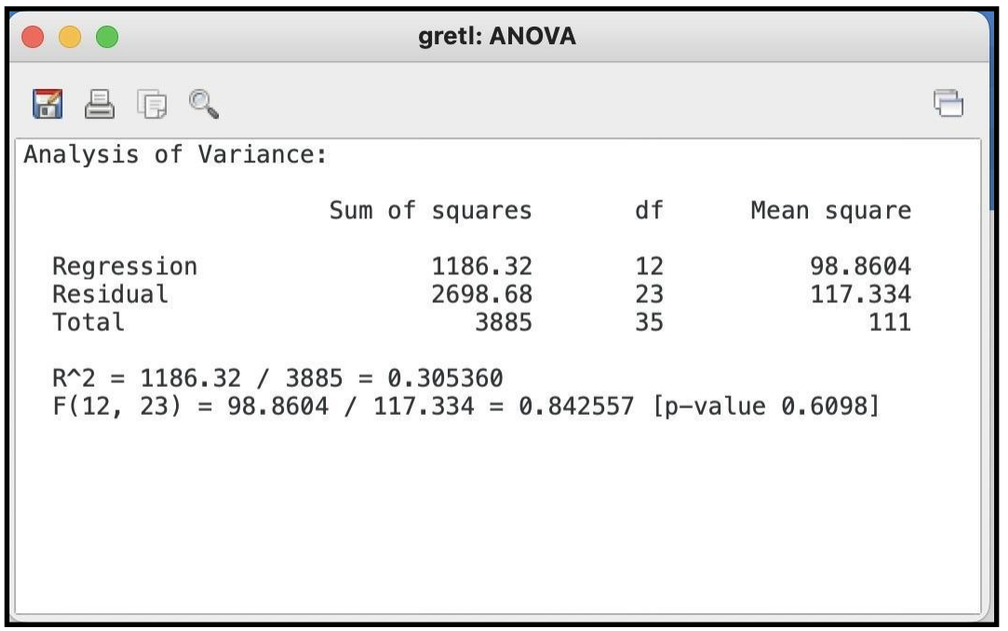
Test for overall significance of the model:

$H0{:}R2$ =0

$Ha{:}R2$ >0

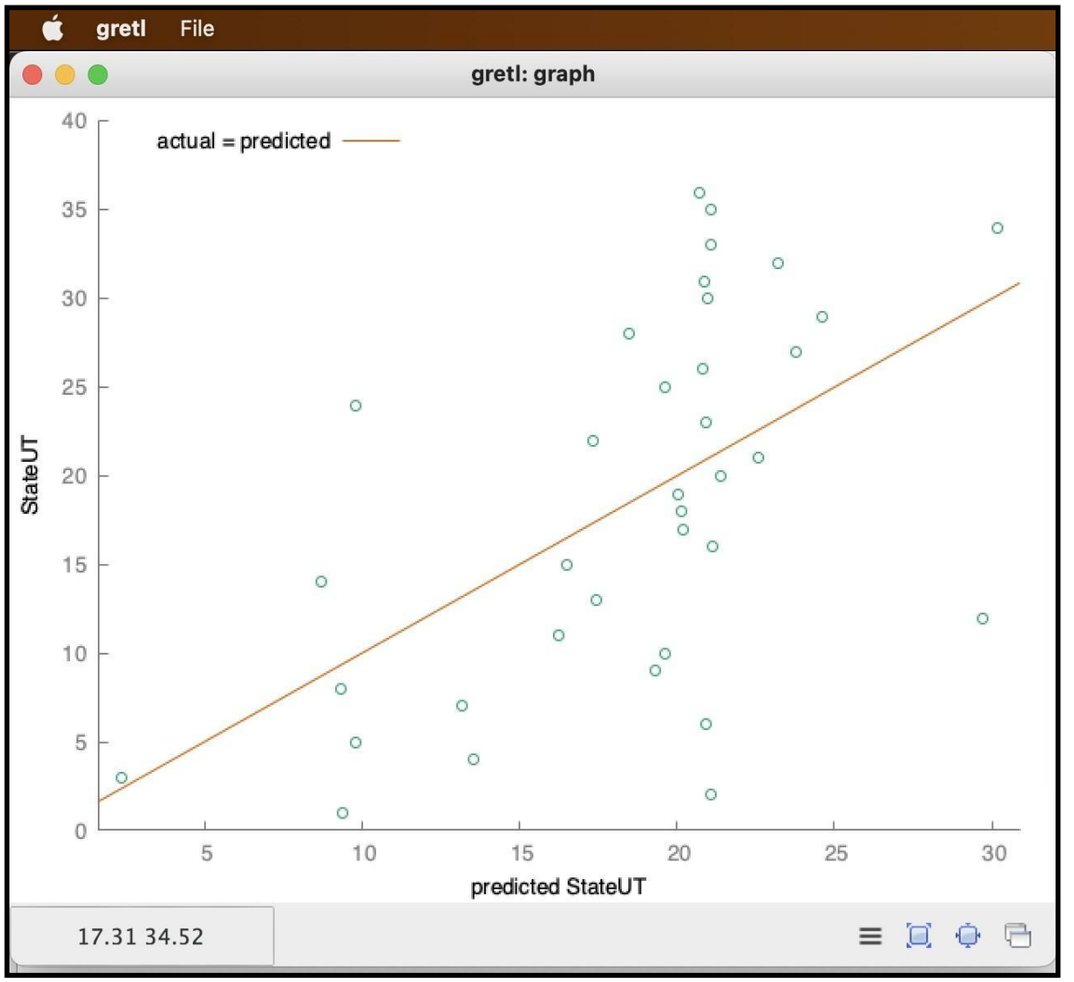
R2 value after estimating the regression using OLS is 0.305360.

$$F = \frac{R^2 / k-1}{1-R^2 / n-k}$$



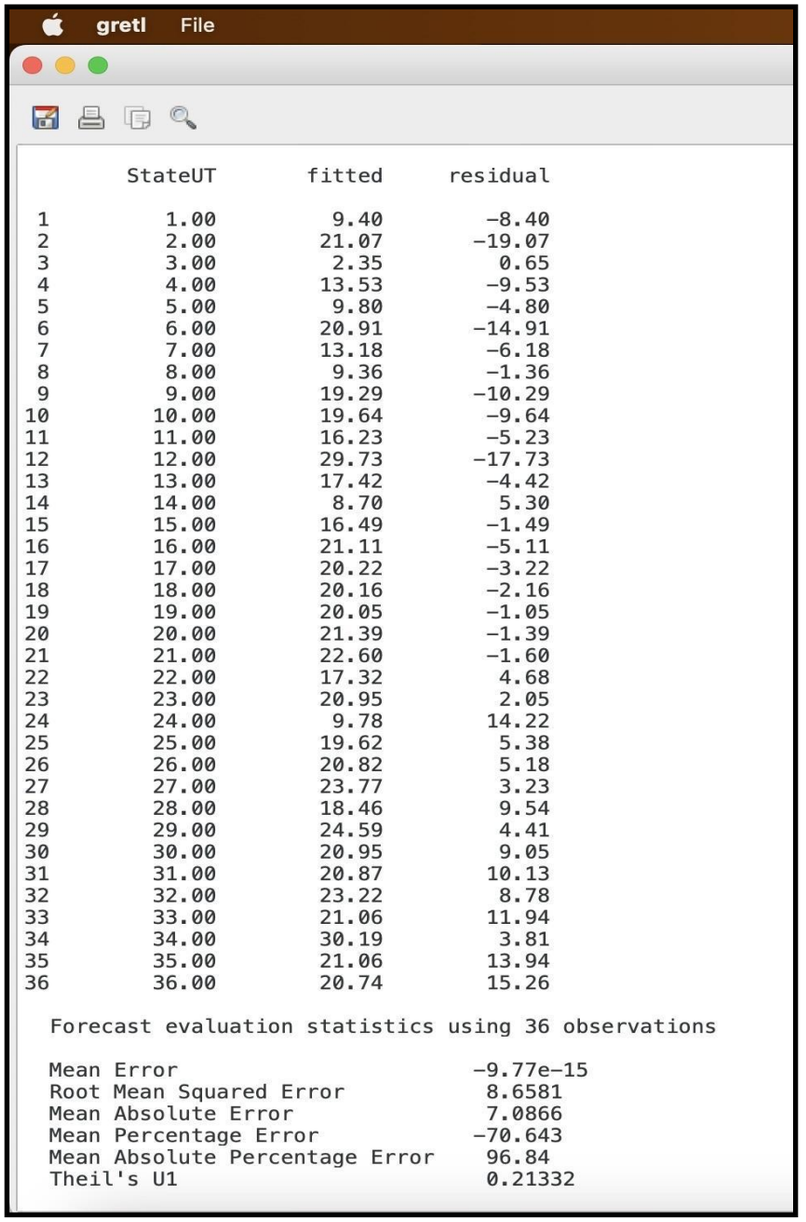
- Statistics: F calculated = F(12,23) = 0.842557
- Critical Value: F tabulated = 2.20
- Since, **F calculated < F tabulated**, we don't reject the null hypothesis at 5% level of significance.This implies that the R^2 value is not significantly different from 0 and this model is not statistically significant. Also, since the R^2 value is equal to 30.5360% therefore; it states that the explanatory variables have no impact on the dependent variable (STATES/UT).

Actual v/s Fitted Graph:



- The predicted model is relatively explicit as from the above graph we can see that the actual values are very highly scattered from the estimated model.
- This shows that there's a non-linear correlation between the model's predictions and its actual values regarding the crime against women state wise/ut.

DISPLAYING THE ACTUAL, FITTED AND RESIDUAL VALUES



	StateUT	fitted	residual
1	1.00	9.40	-8.40
2	2.00	21.07	-19.07
3	3.00	2.35	0.65
4	4.00	13.53	-9.53
5	5.00	9.80	-4.80
6	6.00	20.91	-14.91
7	7.00	13.18	-6.18
8	8.00	9.36	-1.36
9	9.00	19.29	-10.29
10	10.00	19.64	-9.64
11	11.00	16.23	-5.23
12	12.00	29.73	-17.73
13	13.00	17.42	-4.42
14	14.00	8.70	5.30
15	15.00	16.49	-1.49
16	16.00	21.11	-5.11
17	17.00	20.22	-3.22
18	18.00	20.16	-2.16
19	19.00	20.05	-1.05
20	20.00	21.39	-1.39
21	21.00	22.60	-1.60
22	22.00	17.32	4.68
23	23.00	20.95	2.05
24	24.00	9.78	14.22
25	25.00	19.62	5.38
26	26.00	20.82	5.18
27	27.00	23.77	3.23
28	28.00	18.46	9.54
29	29.00	24.59	4.41
30	30.00	20.95	9.05
31	31.00	20.87	10.13
32	32.00	23.22	8.78
33	33.00	21.06	11.94
34	34.00	30.19	3.81
35	35.00	21.06	13.94
36	36.00	20.74	15.26

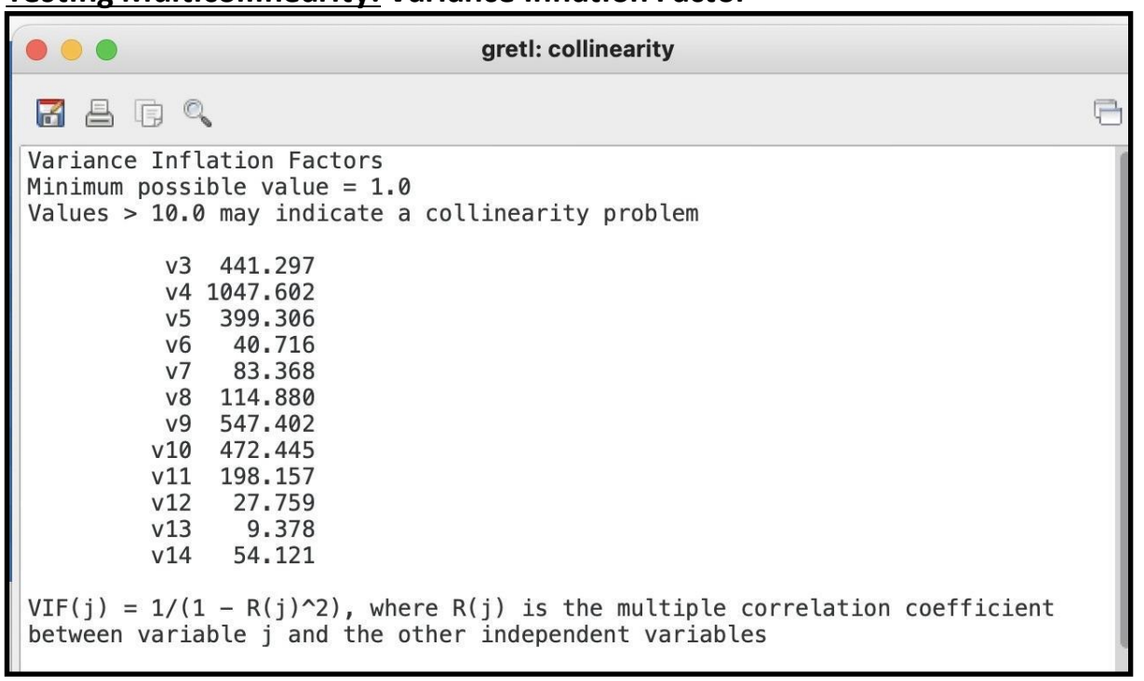
Forecast evaluation statistics using 36 observations

Mean Error	-9.77e-15
Root Mean Squared Error	8.6581
Mean Absolute Error	7.0866
Mean Percentage Error	-70.643
Mean Absolute Percentage Error	96.84
Theil's U1	0.21332

Tests for violations of OLS assumptions:

1) **Multicollinearity:** One of the assumptions of the CLRM is that there is no exact linear relationship among the regressors. If there are one or more such relationships among the regressors we call it multicollinearity.

Testing Multicollinearity: Variance Inflation Factor



VIF values greater than 10 may indicate the problem of collinearity.

There are 2 cases :-

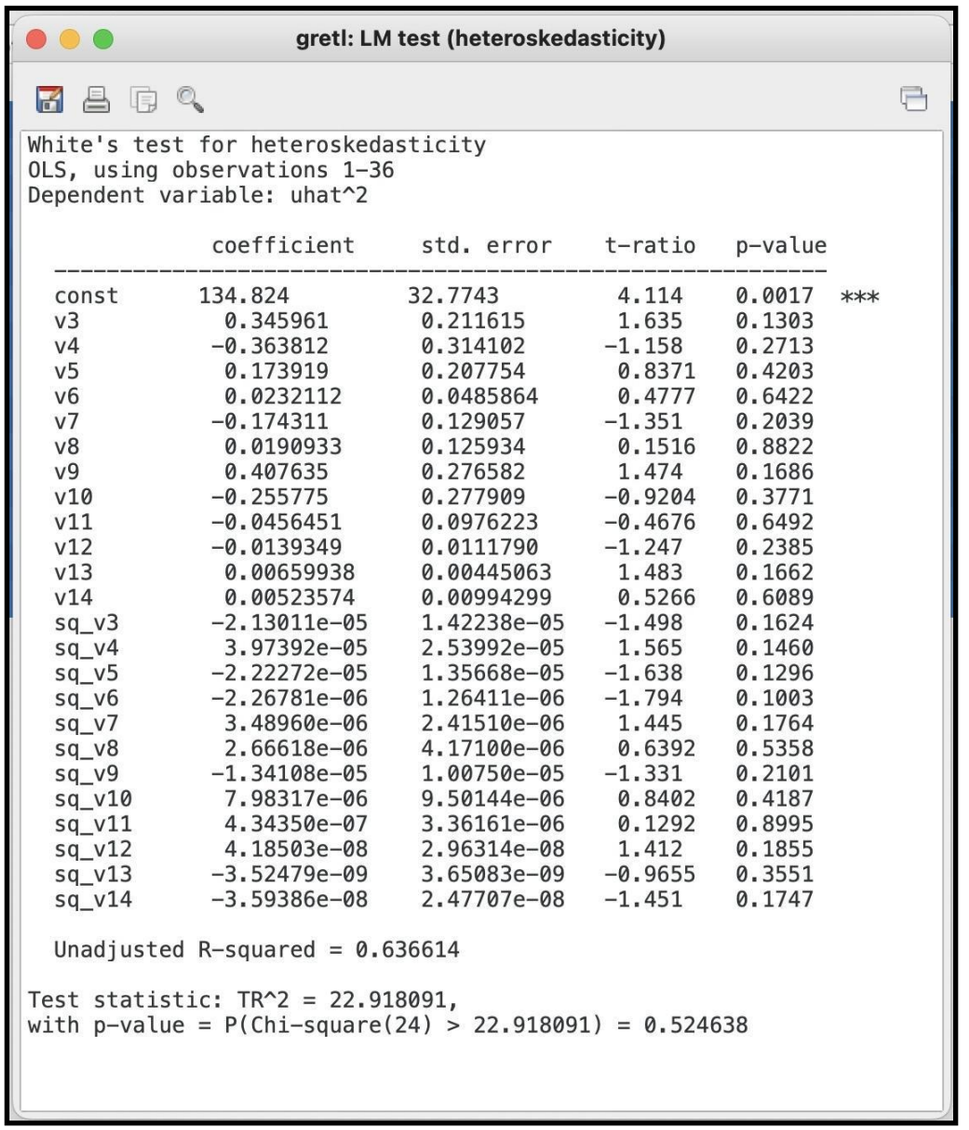
- **High VIF → High V (βi) → Low t values → Chances for not rejecting H0 increases → Variables insignificant.**
- **Low VIF → Low V (βi) → High t values → Chances for rejecting H0 inc → Variables significant.**

As the VIF values for all the independent variables except v13 (9.378) is very much higher than 10. It indicates that there is high multicollinearity present in the model. But variables v13 (9.378) have values less than 10 therefore it reflects no collinearity.

2) Heteroscedasticity: The CLRM assumes that the error term ui in the regression model has homoscedasticity (equal variance) across observations, denoted by σ2. If this assumption is not satisfied, we have the problem of heteroscedasticity.

Testing Heteroscedasticity: White's Test

$H0$: Homoscedasticity Ha: Heteroscedasticity



Test statistic: TR^2 = 22.918091

```
White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 22.9181
  with p-value = P(Chi-square(24) > 22.9181) = 0.524638
```

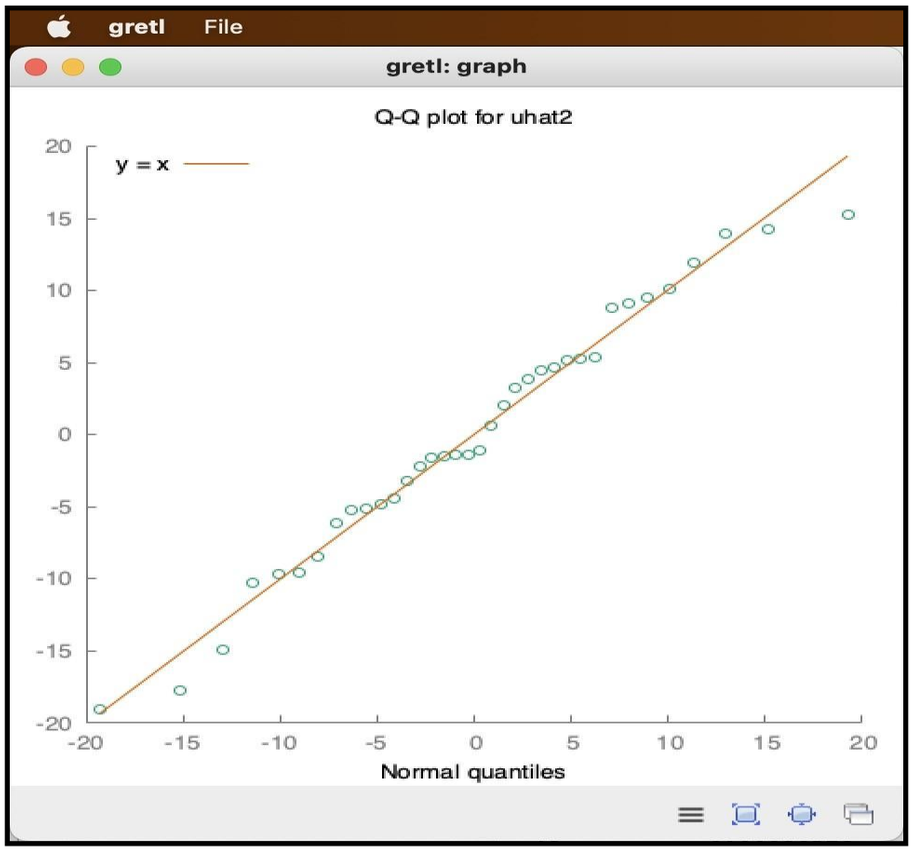
So, we do not reject the null hypothesis. This further signifies that heteroscedasticity is not present in the model.

3) <u>Autocorrelation:</u> One of the assumptions of CLRM is that the error terms, ut, are uncorrelated- that is the error term at time t is not correlated with the error term at time (t- 1) or any other error term in the past. If the error terms are correlated, then we have the problem of autocorrelation.

Therefore in this model: there is no autocorrelation because we have taken cross sectional data.

- **GRAPHS**

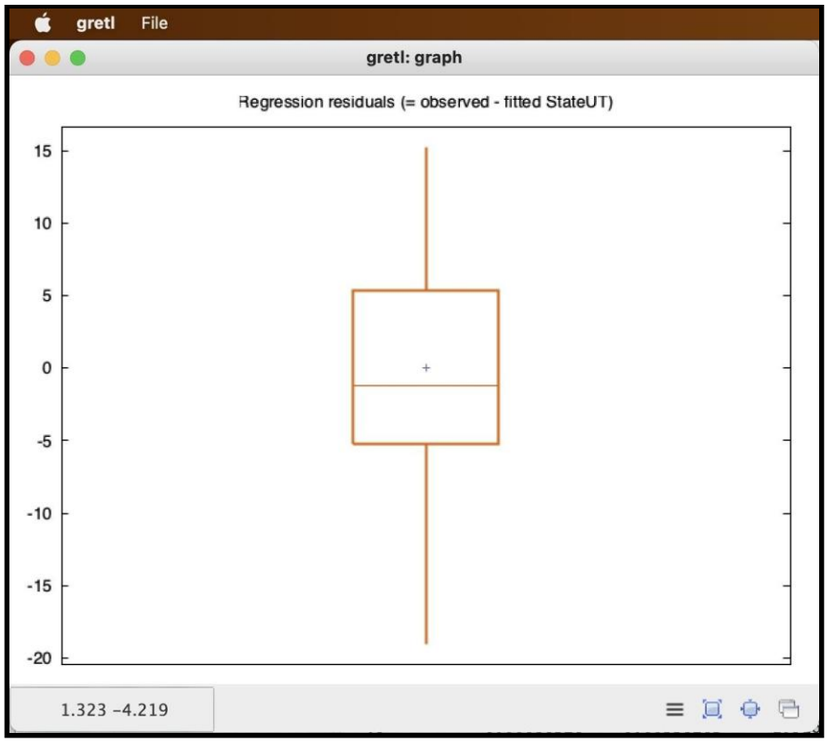
1. RESIDUAL PLOT - (Q-Q PLOT)



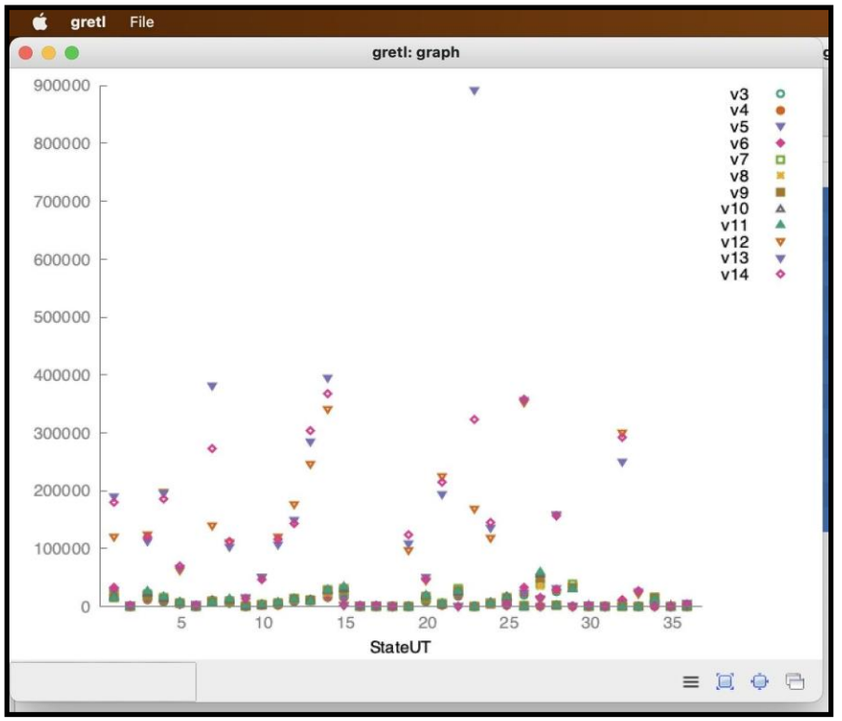
2. BOXPLOT

For some distributions/datasets, we need more information than the simple measures of central tendency (median, mean, and mode). A box plot is a graph that gives you a good indication of how the values in the data are spread out.

THEREFORE ; from this plot we conclude that the model is rightly skewed.



3. X-y scatter plot



CONCLUSION

IN OUR PROJECT WE HAVE DONE ANALYSIS FOR THE CRIME AGAINST WOMEN STATE/UNION TERRITORY WISE UNDER IPC FROM (2010 - 2021). In this model we calculated the measures of goodness of fit for the model which showed the following results:

In R- Square Value k > 1, ADJUSTED R2 < R2; that is, as the number of explanatory variables increases in a model, the adjusted *R*2 becomes increasingly smaller than the unadjusted *R*2. There seems to be a "penalty" involved in adding more explanatory variables to a regression model. Although the unadjusted *R*2 is positive, the adjusted *R*2 turns out to be negative.

In The Confidence Interval Approach to Hypothesis Testing

The partial slope coefficient $\hat{\beta}t$ of V4, V5, V9, V12, V13 are positive which means rates of crime against women in these particular year (state wise) increases**.**

From Anova Table Also since the R^2 value is equal to 30.5360% therefore; it states that the explanatory variables have no impact on the dependent variable (STATES/UT).

Finally, after running regression, we tested for auto-correlation, heteroscedasticity and multicollinearity in the model. We found that there is no heteroscedasticity and no autocorrelation