ECONOMETRICS

DATA ANALYSIS

BY :

SAHIBA KAUR - 27

KOMAL SAINI - 28

SAKSHI AGARWAL - 320

AIM OF THE PROJECT

The aim of our project is to analyse total personal healthcare expenditure of residents in the United States for the year 1998.

SOURCE OF THE PROJECT

The data chosen for the analysis is taken from the Statistical Abstract of the United States, 2000, 2001. It consists of data for 50 states and 1 federal district, District of Columbia.

Y	X1	X2	X3	X4	X5	X6
16056	6618	4609	1552	1064	96.0	4351
2299	986	568	133	42	17.1	615
14782	4977	5135	1397	839	113.0	4667
8463	3324	2225	903	776	53.7	2538
110057	34948	44239	7537	5626	920.5	32683
13669	4850	4314	970	904	119.0	3969
15221	4686	4292	1354	2264	122.2	3273
3106	1166	792	300	290	21.9	744
4258	2585	781	180	245	19.0	521
59724	19742	18985	6204	4880	400.2	14908
27219	10396	8510	2460	1545	197.3	7637
4658	1775	1594	311	204	31.9	1190
3397	1236	935	334	264	27.2	1231
44305	17996	11975	3964	3924	360.3	12070
21259	8515	5613	2058	2337	148.7	5908
10198	4084	2457	945	1186	70.8	2861
9394	3580	2538	854	920	67.4	2639
14414	5731	3785	1564	1283	87.3	3934
16500	7139	4249	1507	1248	96.9	4363
4925	1846	1219	456	476	29.3	1248
19646	7313	5978	1678	1695	156.8	5130
30039	11305	8322	2172	3568	205.8	6144
35647	14641	9186	3885	2459	264.0	9820
20313	6540	7183	1491	1964	138.3	4726
8882	3848	2212	962	687	54.4	2751
20911	8828	5310	1814	2002	136.8	5438
2838	1224	695	234	222	18.7	880
6095	2597	1367	626	697	43.1	1661
5606	1865	1918	478	164	50.9	1744
4658	1559	1405	391	425	35.0	1186
32695	11191	9506	3545	3233	278.3	8096
5344	2317	1415	402	257	36.7	1734
85785	32636	20103	7122	10586	583.1	18159
27327	10987	7106	2566	2347	190.0	7546
2680	1282	612	192	287	14.6	638
42581	16763	11024	3898	4978	293.0	11238
10988	4218	2978	1056	954	73.3	3339
10840	3545	3285	918	838	85.0	3282
51322	20213	13434	5035	5883	329.7	12002
4515	1702	1095	400	468	27.9	988
13204	5597	3254	1315	907	85.9	3840
2842	1257	747	201	286	17.3	731
22021	8276	6719	2129	2001	132.8	5433
67750	25322	20071	6023	4346	500.1	19712
5944	2290	1648	564	300	46.7	2101
2066	712	563	183	177	14.5	591
22261	8689	6265	2130	1546	190.5	6789
19292	6362	5908	1603	1492	163.3	5688
7037	2955	1793	776	515	36.6	1812
19945	7252	5844	1745	2110	137.3	5222
1407	582	343	133	113	11.7	480

Source: Statistical Abstract of the United States, 2000, 2001.

POPULATION

INCOME

NURSING HOME EXP.

DRUG EXP.

PHYSICIAN & OTHER EXP.

HOSPITAL EXP.

HEALTH EXP

DATA

VARIABLES OF THE DATA

healthcare (DEPENDENT VARIABLE) = Total personal health care expenditures by residents of state i in 1998 (million of current dollars)

INDEPENDENT VARIABLES :

hosp / X1 = Hospital care expenditures in 1998 in state i (in million of current dollars)

phys / X2 = Expenditures on physicians and other professional medical services in 1998 in state i (in million of current dollars)

drugs / X3 = Prescription drug expenditures in 1998 in state i in 1998 (in million of current dollars)

nurshome / X4 = Nursing home care expenditures in 1998 in state i(in millions of current dollars)

income / X5 = Personal income in state i in 1998(in billions of current dollars)

pop / X6 = population of state i in 1998 (in thousands)

A PRIORI EXPECTATION

- All the variables should have positive relationship with Y .
- Some variables are quite related thus there is a possibility of multicollinearity.

REGRESSION

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.999751447
R Square	0.999502957
Adjusted R Square	0.999435178
Standard Error	525.2795689
Observations	51

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	2.44E+10	4068855300	14746.58	5.76993E-71
Residual	44	12140420	275918.6255		
Total	50	2.44E+10			

	Coefficients	*andard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-39.64282848	110.303	-0.359399319	0.721016	-261.943948	182.65829	-261.94395	182.658291
hosp	1.060737122	0.116154	9.132141956	1.02E-11	0.826643608	1.2948306	0.82664361	1.29483064
phys	1.028820797	0.08769	11.7324468	3.87E-15	0.852092779	1.2055488	0.85209278	1.20554882
drugs	1.151613498	0.246603	4.669909	2.84E-05	0.654617839	1.6486092	0.65461784	1.64860916
nurshome	1.238324497	0.205889	6.014522251	3.21E-07	0.823382306	1.6532667	0.82338231	1.65326669
income	13.18166504	5.315065	2.480057221	0.017035	2.46985562	23.893474	2.46985562	23.8934745
pop	0.002147892	0.213089	0.010079792	0.992003	-0.42730468	0.4316005	-0.4273047	0.43160047

The fitted model is : **healthcare** = -39.6428+1.060***hosp**+1.0288***phys**+1.1516***drugs**+1.2383***nurshome**+12.1816***income**+0.0021***pop**

The signs of all the coefficients is positive which confirms our a priori expectations.

By analysing the corresponding p values , we can say that only population has come out to be insignificant and all other variables are significant

But regression as a whole is significant.

So we suspect presence of multicollinearity in the sample.

FARRAR GLAUBER TEST

(I) Chi-Square test

	X1	X2	X3	X4	X5	X6
X1	1	0.935146215	0.98461736	0.92914479	0.972877157	0.968708173
X2	0.935146	1	0.919184372	0.78903227	0.984434774	0.98650956
X3	0.984617	0.919184372	1	0.920286097	0.955763366	0.955083692
X4	0.929145	0.78903227	0.920286097	1	0.858176222	0.828347424
X5	0.972877	0.984434774	0.955763366	0.858176222	1	0.993002187
X6	0.968708	0.98650956	0.955083692	0.828347424	0.993002187	1

Value of the standardised determinant	**1.02122E-08**
Calculated Chi-square	**867.8518832**
Tabulated Chi-square	**24.99579014**

Since, calculated Chi > Tabulated Chi , we reject Ho and conclude that multicollinearity exists and is severe.

(II) F-Test

ω	1.02122E-08		
ω11	1.5236E-06	$R^2_{1.23456}$	**0.993297339**
ω22	7.72039E-07	$R^2_{2.13456}$	**0.986772474**
ω33	3.85666E-07	$R^2_{3.12456}$	**0.973520718**
ω44	3.06233E-07	$R^2_{4.12356}$	**0.966652279**
ω55	1.49561E-06	$R^2_{5.12346}$	**0.993171907**
ω66	2.94174E-06	$R^2_{6.12345}$	**0.996528533**

Calculated F
1704.237
857.8991
422.8018
333.3512
1672.718
3301.22

Tabulated F
2.574035

Since for all variables, the calculated F > tabulated F, therefore, we reject Ho and conclude that all the variables are the cause of multicollinearity

(III) t-Test

Partial Correlation		**Calculated t**	**|Calculated t|**	**Decision**	**Conclusion**
$r_{12.3456}$	-0.458	-3.49434667	3.49434667	**Reject Ho**	X1 and X2 are collinear
$r_{13.2456}$	0.2	1.38443731	1.38443731	**Accept Ho**	X1 and X3 are not collinear
$r_{14.2356}$	0.745	7.574733935	7.574733935	**Reject Ho**	X1 and X4 are collinear
$r_{15.2346}$	-0.009	-0.061043442	0.061043442	**Accept Ho**	X1 and X5 are not collinear
$r_{16.2345}$	0.667	6.071763304	6.071763304	**Reject Ho**	X1 and X6 are collinear
$r_{23.1456}$	-0.119	-0.812873339	0.812873339	**Accept Ho**	X2 and X3 are not collinear
$r_{24.1356}$	0.23	1.602908875	1.602908875	**Accept Ho**	X2 and X4 are not collinear
$r_{25.1346}$	0.377	2.760636798	2.760636798	**Reject Ho**	X2 and X5 are collinear
$r_{26.1345}$	0.537	4.317434646	4.317434646	**Reject Ho**	X2 and X6 are collinear
$r_{34.1256}$	0.324	2.322772018	2.322772018	**Reject Ho**	X3 and X4 are collinear
$r_{35.1246}$	-0.235	-1.63976866	1.63976866	**Accept Ho**	X3 and X5 are not collinear
$r_{36.1245}$	0.347	2.509389071	2.509389071	**Reject Ho**	X3 and X6 are collinear
$r_{45.1236}$	0.421	3.147927277	3.147927277	**Reject Ho**	X4 and X5 are collinear
$r_{46.1235}$	-0.779	-8.426209692	8.426209692	**Reject Ho**	X4 and X6 are collinear
$r_{56.1234}$	0.492	3.832902314	3.832902314	**Reject Ho**	X5 and X6 are collinear
		Tabulated t	**2.012895599**		

MULTICOLLINEARITY TEST :

THE FARRAR GLAUBER TEST

Thus , as we expected , multicollinearity is present in this sample. As all forms of expenditure , population and income are dependent on each other, multicollinearity had to be present in the data. Also , as population is an important variable for this model, we can't drop it.

So , we do further analysis keeping in mind that problem of multicollinearity persists in the model.

TABLE OF MULTICOLLINEARITY

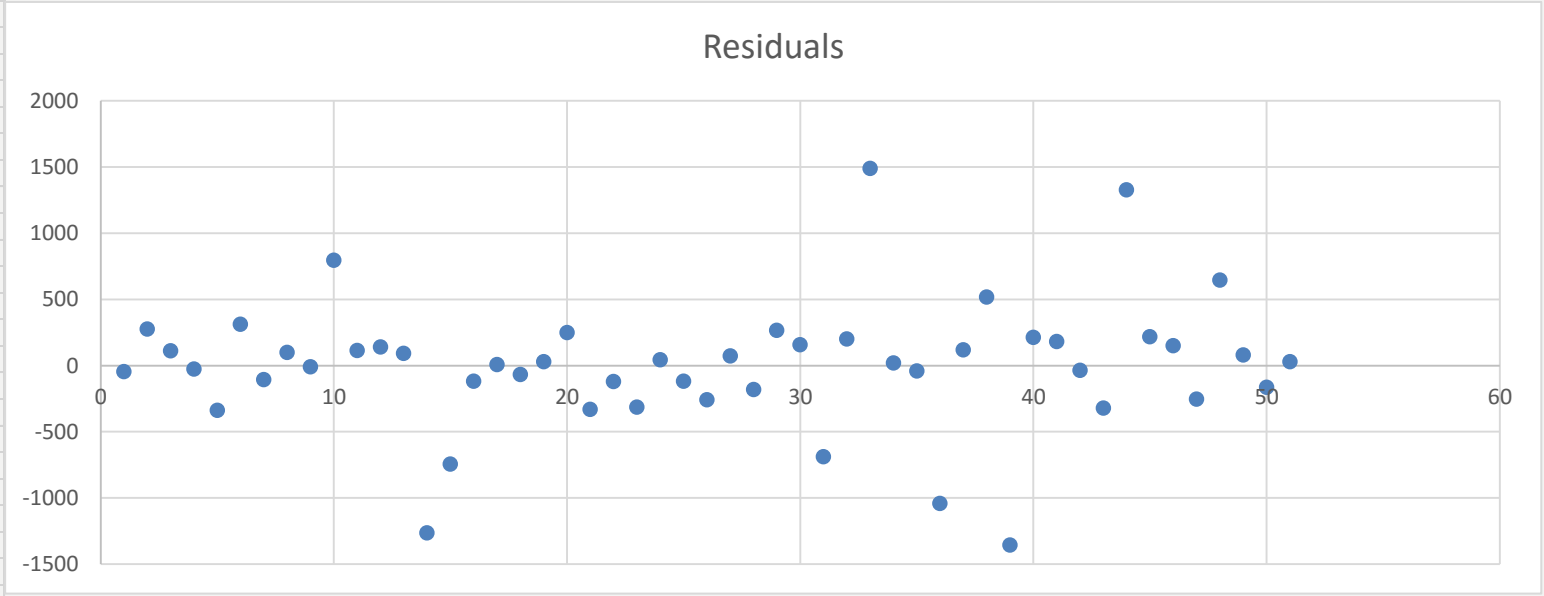
VARIABLES	**HOSP**	**PHYS**	**DRUGS**	**NURSHOM**	**INCOME**	**POPU**
HOSP		COLLINEAR		COLLINEAR		COLLINEAR
PHYS					COLLINEAR	COLLINEAR
DRUGS				COLLINEAR		COLLINEAR
NURSHOM					COLLINEAR	COLLINEAR
INCOME						COLLINEAR
POPU						

RESIDUAL OUTPUT

Observation	*Predicted health*	*Residuals*	*Standard Residuals*
1	16101.81723	-45.81723	-0.092981602
2	2022.515836	276.48416	0.561097606
3	14669.95129	112.04871	0.227391908
4	8489.527199	-26.5272	-0.05383436
5	110395.4481	-338.4481	-0.68684744
6	13356.91869	312.08131	0.633338531
7	15327.35104	-106.351	-0.215829052
8	3006.877377	99.122623	0.201159683
9	4268.122292	-10.12229	-0.020542204
10	58928.54908	795.45092	1.614289956
11	27106.37179	112.62821	0.228567948
12	4516.927015	141.07298	0.286293844
13	3306.117619	90.882381	0.184436915
14	45568.97167	-1263.972	-2.565107045
15	22004.09316	-745.0932	-1.512093805
16	10316.55489	-118.5549	-0.240595568
17	9385.792226	8.2077736	0.016656875
18	14482.63134	-68.63134	-0.139280602
19	16472.00416	27.995839	0.056814822
20	4676.092021	248.90798	0.505134433
21	19977.08971	-331.0897	-0.671914221
22	30159.46664	-120.4666	-0.244475288
23	35961.46746	-314.4675	-0.638181005
24	20269.89799	43.102015	0.087471329
25	8999.397764	-117.3978	-0.238247295
26	21170.66746	-259.6675	-0.526969746
27	2766.502741	71.497259	0.145096704
28	6277.209143	-182.2091	-0.369775659
29	5340.159337	265.84066	0.539497661
30	4500.01403	157.98597	0.320617096
31	33382.85646	-687.8565	-1.395937513
32	5142.556085	201.44392	0.408810752
33	84296.88528	1488.1147	3.019983488
34	27307.5887	19.411305	0.039393347
35	2720.192076	-40.19208	-0.081565894
36	43622.94864	-1041.949	-2.114533002
37	10869.22797	118.77203	0.241036237
38	10322.73462	517.26538	1.049739573
39	52677.42615	-1355.426	-2.750704981
40	4302.362361	212.63764	0.431527319
41	13023.17072	180.82928	0.366975367
42	2877.480903	-35.4809	-0.072005026
43	22343.5316	-322.5316	-0.654546379
44	66422.22109	1327.7789	2.694597616
45	5726.045695	217.9543	0.442316974
46	1917.160365	148.83964	0.302055503
47	22515.73997	-254.74	-0.51696989
48	18645.43971	646.56029	1.312130962
49	6957.241167	79.758833	0.161862761
50	20108.74066	-163.7407	-0.332295683
51	1378.943442	28.056558	0.056938043

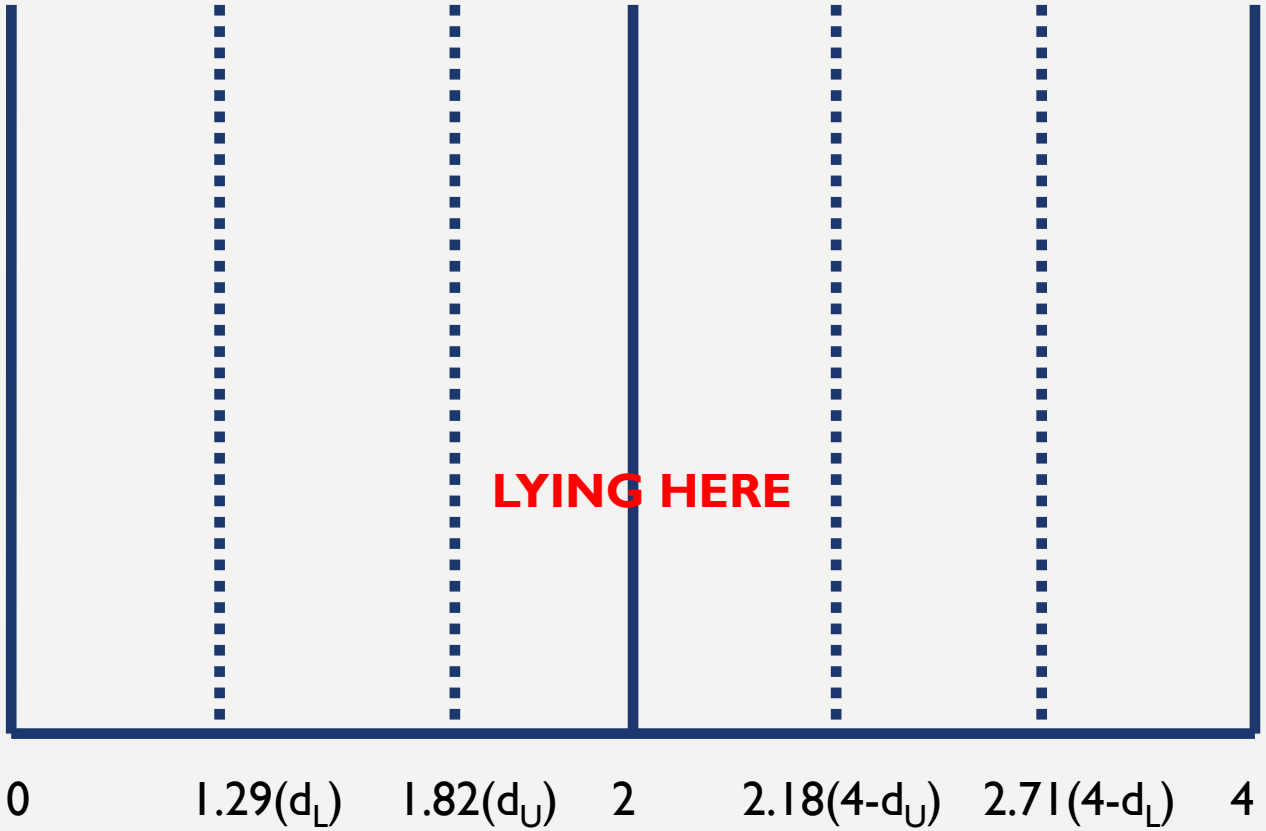
DETECTION OF AUTOCORRELATION

Scatter plot of residuals is not showing any trend but we are further testing using d statistic.



Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	1.000[a]	1.000	.999	525.27957	2.047



TEST FOR AUTOCORRELATION

Thus, **NO AUTOCORRELATION** is present in the data.

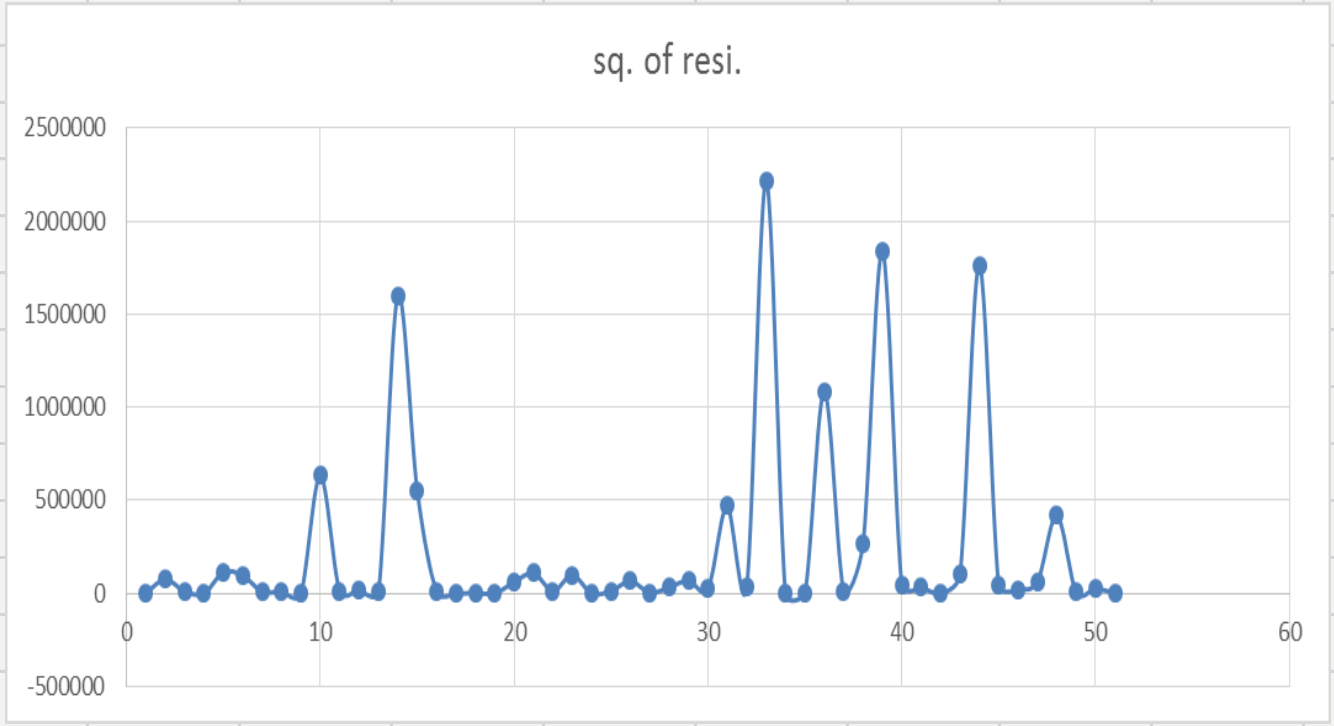
HETEROSCEDASTICITY DETECTION

Detection

For the model,

healthcare=-39.6428+1.060*hosp+1.0288*phys+1.1516*drugs+1.2383*nurshome+12.1816*income+0.0021*pop+e

we plot the square of residuals,



As we see , the variances of the residuals are quite high and show some pattern, we suspect presence of heteroscedasticity in the model.

To confirm this, we proceed for the **Spearman's Rank Correlation Test .**

d_1^2	d_2^2	d_3^2	d_4^2	d_5^2	d_6^2
484	400	441	289	289	361
1089	1089	1225	1225	1024	1024
100	196	121	25	144	196
225	256	256	256	225	225
100	100	100	64	100	100
121	64	169	196	49	81
81	144	121	529	225	49
121	49	49	9	49	64
196	25	1	25	25	0
1	4	9	1	4	4
484	576	529	196	529	576
144	64	196	324	121	144
64	25	16	25	25	4
4	4	4	9	1	1
64	144	64	25	81	49
9	1	4	64	4	4
400	441	324	576	400	361
324	225	441	361	225	256
784	484	625	576	529	625
400	441	324	256	441	361
25	16	36	36	9	49
441	361	324	484	400	289
36	25	36	16	25	36
484	961	400	676	676	529
9	0	16	1	1	4
25	4	4	9	1	4
49	49	25	36	36	16
121	256	121	81	144	196
484	289	400	1024	289	361
256	144	196	121	144	225
4	0	1	1	0	1
196	225	256	441	196	196
1	1	1	0	1	4
1444	1296	1521	1444	1296	1444
0	16	9	9	25	9
4	4	4	1	4	4
9	4	16	25	4	16
484	289	441	441	324	324
4	9	9	0	16	16
400	400	324	225	400	441
1	9	1	9	4	1
0	1	1	9	4	1
9	1	1	9	49	16
0	0	1	9	0	1
289	225	225	324	196	169
484	484	400	400	484	441
25	16	36	0	49	49
169	64	100	144	25	36
25	16	25	16	1	16
64	64	81	144	64	64
25	25	25	16	25	25
10762	**9986**	**10055**	**11182**	**9382**	**9468**

SPEARMAN'S RANK CORRELATION TEST

	$corr(X_1,e)$	$corr(X_2,e)$	$corr(X_3,e)$	$corr(X_4,e)$	$corr(X_5,e)$	$corr(X_6,e)$
	0.51303	0.54814	0.54502	0.49403	0.575475	0.5715837
Cal t	4.05367	4.44497	4.40891	3.85378	4.772524	4.7245126
tab t	2.0129					
Decision:	**Reject Ho**	**Reject Ho**	**Reject Ho**	**Reject Ho**	**Reject Ho**	**Reject Ho**

Since for all combinations Ho is rejected thus heteroscedasticity is present in the data.

health	hosp	phys	drugs	nurshome	income	pop
3.690186	1.52103	1.059297	0.3567	0.244541	0.022064	1
3.738211	1.603252	0.923577	0.21626	0.068293	0.027805	1
3.167345	1.066424	1.100279	0.299336	0.179773	0.024213	1
3.334515	1.309693	0.876675	0.355792	0.305753	0.021158	1
3.367408	1.069302	1.353578	0.230609	0.172138	0.028164	1
3.443941	1.22197	1.086924	0.244394	0.227765	0.029982	1
4.650474	1.431714	1.311335	0.413688	0.69172	0.037336	1
4.174731	1.567204	1.064516	0.403226	0.389785	0.029435	1
8.172745	4.961612	1.49904	0.345489	0.47025	0.036468	1
4.006171	1.324255	1.273477	0.416152	0.327341	0.026845	1
3.564096	1.361268	1.114312	0.322116	0.202305	0.025835	1
3.914286	1.491597	1.339496	0.261345	0.171429	0.026807	1
2.759545	1.004062	0.759545	0.271324	0.21446	0.022096	1
3.670671	1.490969	0.992129	0.328418	0.325104	0.029851	1
3.598341	1.441266	0.950068	0.348341	0.395565	0.025169	1
3.564488	1.427473	0.858791	0.330304	0.41454	0.024747	1
3.559682	1.356574	0.961728	0.323607	0.348617	0.02554	1
3.663955	1.456787	0.962125	0.39756	0.326131	0.022191	1
3.781802	1.636259	0.973871	0.345405	0.286042	0.022209	1
3.946314	1.479167	0.976763	0.365385	0.38141	0.023478	1
3.82963	1.425536	1.165302	0.327096	0.330409	0.030565	1
4.88916	1.840007	1.354492	0.353516	0.580729	0.033496	1
3.630041	1.490937	0.935438	0.395621	0.250407	0.026884	1
4.298138	1.383834	1.51989	0.315489	0.415573	0.029264	1
3.228644	1.398764	0.804071	0.349691	0.249727	0.019775	1
3.845348	1.623391	0.976462	0.333579	0.36815	0.025156	1
3.225	1.390909	0.789773	0.265909	0.252273	0.02125	1
3.669476	1.563516	0.822998	0.376881	0.419627	0.025948	1
3.21445	1.069381	1.099771	0.274083	0.094037	0.029186	1
3.927487	1.314503	1.184654	0.32968	0.358347	0.029511	1
4.038414	1.382288	1.17416	0.437871	0.399333	0.034375	1
3.081892	1.336217	0.816032	0.231834	0.148212	0.021165	1
4.724104	1.797236	1.107054	0.392202	0.582962	0.032111	1
3.621389	1.456003	0.941691	0.340048	0.311026	0.025179	1
4.200627	2.009404	0.959248	0.30094	0.449843	0.022884	1
3.789019	1.491636	0.980957	0.346859	0.442961	0.026072	1
3.290806	1.263252	0.891884	0.316262	0.285714	0.021953	1
3.302864	1.080134	1.000914	0.279707	0.255332	0.025899	1
4.276121	1.684136	1.119313	0.419513	0.490168	0.02747	1
4.569838	1.722672	1.1083	0.404858	0.473684	0.028239	1
3.438542	1.457552	0.847396	0.342448	0.236198	0.02237	1
3.887825	1.719562	1.021888	0.274966	0.391245	0.023666	1
4.053193	1.523284	1.236702	0.391865	0.368305	0.024443	1
3.436993	1.284598	1.018212	0.30555	0.220475	0.02537	1
2.829129	1.089957	0.784388	0.268444	0.142789	0.022228	1
3.49577	1.204738	0.952623	0.309645	0.299492	0.024535	1
3.278981	1.279864	0.922816	0.313743	0.227721	0.02806	1
3.391702	1.118495	1.038678	0.281821	0.262307	0.02871	1
3.883554	1.630795	0.989514	0.428256	0.284216	0.020199	1
3.819418	1.38874	1.119111	0.334163	0.40406	0.026293	1
2.93125	1.2125	0.714583	0.277083	0.235417	0.024375	1

Now we regress per capita healthcare expenditure on the independent variables

TRANSFORMED MODEL

Here, we divided the entire data by population,
As it is giving a solution for heteroscedasticity ,
without creating problem of interpretation.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.993951
R Square	0.987938
Adjusted R	0.986293
Standard E	0.091188
Observatic	51

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	6	29.96694	4.99449	600.6427	1.67E-40
Residual	44	0.365871	0.008315		
Total	50	30.33281			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	0.069662	0.123686	0.563218	0.576145	-0.17961	0.318934	-0.17961	0.318934
hosp	1.035534	0.026975	38.38855	1.79E-35	0.98117	1.089899	0.98117	1.089899
phys	1.027898	0.099627	10.31745	2.53E-13	0.827113	1.228683	0.827113	1.228683
drugs	0.966527	0.310006	3.117771	0.003208	0.341751	1.591303	0.341751	1.591303
nurshome	0.894679	0.155616	5.749288	7.87E-07	0.581056	1.208301	0.581056	1.208301
income	19.03939	4.809946	3.958336	0.000272	9.345576	28.73319	9.345576	28.73319
pop	0	0	65535	#NUM!	0	0	0	0

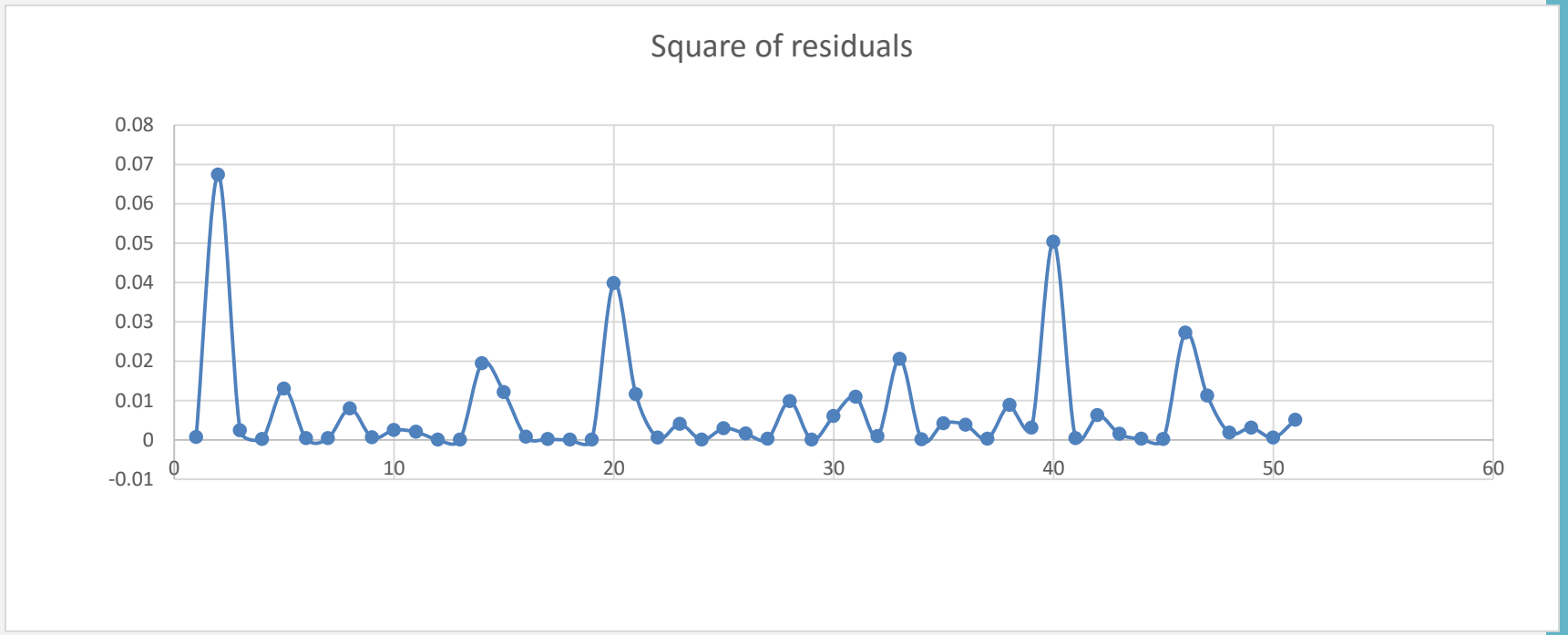
REGRESSION OF TRANSFORMED MODEL

All variables are coming out to be significant, except the variable ‘population’ .

TEST FOR HETEROSCEDASTICITY IN TRASFORMED MODEL

RESIDUAL OUTPUT

Observation	dicted hea	Residuals	Sq. of resi
1	3.717219	-0.02703	0.000731
2	3.478737	0.259474	0.067327
3	3.216102	-0.04876	0.002377
4	3.347302	-0.01279	0.000163
5	3.481435	-0.11403	0.013002
6	3.463137	-0.0192	0.000369
7	4.629728	0.020745	0.00043
8	4.085665	0.089066	0.007933
9	8.197425	-0.02468	0.000609
10	3.956172	0.049999	0.0025
11	3.60891	-0.04481	0.002008
12	3.907481	0.006805	4.63E-05
13	2.764944	-0.0054	2.91E-05
14	3.81005	-0.13938	0.019426
15	3.708508	-0.11017	0.012137
16	3.591897	-0.02741	0.000751
17	3.573941	-0.01426	0.000203
18	3.665723	-0.00177	3.12E-06
19	3.777719	0.004083	1.67E-05
20	3.746795	0.199519	0.039808
21	3.937367	-0.10774	0.011607
22	4.866326	0.022834	0.000521
23	3.693379	-0.06334	0.004012
24	4.298857	-0.00072	5.17E-07
25	3.282542	-0.0539	0.002905
26	3.885192	-0.03984	0.001588
27	3.209101	0.015899	0.000253
28	3.76843	-0.09895	0.009792
29	3.212215	0.002235	4.99E-06
30	3.849699	0.077788	0.006051
31	4.142953	-0.10454	0.010928
32	3.051802	0.030089	0.000905
33	4.580708	0.143396	0.020562
34	3.63169	-0.0103	0.000106
35	4.265508	-0.06488	0.00421
36	3.850583	-0.06156	0.00379
37	3.273833	0.016972	0.000288
38	3.208899	0.093965	0.008829
39	4.331217	-0.0551	0.003036
40	4.34552	0.224318	0.050319
41	3.418258	0.020284	0.000411
42	3.967115	-0.07929	0.006287
43	4.091924	-0.03873	0.0015
44	3.422138	0.014854	0.000221
45	2.815028	0.014101	0.000199
46	3.330763	0.165007	0.027227
47	3.384792	-0.10581	0.011196
48	3.349238	0.042464	0.001803
49	3.8283	0.055254	0.003053
50	3.843159	-0.02374	0.000564
51	3.002282	-0.07103	0.005046



PREDICTION

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.993876
R Square	0.987789
Adjusted R	0.986401
Standard E	0.090628
Observatic	50

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	5	29.23366	5.846732	711.854	6.71E-41
Residual	44	0.361389	0.008213		
Total	49	29.59505			

	Coefficients	*andard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.087226	0.124722	0.699361	0.488007	-0.16413	0.338586	-0.16413499	0.33858638
hosp	1.03585	0.026812	38.63377	1.37E-35	0.981814	1.089887	0.98181428	1.08988653
phys	1.004326	0.102981	9.752494	1.44E-12	0.796781	1.211872	0.79678061	1.21187164
drugs	0.94256	0.309443	3.045992	0.00391	0.318919	1.5662	0.31891923	1.56620034
nurshome	0.895882	0.154666	5.792355	6.81E-07	0.584173	1.207592	0.58417267	1.20759155
income	19.63067	4.832836	4.061935	0.000197	9.890728	29.37061	9.89072755	29.3706112

	phealth	**phosp**	**pphys**	**pdrugs**	**pnurshome**	**pincome**
51'st state	2.93125	1.2125	0.714583	0.277083	0.235417	0.024375
expected health expenditure	3.01144					

Here we have first regressed the model for 50 observations and then using the model and values of X variable we have evaluated the value for the 51st state

Since, the observed value (2.93) is almost equal to the expected value (3.01), we conclude that the model can be used for prediction as well.

phealthcare = 0.069662 + 1.0355*phosp +1.0278*pphys + 0.9665*pdrugs + 0.8946*pnurshome + 19.039*pincome

phealthcare (DEPENDENT VARIABLE) = per capita personal health care expenditures of state i in 1998 (million of current dollars)

INDEPENDENT VARIABLES :

phosp / X1 = per capita hospital care expenditures in 1998 in state i (in million of current dollars)

pphys / X2 =per capita expenditures on physicians and other professional medical services in 1998 in state I (in million of current dollars)

pdrugs / X3 = per capita prescription drug expenditures in 1998 in state i in 1998(in million of current dollars)

pnurshome / X4 = per capita nursing home care expenditures in 1998 in state i(in millions of current dollars)

pincome / X5 = per capita personal income in state i in 1998(in billions of current dollars)

THE BEST REGRESSION EQUATION