

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
In [2]: df = pd.read_csv('MLR-Feature-Elimination.csv')
columns_list = df.columns.tolist()
print(columns_list)

['c1', 'c2', 'c26', 'c27', 'c28', 'c29', 'c30', 'c31', 'c32', 'c33', 'c39', 'c139', 'c142', 'c143', 'c155', 'c157', 'c158', 'c160', 'c161', 'c162', 'c163', 'c7', 'c8', 'c9', 'c10', 'c11', 'c12', 'c13', 'c15', 'c16', 'c17', 'c19', 'c20', 'c21', 'c22', 'c23', 'c34', 'c35', 'c36', 'c52', 'c241']
```

```
In [3]: import statsmodels.api as sm
X = df[['c26', 'c27', 'c28', 'c29', 'c30', 'c31', 'c32', 'c33', 'c39', 'c139', 'c142', 'c143', 'c155', 'c157', 'c158', 'c160', 'c161', 'c162', 'c163', 'c7', 'c8', 'c9', 'c10', 'c11', 'c12', 'c13', 'c15', 'c16', 'c17', 'c19', 'c20', 'c21', 'c22', 'c23', 'c34', 'c35', 'c36', 'c52', 'c241']]
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:          0.785
Model:                  OLS      Adj. R-squared:       0.777
Method:                 Least Squares      F-statistic:       97.27
Date:                   Sat, 02 Sep 2023    Prob (F-statistic): 1.11e-299
Time:                   22:58:20      Log-Likelihood:    -1479.4
No. Observations:      1025      AIC:               3035.
Df Residuals:          987      BIC:               3222.
Df Model:               37
Covariance Type:       nonrobust
=====

```

```

=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const      -136.4882     104.748     -1.303     0.193    -342.042     69.065
c26         0.3634       0.049       7.442     0.000      0.268      0.459
c27        -0.1911       0.895     -0.214     0.831     -1.948      1.565
c28         0.2266       0.044       5.150     0.000      0.140      0.313
c29        -0.4454       0.049     -9.115     0.000     -0.541     -0.349
c30         3.4632       0.458       7.568     0.000      2.565      4.361
c31         0.2667       0.035       7.699     0.000      0.199      0.335
c32         0.1781       0.199       0.895     0.371     -0.212      0.569
c33        -0.6545       0.464     -1.412     0.158     -1.564      0.255
c39        12.9984       1.470       8.845     0.000     10.114     15.882
c139       -0.8439       0.225     -3.745     0.000     -1.286     -0.402
c142        0.0454       0.067       0.682     0.495     -0.085      0.176
c143       -0.1537       0.039     -3.956     0.000     -0.230     -0.077
c155       -0.0342       0.013     -2.684     0.007     -0.059     -0.009
c157        0.2501       0.041       6.100     0.000      0.170      0.331
c158        0.2836       0.023     12.121     0.000      0.238      0.329
c160        0.0040       0.002       2.206     0.028      0.000      0.008
c161        0.0105       0.001       9.632     0.000      0.008      0.013
c162        0.0027       0.002       1.649     0.099     -0.001      0.006
c163        0.0081       0.002       3.724     0.000      0.004      0.012
c7          0.3236       0.292       1.108     0.268     -0.250      0.897
c8         -0.4465       0.137     -3.257     0.001     -0.716     -0.177
c9         -0.6863       0.075     -9.097     0.000     -0.834     -0.538
c10         8.8046       1.541       5.715     0.000      5.781     11.828
c11        -0.1706       0.042     -4.044     0.000     -0.253     -0.088
c12        -0.3028       0.109     -2.765     0.006     -0.518     -0.088
c13         0.0757       0.052       1.455     0.146     -0.026      0.178
c15        -0.4255       0.058     -7.336     0.000     -0.539     -0.312
c16        -0.5046       0.103     -4.919     0.000     -0.706     -0.303
c17        -0.0792       0.021     -3.707     0.000     -0.121     -0.037
c19         0.3822       0.218       1.756     0.079     -0.045      0.809
c20         0.2279       0.042       5.449     0.000      0.146      0.310
c21        -0.1647       0.049     -3.329     0.001     -0.262     -0.068
c22        -0.1258       0.036     -3.455     0.001     -0.197     -0.054
c23        -0.3301       0.048     -6.848     0.000     -0.425     -0.235
c34        -0.5123       1.765     -0.290     0.772     -3.976      2.951
c35         6.3277       1.616       3.917     0.000      3.157      9.498
c36        -1.8280      90.385     -0.020     0.984    -179.197    175.541
=====

```

```

=====
Omnibus:          40.476      Durbin-Watson:          0.546
Prob(Omnibus):    0.000      Jarque-Bera (JB):       113.185
Skew:             -0.049      Prob(JB):               2.64e-25
Kurtosis:         4.625      Cond. No.               3.48e+06
=====

```

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 3.48e+06. This might indicate that there are

strong multicollinearity or other numerical problems.
MSE = 1.0500374274581643

```
In [4]: #dropping c36
X = df[['c26', 'c27', 'c28', 'c29', 'c30', 'c31', 'c32', 'c33', 'c39', 'c139', 'c140', 'c141', 'c142', 'c143', 'c144', 'c145', 'c146', 'c147', 'c148', 'c149', 'c150', 'c151', 'c152', 'c153', 'c154', 'c155', 'c156', 'c157', 'c158', 'c159', 'c160', 'c161', 'c162', 'c163', 'c164', 'c165', 'c166', 'c167', 'c168', 'c169', 'c170', 'c171', 'c172', 'c173', 'c174', 'c175', 'c176', 'c177', 'c178', 'c179', 'c180', 'c181', 'c182', 'c183', 'c184', 'c185', 'c186', 'c187', 'c188', 'c189', 'c190', 'c191', 'c192', 'c193', 'c194', 'c195', 'c196', 'c197', 'c198', 'c199', 'c200', 'c201', 'c202', 'c203', 'c204', 'c205', 'c206', 'c207', 'c208', 'c209', 'c210', 'c211', 'c212', 'c213', 'c214', 'c215', 'c216', 'c217', 'c218', 'c219', 'c220', 'c221', 'c222', 'c223', 'c224', 'c225', 'c226', 'c227', 'c228', 'c229', 'c230', 'c231', 'c232', 'c233', 'c234', 'c235', 'c236', 'c237', 'c238', 'c239', 'c240', 'c241', 'c242', 'c243', 'c244', 'c245', 'c246', 'c247', 'c248', 'c249', 'c250', 'c251', 'c252', 'c253', 'c254', 'c255', 'c256', 'c257', 'c258', 'c259', 'c260', 'c261', 'c262', 'c263', 'c264', 'c265', 'c266', 'c267', 'c268', 'c269', 'c270', 'c271', 'c272', 'c273', 'c274', 'c275', 'c276', 'c277', 'c278', 'c279', 'c280', 'c281', 'c282', 'c283', 'c284', 'c285', 'c286', 'c287', 'c288', 'c289', 'c290', 'c291', 'c292', 'c293', 'c294', 'c295', 'c296', 'c297', 'c298', 'c299', 'c300', 'c301', 'c302', 'c303', 'c304', 'c305', 'c306', 'c307', 'c308', 'c309', 'c310', 'c311', 'c312', 'c313', 'c314', 'c315', 'c316', 'c317', 'c318', 'c319', 'c320', 'c321', 'c322', 'c323', 'c324', 'c325', 'c326', 'c327', 'c328', 'c329', 'c330', 'c331', 'c332', 'c333', 'c334', 'c335', 'c336', 'c337', 'c338', 'c339', 'c340', 'c341', 'c342', 'c343', 'c344', 'c345', 'c346', 'c347', 'c348', 'c349', 'c350', 'c351', 'c352', 'c353', 'c354', 'c355', 'c356', 'c357', 'c358', 'c359', 'c360', 'c361', 'c362', 'c363', 'c364', 'c365', 'c366', 'c367', 'c368', 'c369', 'c370', 'c371', 'c372', 'c373', 'c374', 'c375', 'c376', 'c377', 'c378', 'c379', 'c380', 'c381', 'c382', 'c383', 'c384', 'c385', 'c386', 'c387', 'c388', 'c389', 'c390', 'c391', 'c392', 'c393', 'c394', 'c395', 'c396', 'c397', 'c398', 'c399', 'c400', 'c401', 'c402', 'c403', 'c404', 'c405', 'c406', 'c407', 'c408', 'c409', 'c410', 'c411', 'c412', 'c413', 'c414', 'c415', 'c416', 'c417', 'c418', 'c419', 'c420', 'c421', 'c422', 'c423', 'c424', 'c425', 'c426', 'c427', 'c428', 'c429', 'c430', 'c431', 'c432', 'c433', 'c434', 'c435', 'c436', 'c437', 'c438', 'c439', 'c440', 'c441', 'c442', 'c443', 'c444', 'c445', 'c446', 'c447', 'c448', 'c449', 'c450', 'c451', 'c452', 'c453', 'c454', 'c455', 'c456', 'c457', 'c458', 'c459', 'c460', 'c461', 'c462', 'c463', 'c464', 'c465', 'c466', 'c467', 'c468', 'c469', 'c470', 'c471', 'c472', 'c473', 'c474', 'c475', 'c476', 'c477', 'c478', 'c479', 'c480', 'c481', 'c482', 'c483', 'c484', 'c485', 'c486', 'c487', 'c488', 'c489', 'c490', 'c491', 'c492', 'c493', 'c494', 'c495', 'c496', 'c497', 'c498', 'c499', 'c500', 'c501', 'c502', 'c503', 'c504', 'c505', 'c506', 'c507', 'c508', 'c509', 'c510', 'c511', 'c512', 'c513', 'c514', 'c515', 'c516', 'c517', 'c518', 'c519', 'c520', 'c521', 'c522', 'c523', 'c524', 'c525', 'c526', 'c527', 'c528', 'c529', 'c530', 'c531', 'c532', 'c533', 'c534', 'c535', 'c536', 'c537', 'c538', 'c539', 'c540', 'c541', 'c542', 'c543', 'c544', 'c545', 'c546', 'c547', 'c548', 'c549', 'c550', 'c551', 'c552', 'c553', 'c554', 'c555', 'c556', 'c557', 'c558', 'c559', 'c560', 'c561', 'c562', 'c563', 'c564', 'c565', 'c566', 'c567', 'c568', 'c569', 'c570', 'c571', 'c572', 'c573', 'c574', 'c575', 'c576', 'c577', 'c578', 'c579', 'c580', 'c581', 'c582', 'c583', 'c584', 'c585', 'c586', 'c587', 'c588', 'c589', 'c590', 'c591', 'c592', 'c593', 'c594', 'c595', 'c596', 'c597', 'c598', 'c599', 'c600', 'c601', 'c602', 'c603', 'c604', 'c605', 'c606', 'c607', 'c608', 'c609', 'c610', 'c611', 'c612', 'c613', 'c614', 'c615', 'c616', 'c617', 'c618', 'c619', 'c620', 'c621', 'c622', 'c623', 'c624', 'c625', 'c626', 'c627', 'c628', 'c629', 'c630', 'c631', 'c632', 'c633', 'c634', 'c635', 'c636', 'c637', 'c638', 'c639', 'c640', 'c641', 'c642', 'c643', 'c644', 'c645', 'c646', 'c647', 'c648', 'c649', 'c650', 'c651', 'c652', 'c653', 'c654', 'c655', 'c656', 'c657', 'c658', 'c659', 'c660', 'c661', 'c662', 'c663', 'c664', 'c665', 'c666', 'c667', 'c668', 'c669', 'c670', 'c671', 'c672', 'c673', 'c674', 'c675', 'c676', 'c677', 'c678', 'c679', 'c680', 'c681', 'c682', 'c683', 'c684', 'c685', 'c686', 'c687', 'c688', 'c689', 'c690', 'c691', 'c692', 'c693', 'c694', 'c695', 'c696', 'c697', 'c698', 'c699', 'c700', 'c701', 'c702', 'c703', 'c704', 'c705', 'c706', 'c707', 'c708', 'c709', 'c710', 'c711', 'c712', 'c713', 'c714', 'c715', 'c716', 'c717', 'c718', 'c719', 'c720', 'c721', 'c722', 'c723', 'c724', 'c725', 'c726', 'c727', 'c728', 'c729', 'c730', 'c731', 'c732', 'c733', 'c734', 'c735', 'c736', 'c737', 'c738', 'c739', 'c740', 'c741', 'c742', 'c743', 'c744', 'c745', 'c746', 'c747', 'c748', 'c749', 'c750', 'c751', 'c752', 'c753', 'c754', 'c755', 'c756', 'c757', 'c758', 'c759', 'c760', 'c761', 'c762', 'c763', 'c764', 'c765', 'c766', 'c767', 'c768', 'c769', 'c770', 'c771', 'c772', 'c773', 'c774', 'c775', 'c776', 'c777', 'c778', 'c779', 'c780', 'c781', 'c782', 'c783', 'c784', 'c785', 'c786', 'c787', 'c788', 'c789', 'c790', 'c791', 'c792', 'c793', 'c794', 'c795', 'c796', 'c797', 'c798', 'c799', 'c800', 'c801', 'c802', 'c803', 'c804', 'c805', 'c806', 'c807', 'c808', 'c809', 'c810', 'c811', 'c812', 'c813', 'c814', 'c815', 'c816', 'c817', 'c818', 'c819', 'c820', 'c821', 'c822', 'c823', 'c824', 'c825', 'c826', 'c827', 'c828', 'c829', 'c830', 'c831', 'c832', 'c833', 'c834', 'c835', 'c836', 'c837', 'c838', 'c839', 'c840', 'c841', 'c842', 'c843', 'c844', 'c845', 'c846', 'c847', 'c848', 'c849', 'c850', 'c851', 'c852', 'c853', 'c854', 'c855', 'c856', 'c857', 'c858', 'c859', 'c860', 'c861', 'c862', 'c863', 'c864', 'c865', 'c866', 'c867', 'c868', 'c869', 'c870', 'c871', 'c872', 'c873', 'c874', 'c875', 'c876', 'c877', 'c878', 'c879', 'c880', 'c881', 'c882', 'c883', 'c884', 'c885', 'c886', 'c887', 'c888', 'c889', 'c890', 'c891', 'c892', 'c893', 'c894', 'c895', 'c896', 'c897', 'c898', 'c899', 'c900', 'c901', 'c902', 'c903', 'c904', 'c905', 'c906', 'c907', 'c908', 'c909', 'c910', 'c911', 'c912', 'c913', 'c914', 'c915', 'c916', 'c917', 'c918', 'c919', 'c920', 'c921', 'c922', 'c923', 'c924', 'c925', 'c926', 'c927', 'c928', 'c929', 'c930', 'c931', 'c932', 'c933', 'c934', 'c935', 'c936', 'c937', 'c938', 'c939', 'c940', 'c941', 'c942', 'c943', 'c944', 'c945', 'c946', 'c947', 'c948', 'c949', 'c950', 'c951', 'c952', 'c953', 'c954', 'c955', 'c956', 'c957', 'c958', 'c959', 'c960', 'c961', 'c962', 'c963', 'c964', 'c965', 'c966', 'c967', 'c968', 'c969', 'c970', 'c971', 'c972', 'c973', 'c974', 'c975', 'c976', 'c977', 'c978', 'c979', 'c980', 'c981', 'c982', 'c983', 'c984', 'c985', 'c986', 'c987', 'c988', 'c989', 'c990', 'c991', 'c992', 'c993', 'c994', 'c995', 'c996', 'c997', 'c998', 'c999', 'c1000', 'c1001', 'c1002', 'c1003', 'c1004', 'c1005', 'c1006', 'c1007', 'c1008', 'c1009', 'c1010', 'c1011', 'c1012', 'c1013', 'c1014', 'c1015', 'c1016', 'c1017', 'c1018', 'c1019', 'c1020', 'c1021', 'c1022', 'c1023', 'c1024', 'c1025']
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

=====						
Dep. Variable:	c52	R-squared:	0.785			
Model:	OLS	Adj. R-squared:	0.777			
Method:	Least Squares	F-statistic:	100.1			
Date:	Sat, 02 Sep 2023	Prob (F-statistic):	1.10e-300			
Time:	22:50:44	Log-Likelihood:	-1479.4			
No. Observations:	1025	AIC:	3033.			
Df Residuals:	988	BIC:	3215.			
Df Model:	36					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	-136.5582	104.637	-1.305	0.192	-341.895	68.779
c26	0.3634	0.049	7.447	0.000	0.268	0.459
c27	-0.1912	0.895	-0.214	0.831	-1.947	1.565
c28	0.2266	0.044	5.156	0.000	0.140	0.313
c29	-0.4453	0.049	-9.127	0.000	-0.541	-0.350
c30	3.4636	0.457	7.578	0.000	2.567	4.361
c31	0.2666	0.035	7.718	0.000	0.199	0.334
c32	0.1779	0.199	0.895	0.371	-0.212	0.568
c33	-0.6542	0.463	-1.413	0.158	-1.563	0.254
c39	12.9990	1.469	8.852	0.000	10.117	15.881
c139	-0.8438	0.225	-3.748	0.000	-1.286	-0.402
c142	0.0454	0.066	0.683	0.495	-0.085	0.176
c143	-0.1537	0.039	-3.958	0.000	-0.230	-0.077
c155	-0.0342	0.013	-2.685	0.007	-0.059	-0.009
c157	0.2502	0.041	6.105	0.000	0.170	0.331
c158	0.2836	0.023	12.128	0.000	0.238	0.329
c160	0.0040	0.002	2.209	0.027	0.000	0.008
c161	0.0105	0.001	9.659	0.000	0.008	0.013
c162	0.0027	0.002	1.650	0.099	-0.001	0.006
c163	0.0081	0.002	3.726	0.000	0.004	0.012
c7	0.3241	0.291	1.113	0.266	-0.247	0.895
c8	-0.4465	0.137	-3.258	0.001	-0.715	-0.178
c9	-0.6863	0.075	-9.104	0.000	-0.834	-0.538
c10	8.8058	1.539	5.724	0.000	5.787	11.825
c11	-0.1705	0.042	-4.058	0.000	-0.253	-0.088
c12	-0.3028	0.109	-2.768	0.006	-0.517	-0.088
c13	0.0757	0.052	1.456	0.146	-0.026	0.178
c15	-0.4255	0.058	-7.349	0.000	-0.539	-0.312
c16	-0.5045	0.103	-4.922	0.000	-0.706	-0.303
c17	-0.0792	0.021	-3.709	0.000	-0.121	-0.037
c19	0.3821	0.217	1.757	0.079	-0.045	0.809
c20	0.2280	0.042	5.454	0.000	0.146	0.310
c21	-0.1648	0.049	-3.331	0.001	-0.262	-0.068
c22	-0.1258	0.036	-3.457	0.001	-0.197	-0.054
c23	-0.3301	0.048	-6.852	0.000	-0.425	-0.236
c34	-0.5115	1.763	-0.290	0.772	-3.972	2.949
c35	6.3269	1.614	3.919	0.000	3.159	9.495
=====						
Omnibus:	40.480	Durbin-Watson:	0.546			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	113.199			
Skew:	-0.049	Prob(JB):	2.62e-25			
Kurtosis:	4.625	Cond. No.	3.47e+06			
=====						

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 3.47e+06. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.050037862638325

```
#dropping c27
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c32', 'c33', 'c39', 'c139', 'c142','c143']]
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.785
Model:                  OLS      Adj. R-squared:          0.777
Method:                 Least Squares      F-statistic:          103.0
Date:                   Sat, 02 Sep 2023    Prob (F-statistic):    1.10e-301
Time:                   22:51:00           Log-Likelihood:       -1479.5
No. Observations:      1025           AIC:                  3031.
Df Residuals:          989           BIC:                  3208.
Df Model:               35
Covariance Type:       nonrobust
=====

```

```

=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const      -156.6963     45.447     -3.448     0.001    -245.880    -67.512
c26         0.3634      0.049      7.451     0.000      0.268      0.459
c28         0.2268      0.044      5.166     0.000      0.141      0.313
c29        -0.4452      0.049     -9.130     0.000     -0.541     -0.349
c30         3.4624      0.457      7.579     0.000      2.566      4.359
c31         0.2668      0.035      7.729     0.000      0.199      0.335
c32         0.1783      0.199      0.897     0.370     -0.212      0.568
c33        -0.6554      0.463     -1.416     0.157     -1.564      0.253
c39        13.0024      1.468      8.859     0.000     10.122     15.883
c139        -0.8412      0.225     -3.744     0.000     -1.282     -0.400
c142         0.0445      0.066      0.671     0.502     -0.086      0.175
c143        -0.1536      0.039     -3.959     0.000     -0.230     -0.077
c155        -0.0341      0.013     -2.682     0.007     -0.059     -0.009
c157         0.2503      0.041      6.111     0.000      0.170      0.331
c158         0.2832      0.023     12.154     0.000      0.237      0.329
c160         0.0040      0.002      2.214     0.027      0.000      0.008
c161         0.0105      0.001      9.687     0.000      0.008      0.013
c162         0.0027      0.002      1.657     0.098     -0.001      0.006
c163         0.0080      0.002      3.722     0.000      0.004      0.012
c7           0.3199      0.290      1.102     0.271     -0.250      0.890
c8          -0.4479      0.137     -3.273     0.001     -0.716     -0.179
c9          -0.6866      0.075     -9.114     0.000     -0.834     -0.539
c10         8.8092      1.538      5.729     0.000      5.792     11.827
c11        -0.1699      0.042     -4.055     0.000     -0.252     -0.088
c12        -0.3042      0.109     -2.787     0.005     -0.518     -0.090
c13         0.0757      0.052      1.458     0.145     -0.026      0.178
c15        -0.4252      0.058     -7.349     0.000     -0.539     -0.312
c16        -0.5034      0.102     -4.920     0.000     -0.704     -0.303
c17        -0.0793      0.021     -3.718     0.000     -0.121     -0.037
c19         0.3806      0.217      1.752     0.080     -0.046      0.807
c20         0.2279      0.042      5.456     0.000      0.146      0.310
c21        -0.1654      0.049     -3.350     0.001     -0.262     -0.069
c22        -0.1257      0.036     -3.457     0.001     -0.197     -0.054
c23        -0.3295      0.048     -6.854     0.000     -0.424     -0.235
c34        -0.5183      1.762     -0.294     0.769     -3.977      2.940
c35         6.3356      1.613      3.928     0.000      3.170      9.501
=====

```

```

=====
Omnibus:              40.889      Durbin-Watson:          0.547
Prob(Omnibus):         0.000      Jarque-Bera (JB):       115.091
Skew:                  -0.050      Prob(JB):               1.02e-25
Kurtosis:               4.639      Cond. No.                1.50e+06
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.5e+06. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0500863911450413

```
In [6]: #dropping c142
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c32', 'c33', 'c39', 'c139', 'c143', 'c145']]
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.785
Model:                  OLS      Adj. R-squared:           0.777
Method:                  Least Squares      F-statistic:          109.4
Date:                    Sat, 02 Sep 2023    Prob (F-statistic):    1.32e-303
Time:                    22:51:08           Log-Likelihood:       -1479.7
No. Observations:        1025           AIC:                  3027.
Df Residuals:            991            BIC:                  3195.
Df Model:                 33
Covariance Type:         nonrobust
=====

```

```

=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const      -154.0761      45.193      -3.409      0.001     -242.761     -65.391
c26          0.3613       0.049       7.433      0.000         0.266         0.457
c28          0.2423       0.035       6.992      0.000         0.174         0.310
c29         -0.4397       0.048      -9.152      0.000        -0.534        -0.345
c30          3.4233       0.453       7.551      0.000         2.534         4.313
c31          0.2823       0.026      10.957      0.000         0.232         0.333
c32          0.1559       0.196       0.795      0.427        -0.229         0.541
c33         -0.6193       0.459      -1.349      0.178        -1.520         0.282
c39          12.9088       1.459       8.845      0.000        10.045        15.773
c139         -0.8394       0.223      -3.758      0.000        -1.278        -0.401
c143         -0.1476       0.038      -3.900      0.000        -0.222        -0.073
c155         -0.0319       0.012      -2.588      0.010        -0.056        -0.008
c157          0.2491       0.040       6.250      0.000         0.171         0.327
c158          0.2808       0.023      12.181      0.000         0.236         0.326
c160          0.0041       0.002       2.276      0.023         0.001         0.008
c161          0.0105       0.001       9.823      0.000         0.008         0.013
c162          0.0028       0.002       1.693      0.091        -0.000         0.006
c163          0.0080       0.002       3.765      0.000         0.004         0.012
c7           0.3284       0.290       1.133      0.258        -0.240         0.897
c8          -0.4400       0.136      -3.235      0.001        -0.707        -0.173
c9          -0.6781       0.074      -9.120      0.000        -0.824        -0.532
c10          8.7586       1.527       5.737      0.000         5.763        11.754
c11         -0.1650       0.041      -3.991      0.000        -0.246        -0.084
c12         -0.3034       0.109      -2.783      0.005        -0.517        -0.089
c13          0.0765       0.052       1.476      0.140        -0.025         0.178
c15         -0.4243       0.058      -7.341      0.000        -0.538        -0.311
c16         -0.4899       0.100      -4.874      0.000        -0.687        -0.293
c17         -0.0816       0.021      -3.880      0.000        -0.123        -0.040
c19          0.3875       0.217       1.787      0.074        -0.038         0.813
c20          0.2240       0.041       5.458      0.000         0.143         0.305
c21         -0.1619       0.049      -3.329      0.001        -0.257        -0.066
c22         -0.1243       0.036      -3.449      0.001        -0.195        -0.054
c23         -0.3229       0.047      -6.873      0.000        -0.415        -0.231
c35          6.3578       1.537       4.136      0.000         3.341         9.374
=====

```

```

=====
Omnibus:                40.558      Durbin-Watson:           0.546
Prob(Omnibus):           0.000      Jarque-Bera (JB):        113.650
Skew:                    -0.048      Prob(JB):                2.09e-25
Kurtosis:                 4.628      Cond. No.                 1.49e+06
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.49e+06. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.050672497782554

```

In [7]: #dropping c32
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c

```



```
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.785
Model:                  OLS      Adj. R-squared:           0.778
Method:                 Least Squares      F-statistic:          112.9
Date:                   Sat, 02 Sep 2023    Prob (F-statistic):    1.69e-304
Time:                   22:51:20           Log-Likelihood:       -1480.1
No. Observations:      1025           AIC:                  3026.
Df Residuals:          992           BIC:                  3189.
Df Model:              32
Covariance Type:       nonrobust
=====

```

```

=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const        -121.0921     17.957     -6.743     0.000    -156.330    -85.854
c26           0.3582       0.048      7.395     0.000      0.263      0.453
c28           0.2395       0.034      6.948     0.000      0.172      0.307
c29          -0.4366       0.048     -9.119     0.000     -0.531     -0.343
c30           3.4079       0.453      7.525     0.000      2.519      4.297
c31           0.2758       0.024     11.293     0.000      0.228      0.324
c33          -0.2600       0.083     -3.139     0.002     -0.423     -0.097
c39          12.9132       1.459      8.850     0.000     10.050     15.777
c139         -0.8435       0.223     -3.778     0.000     -1.282     -0.405
c143         -0.1446       0.038     -3.841     0.000     -0.219     -0.071
c155         -0.0321       0.012     -2.607     0.009     -0.056     -0.008
c157          0.2497       0.040      6.269     0.000      0.172      0.328
c158          0.2835       0.023     12.440     0.000      0.239      0.328
c160          0.0041       0.002      2.305     0.021      0.001      0.008
c161          0.0105       0.001      9.817     0.000      0.008      0.013
c162          0.0028       0.002      1.680     0.093     -0.000      0.006
c163          0.0080       0.002      3.736     0.000      0.004      0.012
c7            0.3649       0.286      1.275     0.203     -0.197      0.927
c8           -0.4215       0.134     -3.146     0.002     -0.684     -0.159
c9           -0.6717       0.074     -9.089     0.000     -0.817     -0.527
c10           8.8019       1.525      5.770     0.000      5.809     11.795
c11          -0.1655       0.041     -4.002     0.000     -0.247     -0.084
c12          -0.3057       0.109     -2.804     0.005     -0.520     -0.092
c13           0.0738       0.052      1.426     0.154     -0.028      0.175
c15          -0.4244       0.058     -7.343     0.000     -0.538     -0.311
c16          -0.4799       0.100     -4.813     0.000     -0.676     -0.284
c17          -0.0818       0.021     -3.890     0.000     -0.123     -0.041
c19           0.3957       0.216      1.828     0.068     -0.029      0.821
c20           0.2264       0.041      5.532     0.000      0.146      0.307
c21          -0.1625       0.049     -3.342     0.001     -0.258     -0.067
c22          -0.1256       0.036     -3.490     0.001     -0.196     -0.055
c23          -0.3173       0.046     -6.832     0.000     -0.408     -0.226
c35           6.2123       1.526      4.071     0.000      3.218      9.207
=====

```

```

=====
Omnibus:              47.547      Durbin-Watson:          0.544
Prob(Omnibus):        0.000      Jarque-Bera (JB):       147.399
Skew:                 -0.071      Prob(JB):               9.83e-33
Kurtosis:              4.852      Cond. No.                5.92e+05
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 5.92e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0513431777242574

```

In [8]: #dropping c7
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c157', 'c158', 'c160', 'c161', 'c162', 'c163', 'c7', 'c8', 'c9', 'c10', 'c11', 'c12', 'c13', 'c15', 'c16', 'c17', 'c19', 'c20', 'c21', 'c22', 'c23', 'c35']]
y = df['c52']

```

```
X = sm.add_constant(X)
12
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.784
Model:                  OLS      Adj. R-squared:         0.777
Method:                 Least Squares      F-statistic:         116.4
Date:                   Sat, 02 Sep 2023    Prob (F-statistic):    3.47e-305
Time:                   22:51:26           Log-Likelihood:       -1480.9
No. Observations:      1025           AIC:                  3026.
Df Residuals:          993           BIC:                  3184.
Df Model:               31
Covariance Type:       nonrobust
=====

```

```

=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const      -126.1233      17.524      -7.197      0.000     -160.511     -91.735
c26         0.3694       0.048       7.750      0.000       0.276       0.463
c28         0.2435       0.034       7.089      0.000       0.176       0.311
c29        -0.4457       0.047      -9.413      0.000      -0.539      -0.353
c30         3.4802       0.449       7.743      0.000       2.598       4.362
c31         0.2872       0.023      12.625      0.000       0.243       0.332
c33        -0.2722       0.082      -3.306      0.001      -0.434      -0.111
c39        13.4486       1.398       9.620      0.000      10.705      16.192
c139        -0.8439       0.223      -3.779      0.000      -1.282      -0.406
c143        -0.1472       0.038      -3.913      0.000      -0.221      -0.073
c155        -0.0364       0.012      -3.062      0.002      -0.060      -0.013
c157         0.2576       0.039       6.541      0.000       0.180       0.335
c158         0.2850       0.023      12.520      0.000       0.240       0.330
c160         0.0040       0.002       2.246      0.025       0.001       0.008
c161         0.0105       0.001       9.774      0.000       0.008       0.013
c162         0.0028       0.002       1.684      0.093      -0.000       0.006
c163         0.0079       0.002       3.696      0.000       0.004       0.012
c8          -0.4519       0.132      -3.426      0.001      -0.711      -0.193
c9          -0.7126       0.067     -10.698      0.000      -0.843      -0.582
c10         9.0072       1.517       5.936      0.000       6.030      11.985
c11        -0.1647       0.041      -3.984      0.000      -0.246      -0.084
c12        -0.3294       0.107      -3.067      0.002      -0.540      -0.119
c13         0.0762       0.052       1.472      0.141      -0.025       0.178
c15        -0.4017       0.055      -7.304      0.000      -0.510      -0.294
c16        -0.4728       0.100      -4.748      0.000      -0.668      -0.277
c17        -0.0810       0.021      -3.852      0.000      -0.122      -0.040
c19         0.3925       0.217       1.813      0.070      -0.032       0.817
c20         0.2340       0.040       5.777      0.000       0.154       0.313
c21        -0.1670       0.049      -3.443      0.001      -0.262      -0.072
c22        -0.1252       0.036      -3.476      0.001      -0.196      -0.055
c23        -0.3121       0.046      -6.744      0.000      -0.403      -0.221
c35         6.4675       1.513       4.274      0.000       3.498       9.437
=====

```

```

=====
Omnibus:                52.925      Durbin-Watson:         0.549
Prob(Omnibus):           0.000      Jarque-Bera (JB):      175.012
Skew:                    -0.097      Prob(JB):              9.92e-39
Kurtosis:                5.015      Cond. No.              5.78e+05
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 5.78e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0530667192484815

```

In [9]: #dropping c13
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c157', 'c158', 'c160', 'c161', 'c162', 'c163', 'c8', 'c9', 'c10', 'c11', 'c12', 'c13', 'c15', 'c16', 'c17', 'c19', 'c20', 'c21', 'c22', 'c23', 'c35']]
y = df['c52']
X = sm.add_constant(X)

```

```
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.784
Model:                  OLS      Adj. R-squared:           0.777
Method:                  Least Squares      F-statistic:           120.0
Date:                    Sat, 02 Sep 2023    Prob (F-statistic):      9.18e-306
Time:                    22:51:36    Log-Likelihood:         -1482.0
No. Observations:        1025      AIC:                    3026.
Df Residuals:            994      BIC:                    3179.
Df Model:                 30
Covariance Type:         nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	-127.9672	17.489	-7.317	0.000	-162.287	-93.647
c26	0.3701	0.048	7.761	0.000	0.277	0.464
c28	0.2502	0.034	7.346	0.000	0.183	0.317
c29	-0.4470	0.047	-9.436	0.000	-0.540	-0.354
c30	3.4847	0.450	7.749	0.000	2.602	4.367
c31	0.2939	0.022	13.179	0.000	0.250	0.338
c33	-0.2835	0.082	-3.458	0.001	-0.444	-0.123
c39	14.0710	1.333	10.554	0.000	11.455	16.687
c139	-0.8308	0.223	-3.721	0.000	-1.269	-0.393
c143	-0.1528	0.037	-4.081	0.000	-0.226	-0.079
c155	-0.0386	0.012	-3.274	0.001	-0.062	-0.015
c157	0.2511	0.039	6.413	0.000	0.174	0.328
c158	0.2832	0.023	12.451	0.000	0.239	0.328
c160	0.0040	0.002	2.252	0.025	0.001	0.008
c161	0.0105	0.001	9.753	0.000	0.008	0.013
c162	0.0026	0.002	1.581	0.114	-0.001	0.006
c163	0.0080	0.002	3.760	0.000	0.004	0.012
c8	-0.4393	0.132	-3.336	0.001	-0.698	-0.181
c9	-0.6975	0.066	-10.591	0.000	-0.827	-0.568
c10	9.0999	1.517	5.999	0.000	6.123	12.077
c11	-0.1666	0.041	-4.028	0.000	-0.248	-0.085
c12	-0.2997	0.106	-2.839	0.005	-0.507	-0.093
c15	-0.4263	0.052	-8.128	0.000	-0.529	-0.323
c16	-0.4015	0.087	-4.612	0.000	-0.572	-0.231
c17	-0.0804	0.021	-3.821	0.000	-0.122	-0.039
c19	0.3955	0.217	1.825	0.068	-0.030	0.821
c20	0.2392	0.040	5.926	0.000	0.160	0.318
c21	-0.1653	0.049	-3.407	0.001	-0.261	-0.070
c22	-0.1371	0.035	-3.904	0.000	-0.206	-0.068
c23	-0.2952	0.045	-6.580	0.000	-0.383	-0.207
c35	6.4374	1.514	4.252	0.000	3.466	9.408

```

=====
Omnibus:                54.090      Durbin-Watson:           0.548
Prob(Omnibus):           0.000      Jarque-Bera (JB):        182.954
Skew:                    -0.093      Prob(JB):                1.87e-40
Kurtosis:                 5.061      Cond. No.                 5.76e+05
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 5.76e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0553661011764075

```

In [10]: #dropping c162
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c157', 'c158', 'c160', 'c161', 'c162', 'c163', 'c8', 'c9', 'c10', 'c11', 'c12', 'c15', 'c16', 'c17', 'c19', 'c20', 'c21', 'c22', 'c23', 'c35']]
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()

```

```
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```
=====
Dep. Variable:          c52      R-squared:                0.783
Model:                  OLS      Adj. R-squared:           0.777
Method:                 Least Squares      F-statistic:         123.9
Date:                   Sat, 02 Sep 2023    Prob (F-statistic):    2.82e-306
Time:                   22:51:47           Log-Likelihood:       -1483.3
No. Observations:      1025           AIC:                  3027.
Df Residuals:          995           BIC:                  3175.
Df Model:               29
Covariance Type:       nonrobust
=====
```

	coef	std err	t	P> t	[0.025	0.975]
const	-124.1923	17.338	-7.163	0.000	-158.216	-90.168
c26	0.3611	0.047	7.621	0.000	0.268	0.454
c28	0.2447	0.034	7.216	0.000	0.178	0.311
c29	-0.4380	0.047	-9.306	0.000	-0.530	-0.346
c30	3.3966	0.447	7.606	0.000	2.520	4.273
c31	0.2933	0.022	13.145	0.000	0.249	0.337
c33	-0.2796	0.082	-3.408	0.001	-0.441	-0.119
c39	14.3943	1.319	10.917	0.000	11.807	16.982
c139	-0.8292	0.223	-3.711	0.000	-1.268	-0.391
c143	-0.1442	0.037	-3.890	0.000	-0.217	-0.071
c155	-0.0387	0.012	-3.283	0.001	-0.062	-0.016
c157	0.2565	0.039	6.573	0.000	0.180	0.333
c158	0.2832	0.023	12.445	0.000	0.239	0.328
c160	0.0039	0.002	2.171	0.030	0.000	0.007
c161	0.0109	0.001	10.639	0.000	0.009	0.013
c163	0.0086	0.002	4.082	0.000	0.004	0.013
c8	-0.4561	0.131	-3.472	0.001	-0.714	-0.198
c9	-0.6991	0.066	-10.609	0.000	-0.828	-0.570
c10	8.9883	1.516	5.927	0.000	6.013	11.964
c11	-0.1629	0.041	-3.942	0.000	-0.244	-0.082
c12	-0.3121	0.105	-2.962	0.003	-0.519	-0.105
c15	-0.4319	0.052	-8.249	0.000	-0.535	-0.329
c16	-0.4031	0.087	-4.628	0.000	-0.574	-0.232
c17	-0.0792	0.021	-3.764	0.000	-0.121	-0.038
c19	0.3967	0.217	1.829	0.068	-0.029	0.822
c20	0.2357	0.040	5.843	0.000	0.157	0.315
c21	-0.1610	0.048	-3.320	0.001	-0.256	-0.066
c22	-0.1381	0.035	-3.930	0.000	-0.207	-0.069
c23	-0.3005	0.045	-6.711	0.000	-0.388	-0.213
c35	6.4788	1.515	4.277	0.000	3.506	9.452

```
=====
Omnibus:                53.166      Durbin-Watson:          0.547
Prob(Omnibus):          0.000      Jarque-Bera (JB):       179.242
Skew:                   -0.079      Prob(JB):               1.20e-39
Kurtosis:               5.042      Cond. No.                5.63e+05
=====
```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 5.63e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0580193257108284

```
In [11]: #dropping c19
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c156']]
y = df['c52']
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```


OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.782
Model:                  OLS      Adj. R-squared:         0.776
Method:                 Least Squares      F-statistic:         127.9
Date:                   Sat, 02 Sep 2023    Prob (F-statistic):    1.30e-306
Time:                   22:51:50           Log-Likelihood:       -1485.0
No. Observations:      1025           AIC:                  3028.
Df Residuals:          996           BIC:                  3171.
Df Model:              28
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	-125.8111	17.336	-7.257	0.000	-159.831	-91.791
c26	0.3663	0.047	7.734	0.000	0.273	0.459
c28	0.2515	0.034	7.454	0.000	0.185	0.318
c29	-0.4401	0.047	-9.342	0.000	-0.532	-0.348
c30	3.4022	0.447	7.610	0.000	2.525	4.280
c31	0.2953	0.022	13.236	0.000	0.252	0.339
c33	-0.2699	0.082	-3.293	0.001	-0.431	-0.109
c39	14.6285	1.314	11.134	0.000	12.050	17.207
c139	-0.4288	0.045	-9.546	0.000	-0.517	-0.341
c143	-0.1502	0.037	-4.063	0.000	-0.223	-0.078
c155	-0.0418	0.012	-3.576	0.000	-0.065	-0.019
c157	0.2573	0.039	6.587	0.000	0.181	0.334
c158	0.2832	0.023	12.429	0.000	0.239	0.328
c160	0.0039	0.002	2.152	0.032	0.000	0.007
c161	0.0110	0.001	10.690	0.000	0.009	0.013
c163	0.0087	0.002	4.141	0.000	0.005	0.013
c8	-0.4440	0.131	-3.381	0.001	-0.702	-0.186
c9	-0.6920	0.066	-10.507	0.000	-0.821	-0.563
c10	8.9184	1.518	5.876	0.000	5.940	11.897
c11	-0.1655	0.041	-4.004	0.000	-0.247	-0.084
c12	-0.3099	0.105	-2.939	0.003	-0.517	-0.103
c15	-0.4338	0.052	-8.278	0.000	-0.537	-0.331
c16	-0.4095	0.087	-4.699	0.000	-0.580	-0.238
c17	-0.0800	0.021	-3.797	0.000	-0.121	-0.039
c20	0.2353	0.040	5.827	0.000	0.156	0.315
c21	-0.1557	0.048	-3.213	0.001	-0.251	-0.061
c22	-0.1296	0.035	-3.717	0.000	-0.198	-0.061
c23	-0.3024	0.045	-6.746	0.000	-0.390	-0.214
c35	6.4165	1.516	4.232	0.000	3.441	9.392

```

=====
Omnibus:                53.080      Durbin-Watson:         0.546
Prob(Omnibus):          0.000      Jarque-Bera (JB):      178.982
Skew:                   -0.077      Prob(JB):              1.36e-39
Kurtosis:               5.041      Cond. No.              5.63e+05
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 5.63e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0615776950310878

```

In [12]: #dropping c160
X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c157', 'c158', 'c161', 'c163', 'c8', 'c9', 'c10', 'c11', 'c12', 'c15', 'c16', 'c17', 'c20', 'c21', 'c22', 'c23', 'c35']]
y = df['c52']
19
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())

```

```

y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)

```

OLS Regression Results

```

=====
Dep. Variable:          c52      R-squared:                0.781
Model:                  OLS      Adj. R-squared:           0.775
Method:                 Least Squares      F-statistic:        132.0
Date:                  Sat, 02 Sep 2023      Prob (F-statistic):    1.11e-306
Time:                  22:51:56      Log-Likelihood:       -1487.4
No. Observations:      1025      AIC:                  3031.
Df Residuals:          997      BIC:                  3169.
Df Model:               27
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	-120.2900	17.177	-7.003	0.000	-153.996	-86.584
c26	0.3648	0.047	7.690	0.000	0.272	0.458
c28	0.2579	0.034	7.659	0.000	0.192	0.324
c29	-0.4436	0.047	-9.406	0.000	-0.536	-0.351
c30	3.3909	0.448	7.571	0.000	2.512	4.270
c31	0.2977	0.022	13.334	0.000	0.254	0.341
c33	-0.2494	0.082	-3.058	0.002	-0.409	-0.089
c39	14.9365	1.308	11.416	0.000	12.369	17.504
c139	-0.4333	0.045	-9.638	0.000	-0.521	-0.345
c143	-0.1567	0.037	-4.245	0.000	-0.229	-0.084
c155	-0.0414	0.012	-3.535	0.000	-0.064	-0.018
c157	0.2718	0.039	7.051	0.000	0.196	0.347
c158	0.2805	0.023	12.305	0.000	0.236	0.325
c161	0.0116	0.001	11.651	0.000	0.010	0.014
c163	0.0091	0.002	4.342	0.000	0.005	0.013
c8	-0.4372	0.132	-3.324	0.001	-0.695	-0.179
c9	-0.6982	0.066	-10.592	0.000	-0.828	-0.569
c10	8.7267	1.518	5.749	0.000	5.748	11.705
c11	-0.1709	0.041	-4.134	0.000	-0.252	-0.090
c12	-0.3186	0.106	-3.018	0.003	-0.526	-0.111
c15	-0.4385	0.052	-8.358	0.000	-0.541	-0.336
c16	-0.4218	0.087	-4.842	0.000	-0.593	-0.251
c17	-0.0742	0.021	-3.546	0.000	-0.115	-0.033
c20	0.2305	0.040	5.707	0.000	0.151	0.310
c21	-0.1491	0.048	-3.077	0.002	-0.244	-0.054
c22	-0.1376	0.035	-3.961	0.000	-0.206	-0.069
c23	-0.2999	0.045	-6.681	0.000	-0.388	-0.212
c35	6.6845	1.514	4.415	0.000	3.714	9.655

```

=====
Omnibus:                54.027      Durbin-Watson:          0.548
Prob(Omnibus):           0.000      Jarque-Bera (JB):       188.218
Skew:                    -0.049      Prob(JB):               1.35e-41
Kurtosis:                 5.097      Cond. No.                3.70e+05
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 3.7e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0665144562702324

```

In [14]: X = df[['c26', 'c28', 'c29', 'c30', 'c31', 'c33', 'c39', 'c139', 'c143', 'c155', 'c161', 'c163', 'c8', 'c9', 'c10', 'c11', 'c12', 'c15', 'c16', 'c17', 'c20', 'c21', 'c22', 'c23', 'c35']]
y = df['c52']

```

```
X = sm.add_constant(X)
mlr_model = sm.OLS(y, X).fit()
print(mlr_model.summary())
y_cap = mlr_model.predict(X)
e = y - y_cap
e_squared = e * e
MSE = (e_squared.sum())/1025
print('MSE = ', MSE)
```

OLS Regression Results

```
=====
Dep. Variable:          c52      R-squared:                0.781
Model:                  OLS      Adj. R-squared:           0.775
Method:                  Least Squares      F-statistic:          132.0
Date:                    Sat, 02 Sep 2023     Prob (F-statistic):      1.11e-306
Time:                    22:52:32     Log-Likelihood:         -1487.4
No. Observations:        1025      AIC:                   3031.
Df Residuals:            997      BIC:                   3169.
Df Model:                 27
Covariance Type:         nonrobust
=====
```

	coef	std err	t	P> t	[0.025	0.975]
const	-120.2900	17.177	-7.003	0.000	-153.996	-86.584
c26	0.3648	0.047	7.690	0.000	0.272	0.458
c28	0.2579	0.034	7.659	0.000	0.192	0.324
c29	-0.4436	0.047	-9.406	0.000	-0.536	-0.351
c30	3.3909	0.448	7.571	0.000	2.512	4.270
c31	0.2977	0.022	13.334	0.000	0.254	0.341
c33	-0.2494	0.082	-3.058	0.002	-0.409	-0.089
c39	14.9365	1.308	11.416	0.000	12.369	17.504
c139	-0.4333	0.045	-9.638	0.000	-0.521	-0.345
c143	-0.1567	0.037	-4.245	0.000	-0.229	-0.084
c155	-0.0414	0.012	-3.535	0.000	-0.064	-0.018
c157	0.2718	0.039	7.051	0.000	0.196	0.347
c158	0.2805	0.023	12.305	0.000	0.236	0.325
c161	0.0116	0.001	11.651	0.000	0.010	0.014
c163	0.0091	0.002	4.342	0.000	0.005	0.013
c8	-0.4372	0.132	-3.324	0.001	-0.695	-0.179
c9	-0.6982	0.066	-10.592	0.000	-0.828	-0.569
c10	8.7267	1.518	5.749	0.000	5.748	11.705
c11	-0.1709	0.041	-4.134	0.000	-0.252	-0.090
c12	-0.3186	0.106	-3.018	0.003	-0.526	-0.111
c15	-0.4385	0.052	-8.358	0.000	-0.541	-0.336
c16	-0.4218	0.087	-4.842	0.000	-0.593	-0.251
c17	-0.0742	0.021	-3.546	0.000	-0.115	-0.033
c20	0.2305	0.040	5.707	0.000	0.151	0.310
c21	-0.1491	0.048	-3.077	0.002	-0.244	-0.054
c22	-0.1376	0.035	-3.961	0.000	-0.206	-0.069
c23	-0.2999	0.045	-6.681	0.000	-0.388	-0.212
c35	6.6845	1.514	4.415	0.000	3.714	9.655

```
=====
Omnibus:                54.027      Durbin-Watson:          0.548
Prob(Omnibus):           0.000      Jarque-Bera (JB):       188.218
Skew:                    -0.049      Prob(JB):               1.35e-41
Kurtosis:                 5.097      Cond. No.                3.70e+05
=====
```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 3.7e+05. This might indicate that there are strong multicollinearity or other numerical problems.

MSE = 1.0665144562702324

The variables with higher coefficients are more significant to the model.

That is c30, c39, c10 and c35 represent the independent variables which have the most impact on the model.

This is also evident from the fact that they have the lowest p values.

In [4]:

In []: