# **Customer Churn Linear Regression**

```
import numpy as np
import pandas as pd
import scipy.stats
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn import linear_model
import statsmodels.api as sm
import statsmodels.formula.api as smf
from scipy import stats
```

# **Analysis Goals:**

The goal of this analysis is to determine which, if any, of the variables collected can help predict a customer's longevity with the company. Since relationships may not be immediately clear, I have decided to include a large number of variables in my initial model.

# Reading data

```
In [8]: data= pd.read_csv('C:/Users/cynth/OneDrive/Documents/MS Data Analytics/Clean_Data_Custo
    data.info()
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8950 entries, 0 to 8949
Data columns (total 40 columns);

Data	columns (total 49 columns):		
#	Column	Non-Null Count	Dtype
0	Unnamed: 0	8950 non-null	int64
1	Lat	8950 non-null	float64
2	Lng	8950 non-null	float64
3	Population	8950 non-null	int64
4	Children	8950 non-null	int64
5	Age	8950 non-null	int64
6	Income	8950 non-null	float64
7	Outage_sec_perweek	8950 non-null	float64
8	Email	8950 non-null	int64
9	Contacts	8950 non-null	int64
10	Yearly_equip_failure	8950 non-null	int64
11	Techie	8950 non-null	int64
12	Port_modem	8950 non-null	int64
13	Tablet	8950 non-null	int64
14	Phone	8950 non-null	int64
15	Multiple	8950 non-null	int64
16	OnlineSecurity	8950 non-null	int64
17	OnlineBackup	8950 non-null	int64
18	DeviceProtection	8950 non-null	int64
19	TechSupport	8950 non-null	int64
20	StreamingTV	8950 non-null	int64
21	StreamingMovies	8950 non-null	int64
22	PaperlessBilling	8950 non-null	int64
23	Tenure	8950 non-null	float64
24	MonthlyCharge	8950 non-null	float64
25	Bandwidth_GB_Year	8950 non-null	float64
26	Item1	8950 non-null	int64

localhost:8889/lab 1/78

	27	Item2	8950	non-null	int64
	28	Item3	8950	non-null	int64
	29	Item4	8950	non-null	int64
	30	Item5	8950	non-null	int64
	31	Item6	8950	non-null	int64
	32	Item7	8950	non-null	int64
	33	Item8	8950	non-null	int64
	34	Area_Suburban	8950	non-null	int64
	35	Area_Rural	8950	non-null	int64
	36	Marital_Married	8950	non-null	int64
	37	Marital_Separated	8950	non-null	int64
	38	Marital_NeverMarried	8950	non-null	int64
	39	Marital_Divorced	8950	non-null	int64
	40	Gender_Female	8950	non-null	int64
	41	Gender_Nonbinary	8950	non-null	int64
	42	Contract_Oneyear	8950	non-null	int64
	43	Contract_TwoYear	8950	non-null	int64
	44	InternetService_DSL	8950	non-null	int64
	45	InternetService_None	8950	non-null	int64
	46	PaymentMethod_BankTransferautomatic	8950	non-null	int64
	47	PaymentMethod_MailedCheck	8950	non-null	int64
	48	PaymentMethod_ElectronicCheck	8950	non-null	int64
(	dtype	es: float64(7), int64(42)			

memory usage: 3.3 MB

# Performing Stepwise Backward Multiple Linear Regression

```
lm1 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage sec perweek+Email
In [9]:
         lm1.params
         print(lm1.summary())
```

```
OLS Regression Results
______
Dep. Variable: Tenure R-squared:
                                                  1.000
               OLS Adj. R-squared:
Least Squares F-statistic:
Model:
                                                   1.000
                                               1.153e+07
Method:
Date:
             Sat, 12 Feb 2022 Prob (F-statistic):
Time:
                    23:36:24
                           Log-Likelihood:
                                                  7287.4
No. Observations:
                       8950
                            AIC:
                                               -1.448e+04
Df Residuals:
                       8902
                            BIC:
                                                -1.414e+04
Df Model:
Covariance Type: nonrobust
______
                             coef std err
                                                  P>|t|
                                                            [0.
     0.975]
                            1.1090
                                   0.023 48.233 0.000
Intercept
                                                            1.
064
   1.154
                                   0.000 -0.291 0.771
                         -7.134e-05
                                                            -0.
Lat
001
       0.000
                         -4.473e-05 8.52e-05 -0.525 0.599
Lng
                                                            -0.
000
      0.000
                         -4.127e-08 9.91e-08 -0.416
                                                    0.677 -2.36e
Population
    1.53e-07
Children
                           -0.3758
                                    0.001 -623.691
                                                    0.000
                                                            -0.
377
   -0.375
                                                    0.000
Age
                            0.0400
                                 5.52e-05
                                           724.132
                                                            0.
     0.040
040
```

2/78 localhost:8889/lab

	CustomerChum_	LinearRegression	I		
Income -07 5.57e-08	-3.315e-08	4.53e-08	-0.731	0.465	-1.22e
Outage_sec_perweek	0.0003	0.000	0.829	0.407	-0.
000 0.001 Email	-0.0001	0.000	-0.351	0.726	-0.
001 0.001 Contacts	-0.0003	0.001	-0.252	0.801	-0.
003 0.002		0.002		0.297	
Yearly_equip_failure 006 0.002	-0.0020		-1.044		-0.
Techie 005 0.007	0.0014	0.003	0.472	0.637	-0.
Port_modem 001 0.008	0.0034	0.002	1.505	0.132	-0.
Tablet	0.0003	0.002	0.123	0.902	-0.
005 0.005 Phone	0.0002	0.004	0.043	0.965	-0.
007 0.008 Multiple	0.2650	0.005	55.195	0.000	0.
256 0.274 OnlineSecurity	-0.8309	0.002	-346.068	0.000	-0.
836 -0.826					
OnlineBackup 364 -0.349	-0.3565	0.004	-96.335	0.000	-0.
DeviceProtection 603 -0.592	-0.5972	0.003	-212.643	0.000	-0.
TechSupport	0.3828	0.003	133.496	0.000	0.
377 0.388 StreamingTV	-1.3013	0.006	-219.647	0.000	-1.
313 -1.290 StreamingMovies	-0.7267	0.007	-101.666	0.000	-0.
741 -0.713 PaperlessBilling	-0.0045	0.002	-1.957	0.050	-0.
009 7.48e-06					
MonthlyCharge 035 -0.035	-0.0351	0.000	-271.437	0.000	-0.
Bandwidth_GB_Year 012 0.012	0.0122	5.26e-07	2.32e+04	0.000	0.
Item1 0.006	0.0024	0.002	1.458	0.145	-0.
Item2	-0.0013	0.002	-0.817	0.414	-0.
004 0.002 Item3	0.0012	0.001	0.878	0.380	-0.
002 0.004 Item4	0.0010	0.001	0.794	0.427	-0.
001 0.003					
Item5 003 0.002	-0.0004	0.001	-0.289	0.772	-0.
Item6 004 0.002	-0.0010	0.001	-0.767	0.443	-0.
Item7 004 0.001	-0.0010	0.001	-0.814	0.415	-0.
Item8	-1.331e-05	0.001	-0.011	0.991	-0.
002 0.002 Area_Suburban	-0.0038	0.003	-1.350	0.177	-0.
009 0.002 Area_Rural	0.0033	0.003	1.187	0.235	-0.
002 0.009	-0.0015	0.004			
Marital_Married 009 0.006			-0.425	0.671	-0.
Marital_Separated 004 0.010	0.0027	0.004	0.742	0.458	-0.
Marital_NeverMarried 007 0.007	-0.0003	0.004	-0.075	0.940	-0.
Marital_Divorced	0.0003	0.004	0.076	0.940	-0.

localhost:8889/lab 3/78

007

0 007

007 0.007						
Gender_Female		0.7924	0.002	343.195	0.000	0.
788 0.797						
Gender_Nonbinary		1.0564	0.008	137.711	0.000	1.
041 1.071						
Contract_Oneyear		0.0015	0.003	0.503	0.615	-0.
004 0.007						
Contract_TwoYear		0.0011	0.003	0.403	0.687	-0.
004 0.007						
<pre>InternetService_DSL</pre>		-5.7518	0.004	-1570.145	0.000	-5.
759 -5.745						
<pre>InternetService_None</pre>		-1.1498	0.005	-221.595	0.000	-1.
160 -1.140						
PaymentMethod_BankTransfer	automatic	-0.0009	0.003	-0.245	0.806	-0.
008 0.006						
PaymentMethod_MailedCheck		0.0061	0.003	1.767	0.077	-0.
001 0.013						
PaymentMethod_ElectronicCh	eck	0.0017	0.003	0.530	0.596	-0.
005 0.008						
	========	========	=======	=======	=======	
Omnibus:	31231.936	Durbin-Wa	tson:		2.006	
Prob(Omnibus):	0.000	Jarque-Be	ra (JB):		1459.017	
Skew:	-0.029	Prob(JB):			0.00	
Kurtosis:	1.023	Cond. No.			9.87e+05	
	========	========	=======	=======	=======	

#### Notes:

Dep. Variable:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 9.87e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [10]:

```
#Remove item 8
lm2 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage_sec_perweek+Email
lm2.params
print(lm2.summary())
```

1.000

### OLS Regression Results \_\_\_\_\_\_

Tenure R-squared:

Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	0LS Least Squares Sat, 12 Feb 2022 23:36:24 8950 8903 46 nonrobust	F-statis Prob (F	squared: stic: -statistic):	-1.	1.000 178e+07 0.00 7287.4 448e+04 415e+04	
	=======================================	=======			=======	======
025 0.975]		coef	std err	t	P> t	[0.
Intercept		1.1089	0.023	48.562	0.000	1.
064 1.154 Lat	-7	.128e-05	0.000	-0.291	0.771	-0.
001 0.000 Lng 000 0.000	-4	.473e-05	8.51e-05	-0.525	0.599	-0.
Population -07 1.53e-07	-4	.127e-08	9.91e-08	-0.416	0.677	-2.36e
Children		-0.3758	0.001	-623.759	0.000	-0.

localhost:8889/lab 4/78

		Gustomer Gham_	Linearragicooloi			
377 Age	-0.375	0.0400	5.52e-05	724.246	0.000	0.
040	0.040	0.0400	3.320 03	724,240	0.000	٠.
Income		-3.315e-08	4.53e-08	-0.731	0.465	-1.22e
	.57e-08	0.0003	0.000	0.829	0.407	0
000 000	ec_perweek 0.001	0.0003	0.000	0.829	0.407	-0.
Email		-0.0001	0.000	-0.351	0.726	-0.
001	0.001					
Contacts 003	0.002	-0.0003	0.001	-0.252	0.801	-0.
	quip_failure	-0.0020	0.002	-1.044	0.297	-0.
006	0.002					
Techie	0.007	0.0014	0.003	0.472	0.637	-0.
005 Port_mode	0.007	0.0034	0.002	1.505	0.132	-0.
001	0.008	0.0034	0.002	1.505	0.132	٠.
Tablet		0.0003	0.002	0.123	0.902	-0.
005	0.005	0.0002	0.004	0.043	0.065	0
Phone 007	0.008	0.0002	0.004	0.043	0.965	-0.
Multiple		0.2650	0.005	55.198	0.000	0.
256	0.274					
OnlineSec 836	curity -0.826	-0.8309	0.002	-346.099	0.000	-0.
OnlineBac		-0.3565	0.004	-96.341	0.000	-0.
364	-0.349					
DevicePro		-0.5972	0.003	-212.666	0.000	-0.
603 TechSuppo	-0.592	0.3828	0.003	133.504	0.000	0.
377	0.388	0.3020	0.003	133.304	0.000	0.
Streaming	•	-1.3013	0.006	-219.668	0.000	-1.
313	-1.290	0.7267	0 007	101 674	0.000	0
Streaming 741	-0.713	-0.7267	0.007	-101.674	0.000	-0.
Paperless		-0.0045	0.002	-1.957	0.050	-0.
	.49e-06	0.0054		074 454		
MonthlyCh 035	narge -0.035	-0.0351	0.000	-271.454	0.000	-0.
Bandwidth		0.0122	5.26e-07	2.32e+04	0.000	0.
012	0.012					
Item1 001	0 006	0.0024	0.002	1.464	0.143	-0.
Item2	0.006	-0.0013	0.002	-0.819	0.413	-0.
004	0.002					
Item3	0.004	0.0012	0.001	0.878	0.380	-0.
002 Item4	0.004	0.0010	0.001	0.795	0.427	-0.
001	0.003	0.0010	0.001	0.755	0.427	٠.
Item5		-0.0004	0.001	-0.289	0.772	-0.
003 Item6	0.002	-0.0010	0.001	-0.774	0.439	-0.
004	0.002	-0.0010	0.001	-0.774	0.439	-0.
Item7		-0.0010	0.001	-0.818	0.413	-0.
004	0.001	0.0020	0.003	1 350	0 177	0
Area_Subu 009	urban 0.002	-0.0038	0.003	-1.350	0.177	-0.
Area_Rura		0.0033	0.003	1.187	0.235	-0.
002	0.009					
Marital_M 009	Married 0.006	-0.0015	0.004	-0.425	0.671	-0.
Marital_S		0.0027	0.004	0.743	0.458	-0.
004	0.010					
_	NeverMarried	-0.0003	0.004	-0.075	0.940	-0.
007	0.007					

localhost:8889/lab 5/78

Marital_Divorced	0.0003	0.004	0.076	0.939	-0.
007 0.007	0.7004		242 244		
Gender_Female	0.7924	0.002	343.241	0.000	0.
788 0.797	1 0564	0 000	127 726	0.000	1
Gender_Nonbinary	1.0564	0.008	137.726	0.000	1.
041 1.071	0 0015	0 002	0 504	0 (15	0
Contract_Oneyear	0.0015	0.003	0.504	0.615	-0.
004 0.007	0 0011	0 002	0.404	0 606	0
Contract_TwoYear 004 0.007	0.0011	0.003	0.404	0.686	-0.
InternetService DSL	-5.7518	0.004	-1570.349	0.000	-5.
759 -5.745	-3.7310	0.004	-1370.343	0.000	-5.
InternetService_None	-1.1498	0.005	-221.611	0.000	-1.
160 -1.140	111130	0.003		0.000	
PaymentMethod_BankTransferautomatic	-0.0009	0.003	-0.245	0.806	-0.
008 0.006					
PaymentMethod_MailedCheck	0.0061	0.003	1.767	0.077	-0.
001 0.013					
PaymentMethod_ElectronicCheck	0.0017	0.003	0.530	0.596	-0.
005 0.008					
Omnibus: 31231.935	======== Durbin-Wa1	====== :son:	=======	2.006	
Prob(Omnibus): 0.000				1459.017	
Skew: -0.029		- ().		0.00	
Kurtosis: 1.023	, ,			9.81e+05	
	========		=======	=======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 9.81e+05. This might indicate that there are strong multicollinearity or other numerical problems.

## In [11]:

#Remove phone

lm3 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage\_sec\_perweek+Email
lm3.params
print(lm3.summary())

### OLS Regression Results

Dep. Var Model: Method: Date: Time: No. Obse Df Resid	rvations:	Tenure OLS Least Squares Sat, 12 Feb 2022 23:36:24 8950 8904	Adj. R- F-stati Prob (F Log-Lik AIC:	squared:	-1	1.000 1.000 .204e+07 0.00 7287.4 .448e+04 .416e+04	
Df Model	:	45					
Covarian	ce Type:	nonrobust					
025	  0.975]		coef	std err	t	P> t	[0.
Intercep	 t		1.1091	0.023	49.209	0.000	1.
065	1.153						
Lat		-7	7.129e-05	0.000	-0.291	0.771	-0.
001	0.000						
Lng		- 4	4.472e-05	8.51e-05	-0.525	0.599	-0.
000	0.000						
Populati	on	- 4	4.124e-08	9.91e-08	-0.416	0.677	-2.36e
-07 1	.53e-07						

localhost:8889/lab

	CustomerChurn_	LinearRegressior	1		
Children	-0.3758	0.001	-623.795	0.000	-0.
377 -0.375 Age	0.0400	5.52e-05	724.346	0.000	0.
040 0.040 Income	-3.316e-08	4.53e-08	-0.732	0.464	-1.22e
-07 5.57e-08	0.0003	0.000	0.020	0 407	0
Outage_sec_perweek 000 0.001	0.0003	0.000	0.829	0.407	-0.
Email 0.001	-0.0001	0.000	-0.351	0.725	-0.
Contacts 003 0.002	-0.0003	0.001	-0.252	0.801	-0.
Yearly_equip_failure 006 0.002	-0.0020	0.002	-1.045	0.296	-0.
Techie 005 0.007	0.0014	0.003	0.472	0.637	-0.
Port_modem 001 0.008	0.0034	0.002	1.506	0.132	-0.
Tablet 0.005	0.0003	0.002	0.124	0.902	-0.
Multiple 256 0.274	0.2650	0.005	55.215	0.000	0.
OnlineSecurity 836 -0.826	-0.8309	0.002	-346.123	0.000	-0.
OnlineBackup 364 -0.349	-0.3565	0.004	-96.353	0.000	-0.
DeviceProtection 603 -0.592	-0.5972	0.003	-212.693	0.000	-0.
TechSupport 377 0.388	0.3828	0.003	133.516	0.000	0.
StreamingTV 313 -1.290	-1.3013	0.006	-219.700	0.000	-1.
StreamingMovies 741 -0.713	-0.7267	0.007	-101.694	0.000	-0.
PaperlessBilling 009 5.91e-06	-0.0045	0.002	-1.958	0.050	-0.
MonthlyCharge 035 -0.035	-0.0351	0.000	-271.524	0.000	-0.
Bandwidth_GB_Year 012 0.012	0.0122	5.26e-07	2.32e+04	0.000	0.
Item1 001 0.006	0.0024	0.002	1.464	0.143	-0.
Item2 004 0.002	-0.0013	0.002	-0.818	0.414	-0.
Item3 002 0.004	0.0012	0.001	0.878	0.380	-0.
Item4 001 0.003	0.0010	0.001	0.795	0.427	-0.
Item5 003 0.002	-0.0004	0.001	-0.289	0.773	-0.
Item6 004 0.002	-0.0010	0.001	-0.775	0.439	-0.
Item7 004 0.001	-0.0010	0.001	-0.818	0.413	-0.
Area_Suburban 009 0.002	-0.0038	0.003	-1.350	0.177	-0.
Area_Rural 002 0.009	0.0033	0.003	1.188	0.235	-0.
Marital_Married 009 0.006	-0.0015	0.004	-0.425	0.671	-0.
Marital_Separated 004 0.010	0.0027	0.004	0.743	0.458	-0.
Marital_NeverMarried 007 0.007	-0.0003	0.004	-0.075	0.940	-0.
Marital_Divorced	0.0003	0.004	0.076	0.939	-0.

localhost:8889/lab 7/78

007 0.007						
Gender_Female		0.7924	0.002	343.264	0.000	0.
788 0.797						
Gender_Nonbinary		1.0564	0.008	137.734	0.000	1.
041 1.071						
Contract_Oneyear		0.0015	0.003	0.504	0.614	-0.
004 0.007						
Contract_TwoYear		0.0011	0.003	0.404	0.686	-0.
0.007						_
InternetService_DSL		-5.7518	0.004	-1570.585	0.000	-5.
759 -5.745		1 1400	0 005	221 620	0.000	1
<pre>InternetService_None 160 -1.140</pre>		-1.1498	0.005	-221.630	0.000	-1.
PaymentMethod BankTransfera	utomatic	-0.0009	0.003	-0.245	0.806	-0.
008 0.006	automatic	-0.0003	0.003	-0.243	0.800	-0.
PaymentMethod_MailedCheck		0.0061	0.003	1.767	0.077	-0.
001 0.013		0.0001	0.003	1.707	0.077	٠.
PaymentMethod_ElectronicChe	eck	0.0017	0.003	0.531	0.596	-0.
005 0.008						
=======================================		========	=======	=======	=======	
Omnibus:	31231.907	Durbin-Wa	tson:		2.006	
Prob(Omnibus):	0.000	Jarque-Be	ra (JB):		1459.020	
Skew:	-0.029	Prob(JB):			0.00	
Kurtosis:	1.023	Cond. No.			9.70e+05	
=======================================		========	======	=======	=======	

Dep. Variable:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 9.7e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [12]:

```
#Remove Marital NeverMarried
lm5 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage_sec_perweek+Email
1m5.params
print(lm5.summary())
```

R-squared:

1.000

### OLS Regression Results \_\_\_\_\_\_

Tenure

Model: Method: Date: Time: No. Observations: Df Residuals:	0LS Least Squares Sat, 12 Feb 2022 23:36:24 8950 8905	F-statis Prob (F	stic: -statistic):	-1.	1.000 232e+07 0.00 7287.4 448e+04 417e+04	
Df Model: Covariance Type:	44 nonrobust					
025 0.975]		coef	std err	t	P> t	[0.
Intercept		1.1090	0.022	49.332	0.000	1.
065 1.153 Lat 001 0.000		7.14e-05	0.000	-0.292	0.771	-0.
Lng 000 0.000 Population			8.51e-05 9.91e-08	-0.524 -0.417	0.600 0.677	-0.
-07 1.53e-07 Children		-0.3758	0.001	-623.841	0.000	-0.

localhost:8889/lab 8/78

	_	3			
377 -0.375 Age	0.0400	5.52e-05	724.397	0.000	0.
040 0.040 Income	-3.317e-08	4.53e-08	-0.732	0.464	-1.22e
-07 5.57e-08					
Outage_sec_perweek 000 0.001	0.0003	0.000	0.830	0.407	-0.
Email 0.001	-0.0001	0.000	-0.351	0.725	-0.
Contacts 003 0.002	-0.0003	0.001	-0.252	0.801	-0.
Yearly_equip_failure 006 0.002	-0.0020	0.002	-1.045	0.296	-0.
Techie 0.007	0.0014	0.003	0.474	0.636	-0.
Port_modem 001 0.008	0.0034	0.002	1.506	0.132	-0.
Tablet 0.005 0.005	0.0003	0.002	0.123	0.902	-0.
Multiple 256 0.274	0.2650	0.005	55.221	0.000	0.
OnlineSecurity 836 -0.826	-0.8309	0.002	-346.148	0.000	-0.
OnlineBackup 364 -0.349	-0.3565	0.004	-96.359	0.000	-0.
DeviceProtection 603 -0.592	-0.5972	0.003	-212.707	0.000	-0.
TechSupport 377 0.388	0.3828	0.003	133.527	0.000	0.
StreamingTV 313 -1.290	-1.3013	0.006	-219.712	0.000	-1.
StreamingMovies 741 -0.713	-0.7267	0.007	-101.702	0.000	-0.
PaperlessBilling 009 4.5e-06	-0.0045	0.002	-1.958	0.050	-0.
MonthlyCharge 035 -0.035	-0.0351	0.000	-271.542	0.000	-0.
Bandwidth_GB_Year 012 0.012	0.0122	5.26e-07	2.32e+04	0.000	0.
Item1 0.006	0.0024	0.002	1.463	0.143	-0.
Item2 004 0.002	-0.0013	0.002	-0.817	0.414	-0.
Item3 002 0.004	0.0012	0.001	0.878	0.380	-0.
Item4 001 0.003	0.0010	0.001	0.793	0.428	-0.
Item5 003 0.002	-0.0004	0.001	-0.289	0.773	-0.
Item6 004 0.002	-0.0010	0.001	-0.776	0.438	-0.
Item7	-0.0010	0.001	-0.818	0.413	-0.
Area_Suburban	-0.0038	0.003	-1.350	0.177	-0.
Area_Rural	0.0033	0.003	1.187	0.235	-0.
002 0.009 Marital_Married 008 0.005	-0.0014	0.003	-0.444	0.657	-0.
008 0.005 Marital_Separated 003 0.009	0.0028	0.003	0.896	0.370	-0.
Marital_Divorced  006 0.006	0.0004	0.003	0.131	0.896	-0.
Gender_Female	0.7924	0.002	343.283	0.000	0.
788 0.797					

localhost:8889/lab 9/78

Gender_Nonbinary 041 1.071		1.0564	0.008	137.743	0.000	1.
Contract_Oneyear		0.0015	0.003	0.504	0.614	-0.
Contract_TwoYear		0.0011	0.003	0.404	0.686	-0.
InternetService_DSL 759 -5.745		-5.7518	0.004	-1570.689	0.000	-5.
InternetService_None 160 -1.140		-1.1498	0.005	-221.677	0.000	-1.
PaymentMethod_BankTransfera	utomatic	-0.0009	0.003	-0.245	0.807	-0.
PaymentMethod_MailedCheck 001 0.013		0.0061	0.003	1.768	0.077	-0.
PaymentMethod_ElectronicChe	ck	0.0017	0.003	0.531	0.595	-0.
		=======	======	========	=======	
Omnibus:	31231.878	Durbin-Wa	tson:		2.006	
Prob(Omnibus):	0.000	Jarque-Be	ra (JB):		1459.022	
Skew:	-0.029	Prob(JB):			0.00	
Kurtosis:	1.023	Cond. No.			9.68e+05	
=======================================	=======	=======	======	========	=======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified
- [2] The condition number is large, 9.68e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# **Initial Findings**

The adjusted R2 in the models above, while theoretically possible, most likely indicated a potential issue with multicollinearity. As the correlation matrix below shows Bandwidth is almost a perfect match for Tenure and it may be causing issues with my model. The simple regression below indicates that Bandwidth\_GB\_Year accounts for 98.3% of the variance in Tenure. Since the goal of this project is to find indicators of Tenure, both categorical and continuous, I have elected to remove Bandwidth\_GB\_Year from the model to see if I can find any other potentially important indicators.

# Simple regression

The model below will be utilized as our baseline model. Currently, Bandwidth does an excellent job of predicting customer Tenure.

```
In [13]: #For comparison reasons only, not intended as final model. Final model is Lm32.
    slm = smf.ols('Tenure ~ Bandwidth_GB_Year', data = data).fit()
    slm.params
    print(slm.summary())
```

#### OLS Regression Results

Dep. Variable:	Tenure	R-squared:	0.983
Model:	OLS	Adj. R-squared:	0.983

localhost:8889/lab 10/78

5.199e+05

F-statistic:

Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:		Feb 2022 23:36:25 8950 8948 1 onrobust	Prob (F-stati Log-Likelihoo AIC: BIC:	•	0. -2375 4.752e+ 4.753e+	6. 04
	coef	std err	t	P> t	[0.025	0.975]
Intercept Bandwidth_GB_Year		0.067 1.66e-05		0.000 0.000	-6.258 0.012	-5.996 0.012
Omnibus: Prob(Omnibus): Skew: Kurtosis:		269.275 0.000 -0.315 2.551	Durbin-Watsor Jarque-Bera ( Prob(JB): Cond. No.	• •	1.9 223.4 2.96e- 7.41e+	81 49

Least Squares

Method:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 7.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

# Removing bandwidth from model

```
In [14]:
          lm6 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage_sec_perweek+Email
          lm6.params
          print(lm6.summary())
```

		OLS Regres	sion Resu				
Model: Method Date: Time: No. Ob Df Res Df Mod	: servations: iduals:	Tenure	Adj. R-: F-stati: Prob (F- Log-Like	R-squared: 0.001 stistic: 1.146		0.001 1.146 0.231 -41985. 8.406e+04 8.440e+04	
	========	==========		std err			[0.
025	0.975]						-
Interc	ept 53.109		42.0563	5.638	7.459	0.000	31.
Lat 247	-0.011		-0.1291	0.060	-2.143	0.032	-0.
Lng 027	0.055		0.0144	0.021	0.688	0.492	
Popula -05	5.98e-06			2.44e-05	-1.715	0.086	
Childr 407	en 0.174		-0.1165	0.148	-0.786	0.432	-0.
Age 010	0.043		0.0169	0.014	1.244	0.214	-0.
Income -05	2.18e-05	-8	.544e-08	1.11e-05	-0.008	0.994	-2.19e
Outage <sub>.</sub>	_sec_perweek		0.0470	0.096	0.490	0.624	-0. 11

11/78 localhost:8889/lab

		Customerchum_	LillealReglessio	1		
141	0.235					
Email	0.004	-0.1583	0.093	-1.703	0.089	-0.
341	0.024	0.0106	0 244	0.060	0.053	0
Contacts 628	0.591	-0.0186	0.311	-0.060	0.952	-0.
	uip_failure	0.3869	0.481	0.805	0.421	-0.
555	1.329	0.3803	0.401	0.005	0.421	-0.
Techie		-1.0000	0.750	-1.334	0.182	-2.
470	0.470					
Port_mode	em	0.3444	0.560	0.615	0.539	-0.
753	1.442					
Tablet		0.0315	0.613	0.051	0.959	-1.
170	1.233	0.4004	0.050		0 (75	
Phone	2 200	0.4036	0.962	0.420	0.675	-1.
482 Multiple	2.289	0.2610	1.181	0.221	0.825	-2.
054	2.576	0.2010	1.101	0.221	0.025	-2.
OnlineSec		0.1712	0.590	0.290	0.772	-0.
986	1.329	0.1712	0.330	0.230	0.772	٠.
OnlineBac		1.5183	0.910	1.668	0.095	-0.
265	3.302					
DevicePro	tection	-1.3732	0.691	-1.988	0.047	-2.
727	-0.019					
TechSuppo		0.0344	0.705	0.049	0.961	-1.
348	1.417					
Streaming	•	1.0017	1.457	0.687	0.492	-1.
854	3.858	1 2780	1 750	0 727	0.467	2
Streaming 167	4.725	1.2789	1.758	0.727	0.467	-2.
Paperless		-0.0205	0.569	-0.036	0.971	-1.
136	1.095	0.0203	0.303	0.030	0.571	
MonthlyCh		-0.0218	0.032	-0.686	0.493	-0.
084	0.041					
Item1		-0.3289	0.405	-0.811	0.417	-1.
124	0.466					
Item2		0.1692	0.378	0.447	0.655	-0.
572	0.910					
Item3		0.1221	0.345	0.354	0.723	-0.
554	0.798	0.2011	0.300	0.074	0 220	0
Item4 305	0.007	0.3011	0.309	0.974	0.330	-0.
Item5	0.907	0.5708	0.321	1.777	0.076	-0.
059	1.200	0.3700	0.521	1.///	0.070	-0.
Item6	1.200	-0.1786	0.331	-0.539	0.590	-0.
828	0.470					
Item7		0.2585	0.312	0.828	0.408	-0.
354	0.871					
Item8		-0.4478	0.298	-1.502	0.133	-1.
032	0.137	0.0554	0.606	0.004	0.000	
Area_Subu		-0.0661	0.686	-0.096	0.923	-1.
410	1.278	1 0440	0 697	1 [22	0 120	۵
Area_Rura 301	2.391	1.0449	0.687	1.522	0.128	-0.
Marital_M		0.2657	0.892	0.298	0.766	-1.
484	2.015	0.2037	0.032	0.230	0.700	-•
Marital_S		-0.1246	0.879	-0.142	0.887	-1.
849	1.599					
Marital_N	leverMarried	-0.0511	0.886	-0.058	0.954	-1.
788	1.685					
Marital_D		-0.7450	0.873	-0.854	0.393	-2.
456	0.966	0 8000	0 = 55	4 00-	0.000	_
Gender_Fe		0.5886	0.568	1.037	0.300	-0.
525 Gender_No	1.702	-2.3209	1.886	-1.230	0.219	-6.
019	1.377	-2.3209	1.000	-1.230	0.213	-0.
010						

localhost:8889/lab 12/78

Contract_Oneyear	0.2086	0.718	0.290	0.772	-1.
199 1.616					
Contract_TwoYear	1.3074	0.681	1.919	0.055	-0.
028 2.643					
InternetService_DSL	-0.0014	0.899	-0.002	0.999	-1.
764 1.761					
InternetService_None	-1.1177	1.276	-0.876	0.381	-3.
619 1.384					
PaymentMethod_BankTransferautomatic	-1.1791	0.854	-1.381	0.167	-2.
853 0.495					
PaymentMethod_MailedCheck	0.0784	0.849	0.092	0.926	-1.
586 1.743					
PaymentMethod ElectronicCheck	-0.8120	0.780	-1.041	0.298	-2.
341 0.717					
			========	======	
Omnibus: 37293.275				0.167	
Prob(Omnibus): 0.000		, ,	1	139.958	
Skew: 0.076	5 Prob(JB)	:	2.	89e-248	
Kurtosis: 1.258	3 Cond. No	•	9	.83e+05	
	========	=======	========	======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 9.83e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [15]: #Remove InternetService_DSL
lm7 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Income+Outage_sec_perweek+Email
lm7.params
print(lm7.summary())
```

	OLS Regres	sion Results	
Dep. Variable:	Tenure	R-squared:	0.006
Model:	OLS	Adj. R-squared:	0.001
Method:	Least Squares	F-statistic:	1.172
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.201
Time:	23:36:26	Log-Likelihood:	-41985.
No. Observations:	8950	AIC:	8.406e+04
Df Residuals:	8904	BIC:	8.439e+04
Df Model:	45		
Covariance Type:	nonrobust		

======		coef	std err	t	P> t	[0.
025	0.975]					
Intercep	ot	42.0526	5.063	8.306	0.000	32.
128	51.978					
Lat		-0.1291	0.060	-2.143	0.032	-0.
247	-0.011					
Lng		0.0144	0.021	0.688	0.491	-0.
027	0.055					
Populati	ion	-4.18e-05	2.44e-05	-1.715	0.086	-8.96e
-05	5.98e-06					
Childrer	า	-0.1165	0.148	-0.786	0.432	-0.
407	0.174					
Age		0.0169	0.014	1.244	0.214	-0.
010	0.043					
Income		-8.579e-08	1.11e-05	-0.008	0.994	-2.19e
-05 2	2.18e-05					

localhost:8889/lab 13/78

		CustomerChurn_	LinearRegression	n		
	ec_perweek	0.0470	0.096	0.490	0.624	-0.
141 Email	0.235	-0.1583	0.093	-1.703	0.089	-0.
341 Contacts	0.024	-0.0186	0.311	-0.060	0.952	-0.
628 Yearly_eq	0.591 puip_failure	0.3869	0.481	0.805	0.421	-0.
555 Techie	1.329	-1.0000	0.750	-1.334	0.182	-2.
470 Port_mode	0.470 em	0.3444	0.560	0.615	0.539	-0.
753 Tablet	1.442	0.0315	0.613	0.051	0.959	-1.
169	1.232					
Phone 481	2.289	0.4036	0.962	0.420	0.675	-1.
Multiple 550	2.070	0.2598	0.923	0.281	0.778	-1.
OnlineSec 981	curity 1.323	0.1711	0.588	0.291	0.771	-0.
OnlineBac 033	kup 3.002	1.5176	0.757	2.004	0.045	0.
DevicePro		-1.3736	0.633	-2.171	0.030	-2.
TechSuppo 223		0.0340	0.641	0.053	0.958	-1.
Streaming	gTV	1.0002	1.101	0.908	0.364	-1.
158 Streaming		1.2771	1.304	0.979	0.327	-1.
279 Paperless	_	-0.0205	0.569	-0.036	0.971	-1.
136 MonthlyCh	1.095 narge	-0.0218	0.023	-0.967	0.333	-0.
066 Item1	0.022	-0.3288	0.405	-0.811	0.417	-1.
123 Item2	0.466	0.1691	0.378	0.447	0.655	-0.
572 Item3	0.910	0.1221	0.345	0.354	0.723	-0.
554 Item4	0.798	0.3011	0.309	0.974	0.330	-0.
305 Item5	0.907	0.5708	0.321	1.778	0.075	-0.
059 Item6	1.200		0.331			
828	0.470	-0.1786		-0.539	0.590	-0.
Item7 354	0.871	0.2585	0.312	0.828	0.408	-0.
Item8 032	0.137	-0.4478	0.298	-1.502	0.133	-1.
Area_Subu 410	ırban 1.278	-0.0661	0.686	-0.096	0.923	-1.
Area_Rura 301	1 2.391	1.0449	0.687	1.522	0.128	-0.
Marital_M 484		0.2657	0.892	0.298	0.766	-1.
Marital_S 848		-0.1246	0.879	-0.142	0.887	-1.
Marital_N	leverMarried	-0.0511	0.886	-0.058	0.954	-1.
787 Marital_D		-0.7450	0.873	-0.854	0.393	-2.
456 Gender_Fe		0.5886	0.568	1.037	0.300	-0.
524 Gender_No	1.702 onbinary	-2.3209	1.886	-1.230	0.219	-6.

localhost:8889/lab 14/78

018 1.377						
Contract_Oneyear		0.2085	0.718	0.290	0.771	-1.
199 1.616						
Contract_TwoYear		1.3074	0.681	1.920	0.055	-0.
028 2.642						
InternetService_None		-1.1163	0.870	-1.283	0.200	-2.
822 0.589						
PaymentMethod_BankTransfera	utomatic	-1.1791	0.854	-1.381	0.167	-2.
853 0.494						
PaymentMethod_MailedCheck		0.0783	0.849	0.092	0.926	-1.
586 1.742						
PaymentMethod_ElectronicChe	ck	-0.8120	0.780	-1.041	0.298	-2.
341 0.717						
	========	========	========	========	======	
Omnibus:	37293.265	Durbin-Wa	tson:		0.167	
Prob(Omnibus):	0.000	Jarque-Bei	ra (JB):	1	139.959	
Skew:	0.076	Prob(JB):		2.	89e-248	
Kurtosis:	1.258	Cond. No.		8	.53e+05	
	========	========			======	

Dep. Variable:

Model:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 8.53e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [16]:
```

```
#Remove Income
lm8 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage_sec_perweek+Email+Contac
lm8.params
print(lm8.summary())
```

R-squared:

Adj. R-squared:

0.006

0.001

# OLS Regression Results

Tenure

OLS

Df Res Df Mod	servations: iduals:	Least Squares Sat, 12 Feb 2022 23:36:26 8950 8905 44 nonrobust	Prob (F Log-Like AIC:	-statistic):	8.	1.198 0.173 -41985. 406e+04 438e+04	
	=======================================	=======================================	=======	========		=======	======
	0.975]			std err	t	P> t	[0.
Interc	ept		42.0492	5.043	8.338	0.000	32.
163 Lat	51.935		-0.1291	0.060	-2.144	0.032	-0.
247 Lng	-0.011		0.0144	0.021	0.688	0.491	-0.
027 Popula	0.055 tion	-	4.18e-05	2.44e-05	-1.715	0.086	-8.96e
-05 Childr	5.98e-06 en		-0.1165	0.148	-0.786	0.432	-0.
407 Age	0.174		0.0169	0.014	1.244	0.214	-0.
	0.043 _sec_perweek		0.0470	0.096	0.490	0.624	-0.
141 Email	0.235		-0.1583	0.093	-1.703	0.089	-0.

localhost:8889/lab

	CustomerChum_L	illearRegression			
341 0.024 Contacts	-0.0186	0.311	-0.060	0.952	-0.
628 0.591	-0.0100	0.311	-0.000	0.932	-0.
Yearly_equip_failure	0.3869	0.481	0.805	0.421	-0.
555 1.329 Techie	-1.0000	0.750	-1.334	0.182	-2.
470 0.470 Port_modem	0.3444	0.560	0.615	0.538	-0.
753 1.442	0.3444	0.500	0.015	0.556	-0.
Tablet 169 1.232	0.0315	0.613	0.051	0.959	-1.
Phone	0.4036	0.962	0.420	0.675	-1.
481 2.288 Multiple	0.2598	0.923	0.281	0.778	-1.
550 2.070 OnlineSecurity	0.1711	0.587	0.291	0.771	-0.
981 1.323	0.1/11	0.587	0.291	0.771	-0.
OnlineBackup 033 3.002	1.5176	0.757	2.004	0.045	0.
DeviceProtection	-1.3737	0.633	-2.172	0.030	-2.
614 -0.134 TechSupport	0.0339	0.641	0.053	0.958	-1.
223 1.291					
StreamingTV 158 3.158	1.0002	1.101	0.908	0.364	-1.
StreamingMovies 279 3.833	1.2770	1.304	0.979	0.327	-1.
PaperlessBilling	-0.0205	0.569	-0.036	0.971	-1.
136 1.095 MonthlyCharge	-0.0218	0.023	-0.967	0.333	-0.
066 0.022					
Item1 123 0.465	-0.3288	0.405	-0.811	0.417	-1.
Item2	0.1691	0.378	0.447	0.655	-0.
572 0.910 Item3	0.1221	0.345	0.354	0.723	-0.
554 0.798 Item4	0.3012	0.309	0.975	0.330	-0.
305 0.907					
Item5 059 1.200	0.5707	0.321	1.778	0.075	-0.
Item6	-0.1786	0.331	-0.539	0.590	-0.
827 0.470 Item7	0.2585	0.312	0.828	0.408	-0.
354 0.871	0 4479	a 200	1 502	0 122	1
Item8 032 0.137	-0.4478	0.298	-1.502	0.133	-1.
Area_Suburban 410 1.278	-0.0662	0.686	-0.097	0.923	-1.
Area_Rural	1.0449	0.687	1.522	0.128	-0.
301 2.391 Marital_Married	0.2656	0.892	0.298	0.766	-1.
483 2.015					
Marital_Separated 848 1.599	-0.1246	0.879	-0.142	0.887	-1.
Marital_NeverMarried	-0.0511	0.886	-0.058	0.954	-1.
Marital_Divorced	-0.7450	0.873	-0.854	0.393	-2.
455 0.965 Gender_Female	0.5885	0.568	1.037	0.300	-0.
524 1.701 Gender_Nonbinary	-2.3210	1.886	-1.231	0.219	-6.
018 1.376					
Contract_Oneyear 199 1.616	0.2085	0.718	0.290	0.772	-1.

localhost:8889/lab 16/78

Contract_TwoYear 028 2.642	1.3074	0.681	1.920	0.055	-0.
028 2.642 InternetService_None 822 0.589	-1.1162	0.870	-1.283	0.200	-2.
PaymentMethod_BankTransferautomatic 853 0.494	-1.1791	0.854	-1.381	0.167	-2.
PaymentMethod_MailedCheck 585 1.742	0.0782	0.849	0.092	0.927	-1.
PaymentMethod_ElectronicCheck 341 0.717	-0.8120	0.780	-1.041	0.298	-2.
=======================================			========	======	
Omnibus: 37293.17	79 Durbin-W	Natson:		0.167	
Prob(Omnibus): 0.00	00 Jarque-E	Bera (JB):	1	139.962	
Skew: 0.07	• • • • • • • • • • • • • • • • • • • •	, ,	2.	89e-248	
Kurtosis: 1.25	, ,		2	.66e+05	
				======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.66e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [17]:

```
#Remove Paperless Billing
lm9 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage_sec_perweek+Email+Contac
lm9.params
```

0.006

print(lm9.summary())

Dep. Variable:

# OLS Regression Results

Tenure R-squared:

Df Resid Df Model Covarian	ervations:  uals:  :  ce Type:	OLS Least Squares Sat, 12 Feb 2022 23:36:26 8950 8906 43 nonrobust	Least Squares F-statistic: 2, 12 Feb 2022 Prob (F-statistic): 23:36:26 Log-Likelihood: 8950 AIC: 8906 BIC: 43		0.001 1.226 0.148 -41985. 8.406e+04 8.437e+04			
======								
025	_		coef		t		[0.	
Intercep	ot		42.0406	5.037	8.346	0.000	32.	
166	51.915						_	
Lat 247	-0.011		-0.1291	0.060	-2.145	0.032	-0.	
247 Lng	-0.011		0.0144	0.021	0.688	0.491	-0.	
_	0.055		0.0144	0.021	0.000	0.431	0.	
Populati	.on	-	4.18e-05	2.44e-05	-1.715	0.086	-8.96e	
-05 5								
Children			-0.1166	0.148	-0.787	0.431	-0.	
407 Age	0.174		0.0169	0.014	1.244	0.214	-0.	
010	0.043		0.0103	0.014	1.244	0.214	-0.	
	ec_perweek		0.0470	0.096	0.491	0.624	-0.	
141	0.235							
Email			-0.1583	0.093	-1.703	0.089	-0.	
341	0.024		0.0100	0 211	0.000	0.053	0	
Contacts 628	0.591		-0.0186	0.311	-0.060	0.952	-0.	
020	0.001							

localhost:8889/lab

(	SustomerChurn_Lir	nearRegression			
Yearly_equip_failure	0.3868	0.481	0.805	0.421	-0.
555 1.329 Techie	-1.0000	0.750	-1.334	0.182	-2.
469 0.469 Port_modem	0.3442	0.560	0.615	0.539	-0.
753 <b>1.</b> 442	0.3442	0.500	0.013	0.559	-0.
Tablet	0.0310	0.612	0.051	0.960	-1.
170 1.231 Phone	0.4037	0.961	0.420	0.675	-1.
481 2.288		0.000			
Multiple 549 2.070	0.2605	0.923	0.282	0.778	-1.
OnlineSecurity	0.1711	0.587	0.291	0.771	-0.
980 1.323 OnlineBackup	1.5177	0.757	2.005	0.045	0.
034 3.002					
DeviceProtection 614 -0.134	-1.3737	0.633	-2.172	0.030	-2.
TechSupport	0.0341	0.641	0.053	0.958	-1.
222 1.291 StreamingTV	1.0012	1.101	0.910	0.363	-1.
156 3.159	1.0012	1.101	0.910	0.303	-1.
StreamingMovies	1.2779	1.304	0.980	0.327	-1.
277 3.833 MonthlyCharge	-0.0218	0.023	-0.968	0.333	-0.
066 0.022	0. 2200	0.405	0.013	0 417	4
Item1 123 0.465	-0.3290	0.405	-0.812	0.417	-1.
Item2	0.1691	0.378	0.447	0.655	-0.
572 0.910 Item3	0.1223	0.345	0.355	0.723	-0.
554 0.798					
Item4 305 0.907	0.3011	0.309	0.975	0.330	-0.
Item5	0.5707	0.321	1.778	0.075	-0.
059 1.200 Item6	-0.1784	0.331	-0.539	0.590	-0.
827 0.470					0.
Item7 354 0.870	0.2583	0.312	0.827	0.408	-0.
Item8	-0.4479	0.298	-1.502	0.133	-1.
032 0.136 Area_Suburban	-0.0660	0.685	-0.096	0.923	-1.
410 1.278	-0.0000	0.005	-0.030	0.923	-1.
Area_Rural 301 2.391	1.0450	0.686	1.522	0.128	-0.
301 2.391 Marital_Married	0.2657	0.892	0.298	0.766	-1.
483 2.015	0 1242	0.070	0 141	0.888	1
Marital_Separated 848 1.599	-0.1243	0.879	-0.141	0.888	-1.
Marital_NeverMarried	-0.0513	0.886	-0.058	0.954	-1.
787 1.685 Marital Divorced	-0.7444	0.872	-0.853	0.393	-2.
454 0.966					
Gender_Female 524 1.701	0.5882	0.568	1.036	0.300	-0.
Gender_Nonbinary	-2.3203	1.886	-1.230	0.219	-6.
017 1.376 Contract_Oneyear	0.2087	0.718	0.291	0.771	-1.
199 1.616					
Contract_TwoYear 028 2.642	1.3070	0.681	1.919	0.055	-0.
InternetService_None	-1.1167	0.870	-1.284	0.199	-2.
822 0.589 PaymentMethod_BankTransferautomatic	-1.1791	0.854	-1.381	0.167	-2.
. ayeneric choa_bankir and rer ad comdete	-,-/-	J. UJ-	50-	0.107	٠.

localhost:8889/lab 18/78

852	0.494						
PaymentMet	hod_MailedCheck		0.0785	0.849	0.093	0.926	-1.
585	1.742						
PaymentMethod_ElectronicCheck		-0.8122	0.780	-1.042	0.298	-2.	
341	0.716						
=======	==========	========	========	=======	========	======	
Omnibus:		37293.395	Durbin-Wa	tson:		0.167	
Prob(Omnib	us):	0.000	Jarque-Be	ra (JB):	1	139.957	
Skew:		0.076	Prob(JB):		2.	89e-248	
Kurtosis:		1.258	Cond. No.		2	.66e+05	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.66e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [18]:

### #Remove Tablet

Dep. Variable:

Model:

lm10 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Conta
lm10.params
print(lm10.summary())

Adj. R-squared:

0.006

0.001

# OLS Regression Results

OLS

Tenure R-squared:

Df Resido Df Model Covarian	rvations: uals: : ce Type:	Least Squares Sat, 12 Feb 2022 23:36:26 8950 8907 42 nonrobust	Prob (F- Log-Like AIC: BIC:	-statistic): elihood:	1.256 0.125 -41985. 8.406e+04 8.436e+04		
=======		=======================================	=======	=======	========	:=======	======
005	0 0751		coef	std err	t	P> t	[0.
025							
Intercep			42.0520	5.032	8.357	0.000	32.
188	51.916		0 1201	0.060	2 445	0.022	0
Lat 247	0 011		-0.1291	0.060	-2.145	0.032	-0.
Lng	-0.011		0.0144	0 021	0.689	0.491	-0.
027	0.055		0.0144	0.021	0.005	0.431	0.
Population		-4	.181e-05	2.44e-05	-1.716	0.086	-8.96e
-05 5	.96e-06						
Children			-0.1166	0.148	-0.787	0.431	-0.
	0.174						
Age	0.043		0.0169	0.014	1.244	0.214	-0.
010			0.0470	0.096	0.491	0.624	-0.
	ec_perweek 0.235		0.0470	0.090	0.491	0.024	-0.
Email	0.233		-0.1583	0.093	-1.703	0.089	-0.
	0.024		0.1303	0.055	1.703	0.005	•
Contacts			-0.0186	0.311	-0.060	0.952	-0.
628							
	quip_failure		0.3867	0.481	0.805	0.421	-0.
555	1.329						
Techie	0.470		-0.9996	0.750	-1.334	0.182	-2.
469			0 2442	0 560	0 (15	Q 520	0
Port_mod	elli		0.3443	0.560	0.615	0.539	-0.

localhost:8889/lab

		0		noun togrossion			
753 Phone	1.442		0.4048	0.961	0.421	0.674	-1.
479 Multiple	2.289		0.2603	0.923	0.282	0.778	-1.
549	2.070						
OnlineSed	1.323		0.1715	0.587	0.292	0.770	-0.
OnlineBad 034	ckup 3.002		1.5177	0.757	2.005	0.045	0.
DevicePro	otection		-1.3736	0.632	-2.172	0.030	-2.
613 TechSuppo	-0.134 ort		0.0339	0.641	0.053	0.958	-1.
223 Streaming	1.290		1.0023	1.100	0.911	0.362	-1.
155	3.159						
Streaming 275	gMovies 3.834		1.2792	1.303	0.982	0.326	-1.
MonthlyCh 066	narge 0.022		-0.0218	0.023	-0.969	0.333	-0.
Item1			-0.3287	0.405	-0.812	0.417	-1.
123 Item2	0.465		0.1694	0.378	0.448	0.654	-0.
571 Item3	0.910		0.1221	0.345	0.354	0.723	-0.
554	0.798						
Item4 305	0.907		0.3009	0.309	0.974	0.330	-0.
Item5			0.5703	0.321	1.777	0.076	-0.
059 Item6	1.199		-0.1786	0.331	-0.540	0.590	-0.
827 Item7	0.470		0.2582	0.312	0.827	0.408	-0.
354	0.870						
Item8 032	0.137		-0.4478	0.298	-1.502	0.133	-1.
Area_Subu 410	urban 1.277		-0.0662	0.685	-0.097	0.923	-1.
Area_Rura	al		1.0444	0.686	1.522	0.128	-0.
301 Marital_N	2.390 Married		0.2660	0.892	0.298	0.766	-1.
483 Marital_9	2.015 Separated		-0.1234	0.879	-0.140	0.888	-1.
846	1.600						
787	NeverMarried 1.685		-0.0508	0.886	-0.057	0.954	-1.
Marital_[ 454	Divorced 0.966		-0.7440	0.872	-0.853	0.394	-2.
Gender_F	emale		0.5885	0.567	1.037	0.300	-0.
524 Gender_No	1.701 onbinary		-2.3204	1.886	-1.230	0.219	-6.
017 Contract	1.376 Onevear		0.2092	0.718	0.291	0.771	-1.
198	1.616						
Contract_ 028	2.642		1.3071	0.681	1.920	0.055	-0.
InternetS 822	Service_None 0.589		-1.1167	0.870	-1.284	0.199	-2.
PaymentMe	ethod_BankTransfer	automatic	-1.1793	0.854	-1.382	0.167	-2.
-	0.494 ethod_MailedCheck		0.0784	0.849	0.092	0.926	-1.
585 PaymentMe	1.742 ethod_ElectronicCh	eck	-0.8117	0.780	-1.041	0.298	-2.
340	0.717						-•
Omnibus:		37293.033				0.167	

localhost:8889/lab 20/78

Kurtosis:	1.258	Cond. No.	2.65e+05
Skew:	0.076	Prob(JB):	2.87e-248
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1139.972

Model:

Dep. Variable:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.65e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [19]: #Remo

#Remove Tech support
lm11 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Conta
lm11.params
print(lm11.summary())

OLS Adj. R-squared:

0.006

0.001

OLS	Regres	ssion	Results

Tenure R-squared:

Method: Date: Time: No. Obser Df Residu Df Model: Covarianc	als:	Least Squares F-statistic: Sat, 12 Feb 2022 Prob (F-statistic): 23:36:26 Log-Likelihood: 8950 AIC: 8908 BIC: 41 nonrobust		8.	1.286 0.104 -41985. 8.405e+04 8.435e+04		
		==========	======	=======	========	:======	======
=======	=====		coef	std err	t	P> t	[0.
	0.975]						-
Tntoncont			42 0100	4.971	8.452	0.000	22
Intercept 267	51.754		42.0106	4.971	8.452	0.000	32.
Lat	J1./J4		-0.1291	0.060	-2.146	0.032	-0.
	-0.011						
Lng			0.0144	0.021	0.688	0.491	-0.
027	0.055	_	400 05	2 44 25	4 747	0.004	2 24
Populatio	on 93e-06	-4	.183e-05	2.44e-05	-1.717	0.086	-8.96e
Children	936-00		-0.1166	0.148	-0.787	0.431	-0.
407	0.174		0.1100	0.140	0.707	0.431	٠.
Age			0.0169	0.014	1.245	0.213	-0.
010	0.043						
	c_perweek		0.0470	0.096	0.491	0.624	-0.
141	0.235		0 1500	0.003	1 702	0.000	0
Email 340	0.024		-0.1583	0.093	-1.703	0.089	-0.
Contacts	0.024		-0.0188	0.311	-0.060	0.952	-0.
628	0.591						
	uip_failure		0.3869	0.481	0.805	0.421	-0.
555	1.329						_
Techie	0.470		-0.9995	0.750	-1.334	0.182	-2.
469 Port mode	0.470		0.3446	0.560	0.616	0.538	-0.
753	1.442		0.5440	0.500	0.010	0.550	-0.
Phone	_,,,		0.4050	0.961	0.421	0.674	-1.
479	2.289						
Multiple			0.2434	0.866	0.281	0.779	-1.
454	1.941		0 1705	0 507	0.200	0.773	0
OnlineSec 980	urity 1.321		0.1705	0.587	0.290	0.772	-0.
200	1.741						

localhost:8889/lab 21/78

	Ouc	tomeronum_tm				
OnlineBackup		1.5062	0.725	2.078	0.038	0.
086 2.927 DeviceProtection		-1.3802	0.620	-2.226	0.026	-2.
595 -0.165 StreamingTV		0.9807	1.022	0.960	0.337	-1.
022 2.984 StreamingMovies		1.2521	1.198	1.045	0.296	-1.
096 3.600 MonthlyCharge		-0.0213	0.020	-1.048	0.295	-0.
061 0.019 Item1		-0.3283	0.405	-0.811	0.418	-1.
122 0.466 Item2		0.1691	0.378	0.448	0.655	-0.
571 0.910 Item3		0.1223	0.345	0.355	0.723	-0.
554 0.798 Item4		0.3011	0.309	0.975	0.330	-0.
304 0.907 Item5		0.5704	0.321	1.778	0.076	-0.
059 1.199 Item6 827 0.470		-0.1787	0.331	-0.540	0.589	-0.
827 0.470 Item7 353 0.870		0.2584	0.312	0.828	0.408	-0.
Item8 032 0.137		-0.4477	0.298	-1.502	0.133	-1.
Area_Suburban 410 1.277		-0.0665	0.685	-0.097	0.923	-1.
Area_Rural 301 2.389		1.0442	0.686	1.521	0.128	-0.
Marital_Married 482 2.015		0.2664	0.892	0.299	0.765	-1.
Marital_Separated 847 1.599		-0.1236	0.879	-0.141	0.888	-1.
Marital_NeverMarried 786 1.685		-0.0505	0.885	-0.057	0.954	-1.
Marital_Divorced 454 0.966		-0.7438	0.872	-0.853	0.394	-2.
Gender_Female 524 1.701		0.5884	0.567	1.037	0.300	-0.
Gender_Nonbinary 016 1.377		-2.3195	1.886	-1.230	0.219	-6.
Contract_Oneyear 198 1.616		0.2094	0.718	0.292	0.770	-1.
Contract_TwoYear 028 2.642		1.3070	0.681	1.920	0.055	-0.
<pre>InternetService_None 745     0.537</pre>		-1.1042	0.837	-1.319	0.187	-2.
PaymentMethod_BankTransfera 852 0.494	utomatic	-1.1789	0.854	-1.381	0.167	-2.
PaymentMethod_MailedCheck 585 1.742		0.0786	0.848	0.093	0.926	-1.
PaymentMethod_ElectronicChe 340 0.717	eck	-0.8113	0.780	-1.041	0.298	-2.
Omnibus: Prob(Omnibus): Skew: Kurtosis:	37293.201 0.000 0.076 1.258	Durbin-Wa Jarque-Be	tson: ra (JB):	1	====== 0.167 139.969 88e-248 .61e+05	

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

localhost:8889/lab 22/78

[2] The condition number is large, 2.61e+05. This might indicate that there are strong multicollinearity or other numerical problems.

In [20]:

#Remove Marital Never Married
lm12 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Conta
lm12.params
print(lm12.summary())

		OLS Regres						
Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:		Tenure OLS Least Squares Sat, 12 Feb 2022 23:36:26 8950 8909 40 nonrobust	Tenure R-squared: OLS Adj. R-squared: quares F-statistic: b 2022 Prob (F-statistic): 36:26 Log-Likelihood: 8950 AIC: 8909 BIC: 40		:	0.006 0.001 1.318 0.0862 -41985. 8.405e+04 8.434e+04		
	:=======	==========						
025	0.975]			std err			[0.	
Interce	ept		41.9878	4.954	8.475	0.000	32.	
276 Lat	51.699		-0.1291	0.060	-2.146	0.032	-0.	
Lng	-0.011		0.0144	0.021	0.689	0.491	-0.	
027 Populat	ion:	-4	.185e-05	2.44e-05	-1.718	0.086	-8.96e	
Childre			-0.1165	0.148	-0.787	0.431	-0.	
407 Age	0.174		0.0169	0.014	1.245	0.213	-0.	
010 Outage_ 141	0.043 _sec_perweek 0.235		0.0470	0.096	0.491	0.623	-0.	
Email 340	0.024		-0.1582	0.093	-1.703	0.089	-0.	
Contact 628			-0.0187	0.311	-0.060	0.952	-0.	
	equip_failure 1.329		0.3870	0.481	0.805	0.421	-0.	
Techie 467	0.470		-0.9984	0.749	-1.333	0.183	-2.	
Port_mc			0.3445	0.560	0.616	0.538	-0.	
Phone 478	2.289		0.4055	0.961	0.422	0.673	-1.	
Multipl 454			0.2429	0.866	0.281	0.779	-1.	
	Security 1.321		0.1703	0.587	0.290	0.772	-0.	
OnlineB 086			1.5063	0.725	2.079	0.038	0.	
	Protection -0.165		-1.3799	0.620	-2.226	0.026	-2.	
Streami 022			0.9811	1.022	0.960	0.337	-1.	
	ngMovies 3.600		1.2520	1.198	1.045	0.296	-1.	

localhost:8889/lab 23/78

	otomoronam_Er	noun togroodidin			
MonthlyCharge	-0.0213	0.020	-1.048	0.295	-0.
061 0.019					
Item1	-0.3285	0.405	-0.811	0.417	-1.
122 0.465					
Item2	0.1692	0.378	0.448	0.654	-0.
571 0.910					
Item3	0.1224	0.345	0.355	0.723	-0.
554 0.798					
Item4	0.3007	0.309	0.974	0.330	-0.
305 0.906					
Item5	0.5705	0.321	1.778	0.075	-0.
058 1.199					
Item6	-0.1790	0.331	-0.541	0.589	-0.
828 0.470					
Item7	0.2583	0.312	0.828	0.408	-0.
353 0.870					
Item8	-0.4475	0.298	-1.502	0.133	-1.
032 0.137					
Area_Suburban	-0.0663	0.685	-0.097	0.923	-1.
410 1.277					
Area_Rural	1.0440	0.686	1.521	0.128	-0.
301 2.389					
Marital_Married	0.2911	0.780	0.373	0.709	-1.
238 1.820					
Marital_Separated	-0.0988	0.765	-0.129	0.897	-1.
598 1.400					
Marital_Divorced	-0.7190	0.756	-0.950	0.342	-2.
202 0.764					
Gender_Female	0.5884	0.567	1.037	0.300	-0.
524 1.701					
Gender_Nonbinary	-2.3192	1.885	-1.230	0.219	-6.
015 1.377					
Contract_Oneyear	0.2092	0.718	0.292	0.771	-1.
198 1.616					
Contract_TwoYear	1.3071	0.681	1.920	0.055	-0.
027 2.642					
InternetService_None	-1.1032	0.837	-1.318	0.187	-2.
744 0.537					
PaymentMethod_BankTransferautomatic	-1.1787	0.853	-1.381	0.167	-2.
852 0.494					
PaymentMethod_MailedCheck	0.0793	0.848	0.094	0.925	-1.
1.742					
PaymentMethod_ElectronicCheck	-0.8109	0.780	-1.040	0.298	-2.
339 0.717					
	=======	========		======	
Omnibus: 37293.253	Durbin-Wa	atson:		0.167	
Prob(Omnibus): 0.000	Jarque-B	era (JB):	1	139.968	
Skew: 0.076	Prob(JB)	•	2.	88e-248	
Kurtosis: 1.258	Cond. No	•	2	.60e+05	
				======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.6e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [21]:
```

```
#Remove Contacts
```

lm13 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Yearl
lm13.params
print(lm13.summary())

\_\_\_\_\_\_

Dep. Vari Model: Method: Date: Time: No. Obser Df Residu Df Model Covariano	rvations: uals: :	Tenure OLS Least Squares Sat, 12 Feb 2022 23:36:26 8950 8910 39 nonrobust	F-stati: Prob (F Log-Lik	squared: stic: -statistic)		0.006 0.002 1.352 0.0704 -41985. 8.405e+04 8.433e+04	
		=========	:======:	=======	=======	=======	======
025	0.975]			std err			[0.
Intercept			41.9688	4.944	8.489	0.000	32.
Lat	51.660		-0.1291	0.060	-2.146	0.032	-0.
247 Lng	-0.011		0.0144	0.021	0.689	0.491	-0.
027 Population	0.055	-1	1850-05	2.44e-05	-1.718	0.086	-8.96e
	.89e-06	-4	.1036-03	2.446-03	-1./10	0.000	-0.506
Children 406	0.174		-0.1163	0.148	-0.785	0.432	-0.
Age			0.0169	0.014	1.244	0.213	-0.
	0.043 ec_perweek		0.0469	0.096	0.490	0.624	-0.
141 Email	0.235		-0.1583	0.093	-1.703	0.089	-0.
340	0.024						
Yearly_ed	quip_failure 1.329		0.3869	0.481	0.805	0.421	-0.
Techie			-0.9985	0.749	-1.333	0.183	-2.
467 Port_mode	0.470 em		0.3446	0.560	0.616	0.538	-0.
752 Phone	1.442		0.4054	0.961	0.422	0.673	-1.
	2.289		0.2425		0.280		-1.
455	1.940		0.2423	0.800	0.280	0.773	-1.
OnlineSed	-		0.1700	0.587	0.290	0.772	-0.
981 OnlineBad	1.321		1.5057	0.725	2.078	0.038	0.
085	2.926		1.3037	0.723	2.070	0.030	0.
DevicePro			-1.3798	0.620	-2.226	0.026	-2.
595 Streaming	-0.165 ∍TV		0.9806	1.022	0.960	0.337	-1.
022	2.983						
Streaming 097	gMovies 3.599		1.2510	1.198	1.045	0.296	-1.
MonthlyCh	narge		-0.0213	0.020	-1.048	0.295	-0.
061 Item1	0.019		-0.3285	0.405	-0.811	0.417	-1.
122 Item2	0.465		0.1689	0.378	0.447	0.655	-0.
571	0.909						
Item3 553	0.799		0.1228	0.345	0.356	0.722	-0.
Item4 305	0.906		0.3006	0.309	0.973	0.330	-0.
Item5			0.5702	0.321	1.778	0.076	-0.
059 Item6	1.199		-0.1790	0.331	-0.541	0.589	-0.

localhost:8889/lab 25/78

828

0.470

354       0.870         Item8       -0.4471       0.298       -1.501       0.133       -1         031       0.137       -0.0661       0.685       -0.096       0.923       -1         409       1.277       -0.0661       0.686       1.521       0.128       -0         301       2.389       -0.0940       0.780       0.373       0.709       -1         238       1.820       -0.0980       0.765       -0.128       0.898       -1         597       1.401       -0.0980       0.765       -0.128       0.898       -1         597       1.401       -0.7186       0.756       -0.950       0.342       -2         201       0.764       -0.7186       0.756       -0.950       0.342       -2         201       0.764       -0.788       0.567       1.038       0.299       -0         523       1.701       -0.701       -0.701       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -0.001       -	020	0.470						
Ttem8				0.2580	0.312	0.827	0.408	-0.
031	354	0.870						
Area_Suburban				-0.4471	0.298	-1.501	0.133	-1.
Area_Rural 1.0440 0.686 1.521 0.128 -0  Area_Rural 0.2389  Marital_Married 0.2910 0.780 0.373 0.709 -1  238 1.820  Marital_Separated -0.0980 0.765 -0.128 0.898 -1  597 1.401  Marital_Divorced -0.7186 0.756 -0.950 0.342 -2  201 0.764  Gender_Female 0.5888 0.567 1.038 0.299 -0  523 1.701  Gender_Nonbinary -2.3183 1.885 -1.230 0.219 -6  014 1.377  Contract_Oneyear 0.2088 0.718 0.291 0.771 -1  198 1.615  Contract_TwoYear 1.3066 0.681 1.919 0.055 -0  028 2.641  InternetService_None -1.1024 0.837 -1.318 0.188 -2  743 0.538  PaymentMethod_BankTransferautomatic -1.1791 0.853 -1.382 0.167 -2  852 0.494  PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1  583 1.743  PaymentMethod_MailedCheck -0.8107 0.780 -1.040 0.298 -2  339 0.717								
Area_Rural				-0.0661	0.685	-0.096	0.923	-1.
301       2.389         Marital_Married       0.2910       0.780       0.373       0.709       -1         238       1.820         Marital_Separated       -0.0980       0.765       -0.128       0.898       -1         597       1.401         Marital_Divorced       -0.7186       0.756       -0.950       0.342       -2         201       0.764       0.5888       0.567       1.038       0.299       -0         6ender_Female       0.5888       0.567       1.038       0.299       -0         523       1.701       0.600       0.5888       0.567       1.038       0.299       -0         6ender_Nonbinary       -2.3183       1.885       -1.230       0.219       -6         014       1.377       0.000       0.2088       0.718       0.291       0.771       -1         198       1.615       0.000       0.681       1.919       0.055       -6         028       2.641       0.657       1.318       0.188       -2         743       0.538       0.538       0.848       0.094       0.925       -1         852       0.494       0.949       0.925								
Marital_Married       0.2910       0.780       0.373       0.709       -1         238       1.820         Marital_Separated       -0.0980       0.765       -0.128       0.898       -1         597       1.401       -0.7186       0.756       -0.950       0.342       -2         Marital_Divorced       -0.7186       0.756       -0.950       0.342       -2         201       0.764       0.764       0.5888       0.567       1.038       0.299       -0         523       1.701       0.5888       0.567       1.038       0.299       -0         6ender_Female       0.5888       0.567       1.038       0.299       -0         523       1.701       0.701       -1       0.208       0.219       -6         6ender_Female       0.2888       0.567       1.038       0.299       -0         624       1.377       0.0088       0.718       0.291       0.771       -1         198       1.615       0.064       1.919       0.055       -0         028       2.641       0.066       0.681       1.919       0.055       -0         743       0.538       0.066       <				1.0440	0.686	1.521	0.128	-0.
238       1.820         Marital_Separated       -0.0980       0.765       -0.128       0.898       -1         597       1.401         Marital_Divorced       -0.7186       0.756       -0.950       0.342       -2         201       0.764       0.764       0.5888       0.567       1.038       0.299       -0         523       1.701       0.5888       0.567       1.038       0.299       -0         523       1.701       0.600       0.5888       0.567       1.038       0.299       -0         523       1.701       0.600       0.5888       0.567       1.038       0.299       -0         523       1.701       0.600       0.5888       0.567       1.038       0.299       -0         604       1.377       0.000       0.2088       0.718       0.219       -0         014       1.377       0.000       0.681       1.919       0.055       -0         028       1.615       0.006       0.681       1.919       0.055       -0         028       2.641       0.006       0.837       -1.318       0.188       -2         743       0.538       0.00								
Marital_Separated       -0.0980       0.765       -0.128       0.898       -1         597       1.401       -0.7186       0.756       -0.950       0.342       -2         201       0.764       0.5888       0.567       1.038       0.299       -0         6ender_Female       0.5888       0.567       1.038       0.299       -0         523       1.701       -0.2088       0.567       1.038       0.299       -0         6ender_Nonbinary       -2.3183       1.885       -1.230       0.219       -6         014       1.377       -0.2088       0.718       0.291       0.771       -1         198       1.615       -0.2088       0.718       0.291       0.771       -1         198       1.615       -0.2088       0.718       0.291       0.771       -1         198       1.615       -0.2088       0.718       0.291       0.771       -1         198       1.615       -0.2088       0.681       1.919       0.055       -0         028       2.641       -0.208       0.681       1.919       0.055       -0         743       0.538       -0.208       0.848       0.	_			0.2910	0.780	0.373	0.709	-1.
597 1.401  Marital_Divorced								
Marital_Divorced       -0.7186       0.756       -0.950       0.342       -2         201       0.764       0.5888       0.567       1.038       0.299       -6         523       1.701       0.5888       0.567       1.038       0.299       -6         523       1.701       0.60       0.681       1.230       0.219       -6         604       1.377       0.2088       0.718       0.291       0.771       -1         198       1.615       0.2088       0.718       0.291       0.771       -1         198       1.615       0.006       0.681       1.919       0.055       -6         028       2.641       0.1024       0.837       -1.318       0.188       -2         743       0.538       0.538       0.494       0.837       -1.318       0.188       -2         852       0.494       0.494       0.853       -1.382       0.167       -2         853       1.743       0.740       0.780       -1.040       0.298       -2         339       0.717       0.780       -1.040       0.298       -2         339       0.717       0.780       -1.040				-0.0980	0.765	-0.128	0.898	-1.
201 0.764 Gender_Female 0.5888 0.567 1.038 0.299 -0.523 1.701 Gender_Nonbinary -2.3183 1.885 -1.230 0.219 -0.501				0 7406	0.756	0.050	0.040	
Gender_Female 0.5888 0.567 1.038 0.299 -0.523 1.701 Gender_Nonbinary -2.3183 1.885 -1.230 0.219 -6.5014 1.377 Contract_Oneyear 0.2088 0.718 0.291 0.771 -1.5015 -0.501				-0.7186	0.756	-0.950	0.342	-2.
1.701 Gender_Nonbinary				0 5000	0 567	4 020	0.200	•
Gender_Nonbinary	_			0.5888	0.56/	1.038	0.299	-0.
014				2 2402	1 005	1 220	0.210	_
Contract_Oneyear 0.2088 0.718 0.291 0.771 -1 198 1.615 Contract_TwoYear 1.3066 0.681 1.919 0.055 -0 028 2.641 InternetService_None -1.1024 0.837 -1.318 0.188 -2 743 0.538 PaymentMethod_BankTransferautomatic -1.1791 0.853 -1.382 0.167 -2 852 0.494 PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1 583 1.743 PaymentMethod_ElectronicCheck -0.8107 0.780 -1.040 0.298 -2 339 0.717 ===================================		_		-2.3183	1.885	-1.230	0.219	-6.
198		• -		0 2000	0.710	0 201	0 771	1
Contract_TwoYear 1.3066 0.681 1.919 0.055 -0028 2.641  InternetService_None -1.1024 0.837 -1.318 0.188 -2 743 0.538  PaymentMethod_BankTransferautomatic -1.1791 0.853 -1.382 0.167 -2 852 0.494  PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1 583 1.743  PaymentMethod_ElectronicCheck -0.8107 0.780 -1.040 0.298 -2 339 0.717  ==================================				0.2088	0.718	0.291	0.7/1	-1.
028       2.641         InternetService_None       -1.1024       0.837       -1.318       0.188       -2         743       0.538         PaymentMethod_BankTransferautomatic       -1.1791       0.853       -1.382       0.167       -2         852       0.494         PaymentMethod_MailedCheck       0.0798       0.848       0.094       0.925       -1         583       1.743         PaymentMethod_ElectronicCheck       -0.8107       0.780       -1.040       0.298       -2         339       0.717         ====================================				1 2066	0 601	1 010	0 055	0
InternetService_None -1.1024 0.837 -1.318 0.188 -2 743 0.538  PaymentMethod_BankTransferautomatic -1.1791 0.853 -1.382 0.167 -2 852 0.494  PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1 583 1.743  PaymentMethod_ElectronicCheck -0.8107 0.780 -1.040 0.298 -2 339 0.717  ==================================	_			1.3000	0.001	1.919	0.055	-0.
743  0.538   PaymentMethod_BankTransferautomatic  -1.1791  0.853  -1.382  0.167  -2 852  0.494   PaymentMethod_MailedCheck  0.0798  0.848  0.094  0.925  -1 583  1.743   PaymentMethod_ElectronicCheck  -0.8107  0.780  -1.040  0.298  -2 339  0.717				1 1024	0 927	1 210	A 100	2
PaymentMethod_BankTransferautomatic -1.1791 0.853 -1.382 0.167 -2852 0.494  PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1583 1.743  PaymentMethod_ElectronicCheck -0.8107 0.780 -1.040 0.298 -2839 0.717  ==================================				-1.1024	0.037	-1.310	0.100	-2.
852 0.494  PaymentMethod_MailedCheck 0.0798 0.848 0.094 0.925 -1 583 1.743  PaymentMethod_ElectronicCheck -0.8107 0.780 -1.040 0.298 -2 339 0.717			utomatic	_1 1701	0 853	_1 392	0 167	-2.
PaymentMethod_MailedCheck       0.0798       0.848       0.094       0.925       -1         583       1.743         PaymentMethod_ElectronicCheck       -0.8107       0.780       -1.040       0.298       -2         339       0.717         ====================================		_	utomatic	-1.1/91	0.055	-1.362	0.107	-2.
583       1.743         PaymentMethod_ElectronicCheck       -0.8107       0.780       -1.040       0.298       -2         339       0.717       0.780       -1.040       0.298       -2         Omnibus:       37293.638       Durbin-Watson:       0.167         Prob(Omnibus):       0.000       Jarque-Bera (JB):       1139.955         Skew:       0.076       Prob(JB):       2.90e-248				0 0798	0 8/18	0 001	0 925	-1.
PaymentMethod_ElectronicCheck       -0.8107       0.780       -1.040       0.298       -2         339       0.717       0.780       -1.040       0.298       -2         Omnibus:       37293.638       Durbin-Watson:       0.167         Prob(Omnibus):       0.000       Jarque-Bera (JB):       1139.955         Skew:       0.076       Prob(JB):       2.90e-248	•	_		0.0750	0.040	0.054	0.525	-1.
339 0.717		-	ck	-0 8107	0 780	-1 040	0 298	-2.
Omnibus: 37293.638 Durbin-Watson: 0.167 Prob(Omnibus): 0.000 Jarque-Bera (JB): 1139.955 Skew: 0.076 Prob(JB): 2.90e-248			CK	0.0107	0.700	1.040	0.230	
Prob(Omnibus):       0.000       Jarque-Bera (JB):       1139.955         Skew:       0.076       Prob(JB):       2.90e-248			========	========	=======	========	======	
Prob(Omnibus):       0.000       Jarque-Bera (JB):       1139.955         Skew:       0.076       Prob(JB):       2.90e-248	Omnibus:		37293.638	Durbin-Wa	tson:		0.167	
Skew: 0.076 Prob(JB): 2.90e-248		bus):				1		
·	•			•	<b>\</b> - <b>/</b> ·			
	=======	============	=======	========	=======		======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 2.6e+05. This might indicate that there are strong multicollinearity or other numerical problems.

## In [22]:

### #Remove Mailed Check

lm13 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Yearl lm13.params print(lm13.summary())

## OLS Regression Results

	=======================================		
Dep. Variable:	Tenure	R-squared:	0.006
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	1.388
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.0570
Time:	23:36:27	Log-Likelihood:	-41985.
No. Observations:	8950	AIC:	8.405e+04
Df Residuals:	8911	BIC:	8.432e+04
Df Model:	38		
Covariance Type:	nonrobust		

======= ========== P>|t| coef std err

[0.

2	5	0	. 9	75]	ı

Intercept		42.0207	4.913	8.553	0.000	32.
390	51.651	0 1202	0.000	2 447	0.022	0
Lat 247	-0.011	-0.1292	0.060	-2.147	0.032	-0.
Lng	-0.011	0.0144	0.021	0.690	0.490	-0.
027	0.055	0.0144	0.021	0.090	0.490	-0.
Population		-4.189e-05	2.44e-05	-1.720	0.086	-8.96e
	.86e-06					
Children		-0.1163	0.148	-0.785	0.432	-0.
406	0.174					
Age		0.0169	0.014	1.243	0.214	-0.
010	0.043					
	ec_perweek	0.0468	0.096	0.489	0.625	-0.
141 Email	0.234	Q 1EQ/	0.093	-1.704	0.088	0
341	0.024	-0.1584	0.093	-1.704	0.000	-0.
	quip_failure	0.3861	0.480	0.804	0.422	-0.
556	1.328					
Techie		-0.9994	0.749	-1.334	0.182	-2.
468	0.469					
Port_mode		0.3447	0.560	0.616	0.538	-0.
752	1.442	0.4053	0.064	0.422	0 673	
Phone 478	2.289	0.4053	0.961	0.422	0.673	-1.
478 Multiple		0.2427	0.866	0.280	0.779	-1.
454	1.940	0.2427	0.000	0.200	0.775	-1.
OnlineSed		0.1699	0.587	0.289	0.772	-0.
981	1.320					
OnlineBa	ckup	1.5059	0.724	2.079	0.038	0.
086	2.926					
DevicePro		-1.3796	0.620	-2.226	0.026	-2.
595	-0.165	0.0010	4 000	0.000	0 227	
Streaming 022		0.9810	1.022	0.960	0.337	-1.
Streaming	2.984	1.2512	1.198	1.045	0.296	-1.
096	3.599	1,2312	1.190	1.045	0.290	-1.
MonthlyCh		-0.0213	0.020	-1.048	0.294	-0.
061	•					
Item1		-0.3289	0.405	-0.812	0.417	-1.
123	0.465					
Item2		0.1687	0.378	0.447	0.655	-0.
572	0.909	0 1221	0.245	0.257	0 724	0
Item3 553	A 700	0.1231	0.345	0.357	0.721	-0.
Item4	0.799	0.3003	0.309	0.973	0.331	-0.
305	0.905	0.3003	0.303	0.575	0.331	0.
Item5		0.5701	0.321	1.777	0.076	-0.
059	1.199					
Item6		-0.1784	0.331	-0.539	0.590	-0.
827	0.470					_
Item7	0.000	0.2579	0.312	0.826	0.409	-0.
354	0.869	0 4470	a 209	1 500	0 124	1
Item8 031	0.137	-0.4470	0.298	-1.500	0.134	-1.
Area_Sub		-0.0661	0.685	-0.096	0.923	-1.
409	1.277	3.3331	2.003	0.000	0.525	
Area_Rura		1.0443	0.686	1.522	0.128	-0.
301	2.389					
Marital_N		0.2910	0.780	0.373	0.709	-1.
238	1.820	2 22	0 = 5	0.105	0.000	_
	Separated	-0.0965	0.764	-0.126	0.899	-1.
595	1.402					

localhost:8889/lab 27/78

Marital_Divorced		-0.7169	0.756	-0.948	0.343	-2.
199 0.765 Gender_Female		0.5898	0.567	1.040	0.298	-0.
522 1.702 Gender_Nonbinary		-2.3184	1.885	-1.230	0.219	-6.
014 1.377		0.2097	0.718	0.292	0.770	-1.
Contract_Oneyear 197 1.616						
Contract_TwoYear 028 2.641		1.3067	0.681	1.920	0.055	-0.
<pre>InternetService_None 743      0.537</pre>		-1.1027	0.837	-1.318	0.188	-2.
PaymentMethod_BankTransfe	rautomatic	-1.2209	0.728	-1.677	0.094	-2.
648 0.207 PaymentMethod_ElectronicC	heck	-0.8526	0.640	-1.332	0.183	-2.
107 0.402	=========	========	========	========	======	
Omnibus:		Durbin-Wa			0.167	
Prob(Omnibus):	0.000		, ,		139.966	
Skew: Kurtosis:	0.076 1.258	Prob(JB): Cond. No.			88e-248 .58e+05	
=======================================	========	=======	=======		======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.58e+05. This might indicate that there are strong multicollinearity or other numerical problems.

# In [23]:

#Remove Area Suburban

lm14 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Year1
lm14.params
print(lm14.summary())

# OLS Regression Results

Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	Tenure OLS Least Squares Sat, 12 Feb 2022 23:36:27 8950 8912 37 nonrobust	Adj. R- F-stati Prob (F Log-Lik	squared:		0.006 0.002 1.425 0.0455 -41985. 8.405e+04 8.431e+04	
=======================================						
		coef	std err	t	P> t	[0.
025 0.975]						
Intercept		41.9841	4.898	8.572	0.000	32.
383 51.585						
Lat		-0.1291	0.060	-2.147	0.032	-0.
247 -0.011		0.0111	0.004	0. 400	0.400	
Lng		0.0144	0.021	0.690	0.490	-0.
027 0.055 Population	_ 1	.192e-05	2.44e-05	-1.722	0.085	-8.97e
-05 5.81e-06	-4	.1926-03	2.446-05	-1./22	0.085	-0.576
Children		-0.1163	0.148	-0.786	0.432	-0.
406 0.174						
Age		0.0169	0.014	1.243	0.214	-0.
010 0.043						

localhost:8889/lab 28/78

C	ustorner Churn_Lir	learRegression			
Outage_sec_perweek	0.0468	0.096	0.489	0.625	-0.
141 0.234 Email	-0.1583	0.093	-1.704	0.088	-0.
340 0.024	-0.1363	0.055	-1.704	0.000	-0.
Yearly_equip_failure	0.3857	0.480	0.803	0.422	-0.
556 1.327 Techie	-0.9992	0.749	-1.334	0.182	-2.
468 0.469	0.3332	01743	1.334	0.102	۷.
Port_modem	0.3441	0.560	0.615	0.539	-0.
753 1.441 Phone	0.4046	0.961	0.421	0.674	-1.
479 2.288					
Multiple 454 1.940	0.2426	0.866	0.280	0.779	-1.
OnlineSecurity	0.1690	0.587	0.288	0.773	-0.
981 1.319					
OnlineBackup 087 2.927	1.5065	0.724	2.080	0.038	0.
DeviceProtection	-1.3794	0.620	-2.226	0.026	-2.
594 -0.164					
StreamingTV 022 2.982	0.9801	1.022	0.959	0.337	-1.
StreamingMovies	1.2502	1.198	1.044	0.297	-1.
097 3.598	0.0040		4 040		
MonthlyCharge 061 0.019	-0.0213	0.020	-1.048	0.295	-0.
Item1	-0.3287	0.405	-0.812	0.417	-1.
122 0.465	0.4605	0.270	0.446	0.655	0
Item2 572 0.909	0.1685	0.378	0.446	0.655	-0.
Item3	0.1229	0.345	0.356	0.722	-0.
553 0.799	0 2002	0.200	0 072	0 221	0
Item4 305 0.905	0.3002	0.309	0.972	0.331	-0.
Item5	0.5704	0.321	1.778	0.075	-0.
058 1.199 Item6	-0.1781	0.331	-0.539	0.590	-0.
826 0.470	-0.1781	0.331	-0.555	0.390	-0.
Item7	0.2578	0.312	0.826	0.409	-0.
354 0.869 Item8	-0.4470	0.298	-1.501	0.133	-1.
031 0.137	0.4470	0.230	1.501	0.133	
Area_Rural	1.0775	0.593	1.816	0.069	-0.
086 2.241 Marital_Married	0.2919	0.780	0.374	0.708	-1.
236 1.820	0.2323	01700	0.37	0.700	
Marital_Separated	-0.0967	0.764	-0.127	0.899	-1.
595 1.401 Marital Divorced	-0.7178	0.756	-0.949	0.342	-2.
200 0.764					
Gender_Female 523 1.701	0.5890	0.567	1.039	0.299	-0.
Gender_Nonbinary	-2.3206	1.885	-1.231	0.218	-6.
016 1.374					
Contract_Oneyear 197 1.616	0.2098	0.717	0.292	0.770	-1.
Contract_TwoYear	1.3061	0.681	1.919	0.055	-0.
028 2.640					
<pre>InternetService_None 743      0.537</pre>	-1.1026	0.837	-1.318	0.188	-2.
PaymentMethod_BankTransferautomatic	-1.2202	0.728	-1.676	0.094	-2.
648 0.207	0.0510	0.640	1 224	0 103	2
PaymentMethod_ElectronicCheck 106 0.403	-0.8519	0.640	-1.331	0.183	-2.
				======	

localhost:8889/lab 29/78

Omnibus:		Durbin-Watson:	0.167
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1139.986
Skew:	0.076	Prob(JB):	2.85e-248
Kurtosis:	1.258	Cond. No.	2.57e+05
=======================================	==========		=========

Dep. Variable:

Model:

Method:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 2.57e+05. This might indicate that there are strong multicollinearity or other numerical problems.

## In [24]:

```
#Remove Marital Sep.
lm15 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage_sec_perweek+Email+Yearl
lm15.params
print(lm15.summary())
```

Adj. R-squared:

0.006

0.002

1,465

OLS Regression Results

Least Squares F-statistic:

OLS

Tenure R-squared:

Date: Time: No. Observa Df Residual Df Model: Covariance	ls: Type:	Sat, 12 Feb 2022 23:36:27 8950 8913 36 nonrobust	Prob (F Log-Lik AIC: BIC:	-statistic): elihood:	8. 8.	0.0360 -41985. 404e+04 .431e+04	
		=========	=======	=======	========	:======	
========	====		coef	std err	t	P> t	[0.
025 0.							_
Intercept			41.9595	4.894	8.574	0.000	32.
	L.553		0 1202	0.000	2 147	0.022	0
Lat 247 -0	011		-0.1292	0.060	-2.147	0.032	-0.
Lng	0.011		0.0145	0.021	0.692	0.489	-0.
027 6	055		0.0145	0.021	0.092	0.463	-0.
Population	.033	-Δ	.187e-05	2.43e-05	-1.720	0.085	-8.96e
	5e-06	7	.10/6 05	2.436 03	1.720	0.003	0.300
Children			-0.1163	0.148	-0.786	0.432	-0.
	0.174						
Age			0.0169	0.014	1.243	0.214	-0.
010 6	0.043						
Outage_sec_	_perweek		0.0465	0.096	0.486	0.627	-0.
	0.234						
Email			-0.1584	0.093	-1.704	0.088	-0.
	0.024						
Yearly_equi			0.3875	0.480	0.807	0.420	-0.
	1.329		0.0003	0.740	1 224	0 100	2
Techie 467 0	160		-0.9992	0.749	-1.334	0.182	-2.
Port_modem	0.469		0.3436	0.560	0.614	0.539	-0.
_	L.440		0.3430	0.300	0.014	0.333	-0.
Phone	1.440		0.4036	0.961	0.420	0.674	-1.
	2.287		0.4050	0.501	0.420	0.074	
Multiple			0.2428	0.866	0.280	0.779	-1.
	L.940						
OnlineSecur	rity		0.1694	0.587	0.289	0.773	-0.
	1.320						
OnlineBacku	ıр		1.5072	0.724	2.081	0.037	0.

localhost:8889/lab 30/78

087 DevicePr	2.927 otection		-1.3805	0.620	-2.228	0.026	
595	-0.166		_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0.020		0.020	
Streamin 023	gTV		0.9795	1.021	0.959	0.338	
Streamin	2.982 gMovies		1.2493	1.197	1.043	0.297	
098	3.597						
MonthlyC	_		-0.0213	0.020	-1.048	0.295	
061 Item1	0.019		-0.3291	0.405	-0.813	0.416	
123	0.464		01022	01.05	0.000	01.20	
Item2			0.1687	0.378	0.447	0.655	
571	0.909		0 1224	0.245	0.255	0.722	
Item3 553	0.798		0.1224	0.345	0.355	0.722	
Item4	0.750		0.3002	0.309	0.972	0.331	
305	0.905		<del></del>				
Item5			0.5709	0.321	1.780	0.075	
058	1.200		0.4703	0.334	0 530	0.500	
Item6 827	0.470		-0.1783	0.331	-0.539	0.590	
Item7	0.4/0		0.2579	0.312	0.827	0.408	
354	0.869				- · · · · ·		
Item8			-0.4464	0.298	-1.499	0.134	
030	0.137		4 07	0 500	4 04 4	0.070	
Area_Rur 087			1.0765	0.593	1.814	0.070	
087 Marital_	2.240 Married		0.3244	0.736	0.441	0.660	
119	1.768			0.,50	U	2.300	
_	Divorced		-0.6853	0.711	-0.964	0.335	
079	0.709						
Gender_F			0.5894	0.567	1.039	0.299	
522 Gender N	1.701 onbinary		-2.3186	1.885	-1.230	0.219	
013	1.376		2.5100	1.000	1.250	0.217	
	_Oneyear		0.2098	0.717	0.292	0.770	
197	1.616						
	_TwoYear		1.3055	0.681	1.918	0.055	
029 Internet	2.640 Service_None		-1.1025	0.837	-1.318	0.188	
742	0.537		-1.1023	0.03/	-1.710	0.100	
	ethod_BankTransfer	automatic	-1.2192	0.728	-1.675	0.094	
646	0.208						
-	lethod_ElectronicCh	ieck	-0.8519	0.640	-1.331	0.183	
106	0.403 						
Omnibus:		37293.459				0.167	
Prob(Omn			Jarque-Bei		1	.139.953	
Skew:	•	0.076	*	, ,		90e-248	
Kurtosis	:	1.258	Cond. No.		2	.57e+05	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.57e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [25]: #Remove Multiple
```

```
lm16 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage_sec_perweek+Email+Yearl
lm16.params
print(lm16.summary())
```

localhost:8889/lab 31/78

# OLS Regression Results

Dep. Variable:	Tenure	R-squared:	0.006
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	1.504
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.0284
Time:	23:36:27	Log-Likelihood:	-41985.
No. Observations:	8950	AIC:	8.404e+04
Df Residuals:	8914	BIC:	8.430e+04
Df Model:	35		

Df Mode Covaria	l: nce Type:	35 nonrobust				
	=======================================		========	=======	=======	======
025	0.975]	coef	std err	t	P> t	[0.
		41 5770	4 701	0 045	0 000	22
Interce 364	50.792	41.5778	4.701	8.845	0.000	32.
Lat	30.792	-0.1292	0.060	-2.149	0.032	-0.
	-0.011	-0.1232	0.000	-2.149	0.032	-0.
Lng	0.00=	0.0144	0.021	0.689	0.491	-0.
027	0.055					
Populat	ion	-4.187e-05	2.43e-05	-1.720	0.085	-8.96e
-05	5.85e-06					
Childre	n	-0.1166	0.148	-0.788	0.431	-0.
407	0.174					
Age		0.0168	0.014	1.239	0.215	-0.
010	0.043					_
	sec_perweek	0.0468	0.096	0.489	0.625	-0.
141	0.234	0.1506	0.003	1 707	0.000	0
Email	0.024	-0.1586	0.093	-1.707	0.088	-0.
341 Voanly	0.024 equip_failure	0.3879	0.480	0.808	0.419	-0.
553	1.329	0.3679	0.460	0.000	0.419	-0.
Techie	1.525	-1.0031	0.749	-1.339	0.180	-2.
471	0.465	1.0031	0.745	1.555	0.100	
Port_mo		0.3441	0.559	0.615	0.539	-0.
753	1.441					
Phone		0.4082	0.961	0.425	0.671	-1.
475	2.291					
OnlineS	ecurity	0.1562	0.585	0.267	0.789	-0.
990	1.303					
OnlineB	•	1.4094	0.635	2.220	0.026	0.
165	2.654					
	rotection	-1.4346	0.589	-2.436	0.015	-2.
589 Streami	-0.280	0.7978	0.700	1 010	0 212	0
750	2.345	0.7978	0.790	1.010	0.312	-0.
	ngMovies	1.0241	0.888	1.153	0.249	-0.
717	2.765	1.0241	0.000	1.133	0.243	0.
Monthly		-0.0170	0.013	-1.288	0.198	-0.
043	0.009			_,		
Item1		-0.3296	0.405	-0.814	0.416	-1.
123	0.464					
Item2		0.1663	0.377	0.440	0.660	-0.
574	0.906					
Item3		0.1232	0.345	0.358	0.721	-0.
552	0.799					_
Item4	0.001	0.2991	0.309	0.969	0.333	-0.
306	0.904	0 5700	0 224	4 700	0 07-	^
Item5	1 200	0.5709	0.321	1.780	0.075	-0.
058 T+om6	1.200	0 1703	Q 224	0 520	0 500	0
Item6 827	0.470	-0.1783	0.331	-0.539	0.590	-0.
02/	0.4/0					

localhost:8889/lab 32/78

Item7			0.2587	0.312	0.829	0.407	-0.
353	0.870						
Item8			-0.4458	0.298	-1.497	0.134	-1.
030	0.138						
Area_Ru			1.0756	0.593	1.813	0.070	-0.
087	2.239						
_	_Married		0.3234	0.736	0.439	0.661	-1.
120	1.767						
-	_Divorced		-0.6845	0.711	-0.963	0.336	-2.
078	0.709						_
Gender_I			0.5866	0.567	1.035	0.301	-0.
525	1.698		2 2476	4 005	4 220	0.240	_
	Nonbinary		-2.3176	1.885	-1.230	0.219	-6.
012	1.377		0 2442	0 717	0.204	0.760	4
	t_Oneyear		0.2113	0.717	0.294	0.768	-1.
195	1.618		1 2072	0.600	1 021	0.055	0
027	t_TwoYear 2.641		1.3073	0.680	1.921	0.055	-0.
			1 0022	0.756	1 225	0 105	-2.
484	tService_None 0.480		-1.0022	0.756	-1.325	0.185	-2.
_	Method_BankTransferautomat	ic	-1.2163	0.728	-1.671	0.095	-2.
643	0.211	IC	-1.2103	0.720	-1.0/1	0.033	-2.
	Method_ElectronicCheck		-0.8511	0.640	-1.330	0.184	-2.
106	0.403		-0.0311	0.040	-1.550	0.104	-2.
						======	
Omnibus		. 560	Durbin-Wa	tson:		0.167	
Prob(Om		.000			1	139.870	
Skew:	•	.076				02e-248	
Kurtosi		.258	Cond. No.			.45e+05	
======		=====		=======		======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.45e+05. This might indicate that there are strong multicollinearity or other numerical problems.

## In [26]:

#Remove Online security

lm17 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Year1
lm17.params
print(lm17.summary())

## OLS Regression Results

=====	=========	==========	=======	========	========	======		
Dep. '	Variable:	Tenure	R-square	R-squared:		0.006		
Model: Method: Date:		OLS	Adj. R-squared:			0.002		
		Least Squares	F-statis	tic:		1.547		
		Sat, 12 Feb 2022	Prob (F-	statistic):		0.0222		
Time:		23:36:27	Log-Like	Log-Likelihood:		-41985.		
No. Observations:		8950	AIC:		8.	8.404e+04		
Df Residuals: 8915		BIC:		8.	429e+04			
Df Model:		34						
Covar	iance Type:	nonrobust						
=====	========	=======================================	=======	========	========	========		
=====	=======					- 1.1		
025	0.975]		coef	std err	t	P> t	[0.	
Inter	cept		41.6239	4.697	8.862	0.000	32.	
416	50.831							
Lat			-0.1293	0.060	-2.150	0.032	-0.	
247	-0.011							

localhost:8889/lab 33/78

	Customer Chum_	LillealReglession			
Lng 027 0.055	0.0144	0.021	0.689	0.491	-0.
Population	-4.183e-05	2.43e-05	-1.719	0.086	-8.96e
-05 5.88e-06 Children	-0.1164	0.148	-0.787	0.431	-0.
407 0.174 Age	0.0168	0.014	1.235	0.217	-0.
010 0.043 Outage_sec_perweek	0.0466	0.096	0.487	0.626	-0.
141 0.234 Email	-0.1591	0.093	-1.713	0.087	-0.
341 0.023 Yearly_equip_failure	0.3862	0.480	0.804	0.421	-0.
555 1.327 Techie	-1.0047	0.749	-1.342	0.180	-2.
473 0.463 Port_modem	0.3448	0.559	0.616	0.538	-0.
752 1.441 Phone	0.4062	0.960	0.423	0.672	-1.
477 2.289 OnlineBackup	1.4060	0.635	2.215	0.027	0.
162 2.650 DeviceProtection	-1.4359	0.589	-2.438	0.015	-2.
590 -0.282 StreamingTV	0.7889	0.789	1.000	0.317	-0.
757 2.335 StreamingMovies	1.0137	0.887	1.142	0.253	-0.
726 2.753 MonthlyCharge	-0.0167	0.013	-1.273	0.203	-0.
043 0.009 Item1	-0.3297	0.405	-0.815	0.415	-1.
123 0.464 Item2	0.1660	0.377	0.440	0.660	-0.
574 0.906 Item3	0.1238	0.345	0.359	0.719	-0.
552 0.799 Item4	0.2994	0.309	0.970	0.332	-0.
306 0.904 Item5	0.5701	0.321	1.778	0.075	-0.
059 1.199 Item6	-0.1780	0.331	-0.538	0.590	-0.
826 0.470 Item7	0.2572	0.312	0.825	0.410	-0.
354 0.869 Item8	-0.4452	0.298	-1.495	0.135	-1.
029 0.139 Area_Rural	1.0759	0.593	1.814	0.070	-0.
087 2.239 Marital_Married	0.3207	0.736	0.436	0.663	-1.
122 1.764 Marital_Divorced	-0.6854	0.711	-0.964	0.335	-2.
079 0.708 Gender_Female	0.5842	0.567	1.031	0.303	-0.
527 1.695 Gender_Nonbinary	-2.3234	1.884	-1.233	0.218	-6.
017 1.371 Contract_Oneyear	0.2113	0.717	0.295	0.768	-1.
195 1.617 Contract_TwoYear	1.3095	0.680	1.925	0.054	-0.
024 2.643 InternetService_None	-0.9991	0.756	-1.322	0.186	-2.
481 0.483 PaymentMethod_BankTransferautomatic	-1.2193	0.728	-1.675	0.094	-2.
646 0.207 PaymentMethod_ElectronicCheck	-0.8507	0.640	-1.329	0.184	-2.

localhost:8889/lab 34/78

105	0.404
=====	

Omnibus:	37295.322	Durbin-Watson:	0.167				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1139.896				
Skew:	0.076	Prob(JB):	2.98e-248				
Kurtosis:	1.258	Cond. No.	2.45e+05				

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.45e+05. This might indicate that there are strong multicollinearity or other numerical problems.

## In [27]:

```
#Remove Contract One year
```

Dep. Variable:

lm18 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Yearl
lm18.params
print(lm18.summary())

R-squared:

0.006

OLS Regression Results

Tenure

Df Resido Df Model	rvations: uals: :	OLS Least Squares Sat, 12 Feb 2022 23:36:28 8950 8916 33	Adj. R- F-stati Prob (F Log-Lik	squared: stic: -statistic):		0.002 1.591 0.0171 -41985. 8.404e+04 8.428e+04	
Covarian		nonrobust					
=======		==========	=======	=======	=======	=======	======
			coef	std err	t	P> t	[0.
025							
Intercept			41.6794	4.693	8.881	0.000	32.
480 Lat	50.879		-0.1294	0.060	-2.152	0.031	-0.
	-0.012		-0.1294	0.000	-2.132	0.031	-0.
Lng	0.012		0.0144	0.021	0.690	0.490	-0.
027	0.055						
Population	on	-4	.186e-05	2.43e-05	-1.720	0.086	-8.96e
	.86e-06						
Children			-0.1157	0.148	-0.782	0.434	-0.
406	0.174						
Age			0.0167	0.014	1.233	0.218	-0.
010	0.043		0.0465	0.004		0.607	
	ec_perweek		0.0465	0.096	0.487	0.627	-0.
141	0.234		0 1502	0.003	1 716	0.086	0
Email 341	0.023		-0.1593	0.093	-1.716	0.000	-0.
	quip_failure		0.3875	0.480	0.807	0.420	-0.
553	1.328		0.3073	0.400	0.007	0.420	0.
Techie	1.320		-1.0033	0.749	-1.340	0.180	-2.
471	0.464		_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	• • • • • • • • • • • • • • • • • • • •	_,,,	0.120	
Port_mode			0.3425	0.559	0.612	0.540	-0.
754 <sup>—</sup>	1.439						
Phone			0.4107	0.960	0.428	0.669	-1.
472	2.293						
OnlineBa			1.4044	0.635	2.213	0.027	0.
160	2.648						
DevicePro	otection		-1.4356	0.589	-2.438	0.015	-2.

localhost:8889/lab

590 -0.281 StreamingTV		0.7911	0.789	1.003	0.316	-0.
755 2.337		0.7911	0.769	1.005	0.316	-0.
StreamingMovies 730 2.749		1.0097	0.887	1.138	0.255	-0.
MonthlyCharge		-0.0167	0.013	-1.269	0.204	-0.
042 0.009 Item1		-0.3298	0.405	-0.815	0.415	-1.
123 0.464 Item2		0.1645	0.377	0.436	0.663	-0.
575 0.904 Item3		0.1256	0.345	0.364	0.715	-0.
550 0.801		0.1250	0.545	0.504	0.713	0.
Item4		0.2990	0.309	0.969	0.333	-0.
306 0.904 Item5		0.5705	0.321	1.779	0.075	-0.
058 1.199 Item6		-0.1763	0.331	-0.533	0.594	-0.
824 0.472 Item7		0.2566	0.312	0.823	0.411	-0.
355 0.868		0.2300	0.512	0.023	0.411	0.
Item8 030 0.138		-0.4459	0.298	-1.497	0.134	-1.
Area_Rural		1.0765	0.593	1.815	0.070	-0.
086 2.239 Marital_Married		0.3195	0.736	0.434	0.664	-1.
123 1.762 Marital_Divorced		-0.6883	0.711	-0.968	0.333	-2.
082 0.705		-0.0005	0.711	-0.508	0.555	- 2 •
Gender_Female 526 1.696		0.5852	0.567	1.032	0.302	-0.
Gender_Nonbinary		-2.3222	1.884	-1.232	0.218	-6.
016 1.372 Contract_TwoYear		1.2504	0.650	1.923	0.054	-0.
024 2.525 InternetService_None		-0.9964	0.756	-1.318	0.188	-2.
478 0.485		0.330.	01750	1,310	0.100	-•
PaymentMethod_BankTran 645 0.208	sferautomatic	-1.2189	0.728	-1.675	0.094	-2.
PaymentMethod_Electron 108 0.401	icCheck	-0.8535	0.640	-1.334	0.182	-2.
=======================================					======	
Omnibus:	37292.793				0.167	
Prob(Omnibus):	0.000	•	, ,		139.979	
Skew:	0.076	, ,			86e-248	
Kurtosis:	1.258	Cond. No.			.45e+05	
=======================================	==========	=======	=======	========	======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.45e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [28]:
```

```
#Remove item 3
lm19 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage_sec_perweek+Email+Year1
lm19.params
print(lm19.summary())
```

### OLS Regression Results

Dep. Variable: Tenure R-squared: 0.006
Model: OLS Adj. R-squared: 0.002

Method: Least Squares F-statistic: 1.637 Date: Sat, 12 Feb 2022 Prob (F-statistic): 0.0132 Time: 23:36:28 Log-Likelihood: -41985. AIC: No. Observations: 8950 8.404e+04 Df Residuals: 8917 BIC: 8.427e+04

Df Model: 32 Covariance Type: nonrobust

	nce Type:					
	======== ======	=======================================	========	=======	=======	======
025	0.975]	coef		t		[0.
Interce	•	41.8181	4.677	8.940	0.000	32.
649	50.987	0 1206	0.000	2 155	0.031	0
Lat 247	-0.012	-0.1296	0.060	-2.155	0.031	-0.
Lng	0.012	0.0145	0.021	0.694	0.488	-0.
	0.056					
Populat		-4.193e-05	2.43e-05	-1.723	0.085	-8.96e
	5.78e-06	0 1162	0 140	0.705	0.422	0
Childre 406	n 0.174	-0.1162	0.148	-0.785	0.432	-0.
Age	0.174	0.0167	0.014	1.230	0.219	-0.
_	0.043					
	sec_perweek	0.0461	0.096	0.482	0.630	-0.
141	0.234	0.4504	0.003	4 744	0.007	0
Email 341	0.023	-0.1591	0.093	-1.714	0.087	-0.
	equip_failure	0.3851	0.480	0.802	0.422	-0.
556	1.326					
Techie		-1.0030	0.749	-1.340	0.180	-2.
471	0.465	0.2440	0.550	0.610	0 540	•
Port_mo 755	dem 1.437	0.3410	0.559	0.610	0.542	-0.
Phone	1.43/	0.4103	0.960	0.427	0.669	-1.
	2.293	37.233	0,100	01.2	0,002	_,
OnlineB	ackup	1.4012	0.635	2.208	0.027	0.
157	2.645					_
DeviceP 593	rotection	-1.4394	0.589	-2.445	0.015	-2.
Streami	-0.285 ngTV	0.7853	0.789	0.996	0.319	-0.
760	2.331	017000	07.02	0,120	0.022	
Streami	ngMovies	1.0068	0.887	1.135	0.256	-0.
	2.746					
Monthly 042	Charge 0.009	-0.0166	0.013	-1.266	0.206	-0.
Item1	0.009	-0.2830	0.384	-0.737	0.461	-1.
035	0.469	37233		01.5.	01.02	
Item2		0.1922	0.370	0.520	0.603	-0.
533	0.917	0.0044	0 200	0.050		•
Item4 309	0.901	0.2961	0.308	0.960	0.337	-0.
Item5	0.901	0.5734	0.321	1.789	0.074	-0.
055	1.202	0.373.	0.321	21,703	0.07	•
Item6		-0.1676	0.330	-0.508	0.611	-0.
814	0.479					
Item7	0 072	0.2626	0.311	0.843	0.399	-0.
348 Item8	0.873	-0.4433	0.298	-1.489	0.136	-1.
027	0.140	0.4733	0.200	±ru⊅	0.150	
Area_Ru		1.0827	0.593	1.826	0.068	-0.
080	2.245					_
	_Married	0.3163	0.736	0.430	0.667	-1.
126	1.759					

localhost:8889/lab 37/78

Gender_Female 0.5815 0.567 1.026 0.305 -0.529 1.692 Gender_Nonbinary -2.3260 1.884 -1.234 0.217 -6.020 1.368 Contract_TwoYear 1.2524 0.650 1.927 0.054 -0.022 2.527 InternetService_None -0.9872 0.756 -1.307 0.191 -2.0468 0.494 PaymentMethod_BankTransferautomatic -1.2246 0.728 -1.683 0.092 -2.051 0.202 PaymentMethod_ElectronicCheck -0.8552 0.640 -1.337 0.181 -2.09 0.399	Marital_Divorced		-0.6937	0.711	-0.976	0.329	-2.
Gender_Nonbinary	087 0.699 Gender Female		0.5815	0.567	1.026	0.305	-0.
020       1.368         Contract_TwoYear       1.2524       0.650       1.927       0.054       -0.         022       2.527       -0.9872       0.756       -1.307       0.191       -2.         468       0.494       -0.9872       0.728       -1.683       0.092       -2.         651       0.202       -0.8552       0.640       -1.337       0.181       -2.         109       0.399	529 1.692						
Contract_TwoYear       1.2524       0.650       1.927       0.054       -0.054         022       2.527       0.756       -1.307       0.191       -2.00         468       0.494       0.494       0.728       -1.683       0.092       -2.00         651       0.202       0.202       0.640       -1.337       0.181       -2.00         PaymentMethod_ElectronicCheck       -0.8552       0.640       -1.337       0.181       -2.00         0mnibus:       37289.576       Durbin-Watson:       0.167         Prob(Omnibus):       0.000       Jarque-Bera (JB):       1140.101         Skew:       0.076       Prob(JB):       2.69e-248	_ ,		-2.3260	1.884	-1.234	0.217	-6.
022       2.527         InternetService_None       -0.9872       0.756       -1.307       0.191       -2.         468       0.494         PaymentMethod_BankTransferautomatic       -1.2246       0.728       -1.683       0.092       -2.         651       0.202         PaymentMethod_ElectronicCheck       -0.8552       0.640       -1.337       0.181       -2.         109       0.399							
InternetService_None	<u> </u>		1.2524	0.650	1.927	0.054	-0.
468       0.494         PaymentMethod_BankTransferautomatic       -1.2246       0.728       -1.683       0.092       -2.         651       0.202         PaymentMethod_ElectronicCheck       -0.8552       0.640       -1.337       0.181       -2.         109       0.399         ====================================							_
PaymentMethod_BankTransferautomatic       -1.2246       0.728       -1.683       0.092       -2.         651       0.202       0.202       0.640       -1.337       0.181       -2.         109       0.399       0.399       0.000       0.167       0.167         Prob(Omnibus):       37289.576       Durbin-Watson:       0.167       0.167         Prob(Omnibus):       0.000       Jarque-Bera (JB):       1140.101         Skew:       0.076       Prob(JB):       2.69e-248	_		-0.9872	0.756	-1.307	0.191	-2.
651 0.202  PaymentMethod_ElectronicCheck -0.8552 0.640 -1.337 0.181 -2.  109 0.399			4 2246	0.700	4 603	0.000	2
PaymentMethod_ElectronicCheck       -0.8552       0.640       -1.337       0.181       -2.         109       0.399         ====================================	<del>-</del>	erautomatic	-1.2246	0.728	-1.683	0.092	-2.
109       0.399		Chaal	0 0552	0.640	1 227	0 101	2
Omnibus: 37289.576 Durbin-Watson: 0.167 Prob(Omnibus): 0.000 Jarque-Bera (JB): 1140.101 Skew: 0.076 Prob(JB): 2.69e-248	_	Cneck	-0.8552	0.640	-1.33/	0.181	-2.
Prob(Omnibus):       0.000       Jarque-Bera (JB):       1140.101         Skew:       0.076       Prob(JB):       2.69e-248	109 0.399						
Prob(Omnibus):       0.000       Jarque-Bera (JB):       1140.101         Skew:       0.076       Prob(JB):       2.69e-248	Omnihus:	37289 576	Durhin-Wa	tson:		0 167	
Skew: 0.076 Prob(JB): 2.69e-248					1		
· ·	,		•	, ,			
			, ,				
		========	=======	=======		======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.44e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [29]:
```

```
#Remove phone
```

Dep. Variable:

lm21 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Yearl
lm21.params
print(lm21.summary())

0.006

## OLS Regression Results

Tenure R-squared:

Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	OLS Least Squares Sat, 12 Feb 2022 23:36:28 8950 8918 31 nonrobust	Prob (F- Log-Like AIC: BIC:	•	8.	0.002 1.684 0.0102 -41985. 403e+04 426e+04	
025 0.975]		coef	std err	t	P> t	[0.
Intercept 162 51.197		42.1793	4.600	9.169	0.000	33.
Lat 248 -0.012		-0.1296	0.060	-2.156	0.031	-0.
Lng 026 0.056		0.0145	0.021	0.695	0.487	-0.
Population -05 5.86e-06	-4	.185e-05	2.43e-05	-1.720	0.086	-8.96e
Children 406 0.174		-0.1160	0.148	-0.784	0.433	-0.
Age 010 0.043		0.0168	0.014	1.236	0.217	-0.
Outage_sec_perweek 141 0.234		0.0461	0.096	0.482	0.630	-0.

localhost:8889/lab

Email		-0.1595	0.093	-1.717	0.086	-0.
341 0.023 Yearly_equip_failure		0.3823	0.480	0.797	0.426	-0.
558 1.323						
Techie 470 0.465		-1.0026	0.749	-1.339	0.181	-2.
Port_modem		0.3421	0.559	0.612	0.541	-0.
754 1.438		1 4004	0.635	2 207	0.027	0
OnlineBackup 157 2.644		1.4004	0.635	2.207	0.027	0.
DeviceProtection		-1.4459	0.588	-2.457	0.014	-2.
599 -0.292						
StreamingTV 762 2.329		0.7835	0.789	0.994	0.320	-0.
762 2.329 StreamingMovies		1.0068	0.887	1.135	0.256	-0.
732 2.746						
MonthlyCharge		-0.0167	0.013	-1.267	0.205	-0.
0.009		0. 2060	0.204	0.747	0.455	4
Item1 039 0.465		-0.2868	0.384	-0.747	0.455	-1.
Item2		0.1992	0.369	0.539	0.590	-0.
525 0.923		0.1332	0.505	0.333	0.330	•
Item4		0.2960	0.308	0.959	0.337	-0.
309 0.901						
Item5		0.5747	0.320	1.793	0.073	-0.
054 1.203		A 1600	0.330	-0.512	0.609	0
Item6 815 0.478		-0.1688	0.550	-0.312	0.009	-0.
Item7		0.2635	0.311	0.846	0.397	-0.
347 0.874						
Item8		-0.4425	0.298	-1.487	0.137	-1.
026 0.141		1 005/	0.593	1.831	0.067	0
Area_Rural 077 2.248		1.0854	0.595	1.031	0.007	-0.
Marital_Married		0.3219	0.736	0.437	0.662	-1.
121 1.764						
Marital_Divorced		-0.6918	0.711	-0.973	0.330	-2.
085 0.701						_
Gender_Female 530 1.691		0.5803	0.567	1.024	0.306	-0.
Gender_Nonbinary		-2.3233	1.884	-1.233	0.218	-6.
017 1.370		2.3233	1.004	1.233	0.210	0.
Contract_TwoYear		1.2529	0.650	1.927	0.054	-0.
021 2.527						
InternetService_None		-0.9831	0.755	-1.301	0.193	-2.
464 0.498		1 2240	0.720	1 (0)	0.003	2
PaymentMethod_BankTransfera 650 0.202	automatic	-1.2240	0.728	-1.682	0.093	-2.
PaymentMethod_ElectronicChe	eck	-0.8522	0.640	-1.332	0.183	-2.
106 0.402		<del>-</del>		<del>-</del>		
					=====	
Omnibus:	37289.755			-	0.167	
Prob(Omnibus):	0.000	•	ra (JB):		140.109	
Skew: Kurtosis:	0.076 1.258	Prob(JB): Cond. No.			68e-248 .40e+05	
=======================================	1.230 ========	========	=======	ے =========	======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.4e+05. This might indicate that there are strong multicollinearity or other numerical problems.

localhost:8889/lab 39/78

In [30]:

#Remove Marital Married
lm22 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Outage\_sec\_perweek+Email+Yearl
lm22.params
print(lm22.summary())

#### OLS Regression Results

======		==========			========	======	
Dep. Vai	riable:	Tenure				0.006	
Model:		OLS		squared:		0.002	
Method:		Least Squares				1.734	
Date:		Sat, 12 Feb 2022				0.00773	
Time:		23:36:28	_	elihood:		-41985.	
	ervations:	8950				403e+04	
Df Resid		8919 30	BIC:		8.	425e+04	
	nce Type:	nonrobust					
		===========					
	======						
			coef	std err	t	P> t	[0.
025	0.975]						
Interce			42.2144	4.599	9.178	0.000	33.
	51.230		0.4204	0.000	2 452	0.034	
Lat	-0.012		-0.1294	0.060	-2.153	0.031	-0.
	-0.012		0.0143	0.021	0.686	0.493	-0.
Lng 027	0.055		0.0145	0.021	0.000	0.493	-0.
Populat:		-1	.185e-05	2.43e-05	-1.720	0.086	-8.96e
•	5.86e-06	_	.1050 05	2.436 03	1.720	0.000	0.500
Childre			-0.1161	0.148	-0.785	0.433	-0.
406	0.174						
Age			0.0168	0.014	1.237	0.216	-0.
010	0.043						
Outage_s	sec_perweek		0.0470	0.096	0.492	0.623	-0.
	0.234						
Email			-0.1592	0.093	-1.714	0.087	-0.
341	0.023			0.400	0 704		
	equip_failure		0.3809	0.480	0.794	0.427	-0.
560 Techie	1.321		-1.0003	0.749	-1.336	0.182	-2.
468	0.467		-1.0003	0.749	-1.550	0.102	-2.
Port mod			0.3429	0.559	0.613	0.540	-0.
753	1.439		0.3423	0.555	0.013	0.540	0.
OnlineBa			1.4019	0.634	2.210	0.027	0.
158							
	rotection		-1.4468	0.588	-2.459	0.014	-2.
600	-0.293						
Streamin	0		0.7822	0.788	0.992	0.321	-0.
763	2.328						
	ngMovies		1.0077	0.887	1.136	0.256	-0.
731	2.747		0.0467	0.043	4 270	0.204	
Monthly	•		-0.0167	0.013	-1.270	0.204	-0.
042 Item1	0.009		0 2050	0 201	0 745	0 156	1
038	0.466		-0.2858	0.384	-0.745	0.456	-1.
Item2	0.400		0.1981	0.369	0.536	0.592	-0.
526	0.922		0.1501	0.303	0.550	0.332	0.
Item4	J.J		0.2971	0.308	0.963	0.335	-0.
307	0.902				,		•
Item5			0.5772	0.320	1.801	0.072	-0.
051	1.205						
Item6			-0.1703	0.330	-0.517	0.605	-0.
817	0.476						
Item7			0.2632	0.311	0.845	0.398	-0.

localhost:8889/lab 40/78

347	0.874						
Item8			-0.4412	0.298	-1.482	0.138	-1.
	0.142						
Area_Rur			1.0857	0.593	1.831	0.067	-0.
076	2.248						
Marital_			-0.7690	0.688	-1.117	0.264	-2.
118	0.580						
Gender_F			0.5783	0.567	1.021	0.307	-0.
532	1.689						
Gender_N	_		-2.3252	1.884	-1.234	0.217	-6.
018	1.368						_
Contract			1.2499	0.650	1.923	0.055	-0.
024	2.524			0 755	4 200	0.400	
	Service_None		-0.9837	0.755	-1.302	0.193	-2.
464	0.497		1 2221	0 707	1 600	0.003	2
-	ethod_BankTransfer	automatic	-1.2221	0.727	-1.680	0.093	-2.
648	0.204	o ale	0.0404	0 (40	1 220	0 104	-2.
103	ethod_ElectronicCh 0.404	ieck	-0.8494	0.640	-1.328	0.184	-2.
	0.404 						
Omnibus:		37284.620	Durbin-Wa	tson:		0.167	
Prob(Omn	ibus):	0.000	Jarque-Be	ra (JB):	1	140.303	
Skew:	•	0.076	•	, ,	2.	43e-248	
Kurtosis	•	1.258	Cond. No.		2	.40e+05	
======			========			======	

Dep. Variable:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 2.4e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [31]:

#Remove Outages Sec perweek lm23 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Email+Yearly\_equip\_failure+Tec 1m23.params print(lm23.summary())

0.006

#### OLS Regression Results \_\_\_\_\_\_

Tenure R-squared:

Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	OLS Least Squares Sat, 12 Feb 2022 23:36:28 8950 8920 29 nonrobust	Prob (F-	squared: stic: statistic):	8.	0.003 1.785 0.00588 -41985. 403e+04 424e+04	
025 0.975]		coef	std err	t	P> t	[0.
Intercept 920 51.506 Lat 247 -0.012 Lng 026 0.056 Population -05 6.02e-06	-4	42.7129 -0.1294 0.0145 .168e-05	4.486 0.060 0.021 2.43e-05	9.522 -2.152 0.696 -1.713	0.000 0.031 0.487 0.087	33. -0. -0.
Children		-0.1150	0.148	-0.778	0.437	-0.

localhost:8889/lab 41/78 405

0.175

405	0.1/5						
Age			0.0167	0.014	1.230	0.219	-0.
010	0.043						
Email			-0.1592	0.093	-1.715	0.086	-0.
341	0.023						_
	uip_failure		0.3827	0.480	0.798	0.425	-0.
558	1.323		0.0000	0.740	4 226	0.400	2
Techie	0.467		-0.9999	0.749	-1.336	0.182	-2.
467	0.467		0 2451	0 550	0.617	0 527	-0.
Port_mode	1.441		0.3451	0.559	0.017	0.537	-0.
OnlineBac			1.4014	0.634	2.209	0.027	0.
158	2.645		1.4014	0.054	2.203	0.027	0.
DevicePro			-1.4416	0.588	-2.450	0.014	-2.
	-0.288			0.500	_,,,,,		
Streaming			0.7859	0.788	0.997	0.319	-0.
759	2.331						
Streaming			1.0108	0.887	1.140	0.254	-0.
728	2.750						
MonthlyCh	arge		-0.0167	0.013	-1.269	0.204	-0.
042	0.009						
Item1			-0.2901	0.384	-0.756	0.449	-1.
042	0.462						
Item2			0.1994	0.369	0.540	0.589	-0.
524	0.923						_
Item4	0.004		0.2964	0.308	0.961	0.337	-0.
308	0.901		0 5754	0.330	1 706	0.073	0
Item5	1 202		0.5754	0.320	1.796	0.073	-0.
053 Item6	1.203		-0.1705	0.330	-0.517	0.605	-0.
817	0.476		-0.1703	0.550	-0.517	0.005	-0.
Item7	0.470		0.2609	0.311	0.838	0.402	-0.
349	0.871		0.2003	0.311	0.030	0.402	٠.
Item8	0,0,-		-0.4383	0.298	-1.473	0.141	-1.
022	0.145						
Area_Rura			1.0819	0.593	1.825	0.068	-0.
080	2.244						
Marital_D	ivorced		-0.7655	0.688	-1.112	0.266	-2.
115	0.584						
Gender_Fe			0.5743	0.567	1.014	0.311	-0.
	1.685						
Gender_No			-2.3223	1.884	-1.233	0.218	-6.
015	1.371						_
Contract_			1.2539	0.650	1.929	0.054	-0.
020	2.528		0.0010	0.755	1 200	0.104	2
	ervice_None		-0.9810	0.755	-1.299	0.194	-2.
462	0.500 thod BankTransferau	tomotic	1 2254	0.727	-1.684	0.092	2
651	0.201	Comacic	-1.2254	0.727	-1.004	0.032	-2.
	thod_ElectronicChec	k	-0.8510	0.640	-1.330	0.183	-2.
105	0.403	· K	0.0310	0.040	1.550	0.105	۷.
	=======================================	=======	========	:=======	:=======	======	
Omnibus:		37284.855				0.167	
Prob(Omni		0.000			1	140.310	
Skew:	*	0.076	•	•		43e-248	
Kurtosis:		1.258	Cond. No.		2	.34e+05	
=======		=======	========			=====	
Notes:							

#### Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.34e+05. This might indicate that there are strong multicollinearity or other numerical problems.

localhost:8889/lab 42/78

In [32]:

# Remove item 6
lm24 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Email+Yearly\_equip\_failure+Tec
lm24.params
print(lm24.summary())

### OLS Regression Results

======		===========		========		======	
Dep. Var	riable:	Tenure				0.006	
Model:		OLS		squared:		0.003	
Method:		Least Squares				1.840	
Date:		Sat, 12 Feb 2022				0.00445	
Time:		23:36:28	_	elihood:		-41985.	
	ervations:	8950				403e+04	
Df Resid		8921 28	BIC:		8.	423e+04	
	i. nce Type:	nonrobust					
		==========	:=======	========	========	========	======
	======						
			coef	std err	t	P> t	[0.
025	0.975]						
T			42 4272	4 452	0 534	0.000	22
Intercep 701	51.153		42.42/2	4.452	9.531	0.000	33.
Lat	31.133		-0.1289	0.060	-2.145	0.032	-0.
247	-0.011		-0.1203	0.000	-2.143	0.032	-0.
Lng	0.011		0.0147	0.021	0.704	0.482	-0.
_	0.056		0.02.7	*****		00.02	
Populati		-4	1.172e-05	2.43e-05	-1.715	0.086	-8.94e
-05	5.98e-06						
Childrer	า		-0.1147	0.148	-0.776	0.438	-0.
405	0.175						
Age			0.0166	0.014	1.227	0.220	-0.
	0.043						
Email			-0.1596	0.093	-1.719	0.086	-0.
	0.022		0 2040	0.400	0.000	0 404	
	equip_failure		0.3840	0.480	0.800	0.424	-0.
556 Techie	1.324		-1.0008	0.749	-1.337	0.181	-2.
468	0.467		-1.0000	0.749	-1.33/	0.101	-2.
Port_mod			0.3462	0.559	0.619	0.536	-0.
750	1.442		0.5402	0.333	0.013	0.550	٠.
OnlineBa			1.3989	0.634	2.205	0.027	0.
	2.642						
DevicePr	rotection		-1.4451	0.588	-2.456	0.014	-2.
598	-0.292						
Streamin	ngTV		0.7825	0.788	0.993	0.321	-0.
763	2.328						
	ngMovies		1.0043	0.887	1.132	0.258	-0.
734	2.743		0.0155	0.013	4 044		
Monthly(	•		-0.0166	0.013	-1.261	0.207	-0.
042 T+om1	0.009		0 2201	0 276	0 072	0 202	1
Item1 066	0.410		-0.3281	0.376	-0.872	0.383	-1.
Item2	0.410		0.1770	0.367	0.483	0.629	-0.
542	0.896		0.1770	0.307	0.405	0.025	0.
Item4	0.030		0.2782	0.306	0.908	0.364	-0.
322	0.879			2.300	3.230	3.501	••
Item5	-		0.6090	0.314	1.941	0.052	-0.
006	1.224						
Item7			0.2330	0.307	0.760	0.447	-0.
368	0.834						
Item8			-0.4571	0.295	-1.548	0.122	-1.
036	0.122						
Area_Rur	ral		1.0831	0.593	1.827	0.068	-0.

localhost:8889/lab 43/78

079	2.245						
Marital_Di	vorced		-0.7665	0.688	-1.114	0.265	-2.
116	0.583						
Gender_Fem			0.5768	0.566	1.018	0.309	-0.
	1.687						
Gender_Non	•		-2.3394	1.884	-1.242	0.214	-6.
	1.353		4 0544	0.650	4 020	0.054	
Contract_T			1.2541	0.650	1.930	0.054	-0.
	2.528		0 0747	0.755	1 201	0 107	2
	rvice_None 0.506		-0.9747	0.755	-1.291	0.197	-2.
	b.500 hod_BankTransfera	ıtomatic	-1.2244	0.727	-1.683	0.092	-2.
-	0.202	acomacic	1.22	0.727	1.005	0.032	۷.
	hod ElectronicChe	ck	-0.8491	0.640	-1.328	0.184	-2.
•	0.405						
=======	=======================================		========	=======		======	
Omnibus:		37279.250	Durbin-Wa	tson:		0.167	
Prob(Omnib	us):	0.000		, ,	1	140.519	
Skew:		0.077	` '		2.	19e-248	
Kurtosis:		1.258	Cond. No.		2	.32e+05	
========	===============	========	========	========		======	

Dep. Variable:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.32e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [33]:
```

```
# Remove item 2
lm25 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Email+Yearly_equip_failure+Tec
lm25.params
print(lm25.summary())
```

R-squared:

0.006

## OLS Regression Results

Tenure

Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	OLS Least Squares Sat, 12 Feb 2022 23:36:29 8950 8922 27 nonrobust	F-statis Prob (F- Log-Like AIC:	-statistic):	8.	0.003 1.899 0.00329 -41985. 403e+04 423e+04	
============		=======		=======	=======	:======
=========						
		coef	std err	t	P> t	[0.
025 0.975]						
Intercept		42.5413	4.445	9.570	0.000	33.
828 51.255						
Lat		-0.1287	0.060	-2.141	0.032	-0.
247 -0.011						
Lng		0.0145	0.021	0.695	0.487	-0.
026 0.055						
Population	-	4.17e-05	2.43e-05	-1.714	0.087	-8.94e
-05 6e-06						
Children		-0.1138	0.148	-0.770	0.441	-0.
404 0.176						
Age		0.0167	0.014	1.233	0.218	-0.
010 0.043						
Email		-0.1597	0.093	-1.720	0.085	-0.

localhost:8889/lab

342	0.022		0. 2002	0.400	0.703	0.400	
	uip_failure		0.3803	0.480	0.793	0.428	
560	1.321		4 0000	0 740	4 224	0.400	
Techie	0.467		-1.0000	0.749	-1.336	0.182	
467	0.467		0 0400	0 ==0	0.600	0 500	
Port_moden			0.3482	0.559	0.623	0.533	
748	1.444						
OnlineBack	-		1.4058	0.634	2.217	0.027	
163	2.649						
DeviceProt			-1.4406	0.588	-2.449	0.014	
	-0.288						
Streaming			0.7863	0.788	0.997	0.319	
759	2.331						
StreamingN			1.0109	0.887	1.140	0.254	
727	2.749						
MonthlyCha	_		-0.0167	0.013	-1.271	0.204	
042	0.009					_	
Item1			-0.2195	0.302	-0.727	0.467	
811	0.372						
Item4			0.2763	0.306	0.902	0.367	
	0.877						
Item5			0.6150	0.313	1.962	0.050	
001	1.229						
Item7			0.2491	0.305	0.817	0.414	
348	0.847						
Item8			-0.4466	0.295	-1.517	0.129	
024	0.131						
Area_Rural	L		1.0813	0.593	1.824	0.068	
081	2.243						
Marital_Di	ivorced		-0.7647	0.688	-1.111	0.267	
114	0.584						
Gender_Fen	nale		0.5779	0.566	1.020	0.308	
532	1.688						
Gender_Nor	nbinary		-2.3442	1.884	-1.245	0.213	
036	1.348						
Contract_1	TwoYear		1.2558	0.650	1.933	0.053	
018	2.530						
InternetSe	ervice_None		-0.9801	0.755	-1.298	0.194	
460	0.500						
PaymentMet	hod_BankTransfera	automatic	-1.2237	0.727	-1.682	0.093	
650	0.202						
PaymentMet	hod_ElectronicChe	eck	-0.8433	0.639	-1.319	0.187	
097	0.410						
========		.=======	========	========	========	======	
Omnibus:		37275.854	Durbin-Wa	tson:		0.167	
Prob(Omnib	ous):	0.000			1	140.655	
Skew:	<b>,</b> ·	0.077	•	(/-		04e-248	
Kurtosis:		1.258	Cond. No.			.32e+05	
		1.250					

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.32e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [34]:
```

```
#Remove port modem
lm26 = smf.ols('Tenure ~ Lat+Lng+Population+Children+Age+Email+Yearly_equip_failure+Tec
lm26.params
print(lm26.summary())
```

45/78

OLS Regression Results

Dep. Variable: Tenure R-squared: 0.006 Model: OLS Adj. R-squared: 0.003 Method: Least Squares F-statistic: 1.958 Date: Sat, 12 Feb 2022 Prob (F-statistic): 0.00251 23:36:29 Time: Log-Likelihood: -41985. No. Observations: 8950 AIC: 8.402e+04 Df Residuals: 8923 BIC: 8.422e+04 Df Model: 26

Covariance Type: nonrobust \_\_\_\_\_\_ ========== coef std err t P>|t| [0. 025 0.975] \_\_\_\_\_\_ -----42.6975 4.438 9.621 0.000 Intercept 33. 51.397 998 Lat -0.1292 0.060 -2.150 0.032 -0. 247 -0.011 0.0142 0.682 0.496 Lng 0.021 -0. 0.055 027 -4.166e-05 Population 2.43e-05 -1.712 0.087 -8.94e -05 6.03e-06 -0.1131 Children 0.148 -0.765 0.444 -0. 403 0.177 0.0167 Age 0.014 1.233 0.218 -0. 010 0.043 Email -0.1591 0.093 -1.714 0.087 -0. 341 0.023 Yearly equip failure 0.3833 0.480 0.799 0.424 -0. 557 1.323 Techie -1.0048 0.748 -1.342 0.179 -2. 472 0.462 OnlineBackup 1.4065 0.634 2.218 0.027 0. 163 2.650 DeviceProtection -1.4413 0.588 -2.451 0.014 -2. 594 -0.288 StreamingTV 0.7868 0.788 0.998 0.318 -0. 2.332 1.0157 0.887 1.145 0.252 -0. StreamingMovies 722 2.754 MonthlyCharge -0.0167 0.013 -1.273 0.203 -0. 042 0.009 Item1 -0.2197 0.302 -0.728 0.467 -0. 811 0.372 Item4 0.2776 0.306 0.906 0.365 -0. 323 0.878 Item5 0.6146 0.313 1.961 0.050 0. 000 1.229 Item7 0.2492 0.305 0.818 0.414 -0. 348 0.847 -0.4476 0.295 Item8 -1.520 0.129 -1. 025 0.130 Area\_Rural 1.0859 0.593 1.832 0.067 -0. 2.248 076 Marital Divorced -0.7608 0.688 -1.106 0.269 -2. 110 0.588 Gender Female 0.5712 0.566 1.009 0.313 -0. 1.681 Gender Nonbinary -2.3488 1.883 -1.247 0.212 -6. 041 1.343 Contract\_TwoYear 0.053 1.2579 0.650 1.936 -0. 016 2.532 InternetService None -0.9803 0.755 -1.298 0.194 -2. 460 0.500

localhost:8889/lab 46/78

PaymentMethod_BankTransferautom	atic	-1.2247	0.727	-1.684	0.092	-2.
PaymentMethod_ElectronicCheck 099 0.408		-0.8458	0.639	-1.323	0.186	-2.
Omnibus: 372	====== 60.927	====== Durbin-Wa1	======= :son:	:=======	====== 0.167	
Prob(Omnibus):	0.000	Jarque-Ber	ra (JB):	1	141.173	
Skew:	0.077	Prob(JB):	` '	1.	58e-248	
Kurtosis:	1.257	Cond. No.		2	.32e+05	
=======================================	======				======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 2.32e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [35]:

#### #Remove Lng

lm27 = smf.ols('Tenure ~ Lat+Population+Children+Age+Email+Yearly\_equip\_failure+Techie+ 1m27.params print(lm27.summary())

#### OLS Regression Results \_\_\_\_\_\_

Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	Least Squares Sat, 12 Feb 2022 23:36:29 8950 8924 25 nonrobust	OLS Adj. R-squared: ares F-statistic: 2022 Prob (F-statistic): 5:29 Log-Likelihood: 3950 AIC: 3924 BIC: 25		8. 8.	0.006 0.003 2.017 0.00192 -41986. 8.402e+04 8.421e+04		
=========							
025 0.975]		coef			P> t	[0.	
Intercept		41.3902	4.002	10.343	0.000	33.	
546 49.235 Lat		-0.1293	0.060	-2.151	0.031	-0.	
247 -0.011		0.1200	0.000	2.131	0.031	0.	
Population	-4	.238e-05	2.43e-05	-1.743	0.081	-9e	
-05 5.27e-06							
Children		-0.1116	0.148	-0.755	0.450	-0.	
401 0.178 Age		0.0168	0.014	1.239	0.215	-0.	
010 0.043		0.0100	0.014	1.239	0.215	-0.	
Email		-0.1587	0.093	-1.710	0.087	-0.	
341 0.023							
Yearly_equip_failur	e	0.3765	0.480	0.785	0.432	-0.	
563 1.316		1 0026	0.740	1 240	0 100	2	
Techie 470 0.464		-1.0026	0.748	-1.340	0.180	-2.	
OnlineBackup		1.4127	0.634	2.228	0.026	0.	
170 2.656							
DeviceProtection		-1.4335	0.588	-2.438	0.015	-2.	
586 -0.281		0.7014	0.700	0.001	0 222	0	
StreamingTV 764 2.326		0.7814	0.788	0.991	0.322	-0.	
StreamingMovies		1.0197	0.887	1.150	0.250	-0.	
718 2.758							

localhost:8889/lab 47/78

MonthlyC	•	-0.0167	0.013	-1.272	0.204	-0.
042	0.009					_
Item1		-0.2141	0.302	-0.710	0.478	-0.
805	0.377	0.0700	0 201	0.007	0.044	
Item4	0.070	0.2780	0.306	0.907	0.364	-0.
323	0.879	0 (142	0 242	1 000	0.050	0
Item5	1 220	0.6142	0.313	1.960	0.050	-0.
000 Thom?	1.229	0 2402	0 205	0.015	0 415	0
Item7	0.046	0.2483	0.305	0.815	0.415	-0.
349 Item8	0.846	-0.4487	0.204	1 524	0 120	-1.
026	0.129	-0.4487	0.294	-1.524	0.128	-1.
Area_Rura		1.0870	0.593	1.834	0.067	-0.
075	2.249	1.0070	0.333	1.034	0.007	-0.
Marital_[		-0.7606	0.688	-1.105	0.269	-2.
109	0.588	-0.7000	0.000	-1.105	0.205	-2.
Gender Fe		0.5656	0.566	0.999	0.318	-0.
544	1.676	0.3030	0.300	0.333	0.310	٠.
Gender_No		-2.3564	1.883	-1.251	0.211	-6.
048	1.335	_,	_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		**	
Contract		1.2622	0.650	1.943	0.052	-0.
011	_ 2.536					
	Service_None	-0.9852	0.755	-1.305	0.192	-2.
465	0.495					
PaymentMe	ethod_BankTransferautomatic	-1.2109	0.727	-1.665	0.096	-2.
636	0.214					
PaymentMe	ethod_ElectronicCheck	-0.8339	0.639	-1.305	0.192	-2.
087	0.419					
Omnibus:	37257.47	======== 6	======== +con·	=======	0.167	
Prob(Omn:				1	141.325	
Skew:	0.07	•	, ,		46e-248	
Kurtosis					.09e+05	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.09e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [36]:

#Remove Item 1
lm28 = smf.ols('Tenure ~ Lat+Population+Children+Age+Email+Yearly\_equip\_failure+Techie+
lm28.params
print(lm28.summary())

#### OLS Regression Results

```
______
Dep. Variable:
                    Tenure R-squared:
                                               0.006
Model:
                      OLS
                         Adj. R-squared:
                                               0.003
Method:
               Least Squares
                         F-statistic:
                                               2.081
Date:
             Sat, 12 Feb 2022
                         Prob (F-statistic):
                                             0.00148
Time:
                   23:36:29
                          Log-Likelihood:
                                              -41986.
No. Observations:
                     8950
                          AIC:
                                             8.402e+04
Df Residuals:
                     8925
                          BIC:
                                             8.420e+04
Df Model:
                       24
Covariance Type:
                  nonrobust
______
                           coef
                                std err
                                           t
                                                P>|t|
                                                       [0.
     0.975]
```

localhost:8889/lab 48/78

Intercept			41.1125	3.982	10.323	0.000	33.	
	48.919							
Lat			-0.1297	0.060	-2.158	0.031	-0.	
	-0.012		252 05	2 42 05	4 750	0.000	0.00	
Populatio		-4	.253e-05	2.43e-05	-1.750	0.080	-9.02e	
	11e-06		0 1122	0 140	0.760	0 447	0	
Children 402	0.177		-0.1123	0.148	-0.760	0.447	-0.	
Age	0.1//		0.0169	0.014	1.246	0.213	-0.	
	0.043		0.0105	0.014	1.240	0.213	-0.	
Email	0.043		-0.1584	0.093	-1.706	0.088	-0.	
340	0.024		0,120.	0.023		0.000	•	
	uip_failure		0.3714	0.479	0.775	0.439	-0.	
568	1.311							
Techie			-1.0096	0.748	-1.349	0.177	-2.	
477	0.457							
OnlineBac	kup		1.4224	0.634	2.244	0.025	0.	
180	2.665							
DevicePro			-1.4344	0.588	-2.439	0.015	-2.	
	-0.282							
Streaming			0.7904	0.788	1.003	0.316	-0.	
754	2.335		4 0225	0.006	4 465	0.244		
Streaming			1.0325	0.886	1.165	0.244	-0.	
705	2.770		0.0160	0.013	1 200	0 107	0	
MonthlyCh 043	arge 0.009		-0.0169	0.013	-1.290	0.197	-0.	
Item4	0.009		0.2912	0.306	0.952	0.341	-0.	
308	0.891		0.2312	0.300	0.932	0.541	-0.	
Item5	0.091		0.5947	0.312	1.905	0.057	-0.	
017	1.207		0.33.17	0.312	1.303	0.037	•	
Item7			0.1824	0.290	0.628	0.530	-0.	
387	0.751							
Item8			-0.4974	0.286	-1.737	0.082	-1.	
059	0.064							
Area_Rura	1		1.0899	0.593	1.839	0.066	-0.	
072	2.252							
Marital_D			-0.7595	0.688	-1.104	0.270	-2.	
108	0.589							
Gender_Fe			0.5575	0.566	0.985	0.325	-0.	
552			2 2772	4 003	4 262	0 207		
Gender_No	-		-2.3772	1.883	-1.262	0.207	-6.	
068	1.314		1 2642	0 (50	1 046	0 052	۵	
Contract_ 009	2.538		1.2643	0.650	1.946	0.052	-0.	
	ervice_None		-0.9822	0.755	-1.301	0.193	-2.	
462	0.498		-0.3822	0.755	-1.501	0.193	-2.	
	thod_BankTransferaut	omatic	-1.2165	0.727	-1.673	0.094	-2.	
642	0.209			J., Z,	2.075	0.054	۷.	
	thod_ElectronicCheck		-0.8432	0.639	-1.320	0.187	-2.	
096	0.409		-			-		
=======	=======================================							
Omnibus: 37248.366 Durbin-Watson: 0.167								
Prob(Omni	bus):	0.000	Jarque-E	Bera (JB): 1141.668				
Skew:		0.077	Prob(JB)	:	1.	23e-248		
Kurtosis:		1.257	Cond. No	).	2	.08e+05		
=======	=======================================	======	=======	:=======		======		

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.08e+05. This might indicate that there are strong multicollinearity or other numerical problems.

localhost:8889/lab 49/78

In [37]:

#Remove Item 7
lm29 = smf.ols('Tenure ~ Lat+Population+Children+Age+Email+Yearly\_equip\_failure+Techie+
lm29.params
print(lm29.summary())

#### OLS Regression Results

Dep. Variable: Model:		Tenure OLS Least Squares	R-square Adj. R-s	ed: squared:		0.006 0.003 2.154		
Date:		Sat, 12 Feb 2022				0.00108		
Time:		23:36:29				-41986.		
	onvotions.		_	erriood.				
	ervations:	8950	AIC:			8.402e+04		
Df Resid		8926	BIC:			8.419e+04		
Df Mode		23						
	nce Type:	nonrobust						
		==========	:======:	========	:=======	========	======	
======	======		coef	std err	t	P> t	[0.	
025	0.975]		coei	Stu en	·	PYICI	[o.	
025	_							
Intercep			41.6538	3.888	10.713	0.000	34.	
	49.275		41.0336	3.000	10.713	0.000	54.	
Lat	49.273		0 1200	0.060	-2.147	0.032	-0.	
	0 011		-0.1290	0.000	-2.14/	0.032	-0.	
	-0.011	4	267- 05	2 42- 05	1 756	0.070	0.02-	
Populati		-4	267e-05	2.43e-05	-1.756	0.079	-9.03e	
	1.97e-06		0 4436	0 110	0.764	0.446	•	
Children			-0.1126	0.148	-0.761	0.446	-0.	
402	0.177							
Age			0.0170	0.014	1.252	0.211	-0.	
	0.044							
Email			-0.1577	0.093	-1.699	0.089	-0.	
340	0.024							
Yearly_6	equip_failure		0.3644	0.479	0.760	0.447	-0.	
575	1.304							
Techie			-1.0051	0.748	-1.343	0.179	-2.	
472	0.462							
OnlineBa	ackup		1.4206	0.634	2.241	0.025	0.	
178	2.663							
DevicePr	rotection		-1.4434	0.588	-2.456	0.014	-2.	
596	-0.291							
Streamin	ngTV		0.7801	0.788	0.990	0.322	-0.	
764	2.324							
Streamin	ngMovies		1.0238	0.886	1.155	0.248	-0.	
	2.761							
Monthly(			-0.0168	0.013	-1.282	0.200	-0.	
043	0.009							
Item4			0.3054	0.305	1.002	0.317	-0.	
292	0.903							
Item5			0.5602	0.307	1.823	0.068	-0.	
042	1.163							
Item8			-0.4616	0.281	-1.645	0.100	-1.	
012	0.088							
Area_Rur			1.0917	0.593	1.842	0.065	-0.	
070	2.253		_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0.555	_,,,	0.005		
	Divorced		-0.7554	0.688	-1.098	0.272	-2.	
104	0.593		0.,55.	0.000	1.050	0.2,2		
Gender_F			0.5587	0.566	0.987	0.324	-0.	
551	1.668		0.5507	0.300	0.507	0.324	0.	
	Nonbinary		-2.3713	1.883	-1.259	0.208	-6.	
062	1.320		2.5/15	1.000	1.233	0.200	-0.	
	t_TwoYear		1.2726	0.650	1.959	0.050	-0.	
001	2.546		1.2/20	0.050	1.909	0.000	-0.	
	2.546 tService_None		-0.9811	0.755	-1.300	0.194	-2.	
THEFTHE	rael ATCE MOUE		-0.30TT	0./33	-1.300	0.134		

localhost:8889/lab 50/78

461 0.499						
PaymentMethod_BankTransferautomatic		-1.2191	0.727	-1.677	0.094	-2.
644 0.206						
PaymentMethod_ElectronicCheck		-0.8406	0.639	-1.316	0.188	-2.
093 0.412						
=======================================	========	=======	========		======	
Omnibus:	37245.487	Durbin-Watson:		0.166		
Prob(Omnibus):	0.000	Jarque-Bera (JB): 1141.76		141.763		
Skew:	0.077	7 Prob(JB): 1.17e-248		17e-248		
Kurtosis:	1.257	Cond. No.		2	.03e+05	
=======================================	========	=======	========		======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.03e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [38]:

#### #Remove Yearly equip failure

Dep. Variable:

lm30 = smf.ols('Tenure ~ Lat+Population+Children+Age+Email+Techie+OnlineBackup+DevicePr
lm30.params
print(lm30.summary())

0.005

## OLS Regression Results

Tenure R-squared:

Model: Method: Date: Time: No. Observations: Df Residuals:	OLS Adj. R-squared: Least Squares F-statistic: Sat, 12 Feb 2022 Prob (F-statistic): 23:36:30 Log-Likelihood: 8950 AIC: 8927 BIC:			8.	0.003 2.226 0.000827 -41986. 8.402e+04 8.418e+04		
<pre>Df Model: Covariance Type:</pre>	22 nonrobust						
, ,	:===========	=======	=========	:=======	=======	======	
==========							
025 0.975]			std err			-	
Intercept 193 49.416		41.8041	3.883	10.766	0.000	34.	
Lat 247 -0.011		-0.1289	0.060	-2.145	0.032	-0.	
Population	-4	.286e-05	2.43e-05	-1.764	0.078	-9.05e	
Children		-0.1108	0.148	-0.749	0.454	-0.	
Age		0.0170	0.014	1.251	0.211	-0.	
010 0.044 Email 340 0.024		-0.1584	0.093	-1.707	0.088	-0.	
Techie 475 0.458		-1.0086	0.748	-1.348	0.178	-2.	
OnlineBackup 175 2.660		1.4176	0.634	2.236	0.025	0.	
DeviceProtection 599 -0.294		-1.4465	0.588	-2.461	0.014	-2.	
StreamingTV 762 2.327		0.7827	0.788	0.994	0.320	-0.	
StreamingMovies 711 2.763		1.0260	0.886	1.158	0.247	-0.	
MonthlyCharge		-0.0169	0.013	-1.284	0.199	-0.	

localhost:8889/lab 51/78

043	0.009					
Item4		0.3044	0.305	0.998	0.318	-0
293	0.902					
Item5		0.5602	0.307	1.823	0.068	-0
042	1.163					
Item8		-0.4637	0.281	-1.652	0.098	-1
014	0.086					_
Area_Rur		1.0936	0.593	1.846	0.065	-0
068	2.255	0.7550	0.600	1 000	0 070	2
	_Divorced	-0.7558	0.688	-1.098	0.272	-2
104	0.593	0 5614	0 566	0.002	0 221	0
Gender_F 548	1.671	0.5614	0.566	0.992	0.321	-0
	Vonbinary	-2.3319	1.882	-1.239	0.215	-6
022	1.358	-2.3319	1.002	-1.239	0.213	-0
	t_TwoYear	1.2712	0.650	1.957	0.050	-0
002	2.544	1.2712	0.030	1.337	0.050	Ü
	tService_None	-0.9847	0.755	-1.304	0.192	-2
464	0.495	0120	007.55	_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	01-2-	_
	Method_BankTransferautomatic	-1.2156	0.727	-1.672	0.095	-2
640	0.209					
Payment	Method_ElectronicCheck	-0.8322	0.639	-1.303	0.193	-2
084	$0.4\overline{20}$					
Omnibus	:	======= Durbin-Wa	======= ntson:	========	0.166	
Prob(Omr	nibus): 0.000			1	141.860	
Skew:	0.077			1.12e-248		
Kurtosis	· · ·			.03e+05		

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.03e+05. This might indicate that there are strong multicollinearity or other numerical problems.

### In [39]:

```
#Remove Children
lm31 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm31.params
```

print(lm31.summary())

## OLS Regression Results

Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	Tenure OLS Least Squares Sat, 12 Feb 2022 23:36:30 8950 8928 21 nonrobust	R-squared: Adj. R-squared: F-statistic: Prob (F-statistic): Log-Likelihood: AIC: BIC:		8.	0.005 0.003 2.305 0.000620 -41987. 8.402e+04 8.417e+04		
025 0.975]		coef	std err	t	P> t	[0.	
Intercept 963 49.124 Lat 247 -0.011		41.5434	3.867 0.060	10.743	0.000 0.032	33. -0.	
Population	-4	.289e-05	2.43e-05	-1.765	0.078	-9.05e	

localhost:8889/lab 52/78

05 4	750 06					
o 4	.75e-06		0.0171	0.014	1.266	0.206
)	0.044					
ail			-0.1583	0.093	-1.706	0.088
.0	0.024		1 0001	0 740	4 244	0.400
echie	0.464		-1.0031	0.748	-1.341	0.180
70 nlineBa	0.464		1.4115	0.634	2.227	0.026
69	2.654		1.4115	0.034	2.221	0.020
	otection		-1.4544	0.588	-2.475	0.013
06	-0.302			0.700	5	0.025
treamin	gTV		0.7783	0.788	0.988	0.323
66	2.323					
	gMovies		1.0132	0.886	1.144	0.253
24	2.750		0.06.5-		4	
onthlyC	•		-0.0167	0.013	-1.274	0.203
042 Etem4	0.009		0 3070	0 205	1 010	Q 212
90	0.906		0.3079	0.305	1.010	0.313
tem5	0.500		0.5628	0.307	1.831	0.067
40	1.165		0.0000	0.007		
tem8			-0.4618	0.281	-1.646	0.100
12	0.088					
rea_Rur			1.0859	0.592	1.833	0.067
75	2.247		0 75	0	4	0 0===
	Divorced		-0.7566	0.688	-1.100	0.272
05 ender F	0.592		0 5645	0 566	0.007	Q 210
ender_F 45	1.674		0.5645	0.566	0.997	0.319
_	onbinary		-2.3503	1.882	-1.249	0.212
39	1.339		2.3303	2,002	2,27	V. Z.Z
_	_TwoYear		1.2622	0.649	1.944	0.052
11	2.535					
nternet	Service_None		-0.9858	0.755	-1.306	0.192
-66	0.494					
-	ethod_BankTransfe	erautomatic	-1.2154	0.727	-1.672	0.095
40	0.210	C	0.0343	0.636	4 204	0.400
aymentM 83	ethod_Electronic(	Lneck	-0.8312	0.639	-1.301	0.193
_	0.421 ========					
nibus:	=========	37228.682			======	0.166
rob(Omn:	ibus):	0.000			1	142.385
kew:	<del>- / -</del>	0.077	Prob(JB):	, ,		60e-249
urtosis	•	1.257	Cond. No.			.02e+05

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.02e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [40]: | #Remove streaming tv
```

lm32 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm32.params
print(lm32.summary())

#### OLS Regression Results

\_\_\_\_\_\_ Dep. Variable: Tenure R-squared: 0.005 Model: OLS Adj. R-squared: 0.003 Method: Least Squares F-statistic: 2.372 Date: Sat, 12 Feb 2022 Prob (F-statistic): 0.000524

localhost:8889/lab 53/78

-41987.

23:36:30

Time:

No. Observations: Df Residuals: Df Model:	8950 A 8929 E 20 robust	AIC: BIC:		8.402e+04 8.417e+04			
		coef st	d err	t	P> t  [0	٥.	
025 0.975]							
Tutonout	4.0	. 6257	2 754	10 022	0.000	,	
Intercept 267 47.984	46	0.6257	3.754	10.822	0.000 33	۶.	
Lat	-6	.1281	0.060	-2.133	0.033 -0	∂.	
246 -0.010 Population	_/ 20	95e-05 2.4	130-05	-1.767	0.077 -9.06	50	
-05 4.68e-06	-4.23	756-05 2.2	+36-03	-1.707	0.077 -9.00	Je	
Age	6	0.0171	0.014	1.265	0.206 -0	∂.	
009 0.044							
Email	-6	.1579	0.093	-1.702	0.089 -0	∂.	
340 0.024		0013	0.740	1 220	0 101		
Techie 468 0.465	-1	.0013	0.748	-1.338	0.181 -2	۷٠	
OnlineBackup	1	.2035	0.598	2.013	0.044	ð.	
032 2.375							
DeviceProtection	-1	.5710	0.576	-2.729	0.006 -2	2.	
700 -0.443		. 5200	0.730	0 717	0.474	,	
StreamingMovies 918 1.976	e	.5290	0.738	0.717	0.474 -0	٥.	
MonthlyCharge	-6	0.0076	0.009	-0.815	0.415 -0	۸.	
026 0.011	-						
Item4	6	3100	0.305	1.017	0.309 -0	∂.	
288 0.908	_					_	
Item5	6	.5644	0.307	1.837	0.066 -0	).	
038 1.167 Item8	_0	.4578	0.281	-1.632	0.103 -1	ı	
008 0.092		7.4370	0.201	1.032	0.105	٠.	
Area_Rural	1	.0918	0.592	1.843	0.065 -0	Э.	
069 2.253							
Marital_Divorced	-6	.7543	0.688	-1.096	0.273 -2	2.	
103 0.594 Gender_Female	G	.5551	0.566	0.981	0.327 -0	ð.	
554 1.664	٧	7.3331	0.300	0.901	0.327 -0	٠.	
Gender_Nonbinary	-2	2.3424	1.882	-1.245	0.213 -6	5.	
032 1.347							
Contract_TwoYear	1	2580	0.649	1.937	0.053 -0	∂.	
015 2.531		7574	0.710	1 054	0 202		
<pre>InternetService_None 166     0.651</pre>	-6	7574	0.719	-1.054	0.292 -2	۷.	
PaymentMethod_BankTransferautoma	atic -1	.2204	0.727	-1.679	0.093 -2	2.	
645 0.204							
PaymentMethod_ElectronicCheck	-6	.8331	0.639	-1.304	0.192 -2	2.	
085 0.419							
Omnibus: 372							
Prob(Omnibus):		Ourbin-Watso Jarque-Bera			0.166 3.021		
Skew:		rob(JB):	(/-		e-249		
Kurtosis:		Cond. No.			6e+05		

#### Notes

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.96e+05. This might indicate that there are strong multicollinearity or other numerical problems.

\_\_\_\_\_\_

localhost:8889/lab 54/78

In [41]: #Remove streaming movies

lm33 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm33.params
print(lm33.summary())

### OLS Regression Results

	OLS Regres						
Dep. Variable: Tenum Model: OI Method: Least Square Date: Sat, 12 Feb 202 Time: 23:36:3 No. Observations: 893 Df Residuals: 893 Df Model: Covariance Type: nonrobus		OLS Adj. R-squared: Least Squares F-statistic: Sat, 12 Feb 2022 Prob (F-statistic): 23:36:30 Log-Likelihood: ervations: 8950 AIC: duals: 8930 BIC: l: 19			0.005 0.003 2.470		
025 0.975]		coef		t		[0.	
Intercept		40.1637	3.698	10.861	0.000	32.	
915 47.413 Lat		-0.1282	0.060	-2.135	0.033	-0.	
246 -0.010		0.1202	0.000	2,133	0.033	•	
Population	-4	.273e-05	2.43e-05	-1.759	0.079	-9.04e	
-05 4.9e-06		0.0172	0.014	1.268	0.205	-0.	
Age 009 0.044		0.01/2	0.014	1.200	0.205	-0.	
Email		-0.1579	0.093	-1.701	0.089	-0.	
340 0.024							
Techie 469 0.464		-1.0027	0.748	-1.340	0.180	-2.	
OnlineBackup		1.1091	0.583	1.902	0.057	-0.	
034 2.252		_,	0.1505		0.007		
DeviceProtection		-1.6197	0.572	-2.833	0.005	-2.	
740 -0.499		0 0022	0 007	0 459	0 647	0	
MonthlyCharge 017 0.011		-0.0032	0.007	-0.458	0.647	-0.	
Item4		0.3107	0.305	1.019	0.308	-0.	
287 0.908							
Item5		0.5665	0.307	1.844	0.065	-0.	
036 1.169 Item8		-0.4563	0.281	-1.627	0.104	-1.	
006 0.094		0.4505	0.201	1.027	0.104		
Area_Rural		1.0879	0.592	1.837	0.066	-0.	
073 2.249			0.600	4 007		•	
Marital_Divorced 096 0.601		-0.7477	0.688	-1.087	0.277	-2.	
Gender_Female		0.5514	0.566	0.974	0.330	-0.	
558 1.661							
Gender_Nonbinary		-2.3492	1.882	-1.248	0.212	-6.	
038 1.340		1.2548	0.649	1 022	0 052	0	
Contract_TwoYear 018 2.528		1.2548	0.649	1.932	0.053	-0.	
InternetService_None		-0.6506	0.703	-0.925	0.355	-2.	
028 0.727							
PaymentMethod_BankTrar	nsferautomatic	-1.2319	0.727	-1.695	0.090	-2.	
656 0.193 PaymentMethod_Electror 085 0.419	nicCheck	-0.8331	0.639	-1.304	0.192	-2.	
=======================================		======	========	=======	======		

localhost:8889/lab 55/78

Kurtosis:	1.256	Cond. No.	1.93e+05
Skew:	0.077	Prob(JB):	6.09e-249
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1143.073
Omnibus:	37209.687	Durbin-Watson:	0.166

Model:

Dep. Variable:

Model: Method:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 1.93e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [42]:

```
#Remove monthly charge
lm34 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm34.params
print(lm34.summary())
```

OLS Adj. R-squared:

0.005

0.003

OLS Regression Results

Tenure R-squared:

Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	Least Squares Sat, 12 Feb 2022 23:36:30 8950 8931 18 nonrobust	8931 BIC: 18		0.003 2.595 0.000239 -41988. 8.401e+04 8.415e+04		
=======================================						
025 0.975]			std err		P> t	[0.
Intercept 746 46.537		39.6415	3.517	11.270	0.000	32.
Lat		-0.1282	0.060	-2.135	0.033	-0.
246 -0.010						
Population -05 4.91e-06	-4	.272e-05	2.43e-05	-1.758	0.079	-9.03e
Age		0.0171	0.014	1.261	0.207	-0.
009 0.044						
Email 340 0.024		-0.1582	0.093	-1.705	0.088	-0.
Techie		-1.0047	0.748	-1.343	0.179	-2.
471 0.462				_,_,		
OnlineBackup		1.0369	0.561	1.847	0.065	-0.
063 2.137 DeviceProtection		-1.6637	0.563	-2.953	0.003	-2.
768 -0.559		1.0057	0.303	2.333	0.003	۷.
Item4		0.3110	0.305	1.020	0.308	-0.
287 0.909 Item5		0.5687	0.307	1.851	0.064	-0.
034 1.171		0.3007	0.307	1.851	0.004	-0.
Item8		-0.4568	0.281	-1.628	0.103	-1.
007 0.093		1 0000	0 502	1.838	0.066	0
Area_Rural 073 2.250		1.0886	0.592	1.838	0.000	-0.
Marital_Divorced		-0.7515	0.688	-1.093	0.275	-2.
100 0.597		0 5543	0.566	0.070	0 227	0
Gender_Female 555 1.664		0.5543	0.566	0.979	0.327	-0.
Gender_Nonbinary		-2.3516	1.882	-1.250	0.211	-6.
- -						

localhost:8889/lab 56/78

040 1.337						
Contract_TwoYear		1.2537	0.649	1.931	0.054	-0.
019 2.526						
InternetService_None		-0.5700	0.681	-0.838	0.402	-1.
904 0.764						
PaymentMethod_BankTransferau	tomatic	-1.2299	0.727	-1.693	0.091	-2.
654 0.194						
PaymentMethod_ElectronicChec	k	-0.8292	0.639	-1.298	0.194	-2.
081 0.423						
	=======	========			======	
	37210.191	Durbin-Wat	son:		0.166	
Prob(Omnibus):	0.000	Jarque-Ber	ra (JB):	1:	143.059	
Skew:	0.077	Prob(JB):		6.	14e-249	
Kurtosis:	1.256	Cond. No.		1	.83e+05	
=======================================	=======	========			======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.83e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [43]: #Remove Internet servince none

Model:

Dep. Variable:

lm35 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm35.params
print(lm35.summary())

R-squared:

Adj. R-squared:

0.005

0.003

## OLS Regression Results

Tenure

OLS

Df Resid Df Model Covarian	l: nce Type:	Least Squares Sat, 12 Feb 2022 23:36:31 8950 8932 17 nonrobust	Prob (F- Log-Like AIC: BIC:	-statistic): elihood:	8. 8.	2.707 .000176 -41988. 401e+04 414e+04	
	 :======						
025	_		coef	std err	t 	P> t	[0.
			20 5445	2 545	44.040		2.0
Intercep 651	46.433		39.5417	3.515	11.248	0.000	32.
Lat	40.433		-0.1292	0.060	-2.151	0.032	-0.
	-0.011		0.1252	0.000	2.131	0.032	٠.
Populati		-4	.287e-05	2.43e-05	-1.764	0.078	-9.05e
-05	1.76e-06						
Age			0.0170	0.014	1.253	0.210	-0.
	0.044		0 4530	0.003	4 700	0.000	0
Email 340	0.024		-0.1579	0.093	-1.702	0.089	-0.
Techie	0.024		-0.9978	0.748	-1.334	0.182	-2.
464	0.469		0.3370	017.10	1.33.	0.101	
OnlineBa	ackup		1.0406	0.561	1.854	0.064	-0.
060	2.141						
	rotection		-1.6621	0.563	-2.950	0.003	-2.
767	-0.558		0 2000	0 205	4 047	0.200	0
Item4 288	0.907		0.3099	0.305	1.017	0.309	-0.
Item5	0.307		0.5690	0.307	1.852	0.064	-0.

localhost:8889/lab 57/78

033	1.171						
Item8			-0.4527	0.280	-1.614	0.107	-1.
003	0.097						
Area_Rura			1.0868	0.592	1.835	0.067	-0.
074	2.248						
Marital_[			-0.7483	0.688	-1.088	0.277	-2.
097	0.600						_
Gender_F			0.5544	0.566	0.980	0.327	-0.
555	1.664		2 2672	4 000	4 250	0.000	
Gender_No	_		-2.3673	1.882	-1.258	0.208	-6.
056	1.321		1 2627	0 (40	1 045	0 052	0
Contract_ 010	2.535		1.2627	0.649	1.945	0.052	-0.
	ethod_BankTransfer	automatic	-1.2262	0.727	-1.688	0.092	-2.
650	0.198	automatic	-1.2202	0.727	-1.008	0.032	-2.
	ethod_ElectronicCh	eck	-0.8325	0.639	-1.303	0.192	-2.
085	0.419	CCR	0.0323	0.033	1.303	0.132	۷.
=======		========	=======	========		======	
Omnibus:		37208.884	Durbin-Wa	tson:		0.166	
Prob(Omn	ibus):	0.000	Jarque-Be	ra (JB):	1	143.097	
Skew:	•	0.077	Prob(JB):		6.	02e-249	
Kurtosis	:	1.256	Cond. No.		1	.83e+05	
=======		========	=======	========		======	

Dep. Variable:

Model:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.83e+05. This might indicate that there are strong multicollinearity or other numerical problems.

### In [44]:

```
#Remove Gender female
lm36 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm36.params
print(lm36.summary())
```

R-squared:

Adj. R-squared:

0.005

0.003

## OLS Regression Results

Tenure

OLS

Df Resi Df Mode	ervations: duals:	Least Squares Sat, 12 Feb 2022 23:36:31 8950 8933 16 nonrobust	Prob (F		8.	2.816 .000140 -41988. 401e+04 413e+04	
			coef	std err	t	P> t	[0.
025	0.975] 						
Interce 896	pt 46.647		39.7715	3.508	11.339	0.000	32.
Lat 246	-0.011		-0.1284	0.060	-2.138	0.033	-0.
Populat -05		-4	1.252e-05	2.43e-05	-1.750	0.080	-9.01e
Age	0.043		0.0169	0.014	1.247	0.212	-0.
Email 341	0.023		-0.1593	0.093	-1.716	0.086	-0.
Techie	0.023		-0.9930	0.748	-1.327	0.184	-2.

localhost:8889/lab 58/78

459 0.473						
OnlineBackup 055 2.145		1.0449	0.561	1.861	0.063	-6
DeviceProtection		-1.6743	0.563	-2.972	0.003	-2
779 -0.570		_,,,,	0.000	_,,,	01005	_
Item4		0.3205	0.305	1.052	0.293	-0
277 0.918		0 ===0		4 075	0.064	
Item5 026 1.178		0.5758	0.307	1.875	0.061	-0
1.178 Item8		-0.4552	0.280	-1.623	0.105	-1
005 0.095		00.000	0.100	_,,,	31233	_
Area_Rural		1.0904	0.592	1.841	0.066	-0
071 2.252						_
Marital_Divorced 108 0.588		-0.7596	0.688	-1.105	0.269	-2
Gender_Nonbinary		-2.6547	1.859	-1.428	0.153	-6
298 0.989		_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		0125	
Contract_TwoYear		1.2551	0.649	1.933	0.053	-0
017 2.528		4 0067		4 600	0.001	
PaymentMethod_BankTrans 651 0.198	ferautomatic	-1.2267	0.727	-1.688	0.091	-2
PaymentMethod_Electroni	cCheck	-0.8420	0.639	-1.319	0.187	-2
0.410		010.20	0,000	_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0,10,	
Omnibus:	37193.879	Durbin-Wa	======= tson:		0.166	
Prob(Omnibus):	0.000			1	143.645	
Skew:	0.077	` /			58e-249	
Kurtosis:	1.256	Cond. No.		1	83e+05	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.83e+05. This might indicate that there are strong multicollinearity or other numerical problems.

### In [45]:

```
#Remove item 4
lm37 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm37.params
print(lm37.summary())
```

#### OLS Regression Results

================	===========	=======	========	=======	======	
Dep. Variable:	Tenure	R-squar	ed:		0.005	
Model:	0LS	Adj. R-	squared:		0.003	
Method:	Least Squares	F-stati	stic:		2.930	
Date:	Sat, 12 Feb 2022	Prob (F	-statistic):		0.000115	
Time:	23:36:31	Log-Lik	elihood:		-41989.	
No. Observations:	8950	AIC:		8	.401e+04	
Df Residuals:	8934	BIC:		8	.412e+04	
Df Model:	15					
Covariance Type:	nonrobust					
=======================================	===========	=======	========	=======	========	======
=========						
		coef	std err	t	P> t	[0.
025 0.975]						
Intercept		41.2501	3.214	12.836	0.000	34.
951 47.550						
Lat		-0.1277	0.060	-2.126	0.034	-0.
245 -0.010	_					
Population	-4	.291e-05	2.43e-05	-1.767	0.077	-9.05e

localhost:8889/lab 59/78

-05 4.7e-06						
Age		0.0170	0.014	1.255	0.209	-0.
010 0.044						
Email		-0.1600	0.093	-1.725	0.085	-0.
342 0.022						
Techie		-0.9975	0.748	-1.333	0.182	-2.
464 0.469						
OnlineBackup		1.0421	0.561	1.856	0.063	-0.
058 2.142						
DeviceProtection		-1.6804	0.563	-2.983	0.003	-2.
785 -0.576		0 4420	0 200	4 504	0.444	0
Item5		0.4420	0.280	1.581	0.114	-0.
106 0.990		0 4201	0 270	1 530	0 124	0
Item8 978 0.118		-0.4301	0.279	-1.539	0.124	-0.
978 0.118 Area_Rural		1.0933	0.592	1.846	0.065	-0.
068 2.254		1.0933	0.392	1.040	0.003	-0.
Marital Divorced		-0.7461	0.688	-1.085	0.278	-2.
094 0.602		0.7401	0.000	1.005	0.270	۷.
Gender Nonbinary		-2.6311	1.859	-1.416	0.157	-6.
274 1.012		2.0311	2.033	2.120	0.137	•
Contract_TwoYear		1.2614	0.649	1.943	0.052	-0.
011 2.534						
PaymentMethod_BankTransfer	automatic	-1.2326	0.727	-1.696	0.090	-2.
657 0.192						
PaymentMethod_ElectronicCh	eck	-0.8452	0.639	-1.324	0.186	-2.
097 0.407						
	========	========	========	========	======	
Omnibus:	37174.253				0.165	
Prob(Omnibus):	0.000				144.330	
Skew:	0.077	` '			25e-249	
Kurtosis:	1.255	Cond. No.		1	.67e+05	
=======================================	========	=======	========	=======	======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specifi
- [2] The condition number is large, 1.67e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [46]:

```
#Remove marital divorced
lm38 = smf.ols('Tenure ~ Lat+Population+Age+Email+Techie+OnlineBackup+DeviceProtection+
lm38.params
print(lm38.summary())
```

#### OLS Regression Results

```
______
Dep. Variable:
                     Tenure R-squared:
                                                  0.005
Model:
                       OLS
                          Adj. R-squared:
                                                  0.003
Method:
               Least Squares
                          F-statistic:
                                                 3.055
Date:
              Sat, 12 Feb 2022
                          Prob (F-statistic):
                                               9.57e-05
Time:
                   23:36:31
                           Log-Likelihood:
                                                -41990.
No. Observations:
                      8950
                           AIC:
                                               8.401e+04
Df Residuals:
                      8935
                           BIC:
                                               8.412e+04
Df Model:
                        14
Covariance Type:
                   nonrobust
______
=========
                             coef
                                  std err
                                                   P>|t|
                                                          [0.
     0.975]
Intercept
                          41.0404
                                  3.208 12.794
                                                   0.000
                                                          34.
```

60/78 localhost:8889/lab

			_	ŭ			
752	47.329						
Lat			-0.1275	0.060	-2.123	0.034	-0.
245	-0.010						
Populati		-4	.266e-05	2.43e-05	-1.756	0.079	-9.03e
	.96e-06		0.0160	0.014	1 244	0 212	0
Age 010	0.043		0.0168	0.014	1.244	0.213	-0.
Email	0.045		-0.1594	0.093	-1.718	0.086	-0.
341	0.023		-0.1334	0.055	-1.710	0.000	-0.
Techie	0.023		-0.9938	0.748	-1.329	0.184	-2.
460	0.473		0,7750	017.10	_,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0,10	_,
OnlineBa	ckup		1.0504	0.561	1.871	0.061	-0.
050	2.151						
DevicePr	otection		-1.6682	0.563	-2.962	0.003	-2.
772	-0.564						
Item5			0.4469	0.280	1.599	0.110	-0.
101	0.995						_
Item8	0.404		-0.4235	0.279	-1.516	0.130	-0.
971	0.124		1 0020	0 503	1 020	0.067	0
Area_Rur 077	aı 2.245		1.0839	0.592	1.830	0.067	-0.
Gender N	-		-2.6264	1.859	-1.413	0.158	-6.
270	1.017		-2.0204	1.055	-1.413	0.130	-0.
Contract			1.2665	0.649	1.951	0.051	-0.
006	2.539			0.0.5	_,,,,	0.007	
PaymentM	ethod_BankTransfe	rautomatic	-1.2362	0.727	-1.701	0.089	-2.
661	0.188						
PaymentM	ethod_ElectronicC	heck	-0.8473	0.639	-1.327	0.185	-2.
099	0.405						
	=========				========		
Omnibus:	•1 \	37151.896			4	0.165	
Prob(Omn	ibus):	0.000		Bera (JB):		145.157	
Skew: Kurtosis	•	0.077 1.254	Prob(JB) Cond. No			15e-249 .67e+05	
KUI-C0515	· ===========	1,254	======	=======	T		
				<b></b> _			

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.67e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [47]:

#### #Remove age

lm39 = smf.ols('Tenure ~ Lat+Population+Email+Techie+OnlineBackup+DeviceProtection+Item
lm39.params
print(lm39.summary())

#### OLS Regression Results

==============	===========		========	===
Dep. Variable:	Tenure	R-squared:	0.0	005
Model:	OLS	Adj. R-squared:	0.0	003
Method:	Least Squares	F-statistic:	3.:	171
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	9.00e	-05
Time:	23:36:31	Log-Likelihood:	-419	90.
No. Observations:	8950	AIC:	8.401e	+04
Df Residuals:	8936	BIC:	8.411e	+04
Df Model:	13			
Covariance Type:	nonrobust			
	=======================================		========	========
=========				
		coef std err	t P	> t  [0.
025 0.975]				

localhost:8889/lab 61/78

Intercept			41.9276	3.128	13.405	0.000	35.
	48.059						
Lat			-0.1278	0.060	-2.128	0.033	-0.
	-0.010						
Population		-4	.216e-05	2.43e-05	-1.736	0.083	-8.98e
	45e-06						
Email			-0.1594	0.093	-1.718	0.086	-0.
341	0.022		0.0060	0 740	4 222	0 100	
Techie	0.460		-0.9968	0.748	-1.333	0.183	-2.
463	0.469		1 0507	0 561	1 072	0.061	0
OnlineBack 050	кир 2.151		1.0507	0.561	1.872	0.061	-0.
DevicePro			-1.6562	0.563	-2.941	0.003	-2.
	-0.552		-1.0302	0.303	-2.941	0.003	-2.
Item5	-0.552		0.4443	0.280	1.590	0.112	-0.
104	0.992		0.4443	0.200	1.550	0.112	0.
Item8	0.332		-0.4181	0.279	-1.496	0.135	-0.
966	0.130		01.202	012/2		0.120	•
Area_Rura	1		1.0787	0.592	1.821	0.069	-0.
082	2.240						
Gender_No	nbinary		-2.6839	1.858	-1.444	0.149	-6.
326	0.958						
Contract_	TwoYear		1.2659	0.649	1.950	0.051	-0.
007	2.538						
PaymentMe <sup>-</sup>	thod_BankTransfer	automatic	-1.2187	0.726	-1.678	0.093	-2.
643	0.205						
-	thod_ElectronicCh	eck	-0.8437	0.639	-1.321	0.187	-2.
096	0.408						
	=========				=======		
Omnibus:	hus).	37132.605			1	0.165	
Prob(Omnil Skew:	bus):	0.000	•	Bera (JB):		145.890 49e-249	
Kurtosis:		0.077 1.254	Cond. No			.63e+05	
Kal. (O212)		1.254	conu. No			.036703	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.63e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [48]:

#Remove payment method electronic check

lm40 = smf.ols('Tenure ~ Lat+Population+Email+Techie+OnlineBackup+DeviceProtection+Item
lm40.params
print(lm40.summary())

#### OLS Regression Results

=======================================				
Dep. Variable:	Tenure	R-squared:	0.004	
Model:	OLS	Adj. R-squared:	0.003	
Method:	Least Squares	F-statistic:	3.289	
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	9.01e-05	
Time:	23:36:31	Log-Likelihood:	-41991.	
No. Observations:	8950	AIC:	8.401e+04	
Df Residuals:	8937	BIC:	8.410e+04	
Df Model:	12			
Covariance Type:	nonrobust			
=======================================	=======================================			=======
=========				
		coef std err	t P> t	[0.
025 0.975]				

localhost:8889/lab 62/78

Intercept		41.5285	3.113	13.340	0.000	35.
426 47.631						
Lat		-0.1281	0.060	-2.133	0.033	-0.
246 -0.010						
Population	-4	.181e-05	2.43e-05	-1.721	0.085	-8.94e
-05 5.8e-06						
Email		-0.1572	0.093	-1.694	0.090	-0.
339 0.025						
Techie		-0.9912	0.748	-1.325	0.185	-2.
458 0.475						_
OnlineBackup		1.0445	0.561	1.861	0.063	-0.
056 2.145		4 6605	0 563	2 040	0.000	2
DeviceProtection		-1.6605	0.563	-2.949	0.003	-2.
764 -0.557		0.4472	0 200	1 600	0 110	۵
Item5 101 0.995		0.4472	0.280	1.600	0.110	-0.
Item8		-0.4153	0.279	-1.487	0.137	-0.
963 0.132		-0.4133	0.275	-1.407	0.137	-0.
Area_Rural		1.0686	0.592	1.804	0.071	-0.
092 2.230		1.0000	0.332	1.00	0.072	٠.
Gender_Nonbinary		-2.6967	1.858	-1.451	0.147	-6.
339 0.946						
Contract_TwoYear		1.2578	0.649	1.938	0.053	-0.
015 2.530						
PaymentMethod_BankTransfer	automatic	-0.8490	0.670	-1.266	0.205	-2.
163 0.465						
				:=======	======	
Omnibus:	37116.595				0.164	
Prob(Omnibus):			Bera (JB):		L146.432	
Skew:		Prob(JB)			14e-249	
Kurtosis:	1.253				L.62e+05	
=======================================	========	=======			======	

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.62e+05. This might indicate that there are strong multicollinearity or other numerical problems.

#### In [49]:

```
#Remove PaymentMethod_BankTransferautomatic
lm41 = smf.ols('Tenure ~ Lat+Population+Email+Techie+OnlineBackup+DeviceProtection+Item
lm41.params
print(lm41.summary())
```

#### OLS Regression Results

OLS REGIESSION RESULTS								
Dep. Variable: Model: Method: Date: Time: No. Observations Df Residuals: Df Model: Covariance Type:	Sat, 12 :	Tenure	R-squared: Adj. R-squa F-statistic Prob (F-sta Log-Likelih AIC: BIC:	:: ntistic):	0 3 8.44	992. e+04		
=======================================	coef	std err	t	P> t	[0.025	0.975]		
Intercept Lat Population Email	41.3442 -0.1287 -4.184e-05 -0.1556		13.294 -2.142 -1.723 -1.678	0.000 0.032 0.085 0.093	35.248 -0.246 -8.95e-05 -0.337	47.440 -0.011 5.77e-06 0.026		

localhost:8889/lab 63/78

Techie	-0.9937	0.748	-1.328	0.184	-2.460	0.473
OnlineBackup	1.0460	0.561	1.864	0.062	-0.054	2.146
DeviceProtection	-1.6717	0.563	-2.969	0.003	-2.776	-0.568
Item5	0.4456	0.280	1.594	0.111	-0.102	0.993
Item8	-0.4140	0.279	-1.482	0.138	-0.962	0.134
Area_Rural	1.0797	0.592	1.823	0.068	-0.081	2.241
Gender_Nonbinary	-2.7136	1.858	-1.460	0.144	-6.356	0.929
Contract_TwoYear	1.2557	0.649	1.934	0.053	-0.017	2.528
============	=======	=======	========	========		===
Omnibus:		37090.878	Durbin-Wats	on:	0.	164
Prob(Omnibus):		0.000	Jarque-Bera	(JB):	1147.	408
Skew:		0.077	Prob(JB):		6.98e-	250

Kurtosis:

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

1.62e+05

[2] The condition number is large, 1.62e+05. This might indicate that there are strong multicollinearity or other numerical problems.

1.253 Cond. No.

\_\_\_\_\_\_

#### In [50]:

```
#Remove Techie
```

lm42 = smf.ols('Tenure ~ Lat+Population+Email+OnlineBackup+DeviceProtection+Item5+Item8
lm42.params
print(lm42.summary())

#### OLS Regression Results

=======================================			=========
Dep. Variable:	Tenure	R-squared:	0.004
Model:	OLS	Adj. R-squared:	0.003
Method:	Least Squares	F-statistic:	3.610
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	8.28e-05
Time:	23:36:32	Log-Likelihood:	-41993.
No. Observations:	8950	AIC:	8.401e+04
Df Residuals:	8939	BIC:	8.409e+04
Df Model:	10		
Covariance Type:	nonrobust		

						========
	coef	std err	t	P> t	[0.025	0.975]
Intercept Lat	41.1271 -0.1282	3.106 0.060	13.242 -2.134	0.000 0.033	35.039 -0.246	47.215 -0.010
Population Email	-4.158e-05 -0.1535	2.43e-05 0.093	-1.712 -1.655	0.087 0.098	-8.92e-05 -0.335	6.03e-06 0.028
OnlineBackup	1.0412	0.561	1.855	0.064	-0.059	2.142
DeviceProtection Item5	-1.6863 0.4448	0.563 0.280	-2.995 1.591	0.003 0.112	-2.790 -0.103	-0.583 0.993
Item8	-0.4101	0.279	-1.468	0.142	-0.958	0.138
Area_Rural	1.0805	0.592	1.825	0.068	-0.080	2.241
Gender_Nonbinary	-2.6912 1.2630	1.858 0.649	-1.448 1.945	0.148 0.052	-6.334 -0.010	0.951 2.536
Contract_TwoYear	1.2030	0.049	1.945	0.052	-0.010	2,550

Omnibus:	37056.809	Durbin-Watson:	0.164
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1148.471
Skew:	0.076	Prob(JB):	4.10e-250
Kurtosis:	1.252	Cond. No.	1.62e+05

#### Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.62e+05. This might indicate that there are strong multicollinearity or other numerical problems.

localhost:8889/lab 64/78

```
In [51]: #Remove Gener nonbinary
lm43 = smf.ols('Tenure ~ Lat+Population+Email+OnlineBackup+DeviceProtection+Item5+Item8
lm43.params
print(lm43.summary())
```

#### OLS Regression Results

						:====
Dep. Variable:		Tenure	R-squared:		6	0.004
Model:		OLS	Adj. R-squa	red:	6	0.003
Method:	Leas	st Squares	F-statistic	:	3	3.777
Date:	Sat, 1	2 Feb 2022	Prob (F-sta	tistic):	9.13	Be-05
Time:		23:36:32	Log-Likelih	ood:	-41	L994.
No. Observations	•	8950	AIC:		8.401	Le+04
Df Residuals:		8940	BIC:		8.408	3e+04
Df Model:		9				
Covariance Type:		nonrobust				
=======================================	coef	std err	t	P> t	[0.025	0.975]
Tut-u-ut	44 0554	2 100	42.220		24.060	
Intercept	41.0554				34.968	
Lat	-0.1284		-2.138			
Population		2.43e-05	-1.714			
Email	-0.1528		-1.647		-0.335	0.029
OnlineBackup	1.0220	0.561	1.821	0.069	-0.078	2.122
DeviceProtection	-1.6670	0.563	-2.962	0.003	-2.770	-0.564
Item5	0.4467	0.280	1.598	0.110	-0.101	0.995
Item8	-0.4093	0.279	-1.465	0.143	-0.957	0.138
Area_Rural	1.0766	0.592	1.818	0.069	-0.084	2.238
Contract_TwoYear	1.2691	0.649	1.955	0.051	-0.003	2.542
Omnihus:	=======	======== 37001 335	:========: Durhin-Wats	======= on·		==== 1 163

 Omnibus:
 37001.335
 Durbin-Watson:
 0.163

 Prob(Omnibus):
 0.000
 Jarque-Bera (JB):
 1150.413

 Skew:
 0.076
 Prob(JB):
 1.55e-250

 Kurtosis:
 1.250
 Cond. No.
 1.62e+05

-----

#### Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.62e+05. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [52]:
```

```
#Remove item 8
lm44 = smf.ols('Tenure ~ Lat+Population+Email+OnlineBackup+DeviceProtection+Item5+Area_
lm44.params
print(lm44.summary())
```

#### OLS Regression Results

```
______
Dep. Variable:
                    Tenure R-squared:
                                               0.004
                      OLS Adj. R-squared:
Model:
                                               0.003
Method:
              Least Squares F-statistic:
                                               3.981
            Sat, 12 Feb 2022 Prob (F-statistic): 23:36:32 Log-Likelihood:
Date:
                                            0.000101
Time:
                                              -41995.
                         AIC:
No. Observations:
                     8950
                                             8.401e+04
Df Residuals:
                     8941
                          BIC:
                                             8.407e+04
Df Model:
                       8
Covariance Type: nonrobust
______
              coef std err t P>|t| [0.025 0.975]
```

localhost:8889/lab 65/78

Intercept	39.2803	2.859	13.737	0.000	33.675	44.886
Lat	-0.1268	0.060	-2.111	0.035	-0.245	-0.009
Population	-4.184e-05	2.43e-05	-1.722	0.085	-8.95e-05	5.78e-06
Email	-0.1509	0.093	-1.627	0.104	-0.333	0.031
OnlineBackup	1.0256	0.561	1.827	0.068	-0.075	2.126
DeviceProtection	-1.6751	0.563	-2.976	0.003	-2.778	-0.572
Item5	0.5239	0.275	1.908	0.056	-0.014	1.062
Area_Rural	1.0630	0.592	1.795	0.073	-0.098	2.224
Contract_TwoYear	1.2787	0.649	1.970	0.049	0.006	2.551
Omnibus:	========	 36971.504	Durbin-Wats	======= son:	========	===== 0.162
<pre>Prob(Omnibus):</pre>		0.000	Jarque-Bera	a (JB):	115	1.349
Skew:		0.075	Prob(JB):	, ,	9.72	e-251

Kurtosis:

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

Cond. No.

1.49e+05

[2] The condition number is large, 1.49e+05. This might indicate that there are strong multicollinearity or other numerical problems.

1.249

#### In [53]:

```
#Remove email
```

lm45 = smf.ols('Tenure ~ Lat+Population+OnlineBackup+DeviceProtection+Item5+Area\_Rural+
lm45.params

print(lm45.summary())

#### OLS Regression Results

=======================================	============		
Dep. Variable:	Tenure	R-squared:	0.003
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	4.171
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.000135
Time:	23:36:32	Log-Likelihood:	-41996.
No. Observations:	8950	AIC:	8.401e+04
Df Residuals:	8942	BIC:	8.407e+04
Df Model:	7		

Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
Intercept Lat	37.3712 -0.1245	2.608 0.060	14.331 -2.073	0.000 0.038	32.259 -0.242	42.483 -0.007
Population	-4.243e-05	2.43e-05	-1.747	0.081	-9e-05	5.18e-06
OnlineBackup DeviceProtection	1.0360 -1.6785	0.561 0.563	1.846 -2.982	0.065 0.003	-0.064 -2.782	2.136 -0.575
Item5	0.5265	0.275	1.918	0.055	-0.012	1.065
Area_Rural Contract_TwoYear	1.0582 1.2840	0.592 0.649	1.787 1.978	0.074 0.048	-0.103 0.011	2.219 2.557
============			========	=======	=========	=====

Omnibus:	36932.089	Durbin-Watson:	0.162
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1152.717
Skew:	0.075	Prob(JB):	4.91e-251
Kurtosis:	1.248	Cond. No.	1.36e+05

#### Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified
- [2] The condition number is large, 1.36e+05. This might indicate that there are strong multicollinearity or other numerical problems.

localhost:8889/lab 66/78

```
In [54]:
```

```
#Remove population
```

lm46 = smf.ols('Tenure ~ Lat+OnlineBackup+DeviceProtection+Item5+Area\_Rural+Contract\_Tw
lm46.params

print(lm46.summary())

#### OLS Regression Results

Dep. Variable:	Tenure	R-squared:	0.003
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	4.356
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.000213
Time:	23:36:33	Log-Likelihood:	-41998.
No. Observations:	8950	AIC:	8.401e+04
Df Residuals:	8943	BIC:	8.406e+04
Df Model:	6		

Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
Intercept	36.1641	2.515	14.380	0.000	31.234	41.094
Lat	-0.1024	0.059	-1.745	0.081	-0.218	0.013
OnlineBackup	1.0267	0.561	1.829	0.067	-0.073	2.127
DeviceProtection	-1.6740	0.563	-2.973	0.003	-2.778	-0.570
Item5	0.5260	0.275	1.916	0.055	-0.012	1.064
Area_Rural	1.0544	0.592	1.780	0.075	-0.107	2.215
Contract_TwoYear	1.2663	0.649	1.950	0.051	-0.006	2.539

Omnibus: 36904.363 Durbin-Watson: 0.161
Prob(Omnibus): 0.000 Jarque-Bera (JB): 1153.792
Skew: 0.075 Prob(JB): 2.87e-251
Kurtosis: 1.247 Cond. No. 354.

#### Notes:

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

#### In [55]:

#### #Remove Lat

lm47 = smf.ols('Tenure ~ OnlineBackup+DeviceProtection+Item5+Area\_Rural+Contract\_TwoYea
lm47.params
print(lm47.summary())

## OLS Regression Results

Dep. Variable:	Tenure	R-squared:	0.003
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	4.618
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.000328
Time:	23:36:33	Log-Likelihood:	-41999.
No. Observations:	8950	AIC:	8.401e+04
Df Residuals:	8944	BIC:	8.405e+04
Df Model:	5		

Df Model: 5
Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
Intercept	32.2039	1.082	29.757	0.000	30.082	34.325
OnlineBackup	1.0316	0.561	1.838	0.066	-0.069	2.132
DeviceProtection	-1.6879	0.563	-2.998	0.003	-2.792	-0.584
Item5	0.5252	0.275	1.912	0.056	-0.013	1.064
Area_Rural	1.0518	0.592	1.776	0.076	-0.109	2.213
Contract_TwoYear	1.2600	0.649	1.940	0.052	-0.013	2.533

localhost:8889/lab 67/78

```
Omnibus:
                             36853.829
                                          Durbin-Watson:
                                                                             0.161
                                                                         1155.678
Prob(Omnibus):
                                          Jarque-Bera (JB):
                                  0.000
                                          Prob(JB):
Skew:
                                  0.075
                                                                         1.12e-251
Kurtosis:
                                  1.246
                                          Cond. No.
                                                                              15.6
```

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

### In [56]: #Remove Area Rural

```
lm48 = smf.ols('Tenure ~ OnlineBackup+DeviceProtection+Item5+Contract_TwoYear', data =
lm48.params
print(lm48.summary())
```

#### OLS Regression Results

===========	===============		==========
Dep. Variable:	Tenure	R-squared:	0.002
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	4.983
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.000520
Time:	23:36:33	Log-Likelihood:	-42001.
No. Observations:	8950	AIC:	8.401e+04
Df Residuals:	8945	BIC:	8.405e+04
Df Model:	4		

Covariance Type: nonrobust

===========		========	========		========	=======
	coef	std err	t	P> t	[0.025	0.975]
Intercept	32.5609	1.064	30.616	0.000	30.476	34.646
OnlineBackup	1.0280	0.561	1.831	0.067	-0.073	2.128
DeviceProtection	-1.6986	0.563	-3.017	0.003	-2.802	-0.595
Item5	0.5253	0.275	1.913	0.056	-0.013	1.064
Contract_TwoYear	1.2585	0.649	1.938	0.053	-0.015	2.531

```
      Omnibus:
      36807.122
      Durbin-Watson:
      0.160

      Prob(Omnibus):
      0.000
      Jarque-Bera (JB):
      1157.410

      Skew:
      0.075
      Prob(JB):
      4.70e-252

      Kurtosis:
      1.245
      Cond. No.
      15.2
```

\_\_\_\_\_\_

#### Notes:

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

#### In [57]:

```
#Remove Area Rural
lm49= smf.ols('Tenure ~ DeviceProtection+Item5+Contract_TwoYear', data = data).fit()
lm49.params
print(lm49.summary())
```

#### OLS Regression Results

```
______
Dep. Variable:
                                R-squared:
                         Tenure
                                                            0.002
Model:
                            OLS
                                Adj. R-squared:
                                                            0.002
Method:
                   Least Squares
                                F-statistic:
                                                            5.525
                                Prob (F-statistic):
Date:
                 Sat, 12 Feb 2022
                                                         0.000870
Time:
                        23:36:33
                                Log-Likelihood:
                                                          -42003.
No. Observations:
                           8950
                                AIC:
                                                        8.401e+04
Df Residuals:
                                BIC:
                                                        8.404e+04
                           8946
Df Model:
Covariance Type:
                       nonrobust
```

localhost:8889/lab 68/78

\_\_\_\_\_\_

	coef	std err	t	P> t	[0.025	0.975]
Intercept DeviceProtection Item5 Contract_TwoYear	32.9882 -1.6859 0.5343 1.2557	1.038 0.563 0.275 0.649	31.789 -2.994 1.945 1.933	0.000 0.003 0.052 0.053	30.954 -2.790 -0.004 -0.017	35.022 -0.582 1.073 2.529
Omnibus: Prob(Omnibus): Skew: Kurtosis:	3	6754.208 0.000 0.075 1.243	Durbin-Watso Jarque-Bera Prob(JB): Cond. No.		1159. 1.80e-	_
						===

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

```
In [58]: #Remove Device Contract 2 year
lm50= smf.ols('Tenure ~ DeviceProtection+Item5', data = data).fit()
lm50.params
print(lm50.summary())
```

#### OLS Regression Results

```
______
Dep. Variable:
                      Tenure R-squared:
                                                     0.001
                        OLS Adj. R-squared:
Model:
                                                    0.001
                Least Squares F-statistic:
Method:
                                                    6.416
Date:
               Sat, 12 Feb 2022 Prob (F-statistic):
                                                  0.00164
                            Log-Likelihood:
Time:
                     23:36:33
                                                   -42005.
No. Observations:
                        8950
                            AIC:
                                                  8.402e+04
Df Residuals:
                        8947
                             BIC:
                                                  8.404e+04
Df Model:
                         2
```

Intercept 33.2769 1.027 32.399 0.000 31.264 35.290 DeviceProtection -1.6931 0.563 -3.006 0.003 -2.797 -0.589 Item5 0.5407 0.275 1.968 0.049 0.002 1.079

 Omnibus:
 36703.300
 Durbin-Watson:
 0.158

 Prob(Omnibus):
 0.000
 Jarque-Bera (JB):
 1161.262

 Skew:
 0.075
 Prob(JB):
 6.84e-253

 Kurtosis:
 1.242
 Cond. No.
 14.5

Notes

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

```
In [59]: #Remove Adding back in Bandwidth
lm51= smf.ols('Tenure ~ DeviceProtection+Item5+Bandwidth_GB_Year', data = data).fit()
lm51.params
print(lm51.summary())
```

#### OLS Regression Results

===========			==========
Dep. Variable:	Tenure	R-squared:	0.984
Model:	OLS	Adj. R-squared:	0.984
Method:	Least Squares	F-statistic:	1.778e+05
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.00
Time:	23:36:34	Log-Likelihood:	-23641.

localhost:8889/lab 69/78

No. Observations: Df Residuals: Df Model: Covariance Type:	n	8950 8946 3 onrobust	AIC: BIC:		4.729e+ 4.732e+	
	coef	std err	t	P> t	[0.025	0.975]
Intercept DeviceProtection Item5 Bandwidth_GB_Year	-5.6204 -1.1020 -0.0044 0.0120	0.142 0.072 0.035 1.64e-05		0.000 0.000 0.900 0.000	-5.899 -1.244 -0.074 0.012	-5.341 -0.960 0.065 0.012
Omnibus: Prob(Omnibus): Skew: Kurtosis:		301.677 0.000 -0.318 2.515	Durbin-Watso Jarque-Bera Prob(JB): Cond. No.		1.9 239.0 1.23e- 1.64e+	61 52

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.64e+04. This might indicate that there are strong multicollinearity or other numerical problems.

```
In [60]: #Remove item 5
lm52= smf.ols('Tenure ~ DeviceProtection+Bandwidth_GB_Year', data = data).fit()
lm52.params
print(lm52.summary())
```

#### OLS Regression Results

============	:===========		=========
Dep. Variable:	Tenure	R-squared:	0.984
Model:	OLS	Adj. R-squared:	0.984
Method:	Least Squares	F-statistic:	2.668e+05
Date:	Sat, 12 Feb 2022	<pre>Prob (F-statistic):</pre>	0.00
Time:	23:36:34	Log-Likelihood:	-23641.
No. Observations:	8950	AIC:	4.729e+04
Df Residuals:	8947	BIC:	4.731e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
<pre>Intercept DeviceProtection Bandwidth_GB_Year ==========</pre>	-5.6356 -1.1021 0.0120 =======	0.074 0.072 1.64e-05	-76.593 -15.228 730.098 =======	0.000 0.000 0.000	-5.780 -1.244 0.012	-5.491 -0.960 0.012

Omnibus:	301.736	Durbin-Watson:	1.943				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	239.079				
Skew:	-0.318	Prob(JB):	1.22e-52				
Kurtosis:	2.515	Cond. No.	9.81e+03				

#### Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 9.81e+03. This might indicate that there are strong multicollinearity or other numerical problems.

## **Comparison with Simple regression**

localhost:8889/lab 70/78

```
In [61]: #Remove item 5
lm53= smf.ols('Tenure ~ Bandwidth_GB_Year', data = data).fit()
lm53.params
print(lm53.summary())
```

```
OLS Regression Results
 ______
Dep. Variable: Tenure R-squared:
Model: OLS Adj. R-squared:
                                                                                                                                                                                                   0.983
Model:

Model:

Method:

Date:

Date:

Sat, 12 Feb 2022

Time:

No. Observations:

Df Residuals:

Model:

Model:

Adj. R-squared:

F-statistic:

Prob (F-statistic):

Log-Likelihood:

AIC:

By Residuals:

Model:

Mo
                                                                                                                                                                                                  0.983
                                                                                                                                                                                 5.199e+05
                                                                                                                                                                                                    0.00
                                                                                                                                                                                            -23756.
                                                                                                                                                                                 -23,50.
4.752e+04
                                                                                                                                                                                        4.753e+04
 Df Model:
Covariance Type: nonrobust
 ______
                                            coef std err t P>|t| [0.025 0.975]
  Intercept -6.1272 0.067 -91.494 0.000 -6.258 -5.996 Bandwidth_GB_Year 0.0120 1.66e-05 721.067 0.000 0.012 0.012
 ______
                                                                              269.275 Durbin-Watson:
                                                                                                                                                                                                  1.939
Prob(Omnibus):
Skew:
Kurtosis:
                                                                                0.000 Jarque-Bera (JB):
-0.315 Prob(JB):
2.551 Cond. No.
                                                                                                                                                                                           223.481
                                                                                                                                                                                      2.96e-49
 Kurtosis:
                                                                                                                                                                                           7.41e+03
 ______
```

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 7.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

# 2. Justifying a statistically based variable selection procedure and a model evaluation metric to reduce the initial model in a way that aligns with the research question.

Initially, I began by performing a backwards stepwise regression, however, after running a few models, the adjusted R2 was consistently 1. This indicated potential problems with my model. I ran a correlation matrix on my data and found that Bandwidth\_GB\_Year has a 0.99 correlation with Tenure indicating an almost perfect linear relationship. For the purpose of this project, I decided to remove Bandwidth\_GB\_Year from the regression and try to find any variables that might be significant indicators of Bandwidth. These variables were also reduced by removing the least significant variable at each stage of the model reduction. Ultimately I was left with two variables with p-values less than 0.05. Since these variables accounted for 0.001 of the variability in the data, I added Bandwidth\_GB\_Year into the regression and removed item 5 due to its p-value. This left me with a model that explained 0.984 of the variability in the data.

## 3. Providing a reduced multiple regression model that includes both categorical and continuous variables.

Tenure ~ -1.1021(DeviceProtection) + 0.0120(Bandwidth\_GB\_Year)-5.6356

```
In [62]: linearmodeldata= data[['DeviceProtection','Bandwidth_GB_Year', 'Tenure']]
linearmodeldata.info()
```

localhost:8889/lab 71/78

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8950 entries, 0 to 8949
Data columns (total 3 columns):
    Column
                       Non-Null Count Dtype
0
    DeviceProtection 8950 non-null
                                       int64
    Bandwidth_GB_Year 8950 non-null
                                      float64
 1
    Tenure
                       8950 non-null
                                       float64
dtypes: float64(2), int64(1)
memory usage: 209.9 KB
```

# E. Analyzing the data set using the reduced multiple regression model

- 1. Explaining data analysis process by comparing the initial and reduced multiple regression models, including the following elements:
- the logic of the variable selection technique
- the model evaluation metric

The logic behind selecting so many variables initially was that there may be relationships that were not intuitively related to Tenure. As a result, I included anything related to customer demographics, customer habits, and customer surverys. Initially, I decided to drop 'CaseOrder', 'Customer\_id', 'Interaction', 'UID', 'City', 'TimeZone', 'State', 'County', 'Zip', 'Churn', and 'Job' from the original data set since I felt some of these variables were not necessarily useful such as CaseOrder', 'Customer\_id', 'Interaction', 'UID'. I also removed some columns which I felt duplicated location information such as 'TimeZone', 'City', 'State', 'County', and 'Zip' and I instead kept the location variable with the highest level of granularity, lat and lng. Finally, I removed job and Churn from the data set because I felt like the job column was not useful without further natural language processing and I removed Churn since it was a reponse variable not a predictor.

Once the initial model was constructed I began to refine the model by performing a stepwise regression. At every stage the least significant variable(based on P value) was eliminated and a new model was generated. As explained above, the model appeared to have some issues caused by the high correlation with Bandwidth\_GB\_Year and Tenure. This variable was removed from the regression and the remaining variables were reduced one at a time based on the P-value. Once I was left with only significant variables I re-added Bandwidth\_GB\_Year and I reduced the last remaining variable with a P value above 0.05.

### · a residual plot

Please see the figures below.

2. Providing the output and any calculations of the analysis performed, including the model's residual error.

localhost:8889/lab 72/78

Please see the calculations below for the model's residual error and the code above for the regression outputs.

## 3. Providing the code used to support the implementation of the multiple regression models.

Please see the code above.

## Part V: Data Summary and Implications

## F. Summarizing the findings and assumptions

- 1. Discussing the results of the data analysis, including the following elements:
- a regression equation for the reduced model

y = b1x1 + b2x2

Tenure ~ -1.1021(DeviceProtection)+ 0.0120(Bandwidth\_GB\_Year)-5.6356

 an interpretation of coefficients of the statistically significant variables of the model

For every one unit of Tenure, you can expect to decrease by an average of 1.1021 and you can expect Bandwidth\_GB\_Year to increase by 0.0120.

• the statistical and practical significance of the model

The P-values provided in the final model suggest that all variables in the final model are statistically significant. While comparing the final model with the simple regression, the AIC appears to be better in the simple regression indicating that it may be better at predicting Tenure than the model which includes Device Protection.

• the limitations of the data analysis

As previously mentioned, one limitation of this model is that the AIC for the simple regression is better than the AIC that includes the additional categorical variable. While more variance is explained in the final model which includes device protection, the simple regression will probably perform better with other data. The residuals plotted below indicate that the distribution of erros is fairly normal, with some issues around the tails. The variance is also not constant, it appears to increase and decrease throughout the partial residuals.

## 2. Recommend a course of action based on your results.

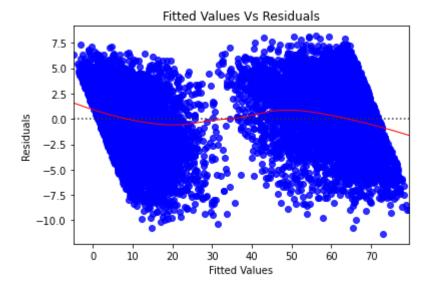
Based on the residuals, I might try to apply a differnt model to see if the behavior can be better predicted, however, for the sake of this company I would suggest that they use the simple

localhost:8889/lab 73/78

regression model to predict Tenure.

```
In [63]: ##3rd party code. See Residual and Prediction Plots In Python.
    uhat=lm52.resid
    yhat = lm52.fittedvalues
```

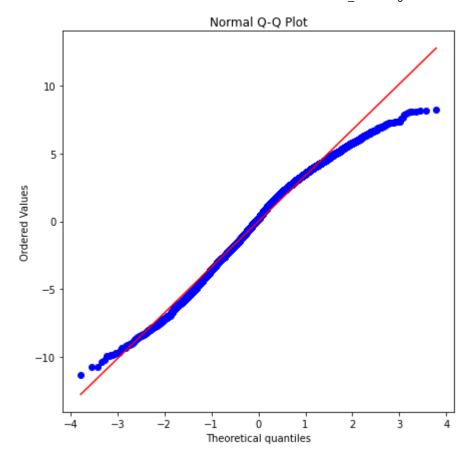
```
In [64]: #3rd party code, see Luhaniwa.
    sns.residplot(x=yhat, y=uhat, lowess=True, color='blue', line_kws={'color': 'red', 'lw'
    plt.xlabel('Fitted Values')
    plt.ylabel('Residuals')
    plt.title("Fitted Values Vs Residuals")
    plt.show()
```



```
In [65]: #3rd party code, see Luhaniwa.

residuals = data.Tenure - yhat
plt.figure(figsize=(7,7))
stats.probplot(residuals, dist="norm", plot=plt)
plt.title("Normal Q-Q Plot")
plt.show()
```

localhost:8889/lab 74/78

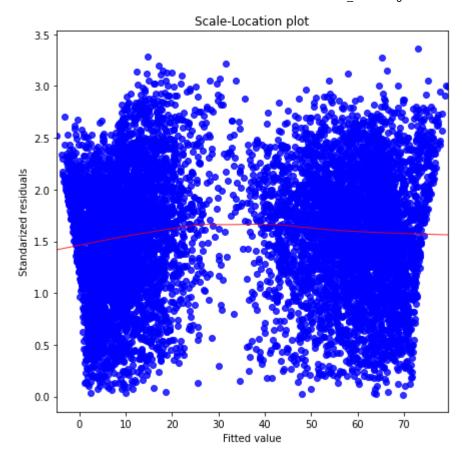


The qq plot above seems to indicate that the assumption of normally distributed residuals is not adequately met. According to the article below, this graph suggest that the residuals have thinner tails than expected.

https://www.ucd.ie/ecomodel/Resources/QQplots\_WebVersion.html

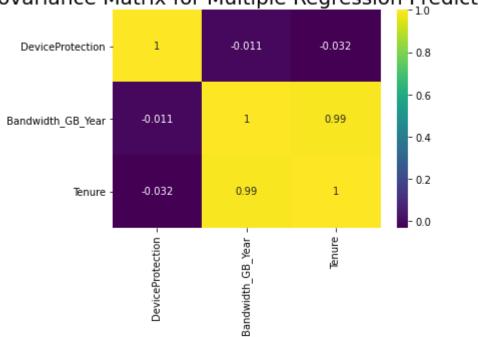
Out[66]: Text(0.5, 1.0, 'Scale-Location plot')

localhost:8889/lab 75/78



```
In [67]: # Multicollonearity
# 3rd party code, see seaborn heatmap
corr = linearmodeldata.corr()
sns.heatmap(corr, annot= True, cmap='viridis')
plt.rcParams['figure.figsize'] = [15, 10]
plt.title('Covariance Matrix for Multiple Regression Predictors', fontsize =20)
plt.show()
# variables not standardized, therefore covariance shown below.
```





localhost:8889/lab 76/78

In [68]: | data.corr()

Out[68]:

	Unnamed:	Lat	Lng	Population	Children	Ag
Unnamed: 0	1.000000	-0.018236	0.003201	-0.016163	-0.009421	0.01250
Lat	-0.018236	1.000000	0.006847	-0.210420	-0.000661	-0.00763
Lng	0.003201	0.006847	1.000000	-0.043112	0.015032	0.00809
Population	-0.016163	-0.210420	-0.043112	1.000000	0.002039	0.01755
Children	-0.009421	-0.000661	0.015032	0.002039	1.000000	-0.01927
Age	0.012500	-0.007636	0.008095	0.017553	-0.019278	1.00000
Income	-0.000956	0.012180	0.003657	-0.012791	0.008053	-0.00432
Outage_sec_perweek	-0.004429	-0.001116	0.017198	0.012753	0.015712	-0.01400
Email	-0.011416	-0.027069	0.004544	0.020289	-0.000699	-0.00056
Contacts	-0.002592	-0.008075	0.006280	0.008257	-0.028512	0.01008
Yearly_equip_failure	0.003645	0.005080	-0.019834	-0.011282	0.016259	-0.00200
Techie	-0.017860	-0.003455	0.005277	-0.007577	-0.009704	-0.00299
Port_modem	0.005150	-0.014387	-0.022159	0.006359	0.008484	0.00025
Tablet	0.000633	0.008252	0.014097	-0.004764	-0.001250	-0.00900
Phone	-0.001871	-0.002295	0.000160	0.008429	0.002409	0.01219
Multiple	0.007000	-0.008121	-0.003038	-0.001752	-0.014049	-0.00479
OnlineSecurity	-0.000618	-0.007688	-0.000286	0.006535	0.003126	-0.01329
OnlineBackup	0.024523	-0.004785	0.017435	0.010121	0.007762	0.00001
DeviceProtection	-0.028176	0.014046	0.023552	-0.007475	0.014090	0.01806
TechSupport	0.002749	-0.007794	-0.003566	-0.015930	-0.001104	0.02351
StreamingTV	0.005045	0.011651	-0.011326	-0.010643	-0.002786	0.00583
StreamingMovies	-0.003230	-0.004550	0.014075	0.009357	0.014210	0.01429
PaperlessBilling	-0.003053	0.018495	0.003579	0.003007	0.010913	-0.00205
Tenure	0.832870	-0.018762	0.006475	-0.013353	-0.007834	0.01191
MonthlyCharge	0.003002	-0.003626	0.015715	0.000007	0.002886	0.01570
Bandwidth_GB_Year	0.825691	-0.019639	0.006555	-0.014148	0.020176	-0.01896
Item1	-0.005678	0.006214	0.026688	0.005561	0.002571	-0.00278
Item2	0.003010	0.009850	0.004409	0.004098	0.009366	0.00822
Item3	0.000503	-0.001275	0.019970	-0.002042	-0.003870	-0.00422
Item4	0.000972	0.011344	0.000181	-0.015931	-0.013314	0.01181
Item5	0.013881	0.001842	-0.000287	0.000937	-0.002719	-0.01024
Item6	-0.009588	-0.008492	-0.005162	0.004876	-0.004582	0.01107

localhost:8889/lab 77/78

	Unnamed: 0	Lat	Lng	Population	Children	Ag
Item7	0.004040	0.014398	0.004845	-0.011001	-0.005531	0.01361
Item8	-0.016094	-0.019631	0.001140	0.009268	-0.009745	0.01768
Area_Suburban	-0.012631	-0.015272	0.003023	0.014162	-0.006477	0.00188
Area_Rural	0.021712	0.002301	0.001952	0.003129	0.016887	-0.00725
Marital_Married	0.010307	0.007252	-0.018527	0.001967	-0.001587	-0.00105
Marital_Separated	0.008118	0.005497	-0.007044	-0.012272	0.001123	-0.00342
Marital_NeverMarried	-0.009006	-0.003240	0.000554	0.018075	-0.006011	0.00211
Marital_Divorced	-0.006747	-0.000093	0.000811	-0.009529	0.001491	0.00978
Gender_Female	0.008082	0.010377	-0.014877	0.011124	-0.010395	-0.00239
Gender_Nonbinary	-0.015772	0.001940	-0.004082	0.000879	0.014595	-0.02508
Contract_Oneyear	-0.007709	-0.006041	-0.000912	-0.006646	0.010080	-0.00757
Contract_TwoYear	0.012564	0.005551	0.009793	0.014128	0.018055	-0.00064
InternetService_DSL	0.004336	-0.014990	-0.012420	-0.003764	-0.001239	-0.00746
InternetService_None	0.000294	0.017852	-0.012695	0.002966	0.008403	0.00885
PaymentMethod_BankTransferautomatic	-0.013118	0.007662	0.019057	-0.000902	-0.000279	0.01909
PaymentMethod_MailedCheck	0.008527	-0.015018	-0.009638	-0.002208	0.000136	-0.01754
PaymentMethod_ElectronicCheck	-0.008028	0.003423	0.018777	-0.010974	-0.001088	-0.00347

49 rows × 49 columns



localhost:8889/lab 78/78