

Feature selection

```
In [2]: import pandas as pd
import numpy as np
from sklearn.linear_model import Lasso
from sklearn.feature_selection import SelectFromModel
```

```
In [3]: data_frame=pd.read_csv('final_data.csv')
```

```
In [5]: data_frame.head()
```

```
Out[5]:
```

	FIPS	45.5_objective	Adj.Death_rate	low_death_rate	upper_death_rate	avg_deaths	rec
0	0	1.0	0.615723	0.695157	0.470742	1.000000	
1	21193	1.0	1.000000	1.000000	0.889692	0.244998	
2	21197	1.0	0.999085	0.970622	0.916463	0.164869	
3	2185	1.0	0.997862	0.858874	1.000000	0.047004	
4	21189	1.0	0.977739	0.904597	0.940306	0.090251	

```
In [6]: x=data_frame.iloc[:,1:]
y=data_frame['FIPS']
```

```
In [18]: feature_sel_model=SelectFromModel(Lasso(alpha=0.005, random_state=0))
feature_sel_model.fit(x,y)
```

```
/opt/anaconda3/lib/python3.8/site-packages/sklearn/linear_model/_c
oordinate_descent.py:529: ConvergenceWarning: Objective did not co
nverge. You might want to increase the number of iterations. Dual i
ty gap: 352735098818.02246, tolerance: 72148110.53876446
model = cd_fast.enet_coordinate_descent(
```

```
Out[18]: SelectFromModel(estimator=Lasso(alpha=0.005, random_state=0))
```

```
In [19]: selected_features=[]
for i,j in zip(x.columns,feature_sel_model.get_support()):
    if j==True:
        selected_features.append(i)
```

```
In [20]: selected_features
```

```
Out[20]: ['45.5_objective',  
          'Adj.Death_rate',  
          'low_death_rate',  
          'upper_death_rate',  
          'avg_deaths',  
          'recent_trend',  
          'recent_5_year_trend',  
          'lower_trend',  
          'upper_trend']
```

```
In [21]: #all features are important.
```