Description

Since the first automobile, the Benz Patent Motor Car in 1886, Mercedes-Benz has stood for important automotive innovations. These include, for example, the passenger safety cell with crumple zone, the airbag and intelligent assistance systems. Mercedes-Benz applies for nearly 2000 patents per year, making the brand the European leader among premium car makers. Daimler’s Mercedes-Benz cars are leaders in the premium car industry. With a huge selection of features and options, customers can choose the customized Mercedes-Benz of their dreams. .

To ensure the safety and reliability of each and every unique car configuration before they hit the road, Daimler’s engineers have developed a robust testing system. But, optimizing the speed of their testing system for so many possible feature combinations is complex and time-consuming without a powerful algorithmic approach. As one of the world’s biggest manufacturers of premium cars, safety and efficiency are paramount on Daimler’s production lines.

Problem Statement / Business Objective

Reduce the time that cars spend on the test bench , resulting in lower carbon dioxide emissions without reducing Daimler’s standards.

Machine Learning Problem

Our job is to predict how long a car on a production line will take to pass the testing phase. This is a classical regression problem.

Performence Metric

- R2 (R square) metric

Exploratory Data Analysis

Let's take a look at the data :

In [1]:

```
import pandas as pd
import numpy as np
from sklearn import preprocessing
from sklearn.metrics import r2_score

import matplotlib.pyplot as plt
```

In [4]:

```
Benz_df = pd.read_csv('train.csv')
print("Data shape : ", Benz_df.shape)
print("Columns : ", Benz_df.columns)
Benz_df.head()
```

```
Data shape :  (4209, 378)
Columns :  Index(['ID', 'y', 'X0', 'X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X
8',
       ...
       'X375', 'X376', 'X377', 'X378', 'X379', 'X380', 'X382', 'X383', 'X3
84',
       'X385'],
      dtype='object', length=378)
```

Out[4]:

	ID	y	X0	X1	X2	X3	X4	X5	X6	X8	...	X375	X376	X377	X378	X379	X380
0	0	130.81	k	v	at	a	d	u	j	o	...	0	0	1	0	0	0
1	6	88.53	k	t	av	e	d	y	l	o	...	1	0	0	0	0	0
2	7	76.26	az	w	n	c	d	x	j	x	...	0	0	0	0	0	0
3	9	80.62	az	t	n	f	d	x	l	e	...	0	0	0	0	0	0
4	13	78.02	az	v	n	f	d	h	d	n	...	0	0	0	0	0	0

5 rows × 378 columns

- 'y' is the time taken at Testing Phase in Seconds
- There are 376 Car features apart from the target variable 'y' and the 'ID' column
- The Car feature names are Anonymised.
- ID column is not equal to rows.

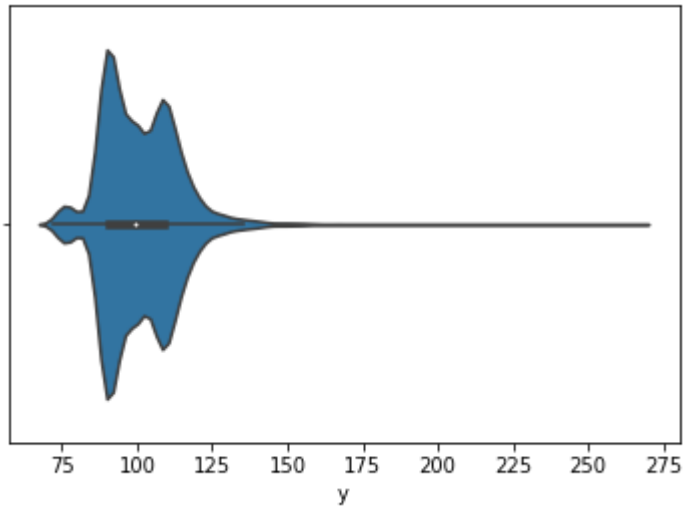
Lets first look at our target varible 'y' ie time taken at Testing Phase (in seconds)

In [3]:

```
import seaborn as sns
sns.violinplot(x=Benz_df['y'])
```

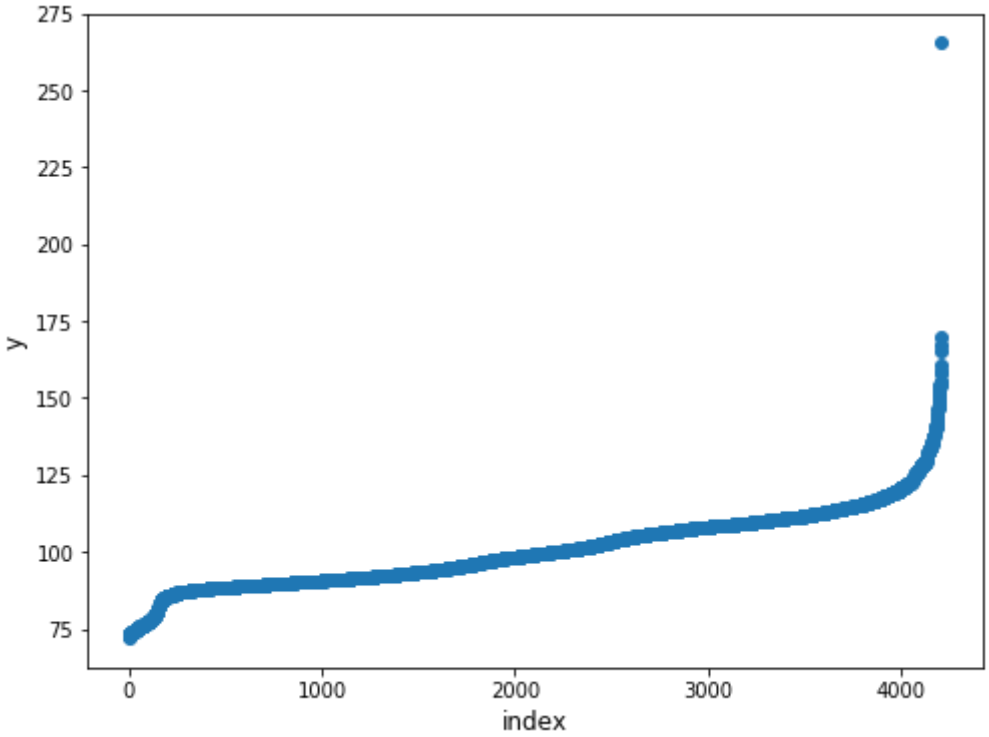
Out[3]:

```
<matplotlib.axes._subplots.AxesSubplot at 0xc0972b0>
```



In [4]:

```
plt.figure(figsize=(8,6))
plt.scatter(range(Benz_df.shape[0]), np.sort(Benz_df.y.values))
plt.xlabel('index', fontsize=12)
plt.ylabel('y', fontsize=12)
plt.show()

print(Benz_df['y'].describe())
```




```
count    4209.000000
mean      100.669318
std        12.679381
min        72.110000
25%        90.820000
50%        99.150000
75%       109.010000
max       265.320000
Name: y, dtype: float64
```

- looking at the scatter plot it seems like a data point(at 265 seconds) is well above the rest.

Missing Values

In [4]:

```
Benz_df.isnull().any().any()
```

Out[4]:

```
False
```

There are no Missing Values in the dataset

Feature Analysis

In [5]:

```
# Seeing datatypes of all the columns Except ID and y Columns
cols = [c for c in Benz_df.columns if 'X' in c]
print('Number of features: {}'.format(len(cols)))

print('Feature types:')
Benz_df[cols].dtypes.value_counts()
```

```
Number of features: 376
Feature types:
```

Out[5]:

```
int64     368
object      8
dtype: int64
```

So we have 376 integer columns and 8 categorical columns

Integer Features

Lets look at the cardinality of all columns

In [6]:

```
## Integer columns
unique_values_dict = {}
for col in Benz_df.columns:
    if col not in ["ID", "y", "X0", "X1", "X2", "X3", "X4", "X5", "X6", "X8"]:
        unique_value = str(np.sort(Benz_df[col].unique()).tolist())
        tlist = unique_values_dict.get(unique_value, [])
        tlist.append(col)
        unique_values_dict[unique_value] = tlist[:]
for unique_val, columns in unique_values_dict.items():
    print("Columns containing the unique values : ",unique_val)
    print("Number of Columns : ",len(columns))
    print(columns)
    print("--------------------------------------------------")
```

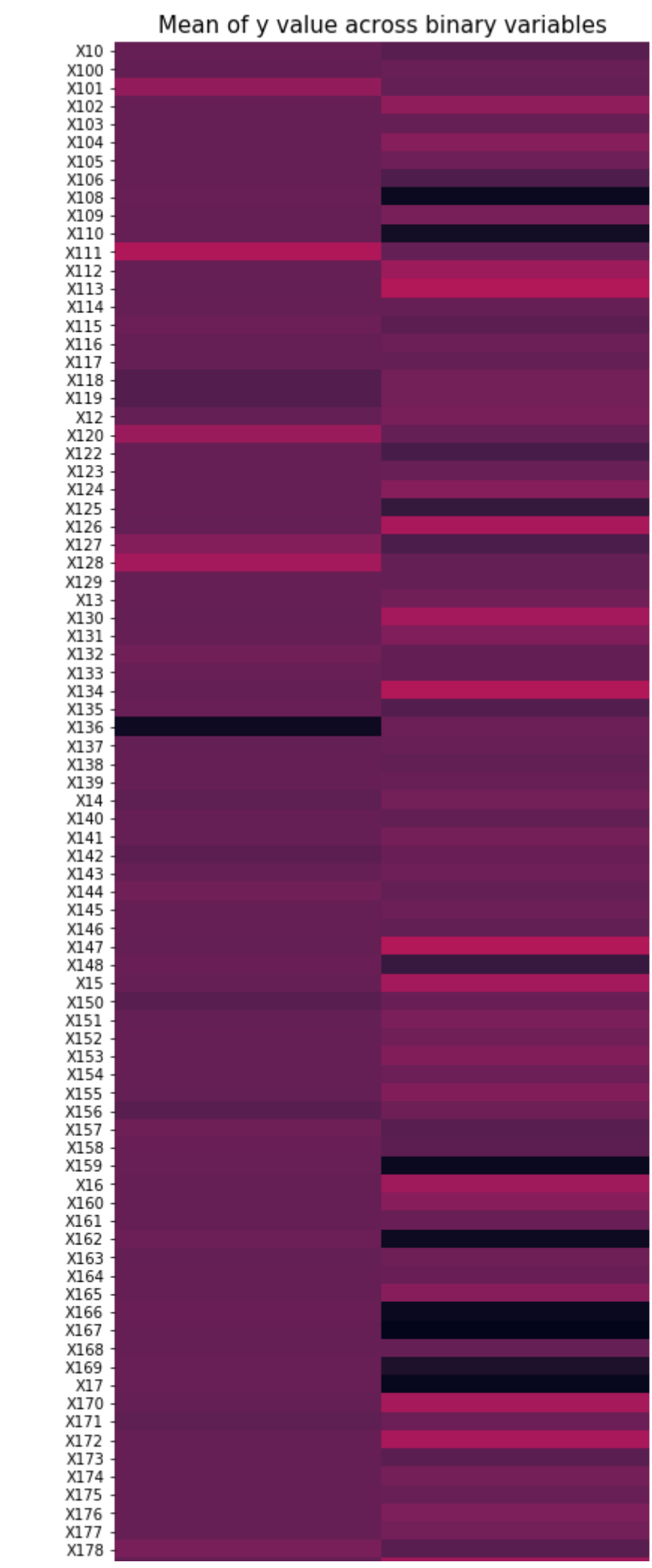
```
Columns containing the unique values :  [0, 1]
Number of Columns :  356
['X10', 'X12', 'X13', 'X14', 'X15', 'X16', 'X17', 'X18', 'X19', 'X20', 'X2
1', 'X22', 'X23', 'X24', 'X26', 'X27', 'X28', 'X29', 'X30', 'X31', 'X32',
'X33', 'X34', 'X35', 'X36', 'X37', 'X38', 'X39', 'X40', 'X41', 'X42', 'X4
3', 'X44', 'X45', 'X46', 'X47', 'X48', 'X49', 'X50', 'X51', 'X52', 'X53',
'X54', 'X55', 'X56', 'X57', 'X58', 'X59', 'X60', 'X61', 'X62', 'X63', 'X6
4', 'X65', 'X66', 'X67', 'X68', 'X69', 'X70', 'X71', 'X73', 'X74', 'X75',
'X76', 'X77', 'X78', 'X79', 'X80', 'X81', 'X82', 'X83', 'X84', 'X85', 'X8
6', 'X87', 'X88', 'X89', 'X90', 'X91', 'X92', 'X94', 'X95', 'X96', 'X97',
'X98', 'X99', 'X100', 'X101', 'X102', 'X103', 'X104', 'X105', 'X106', 'X10
8', 'X109', 'X110', 'X111', 'X112', 'X113', 'X114', 'X115', 'X116', 'X11
7', 'X118', 'X119', 'X120', 'X122', 'X123', 'X124', 'X125', 'X126', 'X12
7', 'X128', 'X129', 'X130', 'X131', 'X132', 'X133', 'X134', 'X135', 'X13
6', 'X137', 'X138', 'X139', 'X140', 'X141', 'X142', 'X143', 'X144', 'X14
5', 'X146', 'X147', 'X148', 'X150', 'X151', 'X152', 'X153', 'X154', 'X15
5', 'X156', 'X157', 'X158', 'X159', 'X160', 'X161', 'X162', 'X163', 'X16
4', 'X165', 'X166', 'X167', 'X168', 'X169', 'X170', 'X171', 'X172', 'X17
3', 'X174', 'X175', 'X176', 'X177', 'X178', 'X179', 'X180', 'X181', 'X18
2', 'X183', 'X184', 'X185', 'X186', 'X187', 'X189', 'X190', 'X191', 'X19
2', 'X194', 'X195', 'X196', 'X197', 'X198', 'X199', 'X200', 'X201', 'X20
2', 'X203', 'X204', 'X205', 'X206', 'X207', 'X208', 'X209', 'X210', 'X21
1', 'X212', 'X213', 'X214', 'X215', 'X216', 'X217', 'X218', 'X219', 'X22
0', 'X221', 'X222', 'X223', 'X224', 'X225', 'X226', 'X227', 'X228', 'X22
9', 'X230', 'X231', 'X232', 'X234', 'X236', 'X237', 'X238', 'X239', 'X24
0', 'X241', 'X242', 'X243', 'X244', 'X245', 'X246', 'X247', 'X248', 'X24
9', 'X250', 'X251', 'X252', 'X253', 'X254', 'X255', 'X256', 'X257', 'X25
8', 'X259', 'X260', 'X261', 'X262', 'X263', 'X264', 'X265', 'X266', 'X26
7', 'X269', 'X270', 'X271', 'X272', 'X273', 'X274', 'X275', 'X276', 'X27
7', 'X278', 'X279', 'X280', 'X281', 'X282', 'X283', 'X284', 'X285', 'X28
6', 'X287', 'X288', 'X291', 'X292', 'X294', 'X295', 'X296', 'X298', 'X29
9', 'X300', 'X301', 'X302', 'X304', 'X305', 'X306', 'X307', 'X308', 'X30
9', 'X310', 'X311', 'X312', 'X313', 'X314', 'X315', 'X316', 'X317', 'X31
8', 'X319', 'X320', 'X321', 'X322', 'X323', 'X324', 'X325', 'X326', 'X32
7', 'X328', 'X329', 'X331', 'X332', 'X333', 'X334', 'X335', 'X336', 'X33
7', 'X338', 'X339', 'X340', 'X341', 'X342', 'X343', 'X344', 'X345', 'X34
6', 'X348', 'X349', 'X350', 'X351', 'X352', 'X353', 'X354', 'X355', 'X35
6', 'X357', 'X358', 'X359', 'X360', 'X361', 'X362', 'X363', 'X364', 'X36
5', 'X366', 'X367', 'X368', 'X369', 'X370', 'X371', 'X372', 'X373', 'X37
4', 'X375', 'X376', 'X377', 'X378', 'X379', 'X380', 'X382', 'X383', 'X38
4', 'X385']
--------------------------------------------------
Columns containing the unique values :  [0]
Number of Columns :  12
['X11', 'X93', 'X107', 'X233', 'X235', 'X268', 'X289', 'X290', 'X293', 'X2
97', 'X330', 'X347']
--------------------------------------------------
```

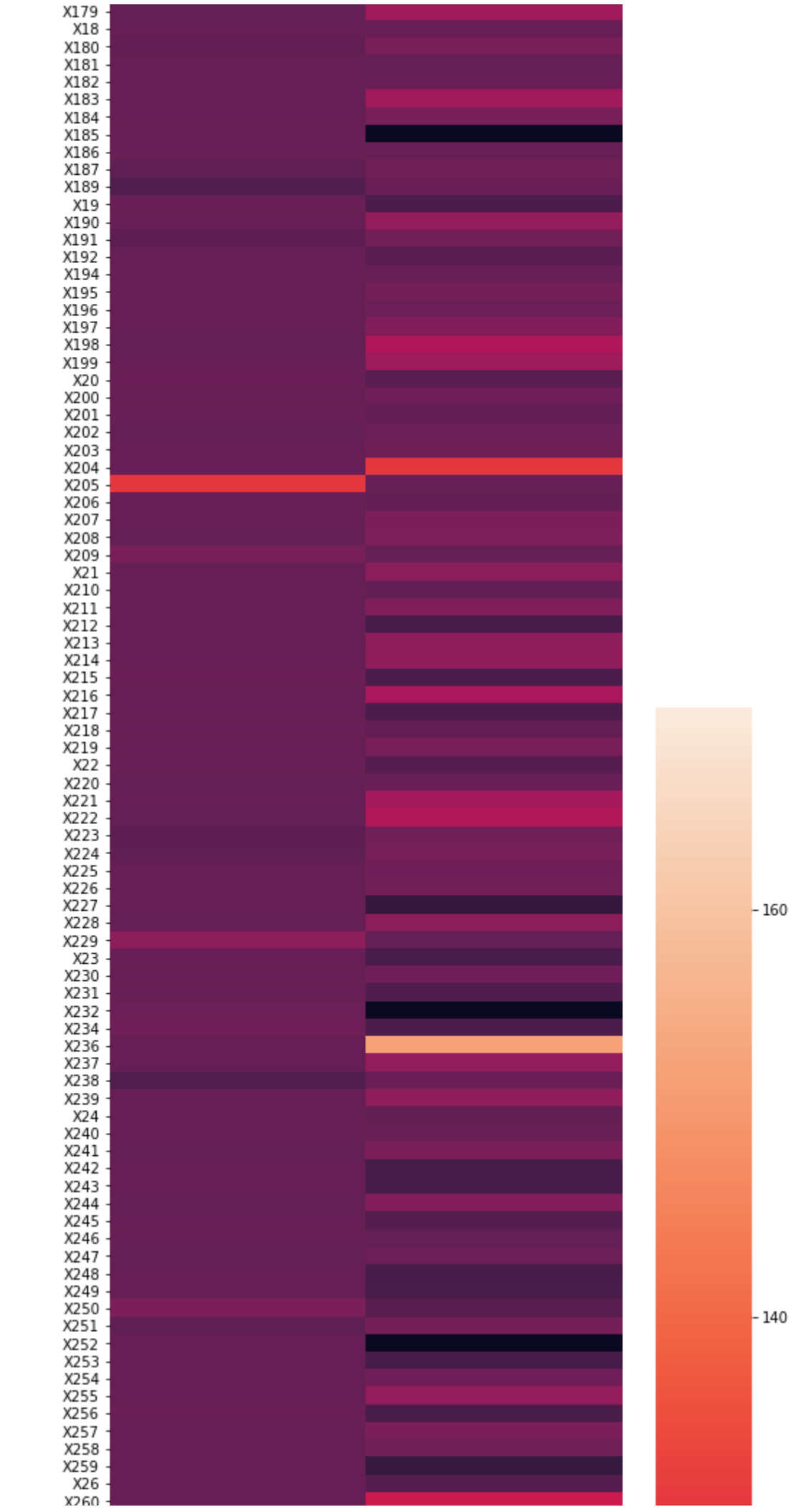
Among the integer columns , 356 are binary feature containing 0 & 1 And 12 feature are constant , containing only 0's We can omit out features containg only one unique value

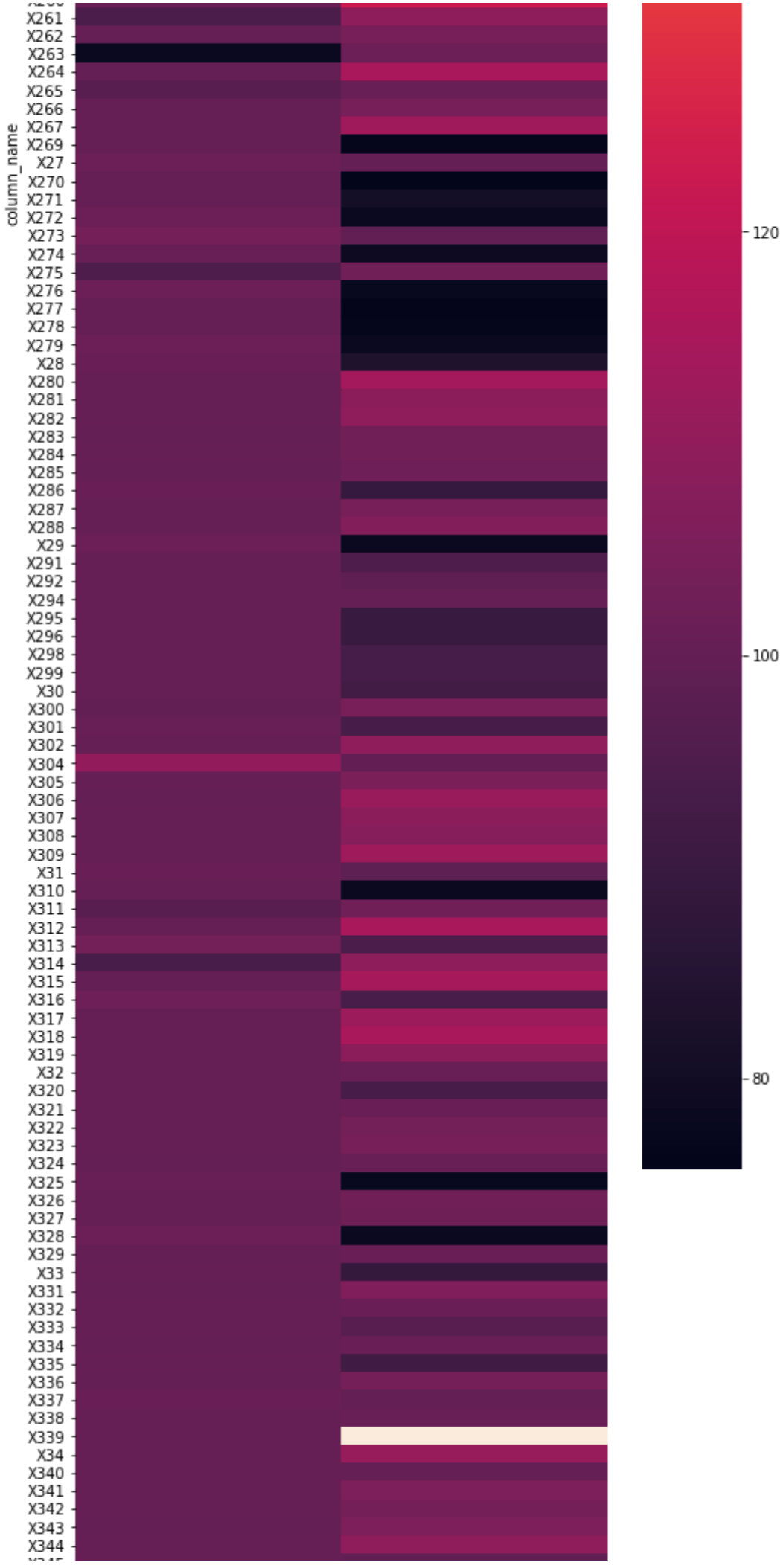
Lets find out the mean value of y for all the binary features , seperatly for 0 & 1

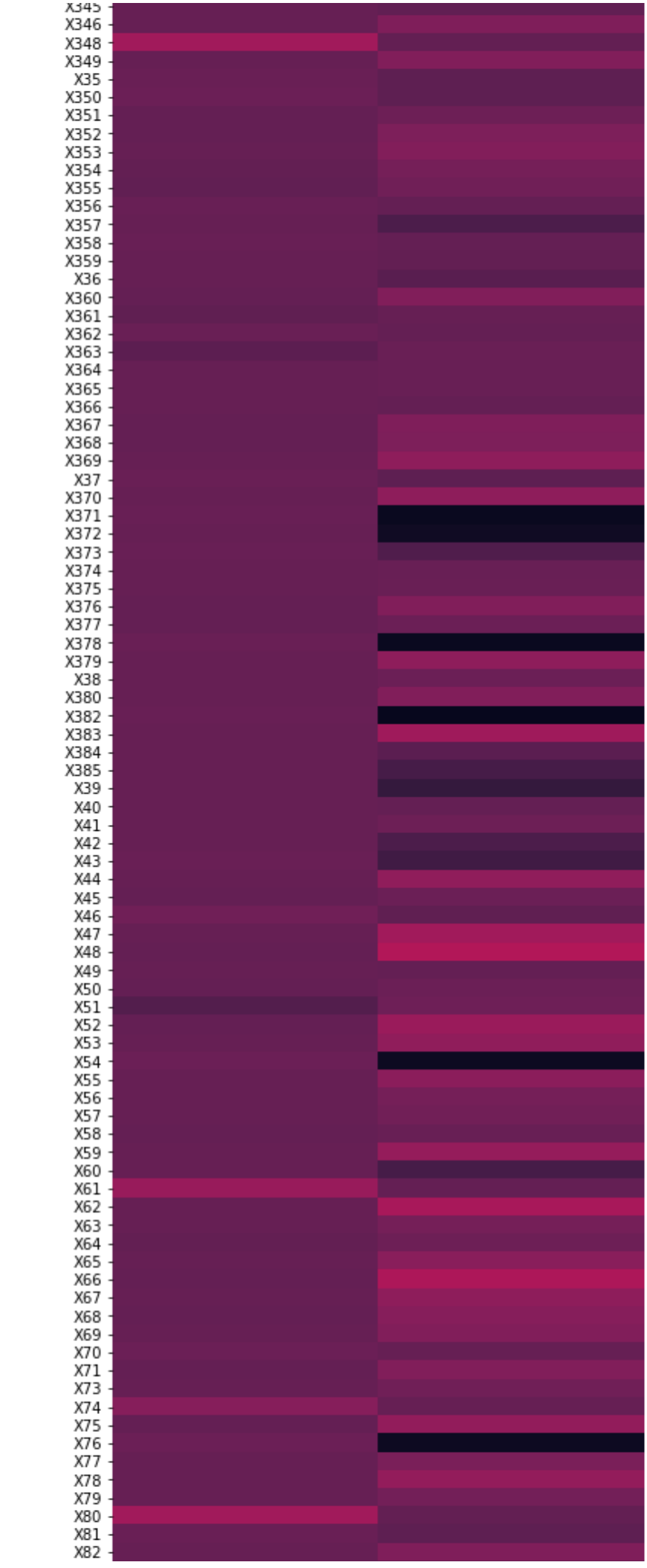
In [7]:

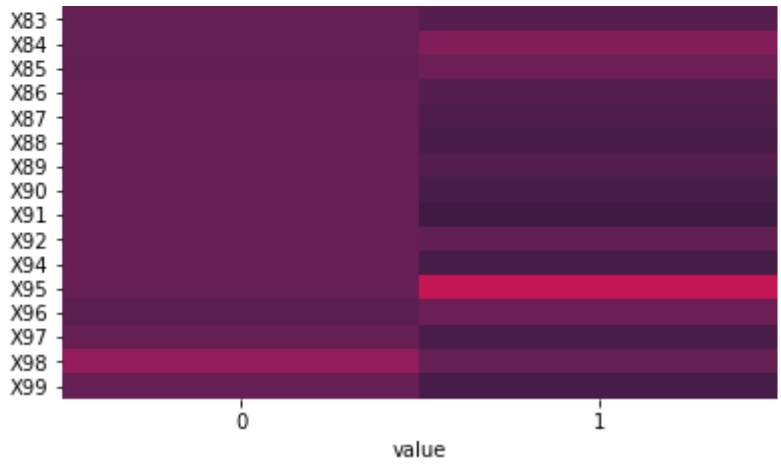
```
zero_mean_list = []
one_mean_list = []
cols_list = unique_values_dict['[0, 1]']
for col in cols_list:
    zero_mean_list.append(Benz_df.loc[Benz_df[col]==0].y.mean())
    one_mean_list.append(Benz_df.loc[Benz_df[col]==1].y.mean())

new_df = pd.DataFrame({"column_name":cols_list+cols_list, "value":[0]*len(cols_list) +
[1]*len(cols_list), "y_mean":zero_mean_list+one_mean_list})
new_df = new_df.pivot('column_name', 'value', 'y_mean')

plt.figure(figsize=(8,80))
sns.heatmap(new_df)
plt.title("Mean of y value across binary variables", fontsize=15)
plt.show()
```


Mean of y value across binary variables
X10
X100
X101
X102
X103
X104
X105
X106
X108
X109
X110
X111
X112
X113
X114
X115
X116
X117
X118
X119
X12
X120
X122
X123
X124
X125
X126
X127
X128
X129
X13
X130
X131
X132
X133
X134
X135
X136
X137
X138
X139
X14
X140
X141
X142
X143
X144
X145
X146
X147
X148
X15
X150
X151
X152
X153
X154
X155
X156
X157
X158
X159
X16
X160
X161
X162
X163
X164
X165
X166
X167
X168
X169
X17
X170
X171
X172
X173
X174
X175
X176
X177
X178


X179
X18
X180
X181
X182
X183
X184
X185
X186
X187
X189
X19
X190
X191
X192
X194
X195
X196
X197
X198
X199
X20
X200
X201
X202
X203
X204
X205
X206
X207
X208
X209
X21
X210
X211
X212
X213
X214
X215
X216
X217
X218
X219
X22
X220
X221
X222
X223
X224
X225
X226
X227
X228
X229
X23
X230
X231
X232
X234
X236
X237
X238
X239
X24
X240
X241
X242
X243
X244
X245
X246
X247
X248
X249
X250
X251
X252
X253
X254
X255
X256
X257
X258
X259
X26
X260
160
140


column_name
X261
X262
X263
X264
X265
X266
X267
X269
X27
X270
X271
X272
X273
X274
X275
X276
X277
X278
X279
X28
X280
X281
X282
X283
X284
X285
X286
X287
X288
X29
X291
X292
X294
X295
X296
X298
X299
X30
X300
X301
X302
X304
X305
X306
X307
X308
X309
X31
X310
X311
X312
X313
X314
X315
X316
X317
X318
X319
X32
X320
X321
X322
X323
X324
X325
X326
X327
X328
X329
X33
X331
X332
X333
X334
X335
X336
X337
X338
X339
X34
X340
X341
X342
X343
X344
120
100
80


X345
X346
X348
X349
X35
X350
X351
X352
X353
X354
X355
X356
X357
X358
X359
X36
X360
X361
X362
X363
X364
X365
X366
X367
X368
X369
X37
X370
X371
X372
X373
X374
X375
X376
X377
X378
X379
X38
X380
X382
X383
X384
X385
X39
X40
X41
X42
X43
X44
X45
X46
X47
X48
X49
X50
X51
X52
X53
X54
X55
X56
X57
X58
X59
X60
X61
X62
X63
X64
X65
X66
X67
X68
X69
X70
X71
X73
X74
X75
X76
X77
X78
X79
X80
X81
X82

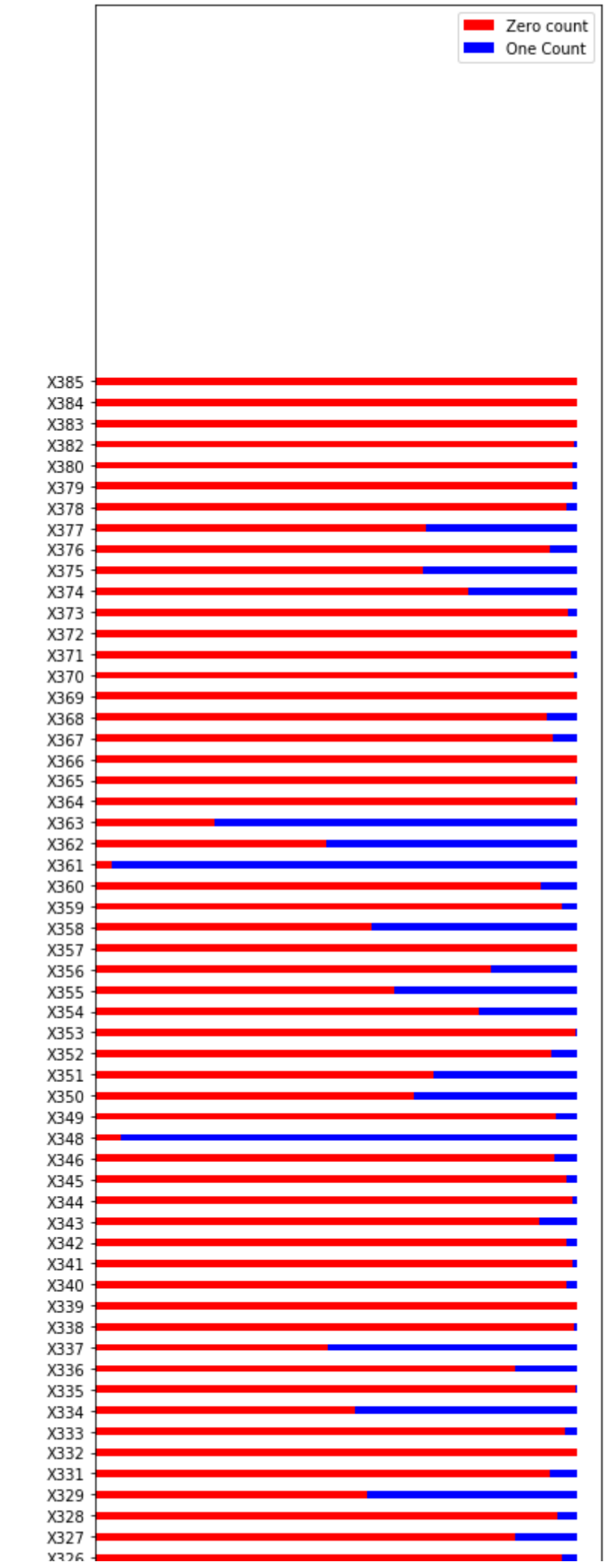


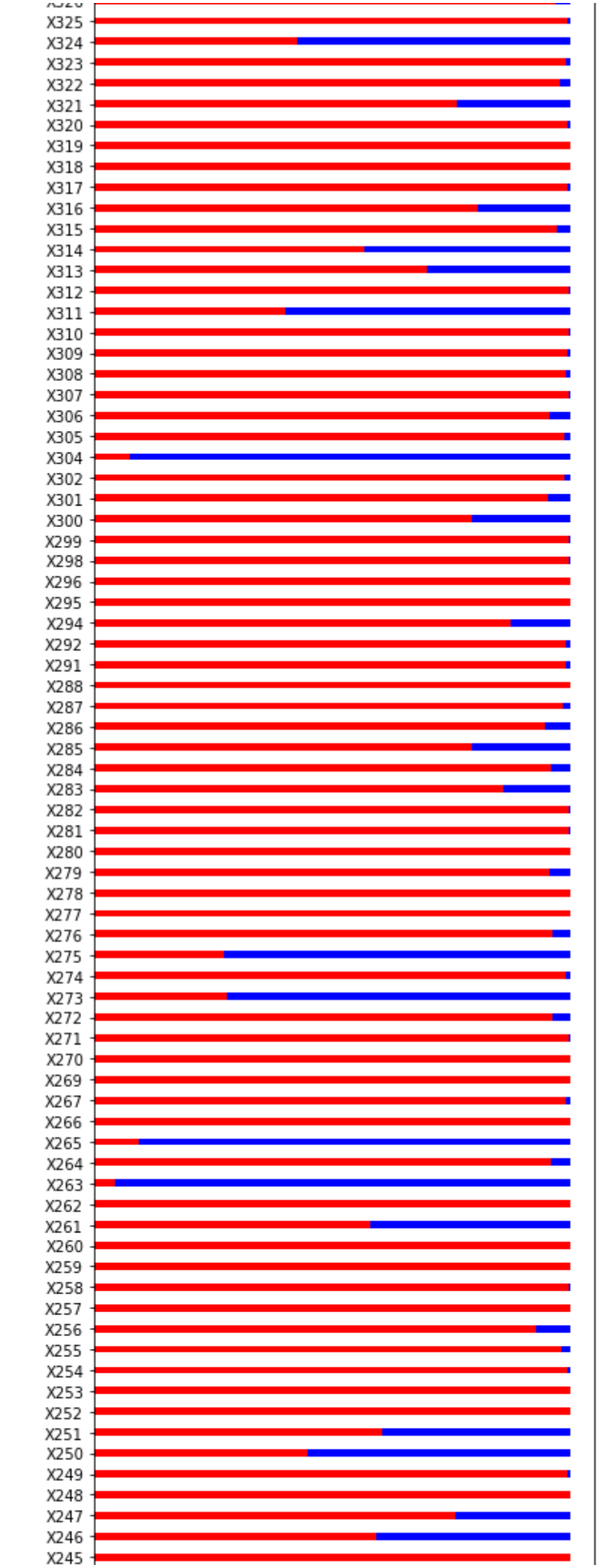
Features showing good colour difference like (X108 , X110 , X136 , X236 , X339) implies that there is a good difference between the mean value of 'y' for each binary value, Hence would be good for predicting our 'y' Variable

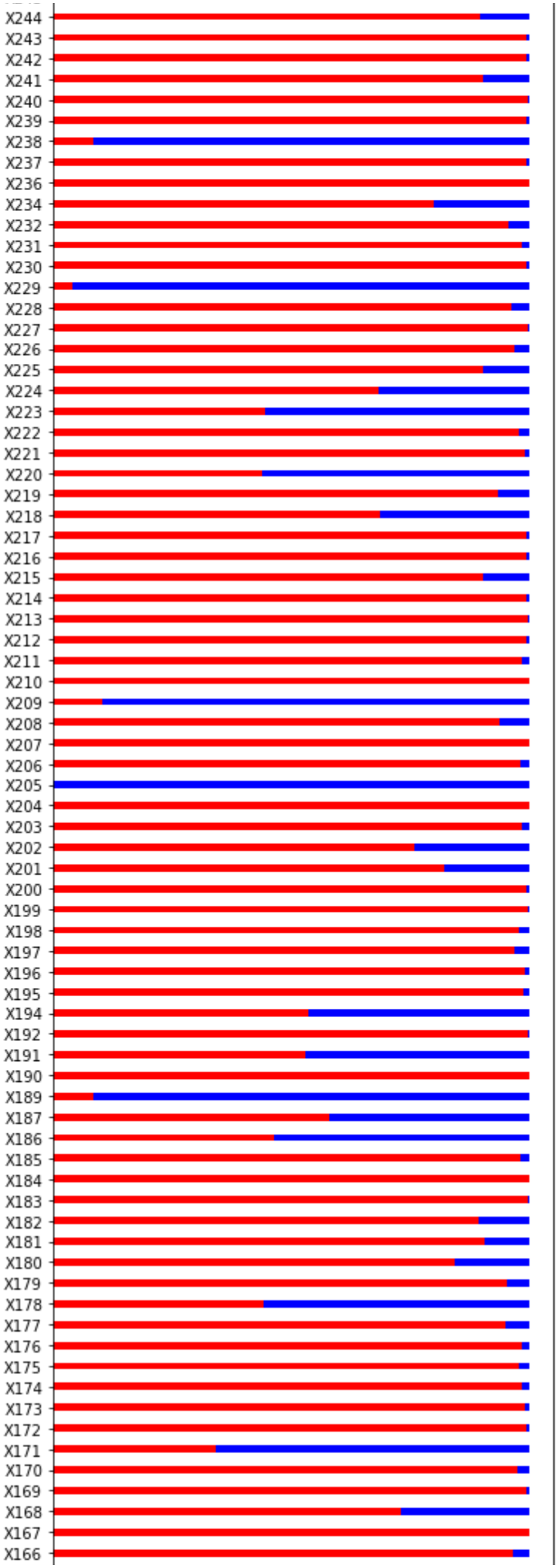
Lets look at the count of 0's and 1's , for each binary feature

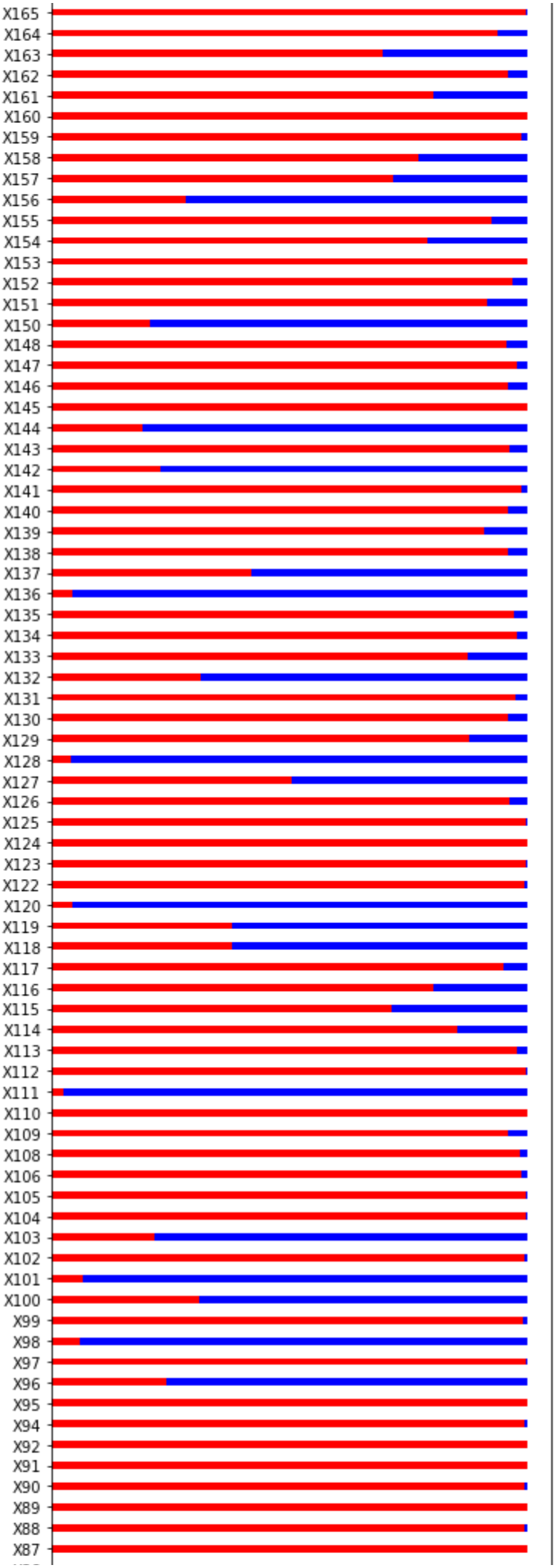
In [8]:

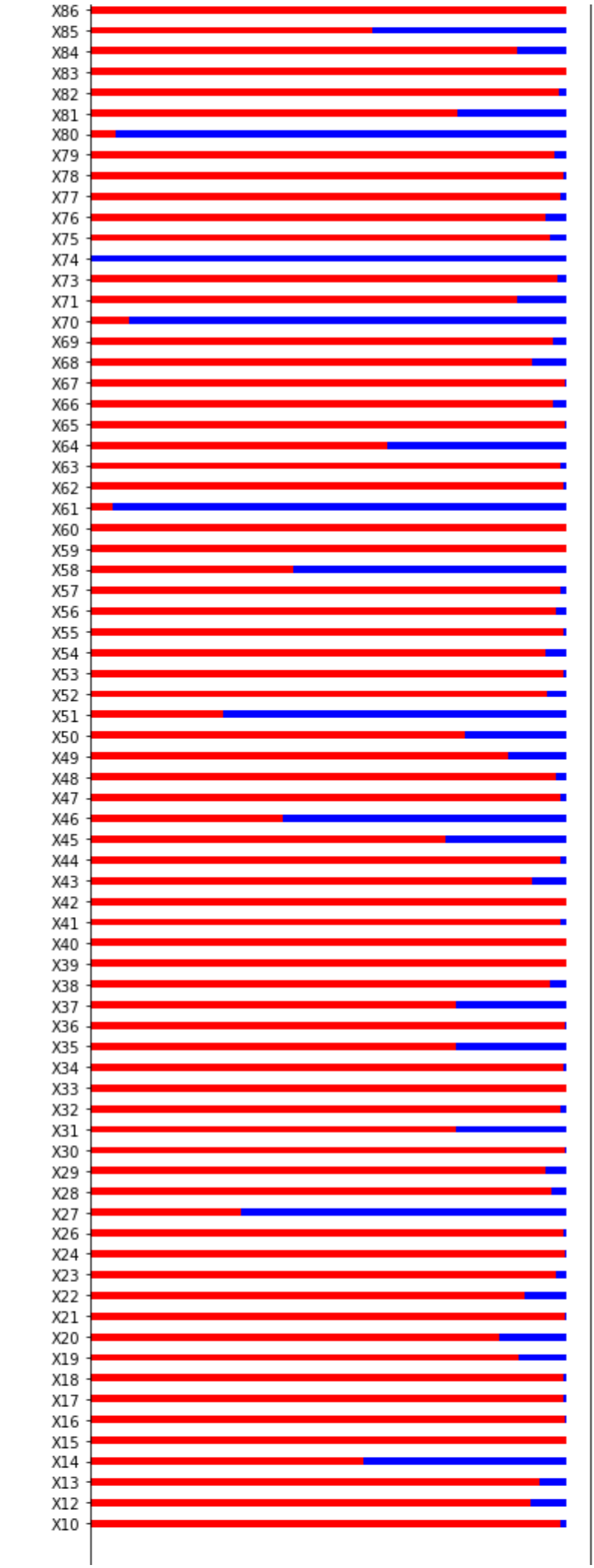
```
zero_count_list = []
one_count_list = []
cols_list = unique_values_dict['[0, 1]']
for col in cols_list:
    zero_count_list.append((Benz_df[col]==0).sum())
    one_count_list.append((Benz_df[col]==1).sum())

N = len(cols_list)
ind = np.arange(N)
width = 0.35

plt.figure(figsize=(6,100))
p1 = plt.barh(ind, zero_count_list, width, color='red')
p2 = plt.barh(ind, one_count_list, width, left=zero_count_list, color="blue")
plt.yticks(ind, cols_list)
plt.legend((p1[0], p2[0]), ('Zero count', 'One Count'))
plt.show()
```

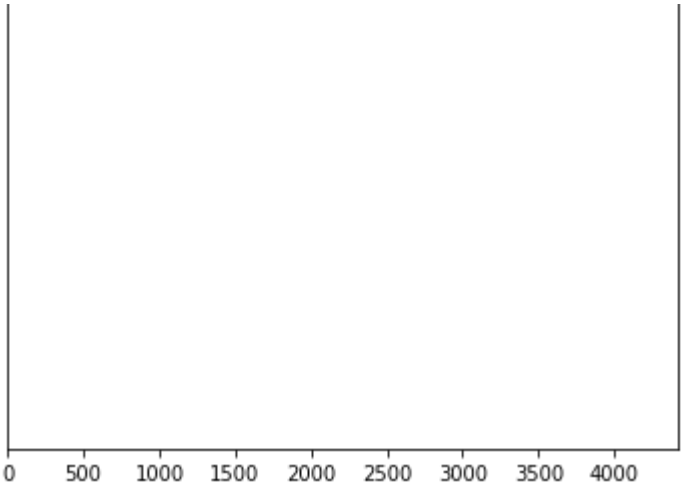

Zero count
One Count
X385
X384
X383
X382
X380
X379
X378
X377
X376
X375
X374
X373
X372
X371
X370
X369
X368
X367
X366
X365
X364
X363
X362
X361
X360
X359
X358
X357
X356
X355
X354
X353
X352
X351
X350
X349
X348
X346
X345
X344
X343
X342
X341
X340
X339
X338
X337
X336
X335
X334
X333
X332
X331
X329
X328
X327
X326


X325
X324
X323
X322
X321
X320
X319
X318
X317
X316
X315
X314
X313
X312
X311
X310
X309
X308
X307
X306
X305
X304
X302
X301
X300
X299
X298
X296
X295
X294
X292
X291
X288
X287
X286
X285
X284
X283
X282
X281
X280
X279
X278
X277
X276
X275
X274
X273
X272
X271
X270
X269
X267
X266
X265
X264
X263
X262
X261
X260
X259
X258
X257
X256
X255
X254
X253
X252
X251
X250
X249
X248
X247
X246
X245


X244
X243
X242
X241
X240
X239
X238
X237
X236
X234
X232
X231
X230
X229
X228
X227
X226
X225
X224
X223
X222
X221
X220
X219
X218
X217
X216
X215
X214
X213
X212
X211
X210
X209
X208
X207
X206
X205
X204
X203
X202
X201
X200
X199
X198
X197
X196
X195
X194
X192
X191
X190
X189
X187
X186
X185
X184
X183
X182
X181
X180
X179
X178
X177
X176
X175
X174
X173
X172
X171
X170
X169
X168
X167
X166


X165
X164
X163
X162
X161
X160
X159
X158
X157
X156
X155
X154
X153
X152
X151
X150
X148
X147
X146
X145
X144
X143
X142
X141
X140
X139
X138
X137
X136
X135
X134
X133
X132
X131
X130
X129
X128
X127
X126
X125
X124
X123
X122
X120
X119
X118
X117
X116
X115
X114
X113
X112
X111
X110
X109
X108
X106
X105
X104
X103
X102
X101
X100
X99
X98
X97
X96
X95
X94
X92
X91
X90
X89
X88
X87


X86
X85
X84
X83
X82
X81
X80
X79
X78
X77
X76
X75
X74
X73
X71
X70
X69
X68
X67
X66
X65
X64
X63
X62
X61
X60
X59
X58
X57
X56
X55
X54
X53
X52
X51
X50
X49
X48
X47
X46
X45
X44
X43
X42
X41
X40
X39
X38
X37
X36
X35
X34
X33
X32
X31
X30
X29
X28
X27
X26
X24
X23
X22
X21
X20
X19
X18
X17
X16
X15
X14
X13
X12
X10



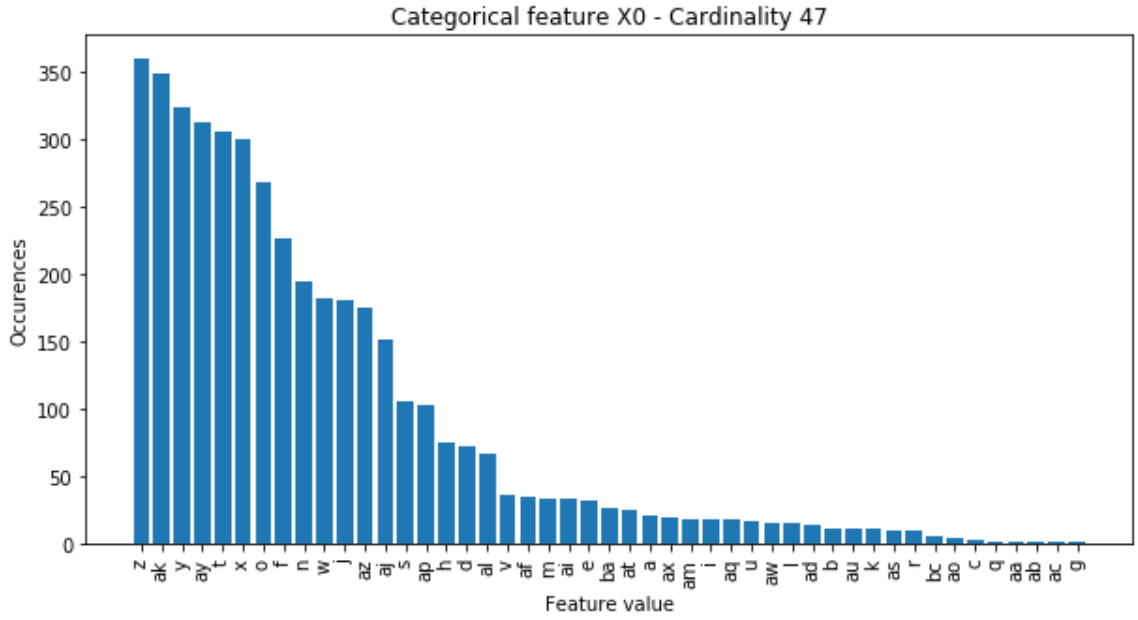
As we can see , there are some feature (like X10 , X18 , X61 , X74) where either of them dominates

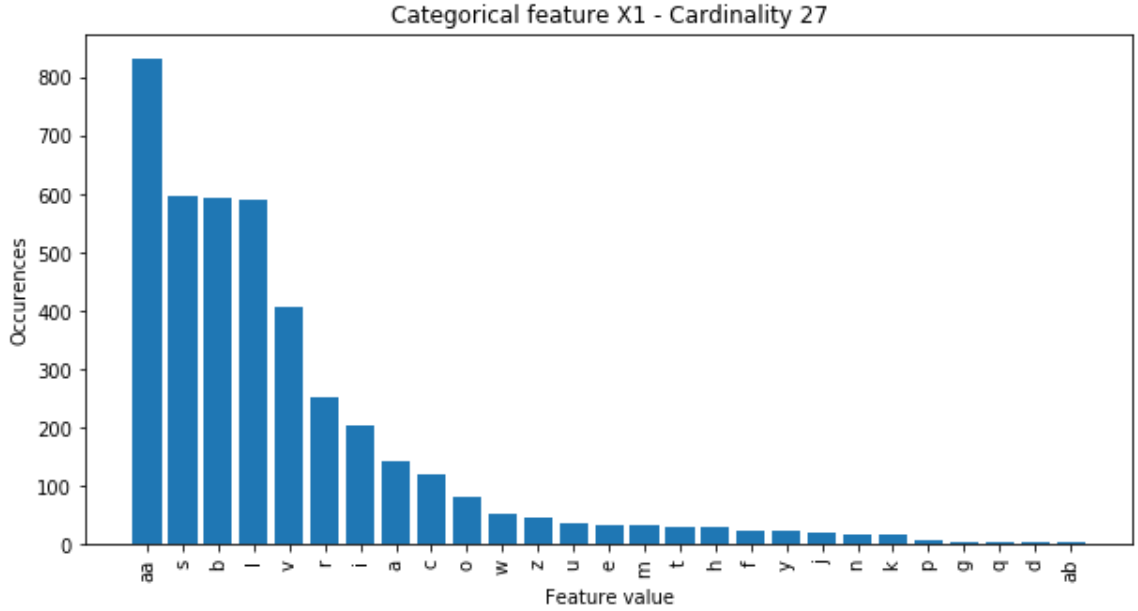
Categorical Features

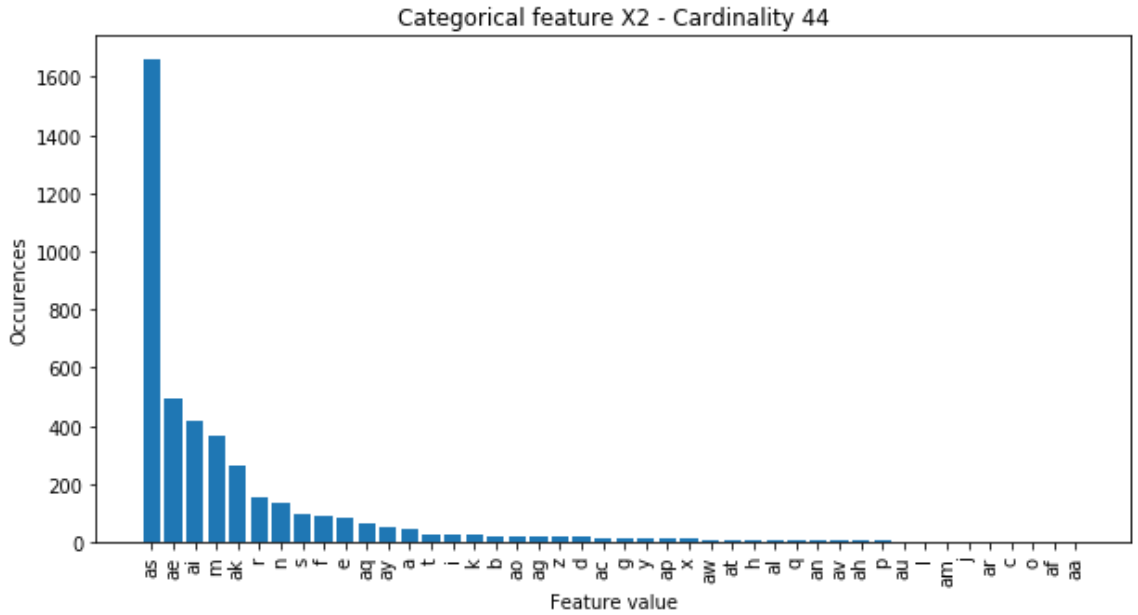
Lets look at the cardinality and occurences of all categories in all columns

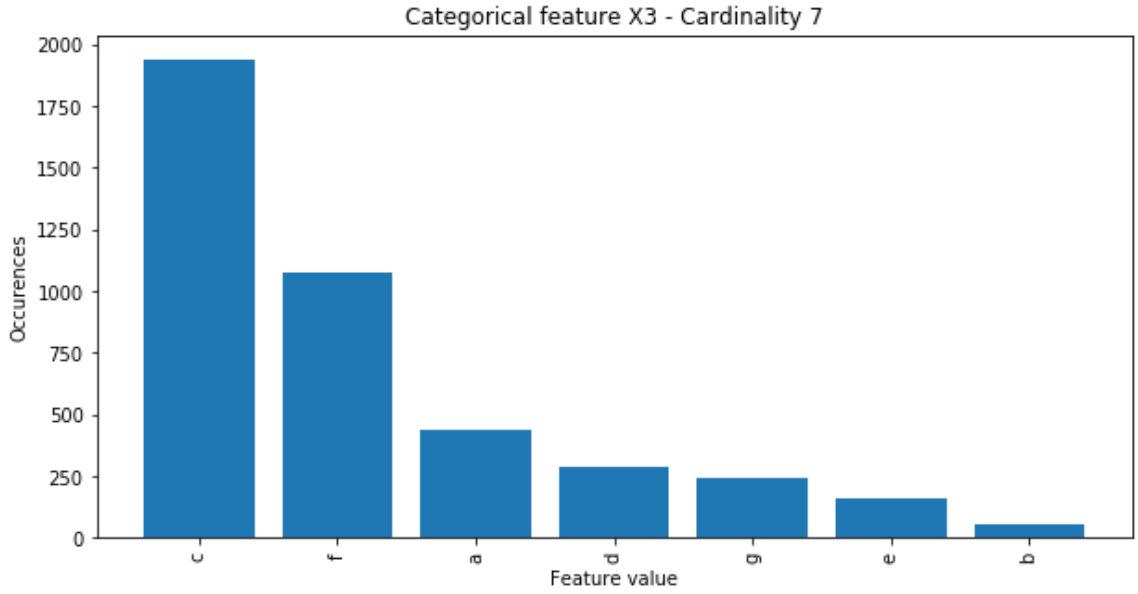
In [11]:

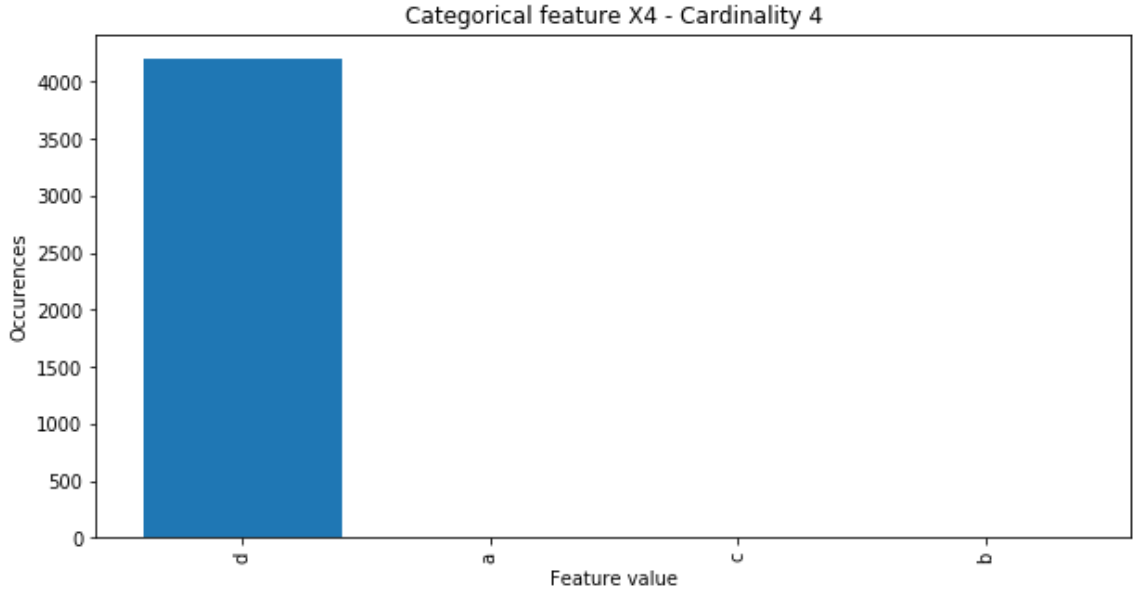
```
Categ_Features = ['X0', 'X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X8']
for c in Categ_Features:
    value_counts = Benz_df[c].value_counts()
    fig, ax = plt.subplots(figsize=(10, 5))
    plt.title('Categorical feature {} - Cardinality {}'.format(c, len(np.unique(Benz_df
[c]))))
    plt.xlabel('Feature value')
    plt.ylabel('Occurences')
    plt.bar(range(len(value_counts)), value_counts.values)
    ax.set_xticks(range(len(value_counts)))
    ax.set_xticklabels(value_counts.index, rotation='vertical')
    plt.show()
```

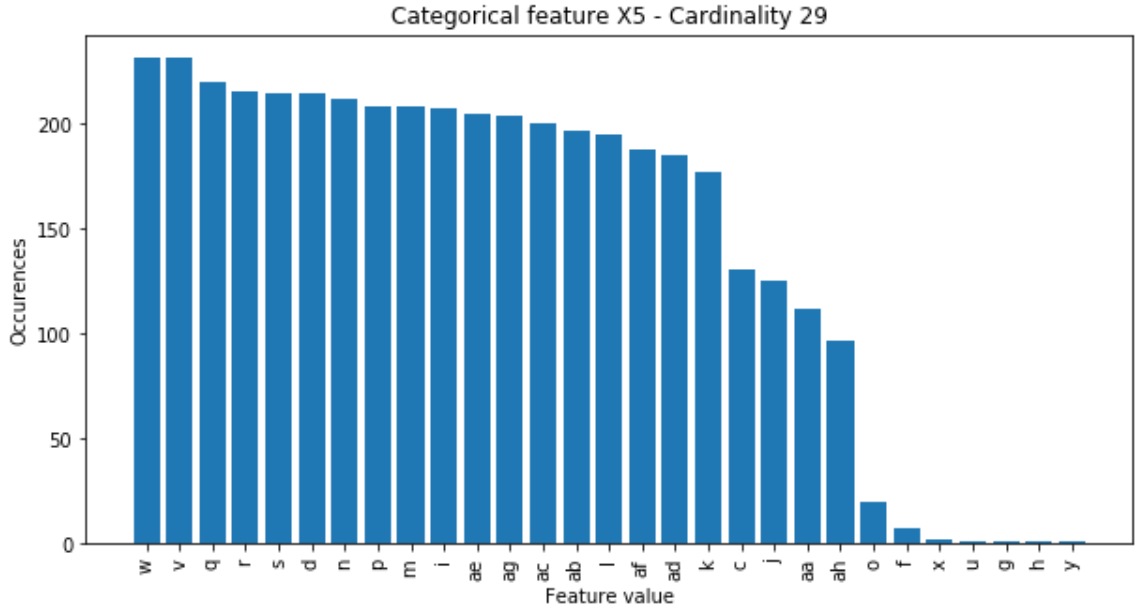

Categorical feature X0 - Cardinality 47
Occurences
Feature value
z ak y ay t x o f n w j az aj s ap h d al v af m ai e ba at a ax am i aq u aw l ad b au k as r bc ao c q aa ab ac g

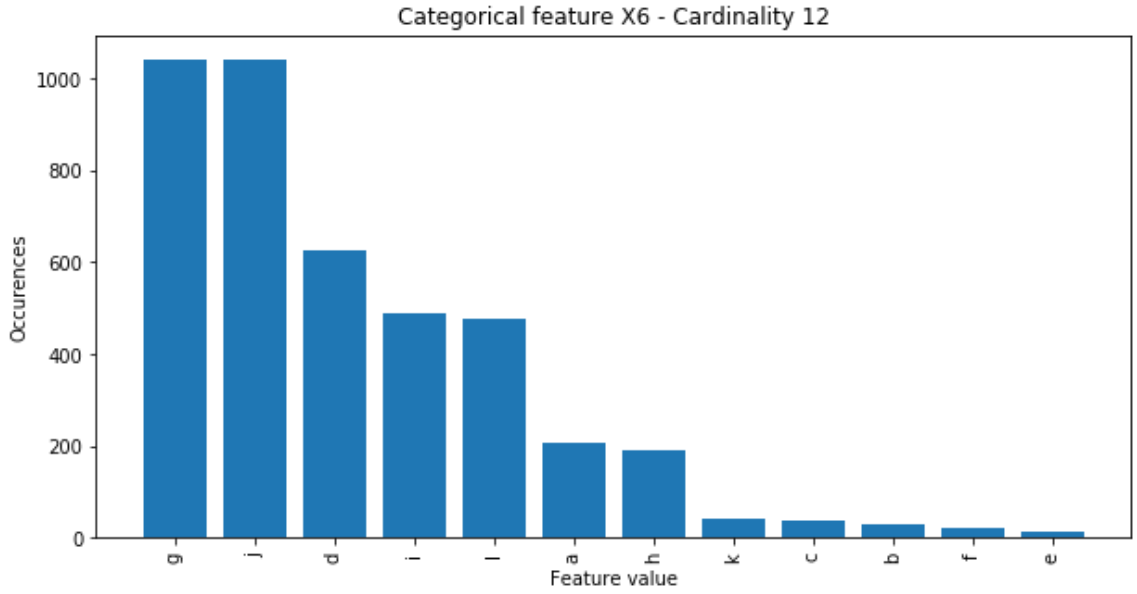

Categorical feature X1 - Cardinality 27
Occurences
Feature value
aa s b l v r i a c o w z u e m t h f y j n k p g q d ab

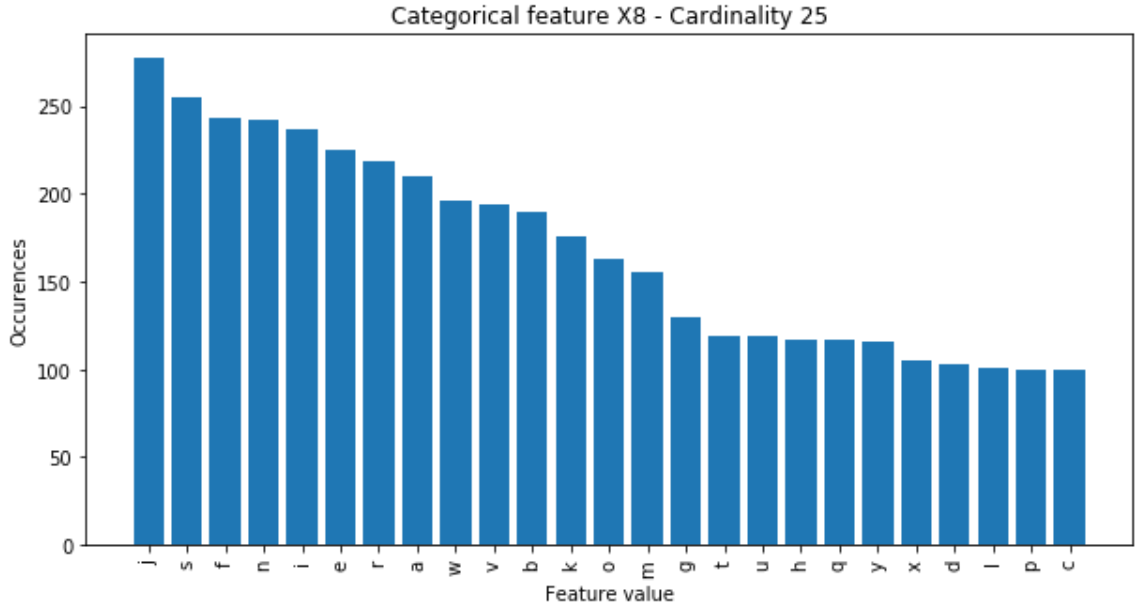

Categorical feature X2 - Cardinality 44
Occurences
Feature value
as ae ai m ak r n s f e aq ay a t i k b ao ag z d ac g y ap x aw at h al q an av ah p au l am j ar c o af aa


Categorical feature X3 - Cardinality 7
Occurences
Feature value
c
f
a
d
g
e
b


Categorical feature X4 - Cardinality 4
Occurences
Feature value
d
a
c
b


Categorical feature X5 - Cardinality 29
Occurences
Feature value
w v q r s d n p m i ae ag ac ab l af ad k c j aa ah o f x u g h y





As we can see , there are some categories which dominates others in some columns.

Modelling

In [81]:

```
Benz_df  = pd.read_csv('train.csv')
Benz_df.drop(Benz_df[Benz_df['y'] > 175].index, axis=0, inplace=True) ## Removing outli
er
```

In [82]:

```
## Test train split
y = Benz_df['y']
X = Benz_df.drop(['y'],axis=1)

# train test split
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

SET 1 , (PCA , ICA with Label Encoder)

Vectorizing categorical columns using Label Encoder

In [83]:

```
from sklearn.preprocessing import LabelEncoder
for c in X_train.columns:
    if X_train[c].dtype == 'object':
        lbl = LabelEncoder()
        lbl.fit(list(X_train[c].values) + list(X_test[c].values))
        X_train[c] = lbl.transform(list(X_train[c].values))
        X_test[c] = lbl.transform(list(X_test[c].values))
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:6: Setti
ngWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy

C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:7: Setti
ngWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
  import sys
```

Adding decomposed features : PCA And ICA features

In [84]:

```
from sklearn.decomposition import PCA, FastICA , TruncatedSVD
n_comp = 12

# PCA
pca = PCA(n_components=n_comp, random_state=42)
pca2_results_train = pca.fit_transform(X_train)
pca2_results_test = pca.transform(X_test)

# ICA
ica = FastICA(n_components=n_comp, random_state=42)
ica2_results_train = ica.fit_transform(X_train)
ica2_results_test = ica.transform(X_test)

# Append decomposition components to datasets
for i in range(1, n_comp+1):
    X_train['pca_' + str(i)] = pca2_results_train[:,i-1]
    X_test['pca_' + str(i)] = pca2_results_test[:, i-1]

    X_train['ica_' + str(i)] = ica2_results_train[:,i-1]
    X_test['ica_' + str(i)] = ica2_results_test[:, i-1]


y_mean = np.mean(y_train)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
  app.launch_new_instance()
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:17: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:19: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:20: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
```

In [85]:

```
print('Train Shape' ,X_train.shape)
print('Test Shape' ,X_test.shape)
```

```
Train Shape (2814, 401)
Test Shape (1387, 401)
```

Elastic Net Regression

In [86]:

```
# Let's perform a cross-validation to find the best combination of alpha and l1_ratio
from sklearn.linear_model import ElasticNetCV, ElasticNet

cv_model = ElasticNetCV(l1_ratio=[.1, .5, .7, .9, .95, .99, .995, 1], eps=0.001, n_alph
as=100, fit_intercept=True,
                        normalize=True, precompute='auto', max_iter=2000, tol=0.0001, c
v=5,
                        copy_X=True, verbose=0, n_jobs=-1, positive=False, random_state
=None, selection='cyclic')
```

In [87]:

```
cv_model.fit(X_train, y_train)
```

Out[87]:

```
ElasticNetCV(alphas=None, copy_X=True, cv=5, eps=0.001, fit_intercept=Tru
e,
       l1_ratio=[0.1, 0.5, 0.7, 0.9, 0.95, 0.99, 0.995, 1], max_iter=2000,
       n_alphas=100, n_jobs=-1, normalize=True, positive=False,
       precompute='auto', random_state=None, selection='cyclic',
       tol=0.0001, verbose=0)
```

In [88]:

```
print('Optimal alpha:' , cv_model.alpha_)
print('Optimal l1_ratio:' , cv_model.l1_ratio_)
print('Number of iterations' , cv_model.n_iter_)
```

```
Optimal alpha: 0.00508696975721008
Optimal l1_ratio: 1.0
Number of iterations 447
```

In [89]:

```
# Let's perform a cross-validation to find the best combination of alpha and l1_ratio
from sklearn.linear_model import ElasticNetCV, ElasticNet

model = ElasticNet(l1_ratio=cv_model.l1_ratio_, alpha = cv_model.alpha_, max_iter=cv_mo
del.n_iter_, fit_intercept=True, normalize = True)
model.fit(X_train, y_train)
```

Out[89]:

```
ElasticNet(alpha=0.00508696975721008, copy_X=True, fit_intercept=True,
      l1_ratio=1.0, max_iter=447, normalize=True, positive=False,
      precompute=False, random_state=None, selection='cyclic', tol=0.0001,
      warm_start=False)
```

In [90]:

```
print('Train Score')
print(r2_score(y_train, model.predict(X_train)))
print('Test Score')
print(r2_score(y_test, model.predict(X_test)))
```

```
Train Score
0.6035282812775089
Test Score
0.5996612482684452
```

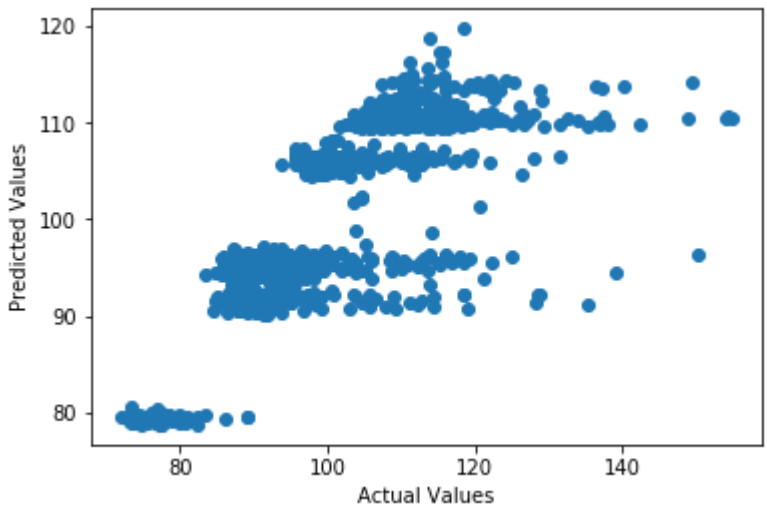
In [91]:

```
print('Actual Vs Predicted')
plt.scatter(y_test, model.predict(X_test))
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[91]:

```
Text(0, 0.5, 'Predicted Values')
```



XGBOOST

In [31]:

```
## Cross validation
import xgboost as xgb
from sklearn.metrics import r2_score
from sklearn.model_selection import GridSearchCV

XGB = xgb.XGBRegressor(objective='reg:squarederror')
XGB_para = {'n_estimators':[100,300,500],'max_depth':[2, 3, 4] , 'learning_rate' : [0.0
01,0.01,0.1,0.2]}
clf = GridSearchCV(XGB, XGB_para, cv=3, scoring='r2')
clf.fit(X_train, y_train)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
```

```
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
```

```
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
```

```
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
```

```
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
```

Out[31]:

```
GridSearchCV(cv=3, error_score='raise-deprecating',
       estimator=XGBRegressor(base_score=0.5, booster='gbtree', colsample_
bylevel=1,
       colsample_bynode=1, colsample_bytree=1, gamma=0,
       importance_type='gain', learning_rate=0.1, max_delta_step=0,
       max_depth=3, min_child_weight=1, missing=None, n_estimators=100,
       n_jobs=1, nthread=None, objective='reg:squarederror',
       random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1,
       seed=None, silent=None, subsample=1, verbosity=1),
       fit_params=None, iid='warn', n_jobs=None,
       param_grid={'n_estimators': [100, 300, 500], 'max_depth': [2, 3,
4], 'learning_rate': [0.001, 0.01, 0.1, 0.2]},
       pre_dispatch='2*n_jobs', refit=True, return_train_score='warn',
       scoring='r2', verbose=0)
```

In [32]:

```
clf.best_params_
```

Out[32]:

```
{'learning_rate': 0.1, 'max_depth': 2, 'n_estimators': 100}
```

In [33]:

```
xg_reg = xgb.XGBRegressor(objective ='reg:linear',learning_rate = 0.1,
                max_depth = 2, n_estimators = 100)
xg_reg.fit(X_train, y_train)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \

[16:56:49] WARNING: src/objective/regression_obj.cu:152: reg:linear is now
deprecated in favor of reg:squarederror.
```

Out[33]:

```
XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1,
       colsample_bynode=1, colsample_bytree=1, gamma=0,
       importance_type='gain', learning_rate=0.1, max_delta_step=0,
       max_depth=2, min_child_weight=1, missing=None, n_estimators=100,
       n_jobs=1, nthread=None, objective='reg:linear', random_state=0,
       reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None,
       silent=None, subsample=1, verbosity=1)
```

In [34]:

```
from sklearn.metrics import r2_score
preds = xg_reg.predict(X_test)


print('Train r2 :',r2_score(y_train, xg_reg.predict(X_train)))
print('Test r2 :', r2_score(y_test, preds))
```

```
Train r2 : 0.6478505129559475
Test r2 : 0.630212368976419
```

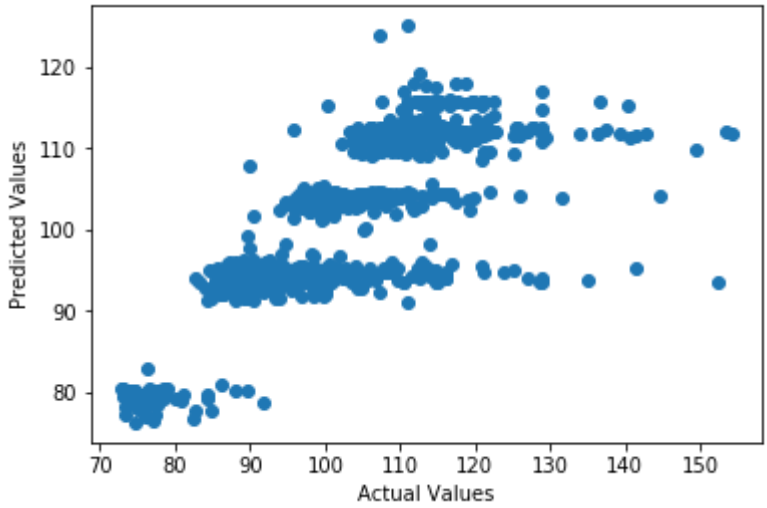
In [35]:

```
print('Actual Vs Predicted')
plt.scatter(y_test,preds)
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[35]:

```
Text(0, 0.5, 'Predicted Values')
```



Stacked Model

In [40]:

```
## # Reference: https://github.com/nilaysen/Mercedes-Benz-Greener-Manufacturing-Kaggle/
blob/master/merc.py
class StackingEstimator(BaseEstimator, TransformerMixin):

    def __init__(self, estimator):
        self.estimator = estimator

    def fit(self, X, y=None, **fit_params):
        self.estimator.fit(X, y, **fit_params)
        return self
    def transform(self, X):
        X = check_array(X)
        X_transformed = np.copy(X)
        # add class probabilities as a synthetic feature
        if issubclass(self.estimator.__class__, ClassifierMixin) and hasattr(self.estim
ator, 'predict_proba'):
            X_transformed = np.hstack((self.estimator.predict_proba(X), X))

        # add class prodiction as a synthetic feature
        X_transformed = np.hstack((np.reshape(self.estimator.predict(X), (-1, 1)), X_tr
ansformed))

        return X_transformed

stacked_pipeline = make_pipeline(
    StackingEstimator(estimator=LassoLarsCV(normalize=True)),
    StackingEstimator(estimator=GradientBoostingRegressor(learning_rate=0.001, loss="hu
ber", max_depth=3, max_features=0.55, min_samples_leaf=18, min_samples_split=14, subsam
ple=0.7)),
    LassoLarsCV()

)


stacked_pipeline.fit(X_train, y_train)
results = stacked_pipeline.predict(X_test)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.037e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.664e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.037e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 4.215e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.664e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.428e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 4.215e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
```

```
ng a regressor, after 14 iterations, i.e. alpha=1.677e-02, with an active
set of 14 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
20 iterations, alpha=1.391e-02, previous alpha=1.349e-02, with an active s
et of 17 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.391e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.391e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.773e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.773e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 7.743e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 6 iterations, i.e. alpha=3.629e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 6 iterations, i.e. alpha=3.629e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.810e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 7.743e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.810e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 4.712e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.799e-02, with an active
```

```
set of 15 regressors, and the smallest cholesky pivot element being 4.829e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.799e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 4.712e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 16 iterations, i.e. alpha=1.740e-02, with an active
set of 16 regressors, and the smallest cholesky pivot element being 9.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 16 iterations, i.e. alpha=1.740e-02, with an active
set of 16 regressors, and the smallest cholesky pivot element being 9.884e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 16 iterations, i.e. alpha=1.740e-02, with an active
set of 16 regressors, and the smallest cholesky pivot element being 9.657e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.667e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 7.146e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.667e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.667e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 6.664e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
21 iterations, alpha=1.771e-02, previous alpha=1.667e-02, with an active s
et of 18 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=6.839e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.555e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 2.581e-0
```

```
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.555e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.412e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
12 iterations, alpha=2.442e-02, previous alpha=2.422e-02, with an active s
et of 11 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 2.980e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 2.788e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 3.161e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.132e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 3.161e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.132e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 3.332e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=2.674e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 2.788e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=2.674e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 2.980e-0
8. Reduce max_iter or increase eps parameters.
```

```
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.332e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.161e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.650e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.799e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.490e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 2.980e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 42 iterations, i.e. alpha=4.479e-03, with an active
set of 40 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 66 iterations, i.e. alpha=3.013e-03, with an active
set of 60 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
```

```
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 4.215e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 6.322e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 4.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 93 iterations, i.e. alpha=2.303e-03, with an active
set of 83 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 95 iterations, i.e. alpha=2.283e-03, with an active
set of 85 regressors, and the smallest cholesky pivot element being 2.980e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 102 iterations, i.e. alpha=2.239e-03, with an active
set of 90 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 102 iterations, i.e. alpha=2.239e-03, with an active
set of 90 regressors, and the smallest cholesky pivot element being 3.650e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 102 iterations, i.e. alpha=2.239e-03, with an active
set of 90 regressors, and the smallest cholesky pivot element being 6.747e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
```

```
ng a regressor, after 104 iterations, i.e. alpha=2.182e-03, with an active
set of 92 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 112 iterations, i.e. alpha=1.955e-03, with an active
set of 98 regressors, and the smallest cholesky pivot element being 7.671e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 146 iterations, i.e. alpha=1.287e-03, with an active
set of 130 regressors, and the smallest cholesky pivot element being 6.747
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 146 iterations, i.e. alpha=1.287e-03, with an active
set of 130 regressors, and the smallest cholesky pivot element being 6.322
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 146 iterations, i.e. alpha=1.287e-03, with an active
set of 130 regressors, and the smallest cholesky pivot element being 6.409
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 146 iterations, i.e. alpha=1.287e-03, with an active
set of 130 regressors, and the smallest cholesky pivot element being 1.490
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 146 iterations, i.e. alpha=1.287e-03, with an active
set of 130 regressors, and the smallest cholesky pivot element being 2.220
e-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 156 iterations, i.e. alpha=1.152e-03, with an active
set of 132 regressors, and the smallest cholesky pivot element being 6.322
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 156 iterations, i.e. alpha=1.152e-03, with an active
set of 132 regressors, and the smallest cholesky pivot element being 4.942
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 156 iterations, i.e. alpha=1.152e-03, with an active
set of 132 regressors, and the smallest cholesky pivot element being 3.161
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 156 iterations, i.e. alpha=1.152e-03, with an active
```

```
set of 132 regressors, and the smallest cholesky pivot element being 7.224
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 156 iterations, i.e. alpha=1.152e-03, with an active
set of 132 regressors, and the smallest cholesky pivot element being 1.825
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
157 iterations, alpha=1.152e-03, previous alpha=1.152e-03, with an active
set of 132 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=1.363e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 2.107e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 20 iterations, i.e. alpha=7.303e-03, with an active
set of 20 regressors, and the smallest cholesky pivot element being 5.268e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 21 iterations, i.e. alpha=7.301e-03, with an active
set of 21 regressors, and the smallest cholesky pivot element being 5.268e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 21 iterations, i.e. alpha=7.301e-03, with an active
set of 21 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 21 iterations, i.e. alpha=7.301e-03, with an active
set of 21 regressors, and the smallest cholesky pivot element being 6.580e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 27 iterations, i.e. alpha=6.670e-03, with an active
set of 25 regressors, and the smallest cholesky pivot element being 5.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 45 iterations, i.e. alpha=4.530e-03, with an active
set of 41 regressors, and the smallest cholesky pivot element being 2.356e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 45 iterations, i.e. alpha=4.530e-03, with an active
set of 41 regressors, and the smallest cholesky pivot element being 4.215e
```

```
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 45 iterations, i.e. alpha=4.530e-03, with an active
set of 41 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 59 iterations, i.e. alpha=3.893e-03, with an active
set of 55 regressors, and the smallest cholesky pivot element being 9.996e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 64 iterations, i.e. alpha=3.637e-03, with an active
set of 60 regressors, and the smallest cholesky pivot element being 6.580e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 64 iterations, i.e. alpha=3.637e-03, with an active
set of 60 regressors, and the smallest cholesky pivot element being 5.268e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 64 iterations, i.e. alpha=3.637e-03, with an active
set of 60 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 65 iterations, i.e. alpha=3.556e-03, with an active
set of 61 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 75 iterations, i.e. alpha=3.095e-03, with an active
set of 71 regressors, and the smallest cholesky pivot element being 5.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 87 iterations, i.e. alpha=2.790e-03, with an active
set of 79 regressors, and the smallest cholesky pivot element being 4.829e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 87 iterations, i.e. alpha=2.790e-03, with an active
set of 79 regressors, and the smallest cholesky pivot element being 7.451e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 88 iterations, i.e. alpha=2.747e-03, with an active
set of 80 regressors, and the smallest cholesky pivot element being 8.560e
-08. Reduce max_iter or increase eps parameters.
```

```
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 101 iterations, i.e. alpha=2.310e-03, with an active
set of 93 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 104 iterations, i.e. alpha=2.264e-03, with an active
set of 94 regressors, and the smallest cholesky pivot element being 4.081e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 104 iterations, i.e. alpha=2.264e-03, with an active
set of 94 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 104 iterations, i.e. alpha=2.255e-03, with an active
set of 94 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
105 iterations, alpha=2.263e-03, previous alpha=2.250e-03, with an active
set of 94 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 8 iterations, i.e. alpha=9.762e-03, with an active s
et of 8 regressors, and the smallest cholesky pivot element being 3.332e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 34 iterations, i.e. alpha=5.842e-03, with an active
set of 34 regressors, and the smallest cholesky pivot element being 8.560e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 35 iterations, i.e. alpha=5.759e-03, with an active
set of 35 regressors, and the smallest cholesky pivot element being 8.560e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 38 iterations, i.e. alpha=5.644e-03, with an active
set of 36 regressors, and the smallest cholesky pivot element being 6.053e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 48 iterations, i.e. alpha=5.185e-03, with an active
set of 46 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 48 iterations, i.e. alpha=5.185e-03, with an active
set of 46 regressors, and the smallest cholesky pivot element being 5.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 49 iterations, i.e. alpha=5.143e-03, with an active
set of 47 regressors, and the smallest cholesky pivot element being 3.161e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 53 iterations, i.e. alpha=4.881e-03, with an active
set of 49 regressors, and the smallest cholesky pivot element being 3.332e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 70 iterations, i.e. alpha=3.995e-03, with an active
set of 64 regressors, and the smallest cholesky pivot element being 8.689e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 77 iterations, i.e. alpha=3.718e-03, with an active
set of 69 regressors, and the smallest cholesky pivot element being 6.409e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 77 iterations, i.e. alpha=3.718e-03, with an active
set of 69 regressors, and the smallest cholesky pivot element being 6.234e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
78 iterations, alpha=3.712e-03, previous alpha=3.696e-03, with an active s
et of 69 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=1.021e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 2.980e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
```

In [42]:

```
print('R2 score on train data:')
print(r2_score(y_train,stacked_pipeline.predict(X_train)))

print('R2 score on test data:')
print(r2_score(y_test,stacked_pipeline.predict(X_test)))
```

```
R2 score on train data:
0.6189375400738832
R2 score on test data:
0.6354401423931675
```

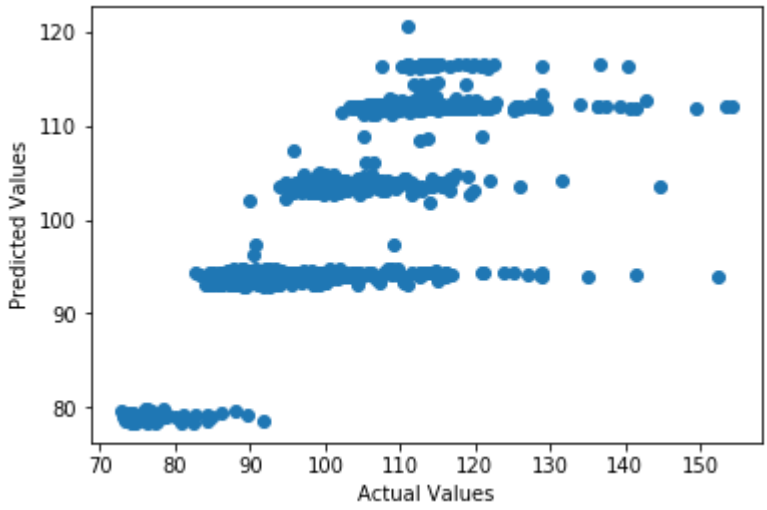
In [43]:

```
print('Actual Vs Predicted')
plt.scatter(y_test,stacked_pipeline.predict(X_test))
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[43]:

```
Text(0, 0.5, 'Predicted Values')
```



Stacked And Averaged Model

In [44]:

```
print('R2 score on train data:') ## Averaged with the XGBoost Model
print(r2_score(y_train,stacked_pipeline.predict(X_train)*0.2855 + xg_reg.predict(X_trai
n)*0.7145))

print('R2 score on test data:')
print(r2_score(y_test,stacked_pipeline.predict(X_test)*0.2855 + xg_reg.predict(X_test)*
0.7145))
```

```
R2 score on train data:
0.6412852821531496
R2 score on test data:
0.6332870652752253
```

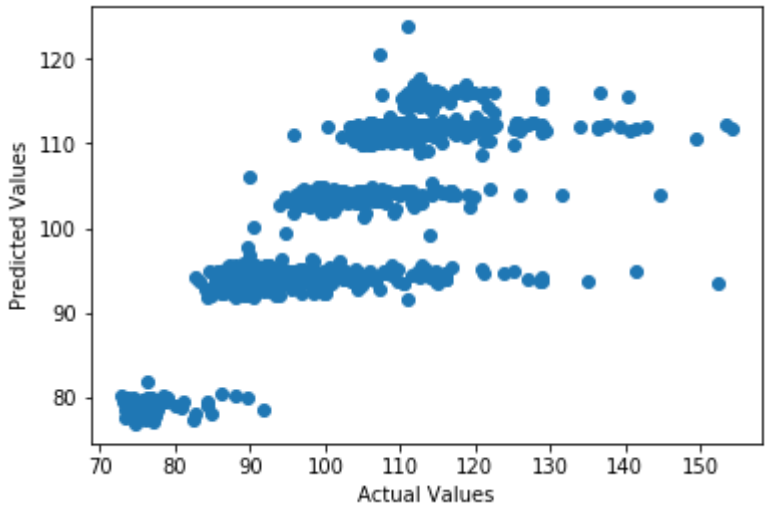
In [46]:

```
print('Actual Vs Predicted')
plt.scatter(y_test,stacked_pipeline.predict(X_test)*0.2855 + xg_reg.predict(X_test)*0.7
145)
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[46]:

```
Text(0, 0.5, 'Predicted Values')
```



Set 2 : SET 1 , (PCA , ICA , TSVD , grp , srp with Label Encoder)

Feature Engineering

In [47]:

```
from sklearn.preprocessing import LabelEncoder
for c in X_train.columns:
    if X_train[c].dtype == 'object':
        lbl = LabelEncoder()
        lbl.fit(list(X_train[c].values) + list(X_test[c].values))
        X_train[c] = lbl.transform(list(X_train[c].values))
        X_test[c] = lbl.transform(list(X_test[c].values))

from sklearn.decomposition import PCA, FastICA , TruncatedSVD
from sklearn.random_projection import GaussianRandomProjection
from sklearn.random_projection import SparseRandomProjection
n_comp = 12

# PCA
pca = PCA(n_components=n_comp, random_state=42)
pca2_results_train = pca.fit_transform(X_train)
pca2_results_test = pca.transform(X_test)

# ICA
ica = FastICA(n_components=n_comp, random_state=42)
ica2_results_train = ica.fit_transform(X_train)
ica2_results_test = ica.transform(X_test)

# TSVD
tsvd = TruncatedSVD(n_components=n_comp, random_state=420)
tsvd_results_train = tsvd.fit_transform(X_train)
tsvd_results_test = tsvd.transform(X_test)

# GRP
grp = GaussianRandomProjection(n_components=n_comp, eps=0.1, random_state=420)
grp_results_train = grp.fit_transform(X_train)
grp_results_test = grp.transform(X_test)

# SRP
srp = SparseRandomProjection(n_components=n_comp, dense_output=True, random_state=420)
srp_results_train = srp.fit_transform(X_train)
srp_results_test = srp.transform(X_test)

# Append decomposition components to datasets
for i in range(1, n_comp+1):
    X_train['pca_' + str(i)] = pca2_results_train[:,i-1]
    X_test['pca_' + str(i)] = pca2_results_test[:, i-1]

    X_train['ica_' + str(i)] = ica2_results_train[:,i-1]
    X_test['ica_' + str(i)] = ica2_results_test[:, i-1]

    X_train['tsvd_' + str(i)] = tsvd_results_train[:, i - 1]
    X_test['tsvd_' + str(i)] = tsvd_results_test[:, i - 1]

    X_train['grp_' + str(i)] = grp_results_train[:, i - 1]
    X_test['grp_' + str(i)] = grp_results_test[:, i - 1]

    X_train['srp_' + str(i)] = srp_results_train[:, i - 1]
    X_test['srp_' + str(i)] = srp_results_test[:, i - 1]


y_mean = np.mean(y_train)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:41: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:42: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:44: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:45: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:47: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:48: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:50: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:51: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:53: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:54: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
```

In [48]:

```
print('Train Shape' ,X_train.shape)
print('Test Shape' ,X_test.shape)
```

```
Train Shape (2814, 437)
Test Shape (1387, 437)
```

Elastic Net Regression

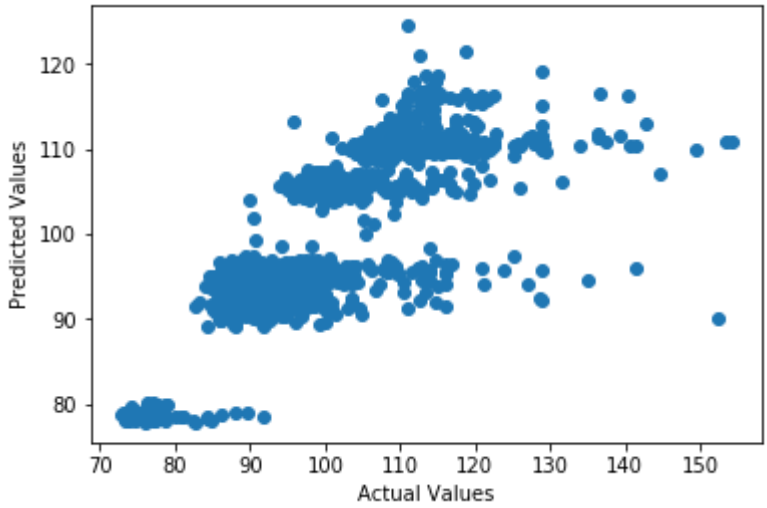
In [49]:

```
# Let's perform a cross-validation to find the best combination of alpha and l1_ratio
from sklearn.linear_model import ElasticNetCV, ElasticNet

cv_model = ElasticNetCV(l1_ratio=[.1, .5, .7, .9, .95, .99, .995, 1], eps=0.001, n_alph
as=100, fit_intercept=True,
                        normalize=True, precompute='auto', max_iter=2000, tol=0.0001, c
v=5,
                        copy_X=True, verbose=0, n_jobs=-1, positive=False, random_state
=None, selection='cyclic')

cv_model.fit(X_train, y_train)

print('Optimal alpha:' , cv_model.alpha_)
print('Optimal l1_ratio:' , cv_model.l1_ratio_)
print('Number of iterations' , cv_model.n_iter_)

from sklearn.linear_model import ElasticNetCV, ElasticNet

model = ElasticNet(l1_ratio=cv_model.l1_ratio_, alpha = cv_model.alpha_, max_iter=cv_mo
del.n_iter_, fit_intercept=True, normalize = True)
model.fit(X_train, y_train)

print('Train Score')
print(r2_score(y_train, model.predict(X_train)))
print('Test Score')
print(r2_score(y_test, model.predict(X_test)))

print('Actual Vs Predicted')
plt.scatter(y_test, model.predict(X_test))
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Optimal alpha: 0.003339585614035846
Optimal l1_ratio: 1.0
Number of iterations 391
Train Score
0.6085805059834002
Test Score
0.6195847628455537
Actual Vs Predicted
```

Out[49]:

```
Text(0, 0.5, 'Predicted Values')
```



XGBOOST

In [50]:

```
## Cross validation
import xgboost as xgb
from sklearn.metrics import r2_score
from sklearn.model_selection import GridSearchCV

XGB = xgb.XGBRegressor(objective='reg:squarederror')
XGB_para = {'n_estimators':[100,300,500],'max_depth':[2, 3, 4] , 'learning_rate' : [0.0
01,0.01,0.1,0.2]}
clf = GridSearchCV(XGB, XGB_para, cv=3, scoring='r2')
clf.fit(X_train, y_train)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
```

```
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
```

```
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
```

```
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
```

```
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
C:\Users\Sharad\Anaconda3\lib\site-packages\xgboost\core.py:587: FutureWar
ning: Series.base is deprecated and will be removed in a future version
  if getattr(data, 'base', None) is not None and \
```

Out[50]:

```
GridSearchCV(cv=3, error_score='raise-deprecating',
       estimator=XGBRegressor(base_score=0.5, booster='gbtree', colsample_
bylevel=1,
       colsample_bynode=1, colsample_bytree=1, gamma=0,
       importance_type='gain', learning_rate=0.1, max_delta_step=0,
       max_depth=3, min_child_weight=1, missing=None, n_estimators=100,
       n_jobs=1, nthread=None, objective='reg:squarederror',
       random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1,
       seed=None, silent=None, subsample=1, verbosity=1),
       fit_params=None, iid='warn', n_jobs=None,
       param_grid={'n_estimators': [100, 300, 500], 'max_depth': [2, 3,
4], 'learning_rate': [0.001, 0.01, 0.1, 0.2]},
       pre_dispatch='2*n_jobs', refit=True, return_train_score='warn',
       scoring='r2', verbose=0)
```

In [51]:

```
clf.best_params_
```

Out[51]:

```
{'learning_rate': 0.1, 'max_depth': 2, 'n_estimators': 100}
```

In [52]:

```
xg_reg = xgb.XGBRegressor(objective ='reg:linear',learning_rate = 0.1,
                max_depth = 2, n_estimators = 100)
xg_reg.fit(X_train, y_train)
```

```
[17:39:17] WARNING: src/objective/regression_obj.cu:152: reg:linear is now
deprecated in favor of reg:squarederror.
```

Out[52]:

```
XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1,
       colsample_bynode=1, colsample_bytree=1, gamma=0,
       importance_type='gain', learning_rate=0.1, max_delta_step=0,
       max_depth=2, min_child_weight=1, missing=None, n_estimators=100,
       n_jobs=1, nthread=None, objective='reg:linear', random_state=0,
       reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None,
       silent=None, subsample=1, verbosity=1)
```

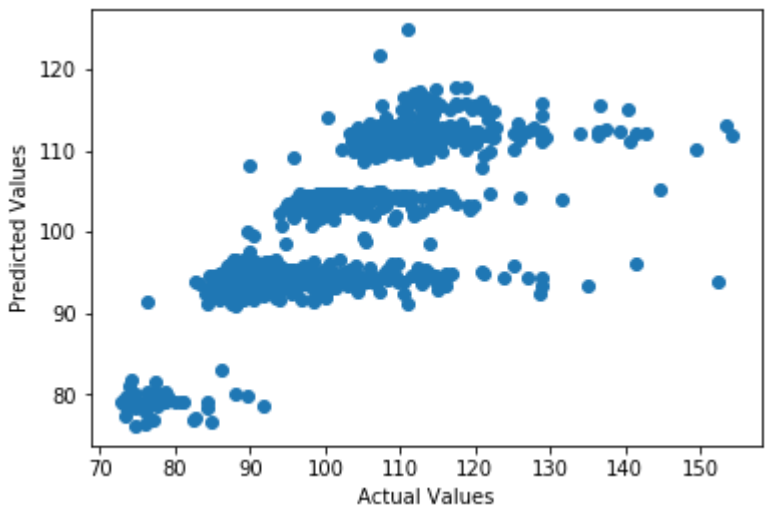
In [53]:

```
from sklearn.metrics import r2_score
preds = xg_reg.predict(X_test)


print('Train r2 :',r2_score(y_train, xg_reg.predict(X_train)))
print('Test r2 :', r2_score(y_test, preds))

print('Actual Vs Predicted')
plt.scatter(y_test,preds)
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Train r2 : 0.6501487419979821
Test r2 : 0.6303809036930859
Actual Vs Predicted
```

Out[53]:

```
Text(0, 0.5, 'Predicted Values')
```



Stacked Model

In [54]:

```
stacked_pipeline = make_pipeline(
    StackingEstimator(estimator=LassoLarsCV(normalize=True)),
    StackingEstimator(estimator=GradientBoostingRegressor(learning_rate=0.001, loss="hu
ber", max_depth=3, max_features=0.55, min_samples_leaf=18, min_samples_split=14, subsam
ple=0.7)),
    LassoLarsCV()

)


stacked_pipeline.fit(X_train, y_train)
results = stacked_pipeline.predict(X_test)

print('R2 score on train data:')
print(r2_score(y_train,stacked_pipeline.predict(X_train)))

print('R2 score on test data:')
print(r2_score(y_test,stacked_pipeline.predict(X_test)))
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.037e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.664e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.037e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 2.220e-1
6. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.519e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.664e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.428e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 6.495e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.699e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 1.825e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
```

```
ng a regressor, after 14 iterations, i.e. alpha=1.677e-02, with an active
set of 14 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
20 iterations, alpha=1.391e-02, previous alpha=1.349e-02, with an active s
et of 17 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.391e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=7.391e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=6.115e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 1.490e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.773e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 6 iterations, i.e. alpha=3.629e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 6 iterations, i.e. alpha=3.629e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.712e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 8 iterations, i.e. alpha=2.587e-02, with an active s
et of 8 regressors, and the smallest cholesky pivot element being 1.490e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.808e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 4.712e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.797e-02, with an active
```

```
set of 15 regressors, and the smallest cholesky pivot element being 4.829e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.797e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 4.712e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 16 iterations, i.e. alpha=1.769e-02, with an active
set of 16 regressors, and the smallest cholesky pivot element being 9.657e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 16 iterations, i.e. alpha=1.769e-02, with an active
set of 16 regressors, and the smallest cholesky pivot element being 9.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.689e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 6.989e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.689e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 6.409e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.689e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 19 iterations, i.e. alpha=1.689e-02, with an active
set of 19 regressors, and the smallest cholesky pivot element being 6.495e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
20 iterations, alpha=1.789e-02, previous alpha=1.667e-02, with an active s
et of 19 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=6.839e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.555e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 2.581e-0
```

```
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=3.555e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.412e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 6.409e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 11 iterations, i.e. alpha=2.422e-02, with an active
set of 11 regressors, and the smallest cholesky pivot element being 6.322e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
12 iterations, alpha=2.442e-02, previous alpha=2.422e-02, with an active s
et of 11 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 1.054e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 2.788e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=5.455e-02, with an active s
et of 3 regressors, and the smallest cholesky pivot element being 3.161e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.132e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 3.161e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.132e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 3.332e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=2.674e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 2.788e-0
8. Reduce max_iter or increase eps parameters.
```

```
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=2.674e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 1.054e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 8 iterations, i.e. alpha=2.316e-02, with an active s
et of 8 regressors, and the smallest cholesky pivot element being 1.054e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 8 iterations, i.e. alpha=2.316e-02, with an active s
et of 8 regressors, and the smallest cholesky pivot element being 3.942e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.332e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.161e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.509e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 3.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 15 iterations, i.e. alpha=1.490e-02, with an active
set of 15 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 18 iterations, i.e. alpha=6.662e-03, with an active
set of 16 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 71 iterations, i.e. alpha=2.704e-03, with an active
set of 65 regressors, and the smallest cholesky pivot element being 6.144e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
```

```
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 71 iterations, i.e. alpha=2.704e-03, with an active
set of 65 regressors, and the smallest cholesky pivot element being 1.825e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 76 iterations, i.e. alpha=2.555e-03, with an active
set of 70 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 79 iterations, i.e. alpha=2.487e-03, with an active
set of 71 regressors, and the smallest cholesky pivot element being 1.054e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 82 iterations, i.e. alpha=2.454e-03, with an active
set of 74 regressors, and the smallest cholesky pivot element being 1.825e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 82 iterations, i.e. alpha=2.454e-03, with an active
set of 74 regressors, and the smallest cholesky pivot element being 4.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 96 iterations, i.e. alpha=2.143e-03, with an active
set of 82 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 98 iterations, i.e. alpha=2.072e-03, with an active
set of 84 regressors, and the smallest cholesky pivot element being 6.664e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 100 iterations, i.e. alpha=2.039e-03, with an active
set of 86 regressors, and the smallest cholesky pivot element being 6.664e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 100 iterations, i.e. alpha=2.039e-03, with an active
set of 86 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 102 iterations, i.e. alpha=2.033e-03, with an active
set of 88 regressors, and the smallest cholesky pivot element being 4.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
```

```
es are small and the current value of alpha is no longer well controlled.
106 iterations, alpha=2.017e-03, previous alpha=2.014e-03, with an active
set of 91 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
4 iterations, alpha=1.371e-02, previous alpha=1.370e-02, with an active se
t of 5 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 22 iterations, i.e. alpha=6.668e-03, with an active
set of 22 regressors, and the smallest cholesky pivot element being 5.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 29 iterations, i.e. alpha=5.968e-03, with an active
set of 29 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 31 iterations, i.e. alpha=5.783e-03, with an active
set of 31 regressors, and the smallest cholesky pivot element being 5.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 40 iterations, i.e. alpha=5.337e-03, with an active
set of 40 regressors, and the smallest cholesky pivot element being 8.229e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 54 iterations, i.e. alpha=4.696e-03, with an active
set of 52 regressors, and the smallest cholesky pivot element being 5.373e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 55 iterations, i.e. alpha=4.683e-03, with an active
set of 53 regressors, and the smallest cholesky pivot element being 5.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 55 iterations, i.e. alpha=4.683e-03, with an active
set of 53 regressors, and the smallest cholesky pivot element being 3.161e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
60 iterations, alpha=4.466e-03, previous alpha=4.384e-03, with an active s
et of 57 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 3 iterations, i.e. alpha=9.490e-03, with an active s
```

```
et of 3 regressors, and the smallest cholesky pivot element being 2.220e-1
6. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)

R2 score on train data:
0.6181109646209477
R2 score on test data:
0.6352489629261115
```

In [55]:

```
print('R2 score on train data:')
print(r2_score(y_train,stacked_pipeline.predict(X_train)))

print('R2 score on test data:')
print(r2_score(y_test,stacked_pipeline.predict(X_test)))
```

```
R2 score on train data:
0.6181109646209477
R2 score on test data:
0.6352489629261115
```

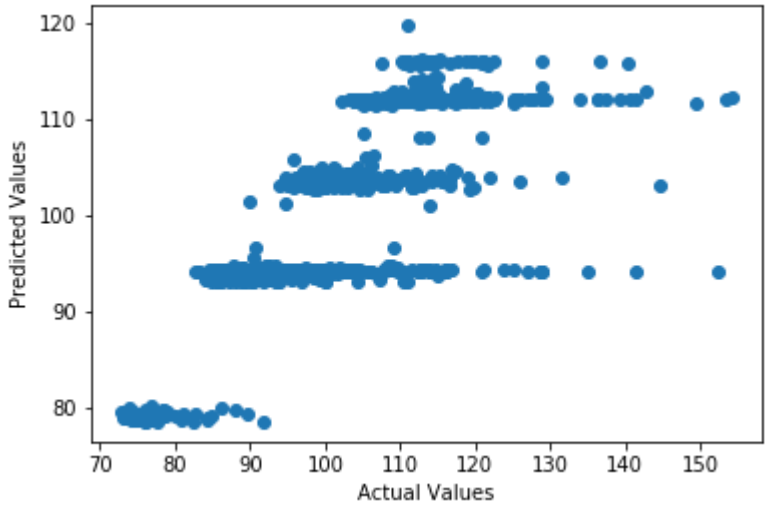
In [56]:

```
print('Actual Vs Predicted')
plt.scatter(y_test,stacked_pipeline.predict(X_test))
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[56]:

```
Text(0, 0.5, 'Predicted Values')
```



Stacked And Averaged Model

In [57]:

```
print('R2 score on train data:') ## Averaged with the XGBoost Model
print(r2_score(y_train,stacked_pipeline.predict(X_train)*0.2855 + xg_reg.predict(X_trai
n)*0.7145))

print('R2 score on test data:')
print(r2_score(y_test,stacked_pipeline.predict(X_test)*0.2855 + xg_reg.predict(X_test)*
0.7145))
```

```
R2 score on train data:
0.6428699894431107
R2 score on test data:
0.6334817432380395
```

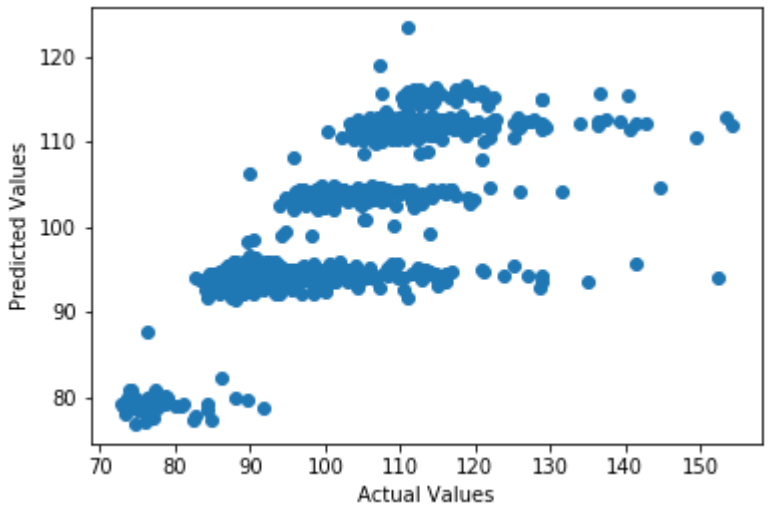
In [58]:

```
print('Actual Vs Predicted')
plt.scatter(y_test,stacked_pipeline.predict(X_test)*0.2855 + xg_reg.predict(X_test)*0.7
145)
plt.xlabel('Actual Values')
plt.ylabel('Predicted Values')
```

```
Actual Vs Predicted
```

Out[58]:

```
Text(0, 0.5, 'Predicted Values')
```



Procedure followed :

- Reading the Data
- Missing Values Check
- Dependent Varaible Analysis
- Feature's Analysis
- Feature Engineering
  - Removing Outliers
  - Vectorising Categorical Columns
  - Adding PCA And ICA Components
  - Adding TSVD , GRP and srp fetures
- Modelling

Conclusion :

In [92]:

```
from prettytable import PrettyTable
x = PrettyTable()
x.field_names = ["SET","Algorithm" , 'R2 Score on Train Data' , 'R2 Score on Test Data'
]
x.add_row(["SET1(PCA,ICA with Label_Encoder)","Elastic Net Regression", 0.603 , 0.599])
x.add_row(["SET1(PCA,ICA with Label_Encoder)","XGBOOST Regression", 0.647 , 0.63])
x.add_row(["SET1(PCA,ICA with Label_Encoder)","Stacked Model", 0.618 , 0.635])
x.add_row(["SET1(PCA,ICA with Label_Encoder)","Stacked And Averaged Model", 0.641 , 0.6
33])
x.add_row(["SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder)","Elastic Net Regression", 0.
608 , 0.619])
x.add_row(["SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder)","XGBOOST Regression", 0.650
, 0.63])
x.add_row(["SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder)","Stacked Model", 0.618 , 0.6
35])
x.add_row(["SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder)","Stacked And Averaged Model"
, 0.642 , 0.633])
print(x)
```

```
+-----------------------------------------------+------------------------
---+-----------------------+----------------------+
|                      SET                      |         Algorithm
| R2 Score on Train Data | R2 Score on Test Data |
+-----------------------------------------------+------------------------
---+-----------------------+----------------------+
|        SET1(PCA,ICA with Label_Encoder)       |   Elastic Net Regression
|         0.603          |         0.599         |
|        SET1(PCA,ICA with Label_Encoder)       |     XGBOOST Regression
|         0.647          |          0.63         |
|        SET1(PCA,ICA with Label_Encoder)       |       Stacked Model
|         0.618          |         0.635         |
|        SET1(PCA,ICA with Label_Encoder)       | Stacked And Averaged Mod
el |         0.641          |         0.633         |
| SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder) |   Elastic Net Regression
|         0.608          |         0.619         |
| SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder) |     XGBOOST Regression
|          0.65          |          0.63         |
| SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder) |       Stacked Model
|         0.618          |         0.635         |
| SET2(PCA,ICA,TSVD,GRP,SRP with Label_Encoder) | Stacked And Averaged Mod
el |         0.642          |         0.633         |
+-----------------------------------------------+------------------------
---+-----------------------+----------------------+
```

- There is not much difference between scores of set 1 and set 2

In [ ]:

```
* Stacked Model Has given better result.
* Making final Model with Stacked Model.
```

Final Model : Modularising Code

In [93]:

```
def final_Model(df) :

    df.drop(df[df['y'] > 175].index, axis=0, inplace=True) ## Removing outlier
    ## Test train split
    y = df['y']
    X = df.drop(['ID','y'],axis=1)

    # train test split
    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)

    from sklearn.preprocessing import LabelEncoder
    for c in X_train.columns:
        if X_train[c].dtype == 'object':
            lbl = LabelEncoder()
            lbl.fit(list(X_train[c].values) + list(X_test[c].values))
            X_train[c] = lbl.transform(list(X_train[c].values))
            X_test[c] = lbl.transform(list(X_test[c].values))

    from sklearn.decomposition import PCA, FastICA , TruncatedSVD
    from sklearn.random_projection import GaussianRandomProjection
    from sklearn.random_projection import SparseRandomProjection
    n_comp = 12

    # PCA
    pca = PCA(n_components=n_comp, random_state=42)
    pca2_results_train = pca.fit_transform(X_train)
    pca2_results_test = pca.transform(X_test)

    # ICA
    ica = FastICA(n_components=n_comp, random_state=42)
    ica2_results_train = ica.fit_transform(X_train)
    ica2_results_test = ica.transform(X_test)

    # TSVD
    tsvd = TruncatedSVD(n_components=n_comp, random_state=420)
    tsvd_results_train = tsvd.fit_transform(X_train)
    tsvd_results_test = tsvd.transform(X_test)

    # GRP
    grp = GaussianRandomProjection(n_components=n_comp, eps=0.1, random_state=420)
    grp_results_train = grp.fit_transform(X_train)
    grp_results_test = grp.transform(X_test)

    # SRP
    srp = SparseRandomProjection(n_components=n_comp, dense_output=True, random_state=4
20)
    srp_results_train = srp.fit_transform(X_train)
    srp_results_test = srp.transform(X_test)

    # Append decomposition components to datasets
    for i in range(1, n_comp+1):
        X_train['pca_' + str(i)] = pca2_results_train[:,i-1]
        X_test['pca_' + str(i)] = pca2_results_test[:, i-1]

        X_train['ica_' + str(i)] = ica2_results_train[:,i-1]
        X_test['ica_' + str(i)] = ica2_results_test[:, i-1]

        X_train['tsvd_' + str(i)] = tsvd_results_train[:, i - 1]
```

```
        X_test['tsvd_' + str(i)] = tsvd_results_test[:, i - 1]

        X_train['grp_' + str(i)] = grp_results_train[:, i - 1]
        X_test['grp_' + str(i)] = grp_results_test[:, i - 1]

        X_train['srp_' + str(i)] = srp_results_train[:, i - 1]
        X_test['srp_' + str(i)] = srp_results_test[:, i - 1]



    stacked_pipeline = make_pipeline(
    StackingEstimator(estimator=LassoLarsCV(normalize=True)),
    StackingEstimator(estimator=GradientBoostingRegressor(learning_rate=0.001, loss="hu
ber", max_depth=3, max_features=0.55, min_samples_leaf=18, min_samples_split=14, subsam
ple=0.7)),
    LassoLarsCV())


    stacked_pipeline.fit(X_train, y_train)
    results = stacked_pipeline.predict(X_test)

    print('R2 score on train data:')
    print(r2_score(y_train,stacked_pipeline.predict(X_train)))

    print('R2 score on test data:')
    print(r2_score(y_test,stacked_pipeline.predict(X_test)))

    print('Actual Vs Predicted')
    plt.scatter(y_test,stacked_pipeline.predict(X_test))
    plt.xlabel('Actual Values')
    plt.ylabel('Predicted Values')
```

In [94]:

```
Benz_df = pd.read_csv('train.csv')
final_Model(Benz_df)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:18: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:19: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\decomposition\fastica
_.py:121: ConvergenceWarning: FastICA did not converge. Consider increasin
g tolerance or the maximum number of iterations.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:53: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:54: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:56: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:57: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:59: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:60: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:62: Sett
```

```
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:63: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:65: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\ipykernel_launcher.py:66: Sett
ingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=6.989e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 2.220e-1
6. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=3.484e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 2.220e-1
6. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=3.484e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 8.941e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 12 iterations, i.e. alpha=1.734e-02, with an active
set of 12 regressors, and the smallest cholesky pivot element being 8.941e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 12 iterations, i.e. alpha=1.734e-02, with an active
set of 12 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
```

```
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.645e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 8.941e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 23 iterations, i.e. alpha=1.002e-02, with an active
set of 23 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 23 iterations, i.e. alpha=1.002e-02, with an active
set of 23 regressors, and the smallest cholesky pivot element being 5.960e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 23 iterations, i.e. alpha=1.002e-02, with an active
set of 23 regressors, and the smallest cholesky pivot element being 6.053e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 23 iterations, i.e. alpha=1.002e-02, with an active
set of 23 regressors, and the smallest cholesky pivot element being 6.144e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 26 iterations, i.e. alpha=9.218e-03, with an active
set of 26 regressors, and the smallest cholesky pivot element being 8.941e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 28 iterations, i.e. alpha=8.419e-03, with an active
set of 28 regressors, and the smallest cholesky pivot element being 8.941e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 29 iterations, i.e. alpha=8.130e-03, with an active
set of 29 regressors, and the smallest cholesky pivot element being 7.885e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 49 iterations, i.e. alpha=4.960e-03, with an active
set of 45 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 49 iterations, i.e. alpha=4.960e-03, with an active
set of 45 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 49 iterations, i.e. alpha=4.960e-03, with an active
set of 45 regressors, and the smallest cholesky pivot element being 5.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 49 iterations, i.e. alpha=4.960e-03, with an active
set of 45 regressors, and the smallest cholesky pivot element being 5.867e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
52 iterations, alpha=4.863e-03, previous alpha=4.809e-03, with an active s
et of 47 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 4 iterations, i.e. alpha=7.588e-02, with an active s
et of 4 regressors, and the smallest cholesky pivot element being 1.490e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=4.561e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 1.490e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=4.561e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 5.771e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 9 iterations, i.e. alpha=2.248e-02, with an active s
et of 9 regressors, and the smallest cholesky pivot element being 5.771e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 9 iterations, i.e. alpha=2.248e-02, with an active s
et of 9 regressors, and the smallest cholesky pivot element being 1.490e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 9 iterations, i.e. alpha=2.248e-02, with an active s
et of 9 regressors, and the smallest cholesky pivot element being 5.162e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 11 iterations, i.e. alpha=1.765e-02, with an active
set of 11 regressors, and the smallest cholesky pivot element being 5.771e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
```

```
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
19 iterations, alpha=1.294e-02, previous alpha=1.244e-02, with an active s
et of 18 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=9.303e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 7.885e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=9.303e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 7.814e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 7 iterations, i.e. alpha=6.207e-02, with an active s
et of 7 regressors, and the smallest cholesky pivot element being 1.054e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
8 iterations, alpha=6.415e-02, previous alpha=6.207e-02, with an active se
t of 7 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=6.099e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.593e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 2 iterations, i.e. alpha=6.099e-02, with an active s
et of 2 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.381e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 5 iterations, i.e. alpha=3.381e-02, with an active s
et of 5 regressors, and the smallest cholesky pivot element being 2.220e-1
6. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 6 iterations, i.e. alpha=3.007e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.829e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
```

```
ng a regressor, after 6 iterations, i.e. alpha=3.007e-02, with an active s
et of 6 regressors, and the smallest cholesky pivot element being 4.593e-0
8. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 10 iterations, i.e. alpha=2.170e-02, with an active
set of 10 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 12 iterations, i.e. alpha=1.637e-02, with an active
set of 12 regressors, and the smallest cholesky pivot element being 4.829e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 13 iterations, i.e. alpha=1.487e-02, with an active
set of 13 regressors, and the smallest cholesky pivot element being 4.829e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
14 iterations, alpha=1.536e-02, previous alpha=1.467e-02, with an active s
et of 13 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\model_selection\_spli
t.py:2053: FutureWarning: You should specify a value for 'cv' instead of r
elying on the default value. The default value will change from 3 to 5 in
version 0.22.
  warnings.warn(CV_WARNING, FutureWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 87 iterations, i.e. alpha=2.659e-03, with an active
set of 85 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 106 iterations, i.e. alpha=2.140e-03, with an active
set of 102 regressors, and the smallest cholesky pivot element being 2.980
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 113 iterations, i.e. alpha=2.000e-03, with an active
set of 107 regressors, and the smallest cholesky pivot element being 9.125
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 117 iterations, i.e. alpha=1.950e-03, with an active
set of 109 regressors, and the smallest cholesky pivot element being 1.490
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 9.186
```

```
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 4.942
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 5.960
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 2.980
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 9.771
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.583e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 8.816
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 134 iterations, i.e. alpha=1.582e-03, with an active
set of 126 regressors, and the smallest cholesky pivot element being 1.490
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 189 iterations, i.e. alpha=9.056e-04, with an active
set of 153 regressors, and the smallest cholesky pivot element being 5.771
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 202 iterations, i.e. alpha=7.995e-04, with an active
set of 156 regressors, and the smallest cholesky pivot element being 7.743
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 8.297
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 5.373
e-08. Reduce max_iter or increase eps parameters.
```

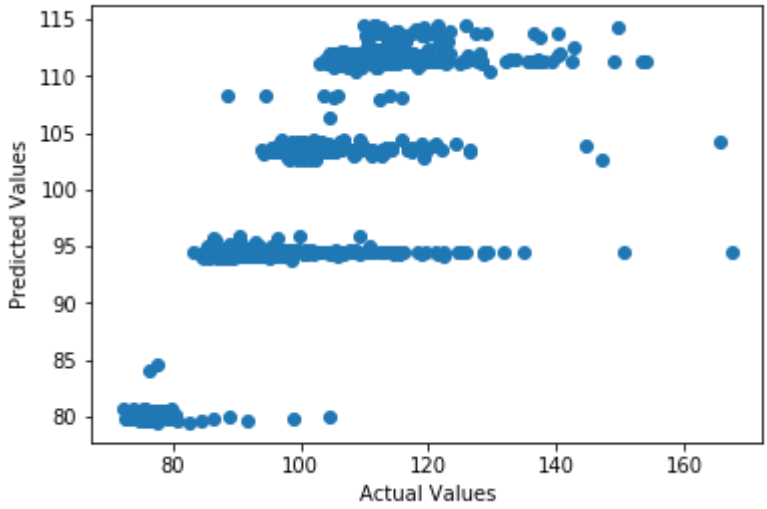
```
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 5.475
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 9.186
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 9.771
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 5.162
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 5.576
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 2.107
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 7.743
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 205 iterations, i.e. alpha=7.916e-04, with an active
set of 159 regressors, and the smallest cholesky pivot element being 1.490
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
209 iterations, alpha=7.751e-04, previous alpha=7.590e-04, with an active
set of 162 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 31 iterations, i.e. alpha=6.089e-03, with an active
set of 31 regressors, and the smallest cholesky pivot element being 8.941e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
```

```
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 36 iterations, i.e. alpha=5.231e-03, with an active
set of 34 regressors, and the smallest cholesky pivot element being 1.825e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 37 iterations, i.e. alpha=5.223e-03, with an active
set of 35 regressors, and the smallest cholesky pivot element being 3.332e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 47 iterations, i.e. alpha=4.203e-03, with an active
set of 45 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 52 iterations, i.e. alpha=3.839e-03, with an active
set of 50 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 52 iterations, i.e. alpha=3.839e-03, with an active
set of 50 regressors, and the smallest cholesky pivot element being 2.220e
-16. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 52 iterations, i.e. alpha=3.839e-03, with an active
set of 50 regressors, and the smallest cholesky pivot element being 5.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
57 iterations, alpha=3.708e-03, previous alpha=3.688e-03, with an active s
et of 54 regressors.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 40 iterations, i.e. alpha=4.374e-03, with an active
set of 40 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 67 iterations, i.e. alpha=3.170e-03, with an active
set of 63 regressors, and the smallest cholesky pivot element being 5.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 75 iterations, i.e. alpha=2.912e-03, with an active
set of 69 regressors, and the smallest cholesky pivot element being 6.747e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
```

```
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 92 iterations, i.e. alpha=2.516e-03, with an active
set of 86 regressors, and the smallest cholesky pivot element being 3.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 92 iterations, i.e. alpha=2.516e-03, with an active
set of 86 regressors, and the smallest cholesky pivot element being 4.942e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 105 iterations, i.e. alpha=2.284e-03, with an active
set of 95 regressors, and the smallest cholesky pivot element being 1.490e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 2.581e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 2.788e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 8.297e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 8.093e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 8.162e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 9.657e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 2.107e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
```

```
ng a regressor, after 107 iterations, i.e. alpha=2.258e-03, with an active
set of 97 regressors, and the smallest cholesky pivot element being 2.980e
-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 114 iterations, i.e. alpha=2.141e-03, with an active
set of 100 regressors, and the smallest cholesky pivot element being 3.799
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 114 iterations, i.e. alpha=2.141e-03, with an active
set of 100 regressors, and the smallest cholesky pivot element being 2.581
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 2.581
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 1.054
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 6.144
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 6.053
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 2.107
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:311: ConvergenceWarning: Regressors in active set degenerate. Droppi
ng a regressor, after 137 iterations, i.e. alpha=1.786e-03, with an active
set of 117 regressors, and the smallest cholesky pivot element being 1.490
e-08. Reduce max_iter or increase eps parameters.
  ConvergenceWarning)
C:\Users\Sharad\Anaconda3\lib\site-packages\sklearn\linear_model\least_ang
le.py:337: ConvergenceWarning: Early stopping the lars path, as the residu
es are small and the current value of alpha is no longer well controlled.
147 iterations, alpha=1.628e-03, previous alpha=1.610e-03, with an active
set of 124 regressors.
  ConvergenceWarning)
```

```
R2 score on train data:
0.6094053167449338
R2 score on test data:
0.579671465810382
Actual Vs Predicted
```



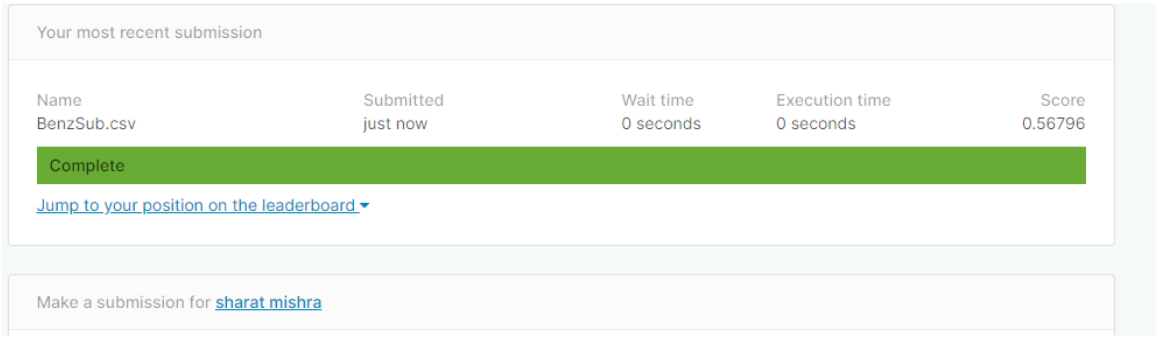
Kaggle Score

In [98]:

```
print('Public Leader Board')
from IPython.display import Image
Image(filename='Benz_KScore.png')
```

```
Public Leader Board
```

Out[98]:



In [ ]:

```
# I Reffered different kernels and discussion for the Competition  and tried incorporat
ing
# Different Feature Engineering Techniques  (1st place solution : Public LB, 0.55799, P
rivate LB: 0.55421)
# My Score is in Top 10% for Private Leader Board
# The score to get into top 10% of Public Leader Board is 0.57736
```

In [99]:

```
from IPython.display import Image
Image(filename='KScore.png')
```

Out[99]:

Please Accept

References

- Kaggle Discussions and Kernels
- https://www.kaggle.com/hakeem/stacked-then-averaged-models-0-5697?scriptVersionId=1236940 (https://www.kaggle.com/hakeem/stacked-then-averaged-models-0-5697?scriptVersionId=1236940)
- https://www.kaggle.com/c/mercedes-benz-greener-manufacturing/discussion/36126 (https://www.kaggle.com/c/mercedes-benz-greener-manufacturing/discussion/36126)