Forest Cover Type Prediction - Internship Project

Data Initialization

Dependencies

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier
from sklearn.metrics import accuracy_score, classification_report,
f1_score

pd.set_option('display.max_columns', None)
```

Merging two data

```
data1 = pd.read_csv('train.csv',index_col="Id") # Got from internship
data2 = pd.read_csv('covtype.csv') # got from internet
forestData = pd.concat([data2,data1],ignore_index=True)
```

Viewing Data

```
forestData.head()

   Elevation  Aspect  Slope  Horizontal_Distance_To_Hydrology  \
0       2596      51      3                               258
1       2590      56      2                               212
2       2804     139      9                               268
3       2785     155     18                               242
4       2595      45      2                               153

   Vertical_Distance_To_Hydrology  Horizontal_Distance_To_Roadways  \
0                               0                              510
1                              -6                              390
2                              65                             3180
3                             118                             3090
4                              -1                              391

   Hillshade_9am  Hillshade_Noon  Hillshade_3pm  \
0            221             232            148
1            220             235            151
```

```
2            234             238            135
3            238             238            122
4            220             234            150

   Horizontal_Distance_To_Fire_Points  Wilderness_Area1
Wilderness_Area2  \
0                                6279                 1
0
1                                6225                 1
0
2                                6121                 1
0
3                                6211                 1
0
4                                6172                 1
0

   Wilderness_Area3  Wilderness_Area4  Soil_Type1  Soil_Type2
Soil_Type3  \
0                 0                 0           0           0
0
1                 0                 0           0           0
0
2                 0                 0           0           0
0
3                 0                 0           0           0
0
4                 0                 0           0           0
0

   Soil_Type4  Soil_Type5  Soil_Type6  Soil_Type7  Soil_Type8
Soil_Type9  \
0           0           0           0           0           0
0
1           0           0           0           0           0
0
2           0           0           0           0           0
0
3           0           0           0           0           0
0
4           0           0           0           0           0
0

   Soil_Type10  Soil_Type11  Soil_Type12  Soil_Type13  Soil_Type14  \
0            0            0            0            0            0
1            0            0            0            0            0
2            0            0            1            0            0
3            0            0            0            0            0
4            0            0            0            0            0
```

```
   Soil_Type15  Soil_Type16  Soil_Type17  Soil_Type18  Soil_Type19  \
0            0            0            0            0            0   
1            0            0            0            0            0   
2            0            0            0            0            0   
3            0            0            0            0            0   
4            0            0            0            0            0   

   Soil_Type20  Soil_Type21  Soil_Type22  Soil_Type23  Soil_Type24  \
0            0            0            0            0            0   
1            0            0            0            0            0   
2            0            0            0            0            0   
3            0            0            0            0            0   
4            0            0            0            0            0   

   Soil_Type25  Soil_Type26  Soil_Type27  Soil_Type28  Soil_Type29  \
0            0            0            0            0            1   
1            0            0            0            0            1   
2            0            0            0            0            0   
3            0            0            0            0            0   
4            0            0            0            0            1   

   Soil_Type30  Soil_Type31  Soil_Type32  Soil_Type33  Soil_Type34  \
0            0            0            0            0            0   
1            0            0            0            0            0   
2            0            0            0            0            0   
3            1            0            0            0            0   
4            0            0            0            0            0   

   Soil_Type35  Soil_Type36  Soil_Type37  Soil_Type38  Soil_Type39  \
0            0            0            0            0            0   
1            0            0            0            0            0   
2            0            0            0            0            0   
3            0            0            0            0            0   
4            0            0            0            0            0   

   Soil_Type40  Cover_Type  
0            0           5  
1            0           5  
2            0           2  
3            0           2  
4            0           5  
```

Shape

```python
print(f"No. of rows: {forestData.shape[0]}")
print(f"No. of cols: {forestData.shape[1]}")
```

```
No. of rows: 596132
No. of cols: 55
```

Data Info

```
forestData.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 596132 entries, 0 to 596131
Data columns (total 55 columns):
 #   Column                              Non-Null Count   Dtype
---  ------                              --------------   -----
 0   Elevation                           596132 non-null  int64
 1   Aspect                              596132 non-null  int64
 2   Slope                               596132 non-null  int64
 3   Horizontal_Distance_To_Hydrology    596132 non-null  int64
 4   Vertical_Distance_To_Hydrology      596132 non-null  int64
 5   Horizontal_Distance_To_Roadways     596132 non-null  int64
 6   Hillshade_9am                       596132 non-null  int64
 7   Hillshade_Noon                      596132 non-null  int64
 8   Hillshade_3pm                       596132 non-null  int64
 9   Horizontal_Distance_To_Fire_Points  596132 non-null  int64
 10  Wilderness_Area1                    596132 non-null  int64
 11  Wilderness_Area2                    596132 non-null  int64
 12  Wilderness_Area3                    596132 non-null  int64
 13  Wilderness_Area4                    596132 non-null  int64
 14  Soil_Type1                          596132 non-null  int64
 15  Soil_Type2                          596132 non-null  int64
 16  Soil_Type3                          596132 non-null  int64
 17  Soil_Type4                          596132 non-null  int64
 18  Soil_Type5                          596132 non-null  int64
 19  Soil_Type6                          596132 non-null  int64
 20  Soil_Type7                          596132 non-null  int64
 21  Soil_Type8                          596132 non-null  int64
 22  Soil_Type9                          596132 non-null  int64
 23  Soil_Type10                         596132 non-null  int64
 24  Soil_Type11                         596132 non-null  int64
 25  Soil_Type12                         596132 non-null  int64
 26  Soil_Type13                         596132 non-null  int64
 27  Soil_Type14                         596132 non-null  int64
 28  Soil_Type15                         596132 non-null  int64
 29  Soil_Type16                         596132 non-null  int64
 30  Soil_Type17                         596132 non-null  int64
 31  Soil_Type18                         596132 non-null  int64
 32  Soil_Type19                         596132 non-null  int64
 33  Soil_Type20                         596132 non-null  int64
 34  Soil_Type21                         596132 non-null  int64
 35  Soil_Type22                         596132 non-null  int64
 36  Soil_Type23                         596132 non-null  int64
 37  Soil_Type24                         596132 non-null  int64
 38  Soil_Type25                         596132 non-null  int64
 39  Soil_Type26                         596132 non-null  int64
 40  Soil_Type27                         596132 non-null  int64
 41  Soil_Type28                         596132 non-null  int64
 42  Soil_Type29                         596132 non-null  int64
 43  Soil_Type30                         596132 non-null  int64
 44  Soil_Type31                         596132 non-null  int64
 45  Soil_Type32                         596132 non-null  int64
 46  Soil_Type33                         596132 non-null  int64
 47  Soil_Type34                         596132 non-null  int64
 48  Soil_Type35                         596132 non-null  int64
 49  Soil_Type36                         596132 non-null  int64
 50  Soil_Type37                         596132 non-null  int64
 51  Soil_Type38                         596132 non-null  int64
 52  Soil_Type39                         596132 non-null  int64
 53  Soil_Type40                         596132 non-null  int64
 54  Cover_Type                          596132 non-null  int64
dtypes: int64(55)
memory usage: 250.1 MB
```

```
forestData.describe()
```

```
           Elevation         Aspect          Slope  \
count  596132.000000  596132.000000  596132.000000   
mean     2954.037879     155.682674      14.164522   
std       286.213696     111.867752       7.523713   
min      1859.000000       0.000000       0.000000   
25%      2801.000000      59.000000       9.000000   
50%      2993.000000     127.000000      13.000000   
75%      3163.000000     260.000000      19.000000   
max      3858.000000     360.000000      66.000000   

       Horizontal_Distance_To_Hydrology 
Vertical_Distance_To_Hydrology  \
count                     596132.000000 
596132.000000   
mean                         268.357052 
46.536990   
std                          212.590510 
58.376281   
min                            0.000000                     -
173.000000   
25%                          108.000000 
7.000000   
50%                          218.000000 
30.000000   
75%                          384.000000 
69.000000   
max                         1397.000000 
601.000000   

       Horizontal_Distance_To_Roadways  Hillshade_9am  Hillshade_Noon 
\
count                    596132.000000  596132.000000   596132.000000
mean                       2334.012289     212.160208      223.208306
std                        1556.966114      26.872779       19.863134
min                           0.000000       0.000000        0.000000
25%                        1092.000000     198.000000      213.000000
50%                        1976.000000     218.000000      226.000000
75%                        3304.000000     231.000000      237.000000
max                        7117.000000     254.000000      254.000000

       Hillshade_3pm  Horizontal_Distance_To_Fire_Points
Wilderness_Area1  \
count  596132.000000                       596132.000000
596132.000000
mean      142.339653                         1968.392089
0.443514
std        38.504181                         1321.038719
0.496800
min         0.000000                            0.000000
0.000000
25%       119.000000                         1015.000000
0.000000
50%       143.000000                         1698.000000
0.000000
75%       168.000000                         2538.000000
1.000000
max       254.000000                         7173.000000
1.000000

       Wilderness_Area2  Wilderness_Area3  Wilderness_Area4
Soil_Type1  \
count     596132.000000     596132.000000     596132.000000
596132.000000
mean           0.050967          0.435664          0.069855
0.005680
std            0.219930          0.495844          0.254903
0.075151
min            0.000000          0.000000          0.000000
0.000000
25%            0.000000          0.000000          0.000000
0.000000
50%            0.000000          0.000000          0.000000
0.000000
```

```
75%             0.000000          1.000000          0.000000
0.000000
max             1.000000          1.000000          1.000000
1.000000

          Soil_Type2     Soil_Type3     Soil_Type4     Soil_Type5  \
count  596132.000000  596132.000000  596132.000000  596132.000000
mean        0.013668       0.009704       0.022208       0.002956
std         0.116109       0.098031       0.147360       0.054286
min         0.000000       0.000000       0.000000       0.000000
25%         0.000000       0.000000       0.000000       0.000000
50%         0.000000       0.000000       0.000000       0.000000
75%         0.000000       0.000000       0.000000       0.000000
max         1.000000       1.000000       1.000000       1.000000

          Soil_Type6     Soil_Type7     Soil_Type8     Soil_Type9  \
count  596132.000000  596132.000000  596132.000000  596132.000000
mean        0.012120       0.000176       0.000302       0.001941
std         0.109421       0.013270       0.017374       0.044012
min         0.000000       0.000000       0.000000       0.000000
25%         0.000000       0.000000       0.000000       0.000000
50%         0.000000       0.000000       0.000000       0.000000
75%         0.000000       0.000000       0.000000       0.000000
max         1.000000       1.000000       1.000000       1.000000

         Soil_Type10    Soil_Type11    Soil_Type12    Soil_Type13  \
count  596132.000000  596132.000000  596132.000000  596132.000000
mean        0.058336       0.021499       0.050657       0.030039
std         0.234378       0.145039       0.219296       0.170694
min         0.000000       0.000000       0.000000       0.000000
25%         0.000000       0.000000       0.000000       0.000000
50%         0.000000       0.000000       0.000000       0.000000
75%         0.000000       0.000000       0.000000       0.000000
max         1.000000       1.000000       1.000000       1.000000

         Soil_Type14    Soil_Type15    Soil_Type16    Soil_Type17  \
count  596132.000000  596132.000000  596132.000000  596132.000000
mean        0.001288       0.000005       0.004964       0.006767
std         0.035870       0.002243       0.070278       0.081983
min         0.000000       0.000000       0.000000       0.000000
25%         0.000000       0.000000       0.000000       0.000000
50%         0.000000       0.000000       0.000000       0.000000
75%         0.000000       0.000000       0.000000       0.000000
max         1.000000       1.000000       1.000000       1.000000

         Soil_Type18    Soil_Type19    Soil_Type20    Soil_Type21  \
count  596132.000000  596132.000000  596132.000000  596132.000000
mean        0.003286       0.006822       0.015765       0.001433
std         0.057231       0.082315       0.124565       0.037822
min         0.000000       0.000000       0.000000       0.000000
```

```
25%          0.000000       0.000000       0.000000       0.000000  
50%          0.000000       0.000000       0.000000       0.000000  
75%          0.000000       0.000000       0.000000       0.000000  
max          1.000000       1.000000       1.000000       1.000000  

          Soil_Type22    Soil_Type23    Soil_Type24    Soil_Type25  \
count  596132.000000  596132.000000  596132.000000  596132.000000   
mean        0.056561       0.098148       0.036125       0.000797   
std         0.231003       0.297515       0.186600       0.028216   
min         0.000000       0.000000       0.000000       0.000000   
25%         0.000000       0.000000       0.000000       0.000000   
50%         0.000000       0.000000       0.000000       0.000000   
75%         0.000000       0.000000       0.000000       0.000000   
max         1.000000       1.000000       1.000000       1.000000   

          Soil_Type26    Soil_Type27    Soil_Type28    Soil_Type29  \
count  596132.000000  596132.000000  596132.000000  596132.000000   
mean        0.004434       0.001847       0.001602       0.195490   
std         0.066437       0.042936       0.039993       0.396578   
min         0.000000       0.000000       0.000000       0.000000   
25%         0.000000       0.000000       0.000000       0.000000   
50%         0.000000       0.000000       0.000000       0.000000   
75%         0.000000       0.000000       0.000000       0.000000   
max         1.000000       1.000000       1.000000       1.000000   

          Soil_Type30    Soil_Type31    Soil_Type32    Soil_Type33  \
count  596132.000000  596132.000000  596132.000000  596132.000000   
mean        0.051826       0.043611       0.089257       0.076778   
std         0.221675       0.204229       0.285115       0.266240   
min         0.000000       0.000000       0.000000       0.000000   
25%         0.000000       0.000000       0.000000       0.000000   
50%         0.000000       0.000000       0.000000       0.000000   
75%         0.000000       0.000000       0.000000       0.000000   
max         1.000000       1.000000       1.000000       1.000000   

          Soil_Type34    Soil_Type35    Soil_Type36    Soil_Type37  \
count  596132.000000  596132.000000  596132.000000  596132.000000   
mean        0.002739       0.003343       0.000216       0.000557   
std         0.052267       0.057724       0.014709       0.023593   
min         0.000000       0.000000       0.000000       0.000000   
25%         0.000000       0.000000       0.000000       0.000000   
50%         0.000000       0.000000       0.000000       0.000000   
75%         0.000000       0.000000       0.000000       0.000000   
max         1.000000       1.000000       1.000000       1.000000   

          Soil_Type38    Soil_Type39    Soil_Type40     Cover_Type  
count  596132.000000  596132.000000  596132.000000  596132.000000  
mean        0.027345       0.024261       0.015448       2.100892  
std         0.163086       0.153860       0.123326       1.447781  
min         0.000000       0.000000       0.000000       1.000000  
```

```
25%           0.000000       0.000000       0.000000       1.000000
50%           0.000000       0.000000       0.000000       2.000000
75%           0.000000       0.000000       0.000000       2.000000
max           1.000000       1.000000       1.000000       7.000000
```

Checking for any Null values

```
forestData.isna().any()
```

```
Elevation                                     False
Aspect                                        False
Slope                                         False
Horizontal_Distance_To_Hydrology              False
Vertical_Distance_To_Hydrology                False
Horizontal_Distance_To_Roadways               False
Hillshade_9am                                 False
Hillshade_Noon                                False
Hillshade_3pm                                 False
Horizontal_Distance_To_Fire_Points            False
Wilderness_Area1                              False
Wilderness_Area2                              False
Wilderness_Area3                              False
Wilderness_Area4                              False
Soil_Type1                                    False
Soil_Type2                                    False
Soil_Type3                                    False
Soil_Type4                                    False
Soil_Type5                                    False
Soil_Type6                                    False
Soil_Type7                                    False
Soil_Type8                                    False
Soil_Type9                                    False
Soil_Type10                                   False
Soil_Type11                                   False
Soil_Type12                                   False
Soil_Type13                                   False
Soil_Type14                                   False
Soil_Type15                                   False
Soil_Type16                                   False
Soil_Type17                                   False
Soil_Type18                                   False
Soil_Type19                                   False
Soil_Type20                                   False
Soil_Type21                                   False
Soil_Type22                                   False
Soil_Type23                                   False
Soil_Type24                                   False
Soil_Type25                                   False
Soil_Type26                                   False
```

```
Soil_Type27                           False
Soil_Type28                           False
Soil_Type29                           False
Soil_Type30                           False
Soil_Type31                           False
Soil_Type32                           False
Soil_Type33                           False
Soil_Type34                           False
Soil_Type35                           False
Soil_Type36                           False
Soil_Type37                           False
Soil_Type38                           False
Soil_Type39                           False
Soil_Type40                           False
Cover_Type                            False
dtype: bool
```

Columns in the data

```
column = forestData.columns
column
```

```
Index(['Elevation', 'Aspect', 'Slope',
'Horizontal_Distance_To_Hydrology',
       'Vertical_Distance_To_Hydrology',
'Horizontal_Distance_To_Roadways',
       'Hillshade_9am', 'Hillshade_Noon', 'Hillshade_3pm',
       'Horizontal_Distance_To_Fire_Points', 'Wilderness_Area1',
       'Wilderness_Area2', 'Wilderness_Area3', 'Wilderness_Area4',
       'Soil_Type1', 'Soil_Type2', 'Soil_Type3', 'Soil_Type4',
'Soil_Type5',
       'Soil_Type6', 'Soil_Type7', 'Soil_Type8', 'Soil_Type9',
'Soil_Type10',
       'Soil_Type11', 'Soil_Type12', 'Soil_Type13', 'Soil_Type14',
       'Soil_Type15', 'Soil_Type16', 'Soil_Type17', 'Soil_Type18',
       'Soil_Type19', 'Soil_Type20', 'Soil_Type21', 'Soil_Type22',
       'Soil_Type23', 'Soil_Type24', 'Soil_Type25', 'Soil_Type26',
       'Soil_Type27', 'Soil_Type28', 'Soil_Type29', 'Soil_Type30',
       'Soil_Type31', 'Soil_Type32', 'Soil_Type33', 'Soil_Type34',
       'Soil_Type35', 'Soil_Type36', 'Soil_Type37', 'Soil_Type38',
       'Soil_Type39', 'Soil_Type40', 'Cover_Type'],
      dtype='object')
```

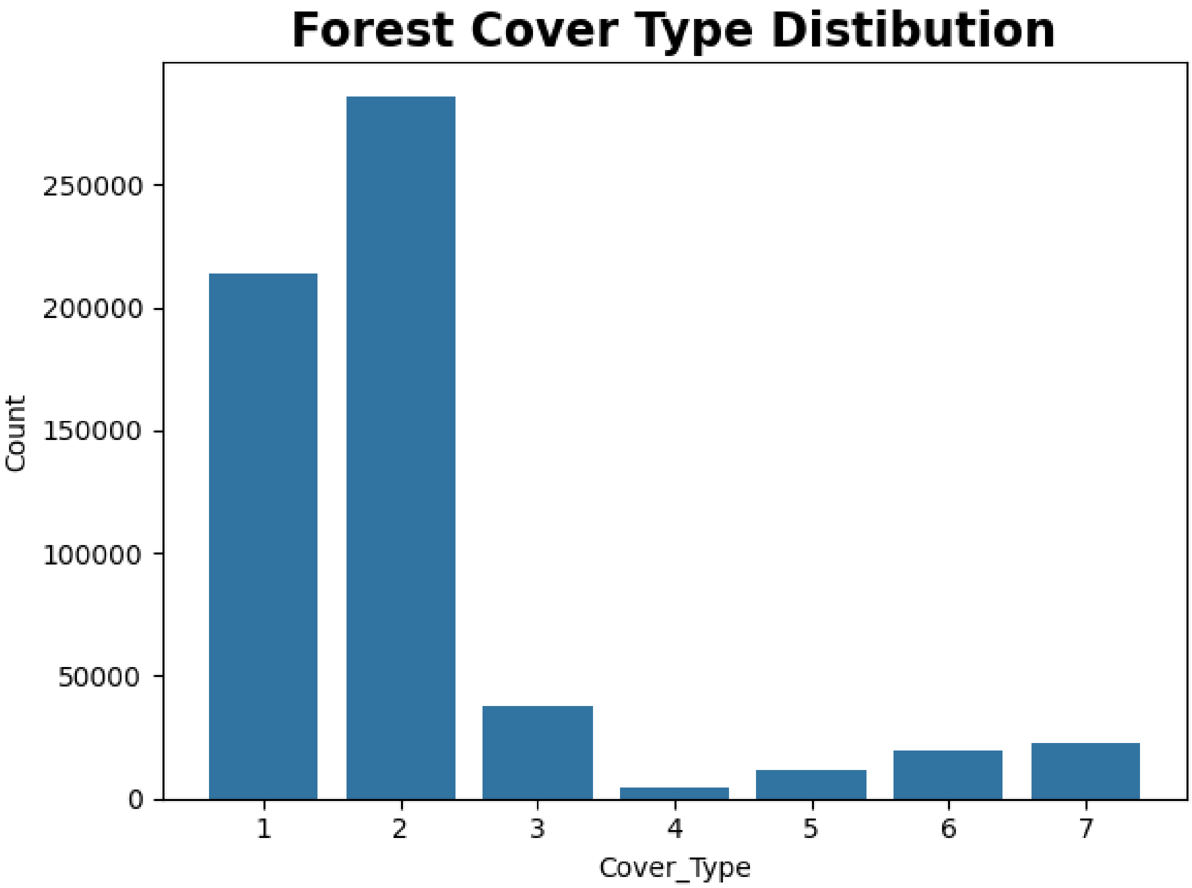
Note: There are no null values hence theres no need to do data cleaning

EDA

Target Variable Analysis

- Plot histogram of the forest cover type distribution.
- Check for class imbalance.

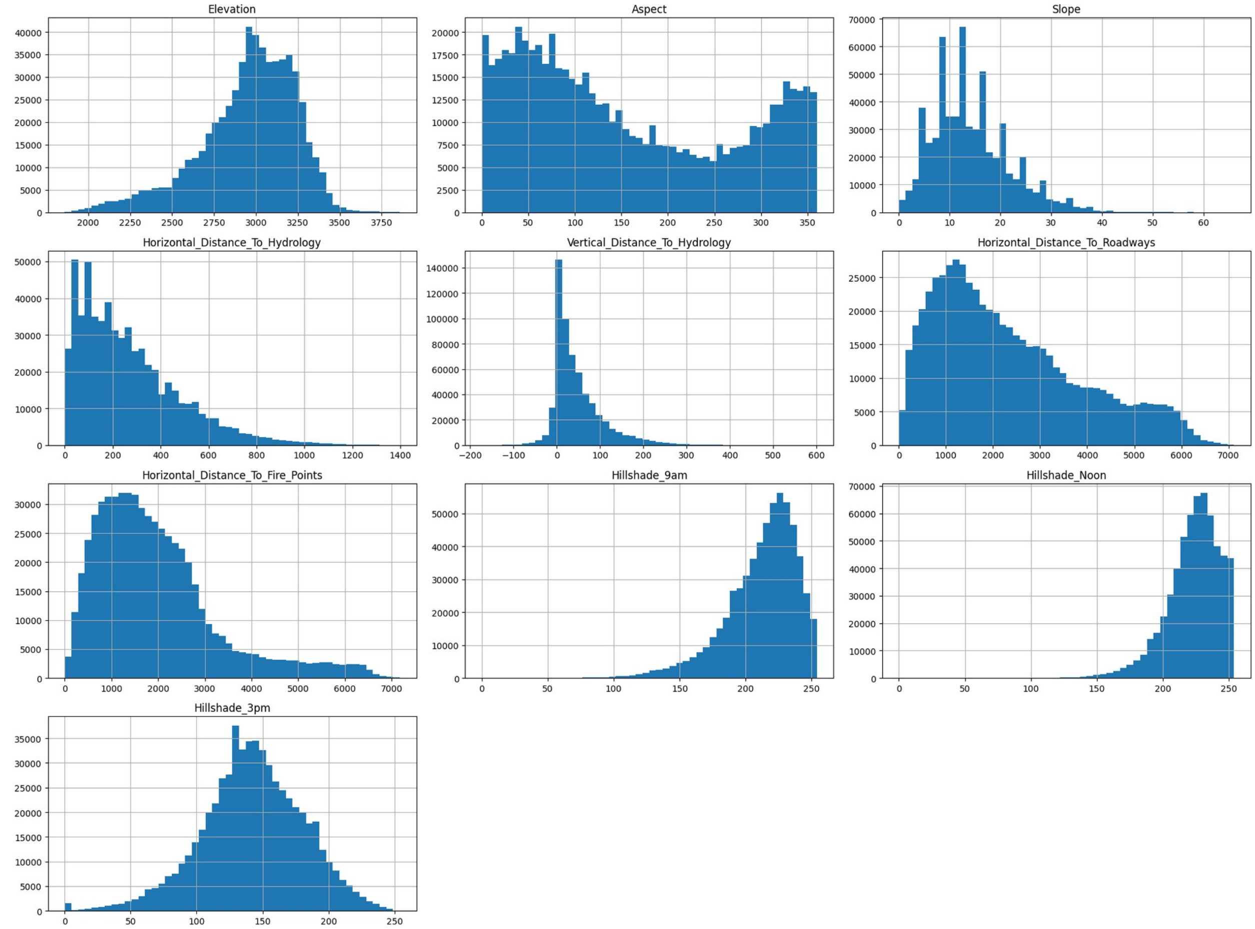
```
ax = sns.countplot(data=forestData,x='Cover_Type')
ax.set_xlabel('Cover_Type')
ax.set_ylabel('Count')
ax.set_title('Forest Cover Type Distibution',fontdict={'weight':'600','size':'17'})
plt.tight_layout()
plt.plot()

[]
```



Feature Distributions

Plot histograms for each necessary numerical feature

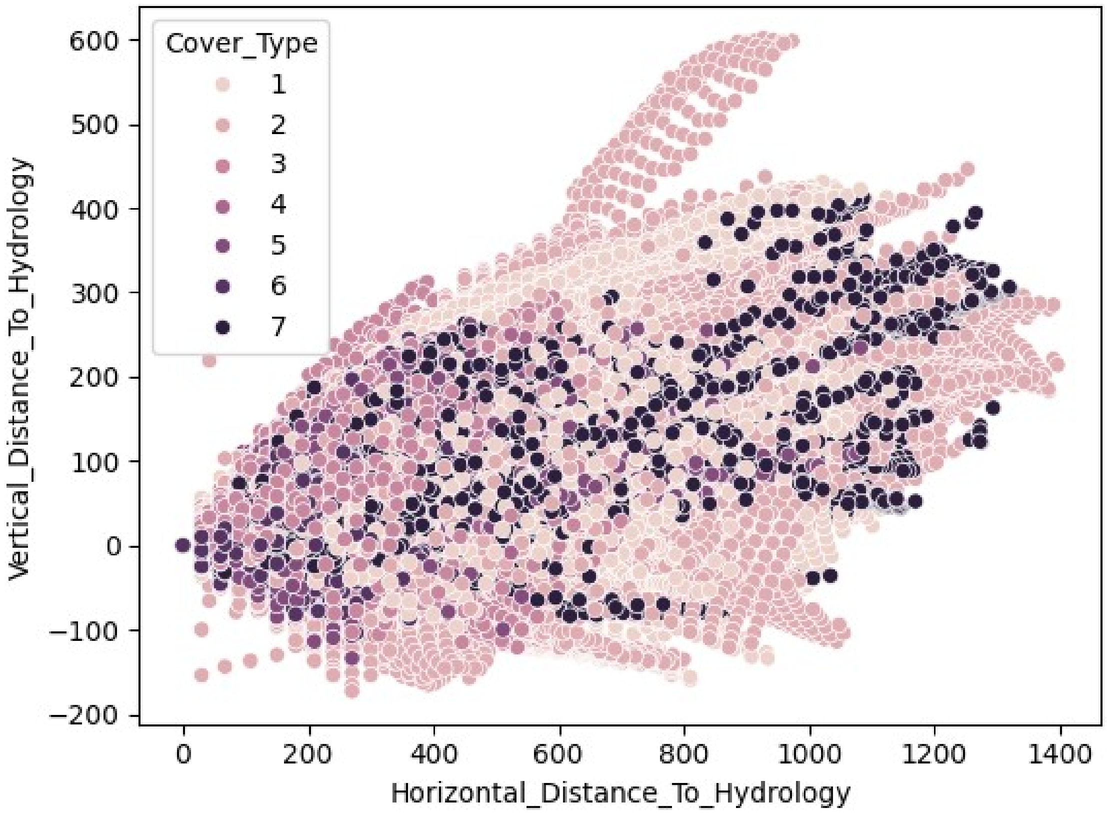
```
forestData[['Elevation','Aspect','Slope','Horizontal_Distance_To_Hydro
logy','Vertical_Distance_To_Hydrology',

'Horizontal_Distance_To_Roadways','Horizontal_Distance_To_Fire_Points'
,'Hillshade_9am','Hillshade_Noon',
           'Hillshade_3pm']].hist(bins=50, figsize=(20,15))
plt.tight_layout()
plt.show()
```



Geospatial Relationships

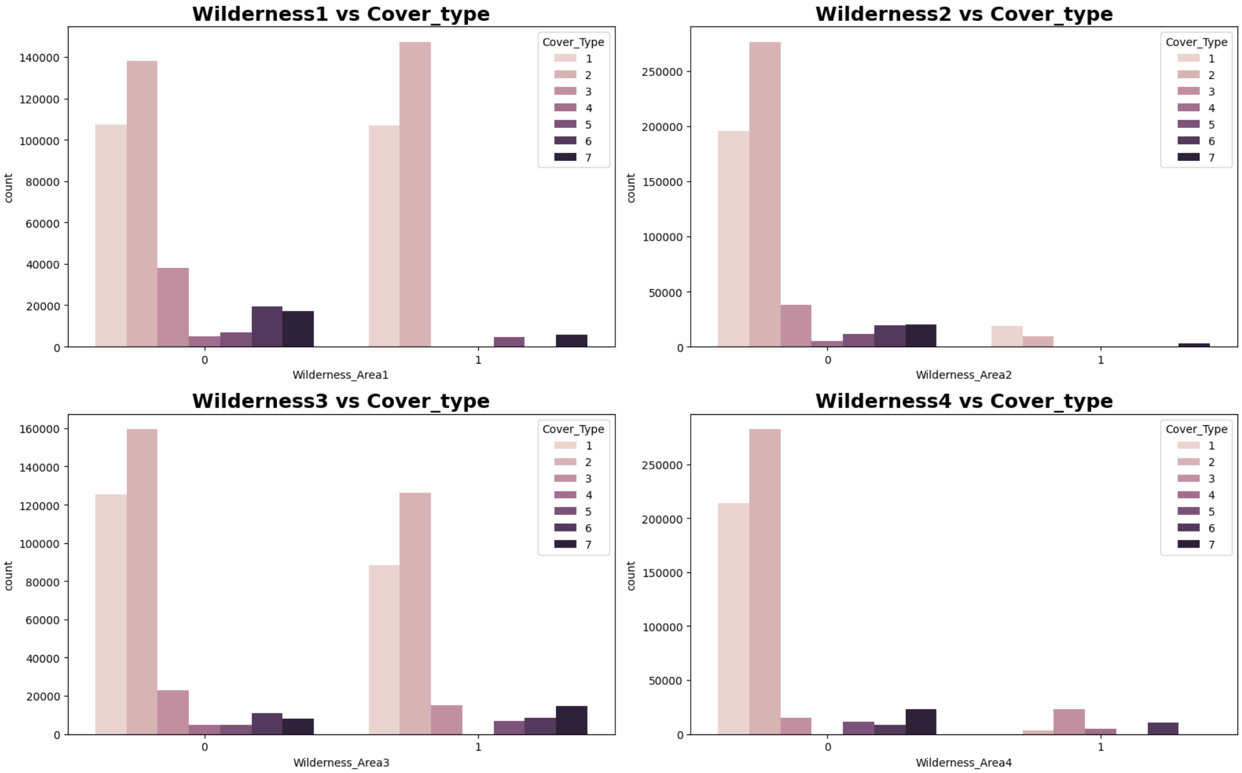
 Since this is geographical data, features like Horizontal_Distance_To_Roadways, Vertical_Distance_To_Hydrology, etc., may relate spatially.

```
sns.scatterplot(x='Horizontal_Distance_To_Hydrology',
y='Vertical_Distance_To_Hydrology', hue='Cover_Type', data=forestData)
plt.show()
```



Wilderness Area vs Cover Type analysis

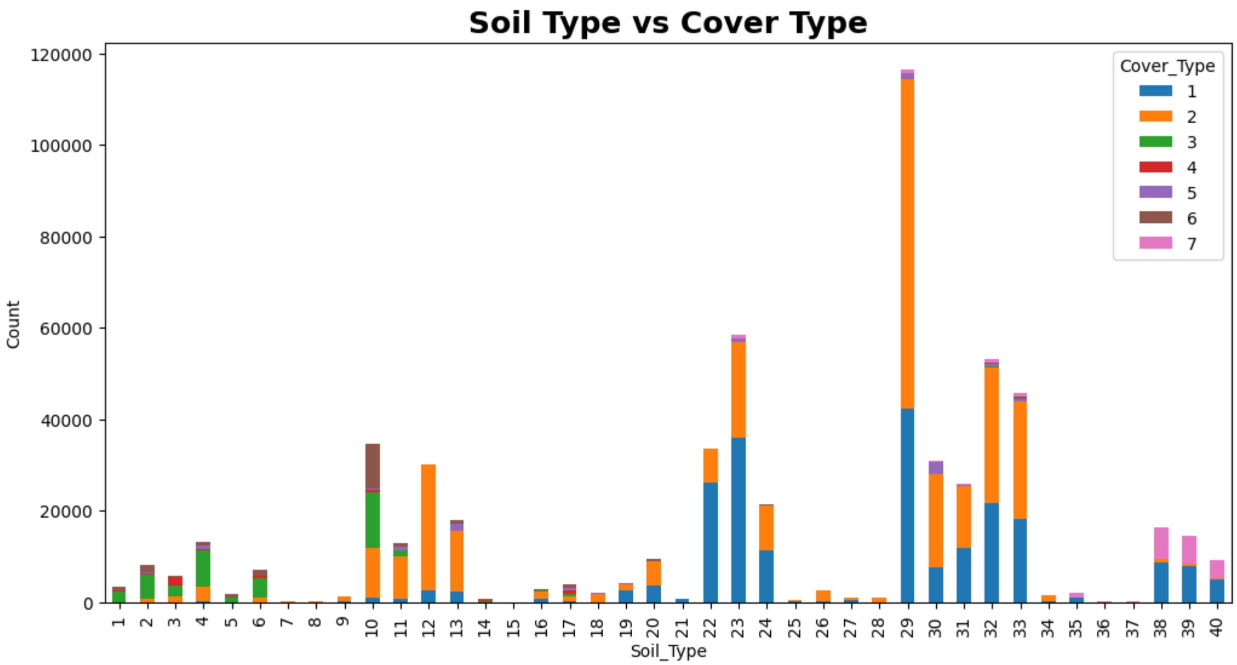
```
# make it for other wilderness area too
fig,axes = plt.subplots(2,2,figsize = (16,10))
ax1 = sns.countplot(x='Wilderness_Area1', hue='Cover_Type',
data=forestData,ax=axes[0,0])
ax2 = sns.countplot(x='Wilderness_Area2', hue='Cover_Type',
data=forestData,ax=axes[0,1])
ax3 = sns.countplot(x='Wilderness_Area3', hue='Cover_Type',
data=forestData,ax=axes[1,0])
ax4 = sns.countplot(x='Wilderness_Area4', hue='Cover_Type',
data=forestData,ax=axes[1,1])
ax1.set_title('Wilderness1 vs
Cover_type',fontdict={'size':'18','weight':'600'})
ax2.set_title('Wilderness2 vs
Cover_type',fontdict={'size':'18','weight':'600'})
ax3.set_title('Wilderness3 vs
Cover_type',fontdict={'size':'18','weight':'600'})
ax4.set_title('Wilderness4 vs
Cover_type',fontdict={'size':'18','weight':'600'})
plt.tight_layout()
plt.show()
```



Soil Type vs Cover type relation

Check relationships between soil types and cover types

```
soil_cols = [f"Soil_Type{i}" for i in range(1,41)]
soil_onehot = forestData[soil_cols]

# get soil type label (e.g. 'Soil_Type7') then convert to integer 7
soil_type_series = soil_onehot.idxmax(axis=1).str.replace('Soil_Type',
'').astype(int)

pd.crosstab(soil_type_series,
forestData['Cover_Type']).plot(kind='bar', stacked=True,
figsize=(12,6))
plt.xlabel('Soil_Type')
plt.ylabel('Count')
plt.title('Soil Type vs Cover
Type',fontdict={'size':'18','weight':'600'})
plt.show()
```



Data Preprocessing

Note: Because we are going to use Tree based models theres no need of Scaling our data

```
X = forestData.drop(['Cover_Type'],axis=1)
y = forestData['Cover_Type']
X_train,X_test,y_train,y_test =
train_test_split(X,y,test_size=0.3,random_state=30)
```

Model Selection

Random Forest Implementation

```
rf_classifier = RandomForestClassifier(n_estimators=100,
random_state=42)

rf_classifier.fit(X_train, y_train)
```

```
RandomForestClassifier(random_state=42)
```

XGBoost Implementation

```
xgb = XGBClassifier()
xgb_ytrain = y_train.apply(lambda x: x-1)
```

```
xgb.fit(X_train,xgb_ytrain)

XGBClassifier(base_score=None, booster=None, callbacks=None,
              colsample_bylevel=None, colsample_bynode=None,
              colsample_bytree=None, device=None,
early_stopping_rounds=None,
              enable_categorical=False, eval_metric=None,
feature_types=None,
              feature_weights=None, gamma=None, grow_policy=None,
              importance_type=None, interaction_constraints=None,
              learning_rate=None, max_bin=None,
max_cat_threshold=None,
              max_cat_to_onehot=None, max_delta_step=None,
max_depth=None,
              max_leaves=None, min_child_weight=None, missing=nan,
              monotone_constraints=None, multi_strategy=None,
n_estimators=None,
              n_jobs=None, num_parallel_tree=None, ...)
```

Model Evaluation

```
y_pred = rf_classifier.predict(X_test)

accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred)
weighted_f1 = f1_score(y_test, y_pred, average='weighted')

print(f"Accuracy: {accuracy:.2f}")
print(f"Weighted F1 Score: {weighted_f1}")
print("\nClassification Report:\n", classification_rep)

Accuracy: 0.96
Weighted F1 Score: 0.955195331162959

Classification Report:
               precision    recall  f1-score   support

           1       0.96      0.94      0.95     64145
           2       0.95      0.97      0.96     85642
           3       0.95      0.96      0.96     11388
           4       0.94      0.97      0.96      1487
           5       0.95      0.84      0.89      3417
           6       0.94      0.91      0.93      5938
           7       0.98      0.96      0.97      6823

    accuracy                           0.96    178840
   macro avg       0.95      0.94      0.94    178840
```

```
weighted avg       0.96      0.96      0.96    178840
```

```
y_pred = xgb.predict(X_test)
xgb_ytest = y_test.apply(lambda x: x-1)
accuracy = accuracy_score(xgb_ytest, y_pred)
classification_rep = classification_report(xgb_ytest, y_pred)
weighted_f1 = f1_score(xgb_ytest, y_pred, average='weighted')

print(f"Accuracy: {accuracy:.2f}")
print(f"Weighted F1 Score: {weighted_f1}")
print("\nClassification Report:\n", classification_rep)
```

```
Accuracy: 0.87
Weighted F1 Score: 0.8737411073859305

Classification Report:
               precision    recall  f1-score   support

           0       0.87      0.84      0.86     64145
           1       0.87      0.90      0.89     85642
           2       0.90      0.90      0.90     11388
           3       0.91      0.95      0.93      1487
           4       0.89      0.63      0.74      3417
           5       0.85      0.82      0.84      5938
           6       0.94      0.91      0.93      6823

    accuracy                           0.87    178840
   macro avg       0.89      0.85      0.87    178840
weighted avg       0.87      0.87      0.87    178840
```

Note: Random Forest perfomed best in this case