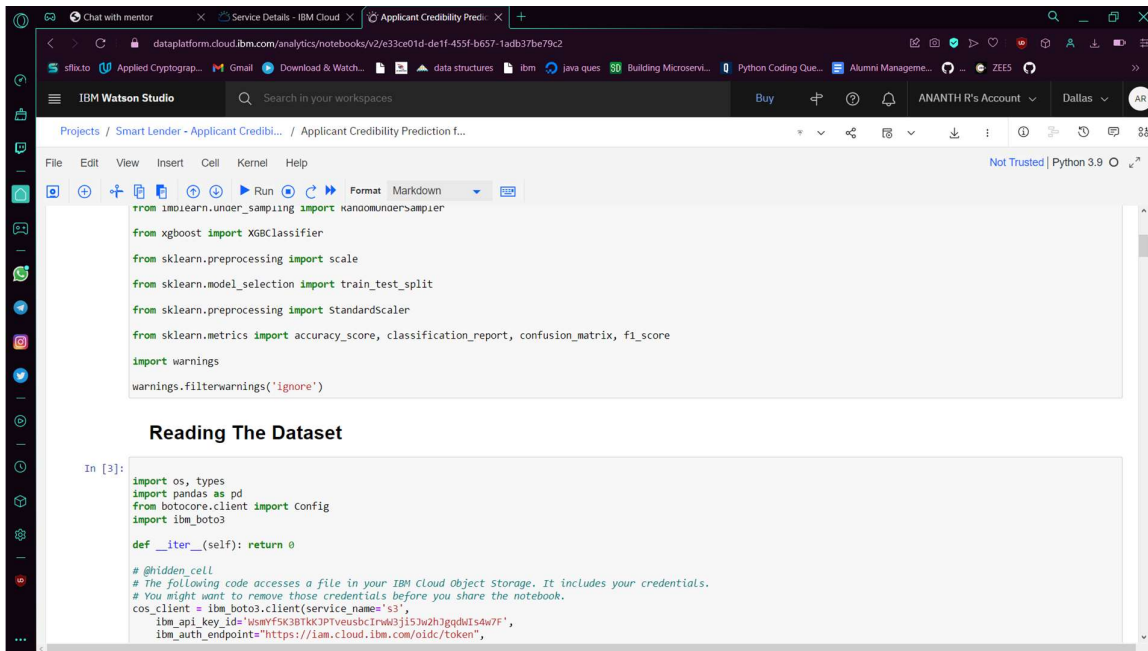
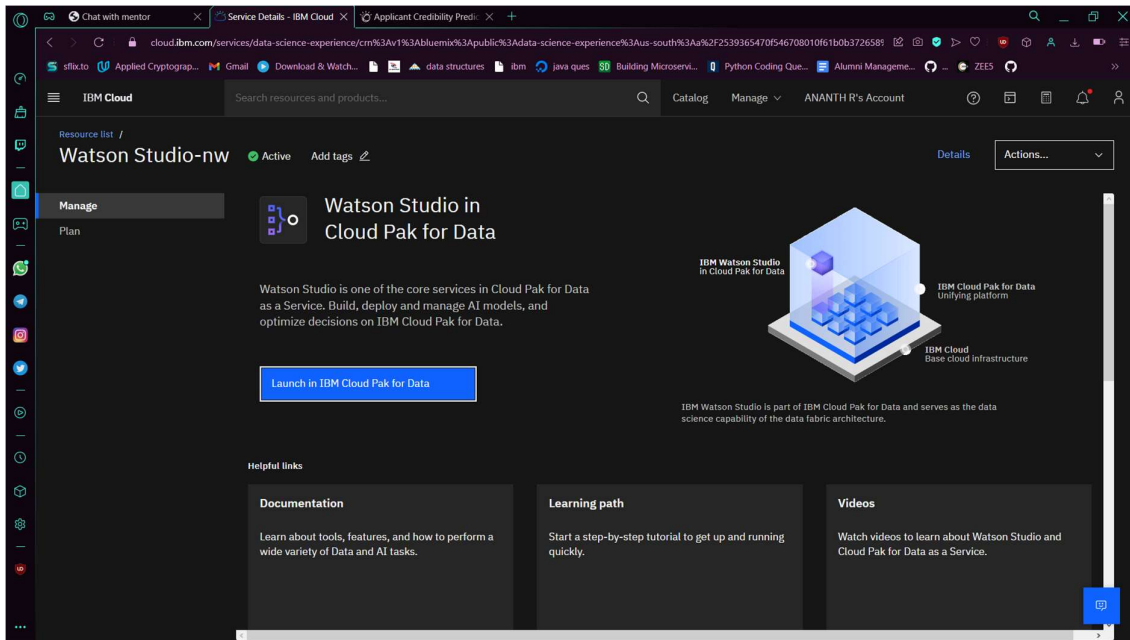


TRAIN THE MODEL ON IBM



Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Predi... | dataplatform.cloud.ibm.com/analyt...

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas | Python 3.9

Projects / Smart Lender - Applicant Credibi... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

```
endpoint_url = 'https://s3.private.us.cloud-object-storage.appdomain.cloud'

bucket = 'smartlenderapplicantcredibilitypr-donotdelete-pr-8epyljwre4kjc'
object_key = 'loan_prediction.csv'

body = cos_client.get_object(bucket=bucket, key=object_key)['body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType(__iter__, body)

df = pd.read_csv(body)
df.head()
```

Out[3]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	Urban	Y
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	Rural	N
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	Urban	Y
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0	Urban	Y
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0	Urban	Y

In [4]: df.head()

Out[4]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	Urban	Y
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	Rural	N
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	Urban	Y
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0	Urban	Y
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0	Urban	Y

Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Predi... | dataplatform.cloud.ibm.com/analyt...

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas | Python 3.9

Projects / Smart Lender - Applicant Credibi... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

```
In [5]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 614 entries, 0 to 613
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Loan_ID               614 non-null   object
1   Gender                601 non-null   object
2   Married               611 non-null   object
3   Dependents            599 non-null   object
4   Education             614 non-null   object
5   Self_Employed         582 non-null   object
6   ApplicantIncome       614 non-null   int64
7   CoapplicantIncome     614 non-null   float64
8   LoanAmount            592 non-null   float64
9   Loan_Amount_Term      600 non-null   float64
10  Credit_History         564 non-null   float64
11  Property_Area         614 non-null   object
12  Loan_Status           614 non-null   object
dtypes: float64(4), int64(1), object(8)
memory usage: 62.5+ KB
```

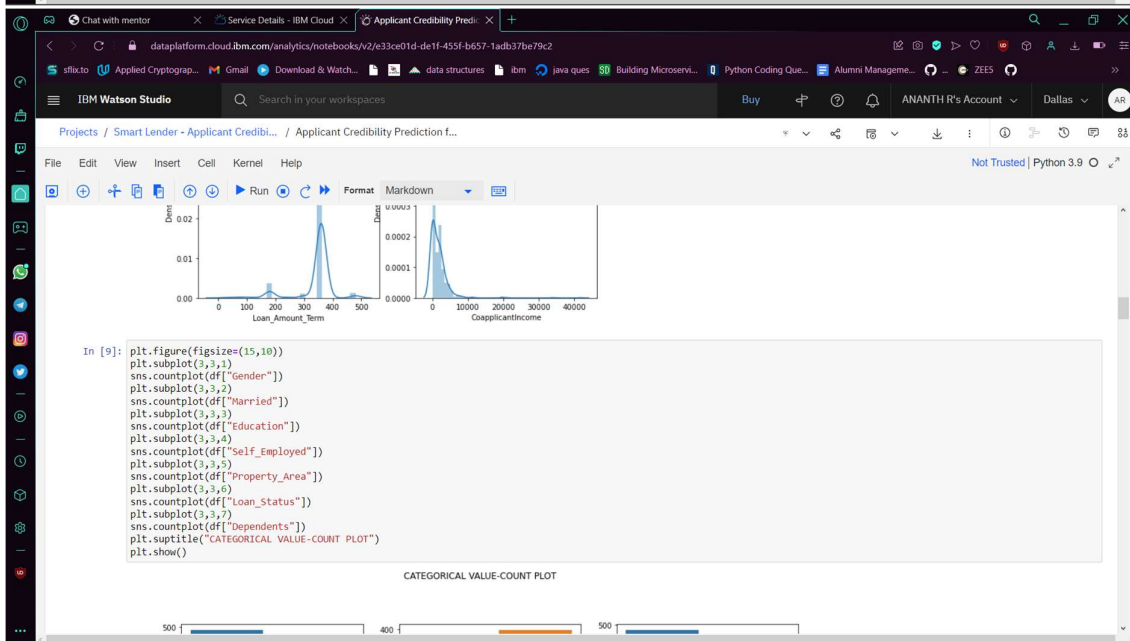
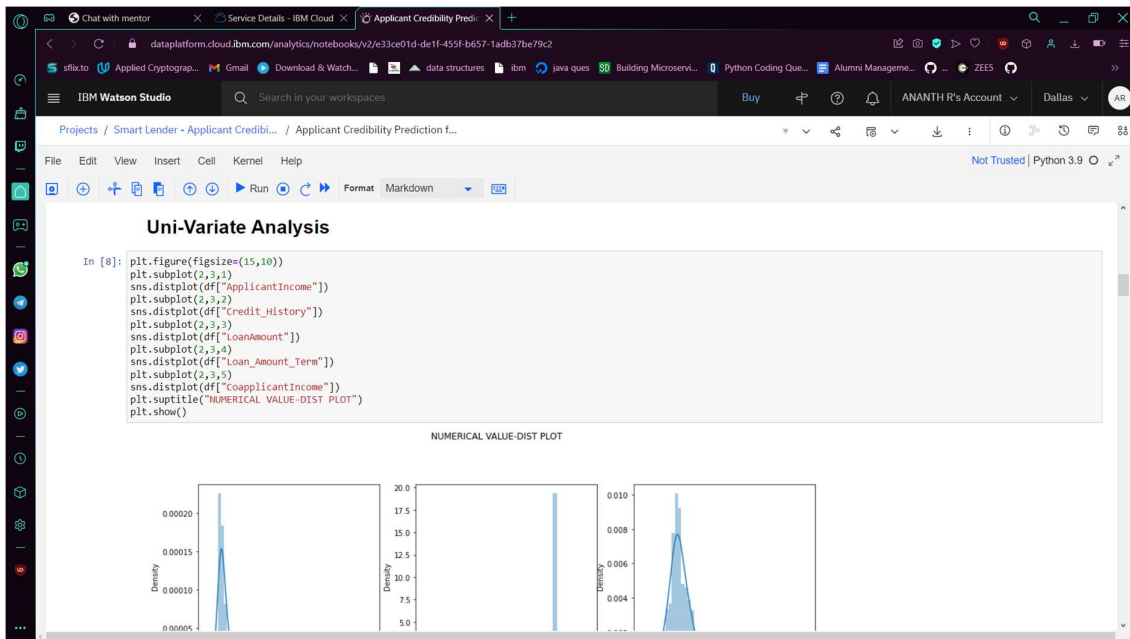
In [6]: df.shape

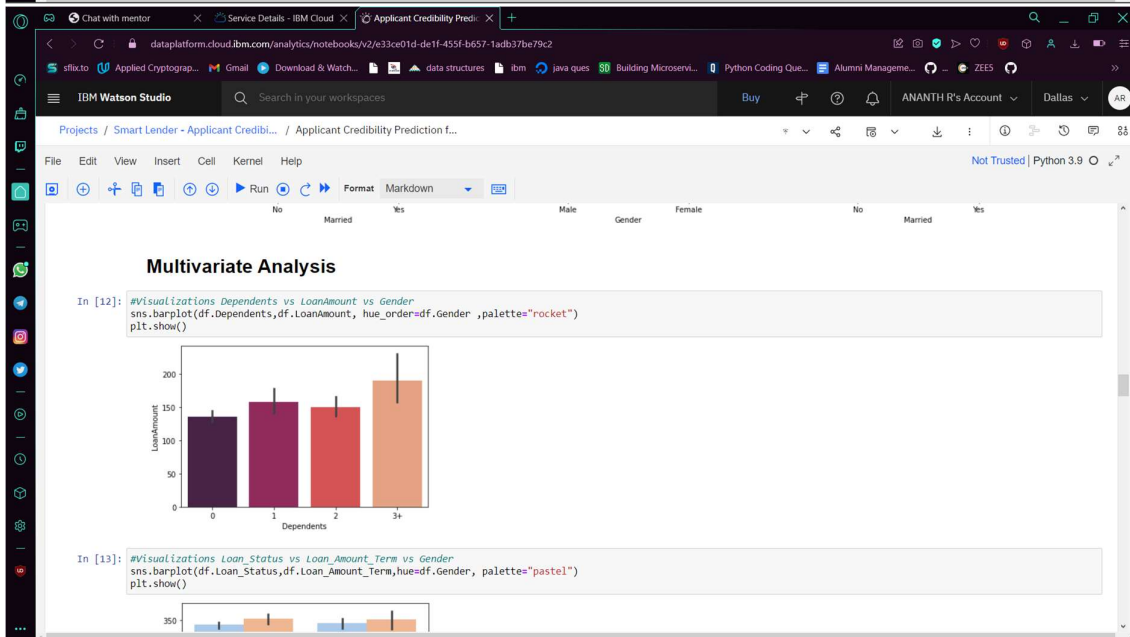
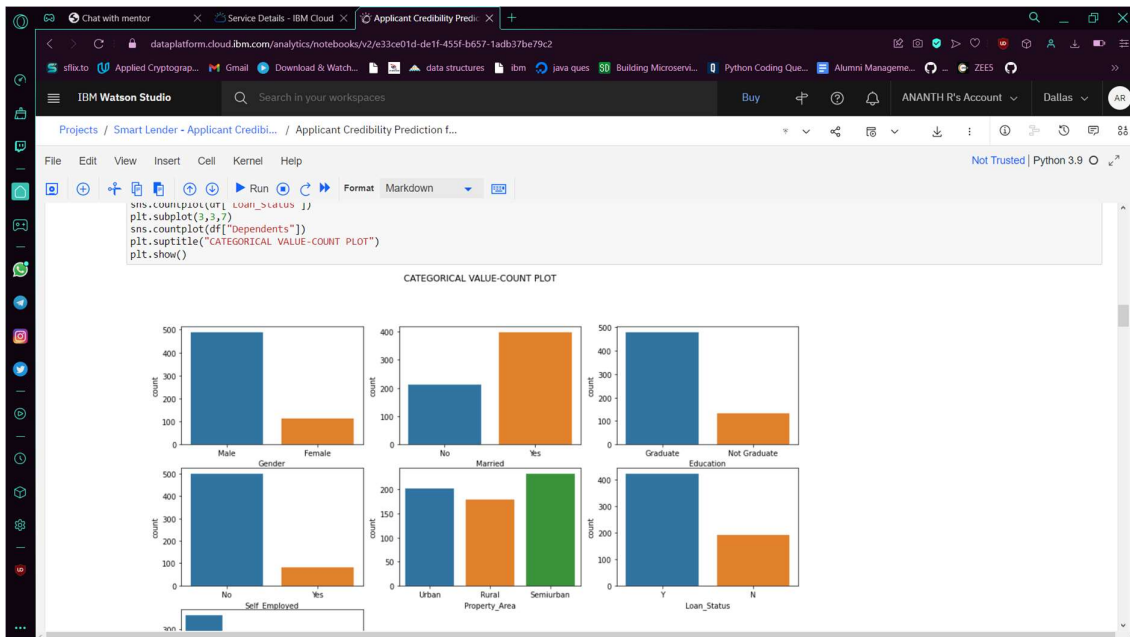
Out[6]: (614, 13)

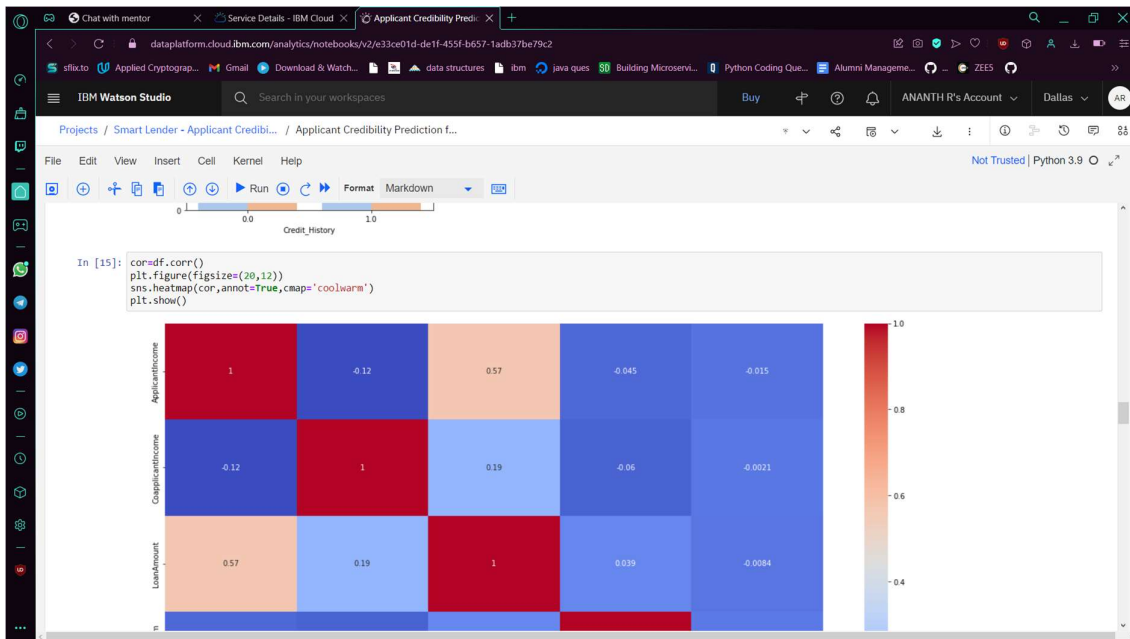
In [7]: df=df.drop(columns=["Loan_ID"],axis=1)

Uni-Variate Analysis

In [8]: plt.figure(figsize=(15,10))







IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas

Projects / Smart Lender - Applicant Crediti... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term Credit_History

Descriptive Analysis

In [16]:

```
df.describe()
```

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	614.000000	614.000000	592.000000	600.000000	564.000000
mean	5403.459283	1621.245798	146.412162	342.000000	0.842199
std	6109.041673	2926.248369	85.587325	65.120411	0.364878
min	150.000000	0.000000	9.000000	12.000000	0.000000
25%	2877.500000	0.000000	100.000000	360.000000	1.000000
50%	3812.500000	1188.500000	128.000000	360.000000	1.000000
75%	5795.000000	2297.250000	168.000000	360.000000	1.000000
max	81000.000000	41867.000000	700.000000	480.000000	1.000000

Data Pre-Processing

Checking For Null Values

In [17]:

```
df.isnull().sum()
```

Out[17]:

IBM Watson Studio interface showing a Jupyter Notebook titled "Applicant Credibility Prediction f...". The notebook displays a pandas DataFrame with the following structure:

```
Out[20]:
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
dtype:	int64	0	0	0	0	0	0	0	0	0	0	0

The notebook contains the following code cells:

```
In [18]: df['LoanAmount']=df['LoanAmount'].fillna(df['LoanAmount'].mean())
df['Loan_Amount_Term']=df['Loan_Amount_Term'].fillna(df['Loan_Amount_Term'].mean())
df['Credit_History']=df['Credit_History'].fillna(df['Credit_History'].mean())

In [19]: df['Gender']=df['Gender'].fillna(df['Gender'].mode()[0])
df['Married']=df['Married'].fillna(df['Married'].mode()[0])
df['Dependents']=df['Dependents'].fillna(df['Dependents'].mode()[0])
df['Self_Employed']=df['Self_Employed'].fillna(df['Self_Employed'].mode()[0])

In [20]: df.isnull().sum()
```

IBM Watson Studio interface showing a Jupyter Notebook titled "Applicant Credibility Prediction f...". The notebook displays a pandas DataFrame with the following structure:

```
Out[21]:
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	Male	No	0	Graduate	No	5849	0.0	146.412162	360.0	1.0	Urban	Y
1	Male	Yes	1	Graduate	No	4563	1508.0	128.000000	360.0	1.0	Rural	N
2	Male	Yes	0	Graduate	Yes	3000	0.0	66.000000	360.0	1.0	Urban	Y
3	Male	Yes	0	Not Graduate	No	2583	2358.0	120.000000	360.0	1.0	Urban	Y
4	Male	No	0	Graduate	No	6000	0.0	141.000000	360.0	1.0	Urban	Y

The notebook contains the following code cells:

```
In [22]: le=LabelEncoder()

In [23]: df.Gender=le.fit_transform(df.Gender)
df.Married=le.fit_transform(df.Married)
df.Education=le.fit_transform(df.Education)
df.Self_Employed=le.fit_transform(df.Self_Employed)
df.Property_Area=le.fit_transform(df.Property_Area)
df.Loan_Status=le.fit_transform(df.Loan_Status)
df.Dependents=le.fit_transform(df.Dependents)

In [24]: df.head()
```


Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Prediction f...

dataplatform.cloud.ibm.com/analystics/notebooks/v2/e33ce01d-de1f-455f-b657-1adb37be79c2

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas | Python 3.9

Projects / Smart Lender - Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

Splitting into dependent and independent data

```
In [25]: df.head()
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	1	0	0	0	0	5849	0.0	146.412162	360.0	1.0	2	1
1	1	1	1	0	0	4583	1508.0	128.000000	360.0	1.0	0	0
2	1	1	0	0	1	3000	0.0	66.000000	360.0	1.0	2	1
3	1	1	0	1	0	2583	2358.0	120.000000	360.0	1.0	2	1
4	1	0	0	0	0	6000	0.0	141.000000	360.0	1.0	2	1

```
In [26]: x=df.iloc[:, :-1]
y=df.Loan_Status
```

```
In [27]: x.head()
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area
0	1	0	0	0	0	5849	0.0	146.412162	360.0	1.0	2
1	1	1	1	0	0	4583	1508.0	128.000000	360.0	1.0	0
2	1	1	0	0	1	3000	0.0	66.000000	360.0	1.0	2
3	1	1	0	1	0	2583	2358.0	120.000000	360.0	1.0	2
4	1	0	0	0	0	6000	0.0	141.000000	360.0	1.0	2

```
In [28]: y.head()
```

Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Prediction f...

dataplatform.cloud.ibm.com/analystics/notebooks/v2/e33ce01d-de1f-455f-b657-1adb37be79c2

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas | Python 3.9

Projects / Smart Lender - Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

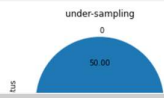
Scaling The Data

```
In [29]: x_scale=pd.DataFrame(scale(x),columns=x.columns)
x_scale.head()
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area
0	0.472343	-1.372089	-0.737806	-0.528362	-0.392601	0.072991	-0.554487	0.000000	0.279851	0.45164	1.223298
1	0.472343	0.728816	0.253470	-0.528362	-0.392601	-0.134412	-0.038732	-0.219273	0.279851	0.45164	-1.318513
2	0.472343	0.728816	-0.737806	-0.528362	2.547117	-0.393747	-0.554487	-0.967641	0.279851	0.45164	1.223298
3	0.472343	0.728816	-0.737806	1.892641	-0.392601	-0.482062	0.251980	-0.314547	0.279851	0.45164	1.223298
4	0.472343	-1.372089	-0.737806	-0.528362	-0.392601	0.097728	-0.554487	-0.064454	0.279851	0.45164	1.223298

Balancing The Dataset

```
In [30]: rus=RandomUnderSampler(sampling_strategy=1)
x_res,y_res=rus.fit_resample(x,y)
ax=y_res.value_counts().plot.pie(autopct='%2f')
_ =ax.set_title("under-sampling")
```



Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Predi... | dataplatform.cloud.ibm.com/analytics/notebooks/v2/633ce01d-de1f-455f-b657-1adb37be79c2

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas

Projects / Smart Lender - Applicant Credibi... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

Splitting Data Into Train and Test

```
In [31]: xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.3,random_state=10)
```

```
In [32]: xtrain.head()
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area
245	1	0	0	0	0	6050	4333.0	120.0	180.0	1.0	2
413	1	1	0	1	0	2253	2033.0	110.0	360.0	1.0	0
126	1	1	3	0	0	23803	0.0	370.0	360.0	1.0	0
531	1	1	3	0	0	4281	0.0	100.0	360.0	1.0	2
188	1	1	0	0	1	674	5296.0	168.0	360.0	1.0	0

```
In [33]: xtest.head()
```

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area
285	1	0	0	0	0	3158	3053.0	89.0	360.0	1.000000	0
323	0	0	0	0	0	3166	2985.0	132.0	360.0	0.842199	0
482	1	1	0	0	0	2083	3150.0	128.0	360.0	1.000000	1
173	1	1	0	0	0	5708	5625.0	187.0	360.0	1.000000	1
518	1	0	0	0	0	4883	1915.0	185.0	360.0	1.000000	1

```
In [34]: ytrain.head()
```

Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Predi... | dataplatform.cloud.ibm.com/analytics/notebooks/v2/633ce01d-de1f-455f-b657-1adb37be79c2

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas

Projects / Smart Lender - Applicant Credibi... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

```
Name: Loan_Status, dtype: int64
```

```
In [36]: !pip install ibm_watson_machine_learning
```

Requirement already satisfied: ibm_watson_machine_learning in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (1.0.257)
Requirement already satisfied: pandas<1.5.0,>=0.24.2 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (1.3.4)
Requirement already satisfied: urllib3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (1.26.7)
Requirement already satisfied: importlib-metadata in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (4.8.2)
Requirement already satisfied: lomond in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (0.3.3)
Requirement already satisfied: packaging in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (21.3)
Requirement already satisfied: tabulate in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (0.8.9)
Requirement already satisfied: ibm-cos-sdk==2.11.* in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2.11.0)
Requirement already satisfied: certifi in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2022.9.24)
Requirement already satisfied: requests in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2.26.0)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learning) (0.10.0)
Requirement already satisfied: ibm-cos-sdk-s3transfer==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learning) (2.11.0)
Requirement already satisfied: ibm-cos-sdk-core==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learning) (2.11.0)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk-core==2.11.0->ibm_watson_machine_learning) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from pandas<1.5.0,>=0.24.2->ibm_watson_machine_learning) (2021.3)
Requirement already satisfied: numpy>=1.17.3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from pandas<1.5.0,>=0.24.2->ibm_watson_machine_learning) (1.20.3)
Requirement already satisfied: six>=1.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from python-dateutil<3.0.0,>=2.1->ibm-cos-sdk-core==2.11.0->ibm_watson_machine_learning) (1.15.0)
Requirement already satisfied: charset-normalizer==2.0.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from requests->ibm_watson_machine_learning) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from requests->ibm_watson_machine_learning) (3.3)
Requirement already satisfied: zipp>=0.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from importlib-metadata->ibm_watson_machine_learning) (3.6.0)
Requirement already satisfied: pyparsing<3.0.5,>=2.0.2 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from packaging->ibm_watson_machine_learning) (3.0.4)

Chat with mentor | Service Details - IBM Cloud | Applicant Credibility Prediction f...

dataplatform.cloud.ibm.com/analystics/notebooks/v2/633ce01d-de1f-455f-b657-1adb37be79c2

Applied Cryptograp... | Gmail | Download & Watch... | data structures | ibm | java ques | Building Microservi... | Python Coding Que... | Alumni Manage... | ZEE5

IBM Watson Studio | Search in your workspaces | Buy | ANANTH R's Account | Dallas

Projects / Smart Lender - Applicant Credib... / Applicant Credibility Prediction f...

File Edit View Insert Cell Kernel Help | Not Trusted | Python 3.9

```
In [90]: space_uid = guid_from_space_name(client,'models')
print("Space UID = " + space_uid)

Space UID = f83ca0d5-ad11-4cd2-b365-8545fb02e434

In [91]: client.set.default_space(space_uid)

Out[91]: 'SUCCESS'

In [95]: client.software_specifications.list(limit=500)

pytorch-onnx_1.2-py3.6 10c80299-cc97-56da-b860-29c8e80000e7 base
kernel-spark3.3-r3.6 1c9e5454-f216-59dd-a20e-474a5cdf5988 base
pytorch-onnx_rt22.1-py3.9-edt 1d362186-7ad5-5b59-8b6c-9d0880bde37f base
tensorflow_2.1-py3.6 1eb25b84-d6ed-5dde-b6a5-3fbdff166566 base
spark-mllib_3.2 20047f72-0a98-58c7-9ff5-a77b012e08f5 base
tensorflow_2.4-py3.8-horovod 217c16f6-178f-56bf-824a-b19f2056dca9 base
runtime-22.1-py3.9-cuda 26215f05-08c3-5a41-a1b0-da66306ce658 base
do_py3.8 295addb5-9ef9-547e-9bf4-92ae3563e720 base
autoai-ts_3.8-py3.8 2aa0c932-798f-5ae9-ab06-15e0c2402f05 base
tensorflow_1.15-py3.6 2b73a275-7cbf-420b-a912-eae7f436e0bc base
kernel-spark3.3-py3.9 2b7961e2-e3b1-5a8c-a491-482c8368839a base
pytorch_1.2-py3.6 2c8ef57d-2687-4b7d-acce-01f94976dac1 base
spark-mllib_2.3 2e51f700-bca0-4b0d-88dc-5c6791338875 base
pytorch-onnx_1.1-py3.6-edt 32983cea-3f32-4400-8965-d0e87a86d67e base
spark-mllib_3.0-py37 36507ebe-0770-550a-ab2a-eafe787600e9 base
spark-mllib_2.4 390d21f8-e58b-4fac-9c55-d7ceda621326 base
autoai-ts_rt22.2-py3.10 396b2e83-0953-5b66-9a55-7ce1628a406f base
xgboost_0.82-py3.6 39e31acd-5f30-41dc-ae44-60233c80306e base
pytorch-onnx_1.2-py3.6-edt 4058900e-7019-4e28-8daa-fb03b6f4fe12 base
pytorch-onnx_rt22.2-py3.10 40e73f55-783a-5535-b3fa-0c8094291431 base

In [96]: model_props = {
client.software_specifications.ModelMetadata.NAME: "loan_prediction"
```