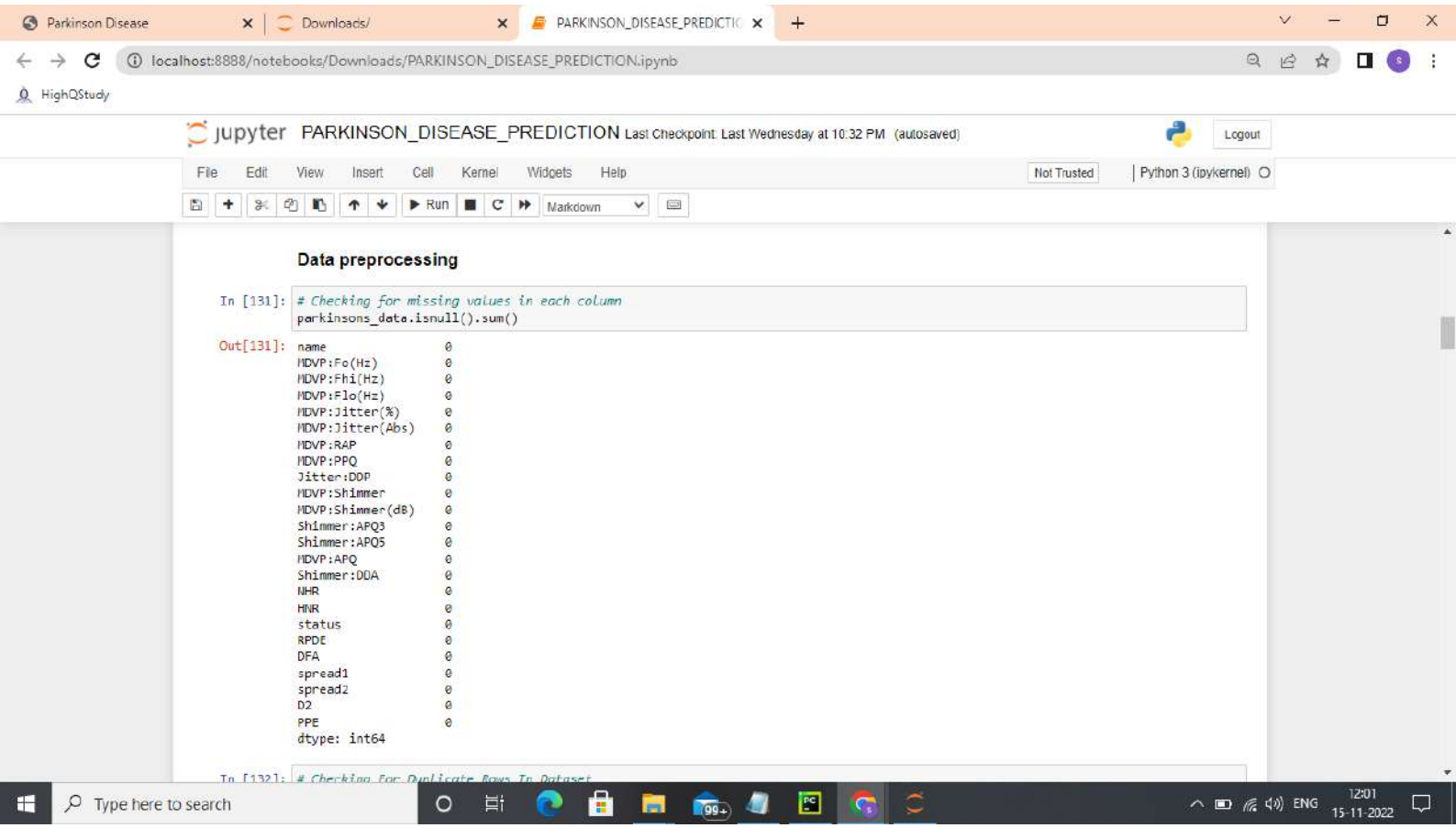

Parkinson Disease
Downloads/
PARKINSON_DISEASE_PREDICTIC
localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb
HighQStudy
jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Data preprocessing
In [131]:
# Checking for missing values in each column
parkinsons_data.isnull().sum()
Out[131]:
name 0
MDVP:Fo(Hz) 0
MDVP:Fhi(Hz) 0
MDVP:Flo(Hz) 0
MDVP:Jitter(%) 0
MDVP:Jitter(Abs) 0
MDVP:RAP 0
MDVP:PPQ 0
Jitter:DDP 0
MDVP:Shimmer 0
MDVP:Shimmer(dB) 0
Shimmer:APQ3 0
Shimmer:APQ5 0
MDVP:APQ 0
Shimmer:DDA 0
NHR 0
HNR 0
status 0
RPDE 0
DFA 0
spread1 0
spread2 0
D2 0
PPE 0
dtype: int64
In [132]:
Checking For Duplicate Rows In Dataset
Type here to search
ENG
12:01
15-11-2022

Parkinson Disease | Downloads/ | PARKINSON_DISEASE_PREDICTI | localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb

HighQStudy

jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

Run Markdown

In [132]:
```
# Checking For Duplicate Rows In Dataset

print('duplicated rows are',parkinsons_data.duplicated().sum())
```

```
duplicated rows are 0
```

In [133]:
```
# Getting some statistical measures about the data
parkinsons_data.describe()
```

Out[133]:

	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP	MDVP:Shimmer	MDVP:Shimmer(dB)	...
count	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	...
mean	154.228641	197.104918	116.324631	0.006220	0.000044	0.003306	0.003446	0.009920	0.029709	0.282251	...
std	41.390065	91.491548	43.521413	0.004848	0.000035	0.002968	0.002759	0.008903	0.018857	0.194877	...
min	88.333000	102.145000	65.476000	0.001680	0.000007	0.000680	0.000920	0.002040	0.009540	0.085000	...
25%	117.572000	134.862500	84.291000	0.003460	0.000020	0.001660	0.001860	0.004985	0.016505	0.148500	...
50%	148.790000	175.829000	104.315000	0.004940	0.000030	0.002500	0.002690	0.007490	0.022970	0.221000	...
75%	182.769000	224.205500	140.018500	0.007365	0.000060	0.003835	0.003955	0.011505	0.037885	0.350000	...
max	260.105000	592.030000	239.170000	0.033160	0.000260	0.021440	0.019580	0.064330	0.119080	1.302000	...

8 rows × 23 columns

In [134]:
```
# Distribution of target Variable
parkinsons_data['status'].value_counts()
```

Out[134]:
```
1    147
0     48
Name: status, dtype: int64
```

Type here to search ENG 12:01 15-11-2022

Parkinson Disease | Downloads/ | PARKINSON_DISEASE_PREDICTI | +

localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb

HighQStudy

jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

Run Markdown

In [134]:
```
# Distribution of target Variable
parkinsons_data['status'].value_counts()
```

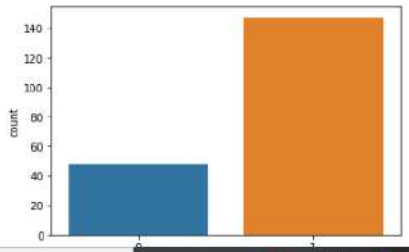
Out[134]:
```
1    147
0     48
Name: status, dtype: int64
```

In [135]:
```
# Dropping The Name Column
parkinsons_data.drop(['name'], axis=1, inplace=True)
```

Exploratry Data Analysis

In [136]:
```
#Balance of Data
sns.countplot(x='status',data=parkinsons_data)
```

Out[136]: <AxesSubplot:xlabel='status', ylabel='count'>



Type here to search ENG 12:01 15-11-2022

Parkinson Disease | Downloads/ | PARKINSON_DISEASE_PREDICTI

localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb

HighQStudy

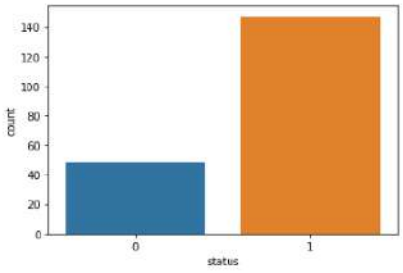
jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

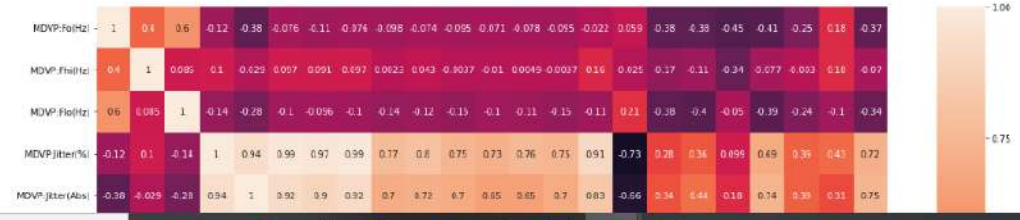
Not Trusted | Python 3 (ipykernel)

Run Markdown



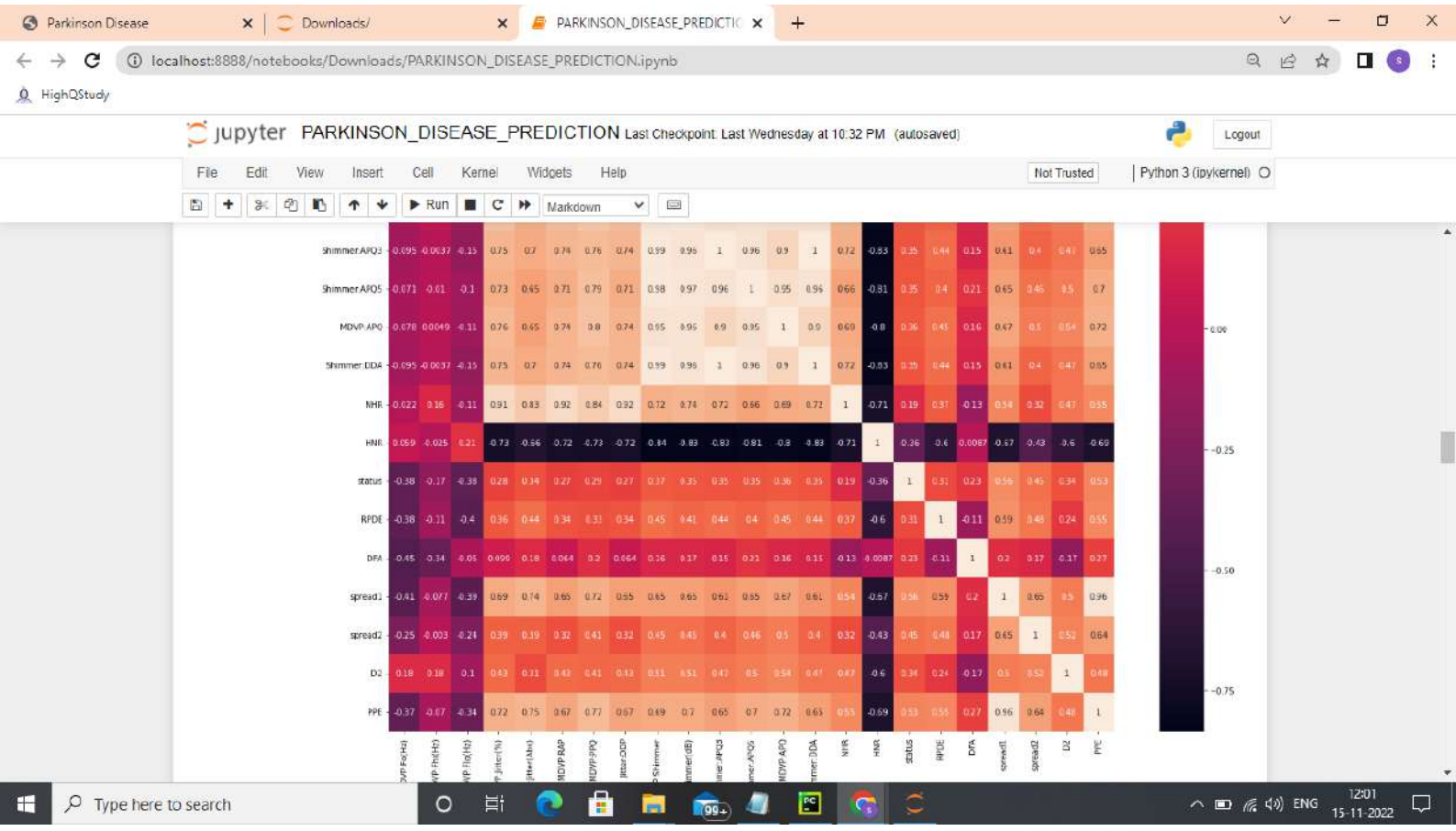
In [137]:
```
fig, ax = plt.subplots(figsize=(20,20))
sns.heatmap(parkinsons_data.corr(),annot=True,ax=ax)
```

Out[137]: <AxesSubplot:>



Type here to search

ENG 12:01 15-11-2022


Parkinson Disease
Downloads/
PARKINSON_DISEASE_PREDICTI
localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb
HighQStudy
jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Shimmer:APQ3
Shimmer:APQ5
MDVP:APQ
Shimmer:DDA
NHR
HNR
status
RPDE
DFA
spread1
spread2
D2
PPE
0.09
0.00
−0.25
−0.50
−0.75
Type here to search
ENG
12:01
15-11-2022

Parkinson Disease | Downloads/ | PARKINSON_DISEASE_PREDICTI | localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb

HighQStudy

jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted | Python 3 (ipykernel)

Run Markdown

Balancing Dataset

In this section, as it is observed that the Dataset Is Heavily Imbalanced, with Number of Samples of Parkinson Disease Samples being 147, and Non-Parkinson Being only 48. Hence, in this section, we make use of SMOTE to Oversample and Balance the dataset

In [138]:
```
# Exploring Imabalance In Dataset

parkinsons_data['status'].value_counts()
```

Out[138]:
```
1    147
0     48
Name: status, dtype: int64
```

In [139]:
```
# Extracting Features Into Features & Target


x = parkinsons_data.drop(['status'], axis=1)
y = parkinsons_data['status']

print('Feature (x) Shape Before Balancing :', x.shape)
print('Target (y) Shape Before Balancing :', y.shape)
```

```
Feature (x) Shape Before Balancing : (195, 22)
Target (y) Shape Before Balancing : (195,)
```

In [140]:
```
sm = SMOTE(random_state=300)

# Resampling Data

x, y = sm.fit_resample(x, y)
```

In [141]:
```
print('Feature (x) Shape After Balancing :', x.shape)
```

Type here to search ENG 12:01 15-11-2022

Parkinson Disease | Downloads/ | PARKINSON_DISEASE_PREDICTI | localhost:8888/notebooks/Downloads/PARKINSON_DISEASE_PREDICTION.ipynb

HighQStudy

jupyter PARKINSON_DISEASE_PREDICTION Last Checkpoint: Last Wednesday at 10:32 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

Run Markdown

In [141]:
```
print('Feature (x) Shape After Balancing :', x.shape)
print('Target (y) Shape After Balancing :', y.shape)
```

```
Feature (x) Shape After Balancing : (294, 22)
Target (y) Shape After Balancing : (294,)
```

In [142]:
```
# Scaling features between -1 and 1  for normalization

from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler((-1,1))

x = scaler.fit_transform(x)
Y = y
```

In [143]:
```
# splitting the dataset into traning and testing
x_train , x_test , y_train , y_test = train_test_split(x, Y , test_size=0.20, random_state=20)
```

In [144]:
```
x_train.shape
```

Out[144]: (235, 22)

In [145]:
```
x_test.shape
```

Out[145]: (59, 22)

In [146]:
```
y_train.shape
```

Out[145]: (235,)

In [147]:
```
y_test.shape
```

Type here to search ENG 12:01 15-11-2022