


```
import pandas

real = pandas.read_csv("/content/real_feature_values.csv")

model = pandas.read_csv("/content/model_feature_values.csv")

generated = pandas.read_csv("/content/generated_feature_values.csv")

real
```



Unnamed: 0				Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Hi
0	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
1	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
2	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
3	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
4	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
...	...			...		...
195	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
196	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
197	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
198	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	
199	D:\UU\Term 6\Sound and Music Technology\Projec...			0.0	0.0	

200 rows × 1496 columns

model



	Unnamed: 0	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_His
0	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
1	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
2	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
3	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
4	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
...	...	...	...	
95	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
96	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
97	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
98	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	
99	D:\UU\Term 6\Sound and Music Technology\Projec...	0.0	0.0	

100 rows × 1496 columns

```
real['Category']="1"
```

```
model['Category']="0"
```

```
generated['Category']="0"
```

```
real = real.drop(columns="Unnamed: 0")
```

```
model = model.drop(columns="Unnamed: 0")

generated = generated.drop(columns="Unnamed: 0")
```

real

↳

	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_P
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	
...	...	...	...	
195	0.0	0.0	0.0	
196	0.0	0.0	0.0	
197	0.0	0.0	0.0	
198	0.0	0.0	0.0	
199	0.0	0.0	0.0	

200 rows × 1496 columns

model



	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_Pi
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	
...	...	...	...	
95	0.0	0.0	0.0	
96	0.0	0.0	0.0	
97	0.0	0.0	0.0	
98	0.0	0.0	0.0	
99	0.0	0.0	0.0	

100 rows × 1496 columns

generated



	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_Pi
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	
...	...	...	...	
95	0.0	0.0	0.0	
96	0.0	0.0	0.0	
97	0.0	0.0	0.0	
98	0.0	0.0	0.0	
99	0.0	0.0	0.0	

100 rows × 1496 columns

```
combined_dataset = pandas.concat([generated, model],sort=False)
```

combined\_dataset

↗

	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_Pi
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	
...	...	...	...	
95	0.0	0.0	0.0	
96	0.0	0.0	0.0	
97	0.0	0.0	0.0	
98	0.0	0.0	0.0	
99	0.0	0.0	0.0	

200 rows × 1496 columns

```
combined_dataset = combined_dataset.sample(frac=1)
```

combined\_dataset



	Number_of_Pitches	Number_of_Pitch_Classes	Number_of_Common_Pitches	Number_of_Common_Pitches
<b>79</b>	8.0	6.0	4.0	
<b>34</b>	20.0	11.0	5.0	
<b>56</b>	12.0	7.0	4.0	
<b>76</b>	7.0	6.0	4.0	
<b>17</b>	10.0	7.0	6.0	
...	...	...	...	
<b>34</b>	7.0	6.0	6.0	
<b>5</b>	11.0	7.0	6.0	
<b>90</b>	13.0	7.0	5.0	
<b>89</b>	33.0	11.0	0.0	
<b>36</b>	10.0	7.0	4.0	

200 rows × 227 columns

```
combined_feature_dataset = combined_dataset.to_csv("combined_generated.csv", index=False, encoding='utf-8')
```

```
preprocessed_model_feature_values = model.to_csv("preprocessed_model_feature_values.csv", index=False, encoding='utf-8')
```

```
preprocessed_real_feature_values = real.to_csv("preprocessed_real_feature_values.csv", index=False, encoding='utf-8')
```

```
preprocessed_model_feature_values = model.to_csv("preprocessed_model_feature_values.csv", index=False, encoding='utf-8')
```

```
preprocessed_generated_feature_values = generated.to_csv("preprocessed_generated_feature_values.csv", index=False, encoding='utf-8')
```

```
combined_all = pandas.concat([generated, model, real], sort=False)
```

```
combined_all
```



	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_P
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	
...	...	...	...	
195	0.0	0.0	0.0	
196	0.0	0.0	0.0	
197	0.0	0.0	0.0	
198	0.0	0.0	0.0	
199	0.0	0.0	0.0	

400 rows × 1496 columns

```
combined_all = combined_all.sample(frac=1)

combined_feature_dataset = combined_all.to_csv("combined_all.csv", index=False, encoding='utf

data = pandas.read_csv("/content/preprocessed_combined_all.csv")

data
```



	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_P
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	
...	...	...	...	
395	0.0	0.0	0.0	
396	0.0	0.0	0.0	
397	0.0	0.0	0.0	
398	0.0	0.0	0.0	
399	0.0	0.0	0.0	

400 rows × 1496 columns

data.isnull()



	Basic_Pitch_Histogram_0	Basic_Pitch_Histogram_1	Basic_Pitch_Histogram_2	Basic_P
0	False	False	False	
1	False	False	False	
2	False	False	False	
3	False	False	False	
4	False	False	False	
...	...	...	...	
395	False	False	False	
396	False	False	False	
397	False	False	False	
398	False	False	False	
399	False	False	False	

400 rows × 1496 columns

data = data.loc[:, (data != data.iloc[0]).any()]

data

	Basic_Pitch_Histogram_48	Basic_Pitch_Histogram_49	Basic_Pitch_Histogram_50	Basi
<b>0</b>	0.0	0.0	0.0	0.0
<b>1</b>	0.0	0.0	0.0	0.0
<b>2</b>	0.0	0.0	0.0	0.0
<b>3</b>	0.0	0.0	0.0	0.0
<b>4</b>	0.0	0.0	0.0	0.0
...	...	...	...	...
<b>395</b>	0.0	0.0	0.0	0.0
<b>396</b>	0.0	0.0	0.0	0.0
<b>397</b>	0.0	0.0	0.0	0.0
<b>398</b>	0.0	0.0	0.0	0.0
<b>399</b>	0.0	0.0	0.0	0.0

400 rows × 684 columns

```

from sklearn import preprocessing
min_max_scaler = preprocessing.MinMaxScaler()
scaled_array = min_max_scaler.fit_transform(data)

```

scaled\_array

```

array([[0.      , 0.      , 0.      , ..., 0.7676 , 0.54117063,
        0.      ],
       [0.      , 0.      , 0.      , ..., 0.8544 , 0.84672619,
        1.      ],
       [0.      , 0.      , 0.      , ..., 0.8456 , 0.56349206,
        1.      ],
       ...,
       [0.      , 0.      , 0.      , ..., 0.7792 , 0.5952381 ,
        1.      ],
       [0.      , 0.      , 0.      , ..., 0.72848 , 0.49107143,
        1.      ],
       [0.      , 0.      , 0.      , ..., 0.79416 , 0.66170635,
        0.      ]])

```

```

data_scaled = pandas.DataFrame(scaled_array)
data_scaled

```

```

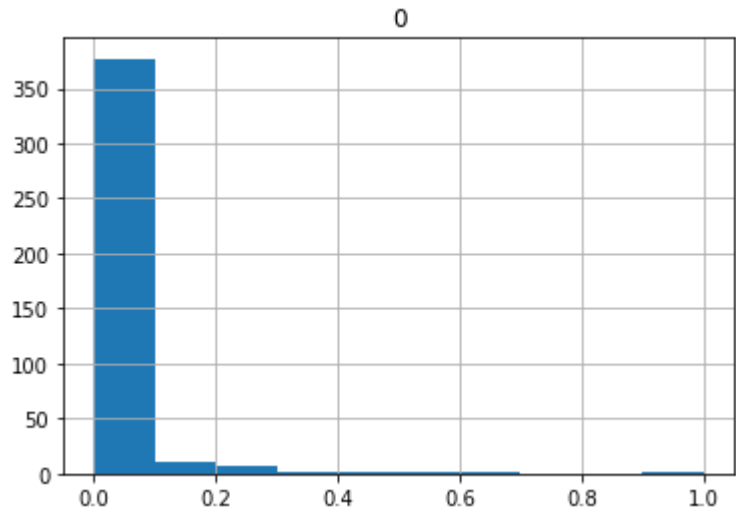

```

	0	1	2	3	4	5	6	7	8	9	10	11
0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.0	0.0	0.0	0.0	0.0	0.502993	0.0	0.600000	0.000000	0.000000	0.000000	0.468393
2	0.0	0.0	0.0	0.0	0.0	0.123151	0.0	0.000000	0.000000	0.048576	0.000000	0.641967
3	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000
4	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000
...	...	...	...	...	...	...	...	...	...	...	...	...
395	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.157243	0.000000	0.381983	0.000000
396	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.064363	0.149049
397	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000
398	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.377953	0.000000	1.000000	0.000000	0.839344
399	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000

400 rows × 684 columns

```
data_scaled.hist([0])
```

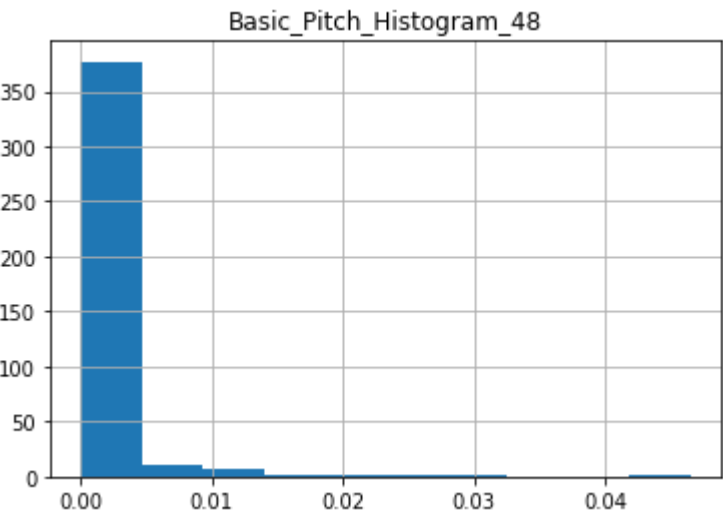
```
↳ array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f74e6b1fe80>]],
      dtype=object)
```



```
data.hist('Basic_Pitch_Histogram_48')
```

```
↳
```

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f74e65356d8>]],
      dtype=object)
```



```
columns = list(data.columns)
```

```
data_scaled.columns = columns
```

data

	Basic_Pitch_Histogram_48	Basic_Pitch_Histogram_49	Basic_Pitch_Histogram_50	Basi
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	
...	...	...	...	
395	0.0	0.0	0.0	
396	0.0	0.0	0.0	
397	0.0	0.0	0.0	
398	0.0	0.0	0.0	
399	0.0	0.0	0.0	

400 rows × 684 columns

data\_scaled



	Basic_Pitch_Histogram_48	Basic_Pitch_Histogram_49	Basic_Pitch_Histogram_50	Basi
<b>0</b>	0.0	0.0	0.0	
<b>1</b>	0.0	0.0	0.0	
<b>2</b>	0.0	0.0	0.0	
<b>3</b>	0.0	0.0	0.0	
<b>4</b>	0.0	0.0	0.0	
...	...	...	...	...
<b>395</b>	0.0	0.0	0.0	
<b>396</b>	0.0	0.0	0.0	
<b>397</b>	0.0	0.0	0.0	
<b>398</b>	0.0	0.0	0.0	
<b>399</b>	0.0	0.0	0.0	

400 rows × 684 columns

```
data.to_csv("combined_all_final.csv", index=False, encoding='utf8')
```

```
data_scaled.to_csv("combined_all_final_scaled.csv", index=False, encoding='utf8')
```

