

## process

November 25, 2024

```
[1]: from astropy.io import fits
import numpy as np
import pandas as pd
fits_file = fits.open('atlas-cdfs.fits')
print(fits_file)
```

[<astropy.io.fits.hdu.image.PrimaryHDU object at 0x7cd77ef95780>]

WARNING: VerifyWarning: Invalid 'BLANK' keyword in header. The 'BLANK' keyword is only applicable to integer data, and will be ignored in this HDU.

[astropy.io.fits.hdu.image]

```
[2]: header = fits_file[0].header
print(header)
```

```
SIMPLE = T /
BITPIX = -32 /
NAXIS = 4 /
NAXIS1 = 3937 /
NAXIS2 = 3057 /
NAXIS3 = 1 /
NAXIS4 = 1 /
EXTEND = T /
BSCALE = 1.000000000000E+00 /
BZERO = 0.000000000000E+00 /
BLANK = -1 /
TELESCOP= 'ATCA' /
CDELTA1 = -5.55555561268E-04 /
CRPIX1 = 2.924000000000E+03 /
CRVAL1 = 5.21524152154E+01 /
CTYPE1 = 'RA---NCP' /
CDELTA2 = 5.55555561268E-04 /
CRPIX2 = 4.890000000000E+02 /
CRVAL2 = -2.86959714237E+01 /
CTYPE2 = 'DEC--NCP' /
CDELTA3 = 8.000000000000E+06 /
CRPIX3 = 1.000000000000E+00 /
CRVAL3 = 1.36736571789E+09 /
CTYPE3 = 'FREQ-LSR' /
```



Processing file /DATA/ATHENA_7/c1241/data/2004-01-07_0658.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 37320	HISTORY ATLOD:
Summary of spectra flagged	HISTORY ATLOD:
Flagging Reason                      Fraction	HISTORY ATLOD:
-----	HISTORY ATLOD:
Antenna off-source/off-line    4.18%	HISTORY ATLOD:
Bad SYSCAL values                1.48%	HISTORY ATLOD:
-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-07_0752.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 76560	HISTORY ATLOD:
Summary of spectra flagged	HISTORY ATLOD:
Flagging Reason                      Fraction	HISTORY ATLOD:
-----	HISTORY ATLOD:
Antenna off-source/off-line    7.11%	HISTORY ATLOD:
Bad SYSCAL values                0.13%	HISTORY ATLOD:
-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-07_0943.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 76680	HISTORY ATLOD:
Summary of spectra flagged	HISTORY ATLOD:
Flagging Reason                      Fraction	HISTORY ATLOD:
-----	HISTORY ATLOD:
Antenna off-source/off-line    18.8%	HISTORY ATLOD:
Bad SYSCAL values                0.08%	HISTORY ATLOD:
-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-07_1133.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 39840	HISTORY ATLOD:
Summary of spectra flagged	HISTORY ATLOD:
Flagging Reason                      Fraction	HISTORY ATLOD:
-----	HISTORY ATLOD:
Antenna off-source/off-line    5.42%	HISTORY ATLOD:
Bad SYSCAL values                0.10%	HISTORY ATLOD:
-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-07_1234.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 147480	HISTORY ATLOD:
Summary of spectra flagged	HISTORY ATLOD:
Flagging Reason                      Fraction	HISTORY ATLOD:
-----	HISTORY ATLOD:
Antenna off-source/off-line    2.28%	HISTORY ATLOD:
Bad SYSCAL values                0.40%	HISTORY ATLOD:
-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-08_0434.c124	HISTORY ATLOD:
RPFITS file version is unknown	HISTORY ATLOD:
Total number of spectra selected: 117840	HISTORY ATLOD:

Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Antenna off-source/off-line	3.05%	HISTORY ATLOD:
Bad SYSCAL values	0.10%	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-08_0722.c124		HISTORY ATLOD:
RPFITS file version is unknown		HISTORY ATLOD:
Total number of spectra selected: 105240		HISTORY ATLOD:
Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Antenna off-source/off-line	7.07%	HISTORY ATLOD:
Bad SYSCAL values	0.06%	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-08_0957.c124		HISTORY ATLOD:
RPFITS file version is unknown		HISTORY ATLOD:
Total number of spectra selected: 105000		HISTORY ATLOD:
Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Antenna off-source/off-line	16.2%	HISTORY ATLOD:
Bad SYSCAL values	0.10%	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-08_1227.c124		HISTORY ATLOD:
RPFITS file version is unknown		HISTORY ATLOD:
Total number of spectra selected: 164880		HISTORY ATLOD:
Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Antenna off-source/off-line	2.69%	HISTORY ATLOD:
Bad SYSCAL values	0.21%	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-12_0404.c124		HISTORY ATLOD:
RPFITS file version is unknown		HISTORY ATLOD:
Total number of spectra selected: 110160		HISTORY ATLOD:
Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Antenna off-source/off-line	6.50%	HISTORY ATLOD:
Bad SYSCAL values	0.28%	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-12_0642.c124		HISTORY ATLOD:
RPFITS file version is unknown		HISTORY ATLOD:
Total number of spectra selected: 102840		HISTORY ATLOD:
Summary of spectra flagged		HISTORY ATLOD:
Flagging Reason	Fraction	HISTORY ATLOD:
-----	-----	HISTORY ATLOD:

Antenna off-source/off-line	6.81%			HISTORY ATLOD:
Bad SYSCAL values	0.40%			HISTORY ATLOD:
-----				HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-12_0908.c124				HISTORY ATLOD:
RPFITS file version is unknown				HISTORY ATLOD:
Total number of spectra selected: 141480				HISTORY ATLOD:
Summary of spectra flagged				HISTORY ATLOD:
Flagging Reason		Fraction		HISTORY ATLOD:
-----				HISTORY ATLOD:
Antenna off-source/off-line	15.8%			HISTORY ATLOD:
Bad SYSCAL values	0.10%			HISTORY ATLOD:
-----				HISTORY ATLOD:
Processing file /DATA/ATHENA_7/c1241/data/2004-01-12_1229.c124				HISTORY ATLOD:
RPFITS file version is unknown				HISTORY ATLOD:
Total number of spectra selected: 139680				HISTORY ATLOD:
Summary of spectra flagged				HISTORY ATLOD:
Flagging Reason		Fraction		HISTORY ATLOD:
-----				HISTORY ATLOD:
Antenna off-source/off-line	6.67%			HISTORY ATLOD:
Bad SYSCAL values	0.54%			HISTORY ATLOD:
-----				HISTORY UVFLAG:
Executed on: 05JAN04:12:09:06.0				HISTORY UVFLAG:
Command line inputs follow:				HISTORY UVFLAG:
vis=cdfs-jan				HISTORY UVFLAG:
flagval=flag				HISTORY UVFLAG:
select=window(2)				HISTORY UVFLAG:
edge=2				HISTORY UVFLAG:
Overview of flagging on visibility file cdfs-jan				HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)				HISTORY UVFLAG:
Total number of records selected: 646804; out of 646804 records				HISTORY UVFLAG:
Antennas used: 1,2,3,4,5,6				HISTORY UVFLAG:
Counts of correlations within selected channels				HISTORY UVFLAG:
channel	Originally	Currently		HISTORY UVFLAG:
Good:	7745088	5163392	Changed to bad: 2581696	HISTORY UVFLAG:
Bad:	663364	3245060		HISTORY UVFLAG:
Executed on: 05JAN04:12:11:40.0				HISTORY UVFLAG:
Command line inputs follow:				HISTORY UVFLAG:
vis=cdfs-jan				HISTORY UVFLAG:
flagval=flag				HISTORY UVFLAG:
select=window(2)				HISTORY UVFLAG:
line=channel,2,7,1				HISTORY UVFLAG:
Overview of flagging on visibility file cdfs-jan				HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)				HISTORY UVFLAG:
Total number of records selected: 646804; out of 646804 records				HISTORY UVFLAG:
Antennas used: 1,2,3,4,5,6				HISTORY UVFLAG:
Counts of correlations within selected channels				HISTORY UVFLAG:
channel	Originally	Currently		HISTORY UVFLAG:
Good:	1290848	0	Changed to bad: 1290848	HISTORY UVFLAG:

Bad:	2760	1293608			HISTORY UVFLAG:
Executed on:	05JAN04:12:17:07.0				HISTORY UVFLAG:
Command line inputs follow:					HISTORY UVFLAG:
vis=cdfs-jan					HISTORY UVFLAG:
flagval=flag					HISTORY UVFLAG:
select=@janflags					HISTORY UVFLAG:
Overview of flagging on visibility file cdfs-jan					HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)					HISTORY UVFLAG:
Total number of records selected:	960; out of 646804 records				HISTORY UVFLAG:
Antennas used:	1,2,3,4,5,6				HISTORY UVFLAG:
Counts of correlations within selected channels					HISTORY UVFLAG:
channel	Originally	Currently			HISTORY UVFLAG:
Good:	18240	0	Changed to bad:	18240	HISTORY UVFLAG:
Bad:	6720	24960			HISTORY UVFLAG:
wide	Originally	Currently			HISTORY UVFLAG:
Good:	0	0	Changed to bad:	0	HISTORY UVFLAG:
Bad:	0	0			HISTORY UVSPLIT:
Miriad UvSplit: version 1.0 19-Sep-04					HISTORY UVSPLIT:
Executed on:	05JAN04:12:22:24.0				HISTORY UVSPLIT:
Command line inputs follow:					HISTORY UVSPLIT:
options=mosaic					HISTORY UVFLAG:
vis=cdfs-jan					HISTORY UVFLAG:
Executed on:	05JAN04:12:30:21.0				HISTORY UVFLAG:
Command line inputs follow:					HISTORY UVFLAG:
vis=cdf.1344jan					HISTORY UVFLAG:
flagval=flag					HISTORY UVFLAG:
select=@jancdfflags					HISTORY UVFLAG:
Overview of flagging on visibility file cdf.1344jan					HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)					HISTORY UVFLAG:
Total number of records selected:	7288; out of 527612 records				HISTORY UVFLAG:
Antennas used:	1,2,3,4,5,6				HISTORY UVFLAG:
Counts of correlations within selected channels					HISTORY UVFLAG:
channel	Originally	Currently			HISTORY UVFLAG:
Good:	94536	0	Changed to bad:	94536	HISTORY UVFLAG:
Bad:	208	94744			HISTORY UVFLAG:
wide	Originally	Currently			HISTORY UVFLAG:
Good:	0	0	Changed to bad:	0	HISTORY UVFLAG:
Bad:	0	0			HISTORY UVFLAG:
Executed on:	05JAN04:12:56:03.0				HISTORY UVFLAG:
Command line inputs follow:					HISTORY UVFLAG:
vis=cdf.1344jan					HISTORY UVFLAG:
flagval=flag					HISTORY UVFLAG:
select=@cdf1.flag01					HISTORY UVFLAG:
Overview of flagging on visibility file cdf.1344jan					HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)					HISTORY UVFLAG:
Total number of records selected:	88; out of 527612 records				HISTORY UVFLAG:
Antennas used:	4,5				HISTORY UVFLAG:
Counts of correlations within selected channels					HISTORY UVFLAG:

channel	Originally	Currently			HISTORY UVFLAG:
Good:	1144	0	Changed to bad:	1144	HISTORY UVFLAG:
Bad:	0	1144			HISTORY UVFLAG:
wide	Originally	Currently			HISTORY UVFLAG:
Good:	0	0	Changed to bad:	0	HISTORY UVFLAG:
Bad:	0	0			HISTORY UVFLAG:
Executed on: 05JAN04:13:03:04.0					HISTORY UVFLAG:
Command line inputs follow:					HISTORY UVFLAG:
vis=cdf.1344jan					HISTORY UVFLAG:
flagval=flag					HISTORY UVFLAG:
select=@cdf1a.flag01					HISTORY UVFLAG:
Overview of flagging on visibility file cdf.1344jan					HISTORY UVFLAG:
Changed flags set to FALSE (data flagged as bad)					HISTORY UVFLAG:
Total number of records selected: 204; out of 527612 records					HISTORY UVFLAG:
Antennas used: 1,2,3,4,5,6					HISTORY UVFLAG:
Counts of correlations within selected channels					HISTORY UVFLAG:
channel	Originally	Currently			HISTORY UVFLAG:
Good:	2652	0	Changed to bad:	2652	HISTORY UVFLAG:
Bad:	0	2652			HISTORY UVFLAG:
wide	Originally	Currently			HISTORY UVFLAG:
Good:	0	0	Changed to bad:	0	HISTORY UVFLAG:
Bad:	0	0			HISTORY UVAFLAG:
Miriad Uvaflag: version 1.0 22-Oct-98					HISTORY UVAFLAG:
Executed on: 05JAN04:13:44:48.0					HISTORY UVAFLAG:
Command line inputs follow:					HISTORY UVAFLAG:
tvis=cdf.1344jan.stokes					HISTORY UVAFLAG:
vis=cdf.1344jan					HISTORY UVAFLAG:
options=nopol					HISTORY UVAFLAG:
select=time(4:00,6:00)					HISTORY UVAFLAG:
Miriad Uvaflag: version 1.0 22-Oct-98					HISTORY UVAFLAG:
Executed on: 05JAN04:13:47:01.0					HISTORY UVAFLAG:
Command line inputs follow:					HISTORY UVAFLAG:
tvis=cdf.1344jan.stokes					HISTORY UVAFLAG:
vis=cdf.1344jan					HISTORY UVAFLAG:
options=nopol					HISTORY UVAFLAG:
select=time(6:00,8:00)					HISTORY UVAFLAG:
Miriad Uvaflag: version 1.0 22-Oct-98					HISTORY UVAFLAG:
Executed on: 05JAN04:13:49:36.0					HISTORY UVAFLAG:
Command line inputs follow:					HISTORY UVAFLAG:
tvis=cdf.1344jan.stokes					HISTORY UVAFLAG:
vis=cdf.1344jan					HISTORY UVAFLAG:
options=nopol					HISTORY UVAFLAG:
select=time(8:00,10:00)					HISTORY UVAFLAG:
Miriad Uvaflag: version 1.0 22-Oct-98					HISTORY UVAFLAG:
Executed on: 05JAN04:13:51:59.0					HISTORY UVAFLAG:
Command line inputs follow:					HISTORY UVAFLAG:
tvis=cdf.1344jan.stokes					HISTORY UVAFLAG:
vis=cdf.1344jan					HISTORY UVAFLAG:





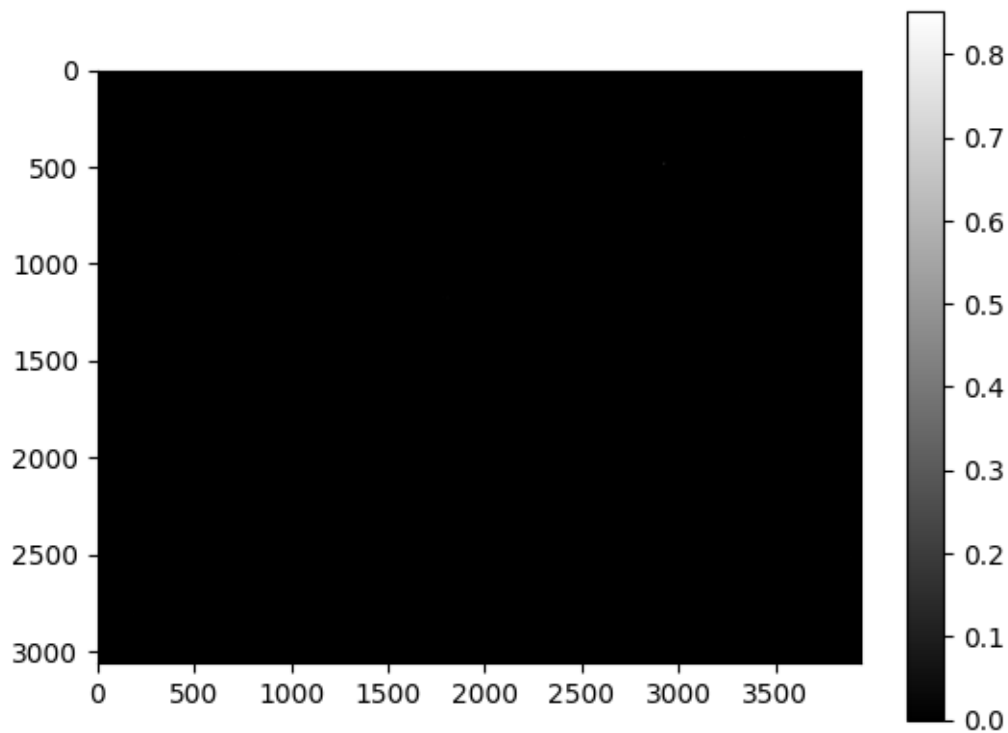








```
plt.imshow(image_data, cmap='gray' )
plt.colorbar()
plt.show()
```



```
[ ]: #Preprocessing data
```

```
# flattened_data = image_data.flatten()
# normalised_data = flattened_data / np.max(flattened_data)
# print(normalised_data)

print(image_data_df)
image_data_df.to_csv("Atlas_data.xlsx", index=False)
```

	0	1	2	3	4	5	6	
0	0.000235	0.000212	0.000125	-0.000064	-0.000229	-0.000110	0.000161	\
1	0.000204	0.000170	0.000105	-0.000052	-0.000218	-0.000133	0.000139	
2	0.000142	0.000121	0.000106	-0.000007	-0.000186	-0.000159	0.000098	
3	0.000049	0.000071	0.000133	0.000073	-0.000125	-0.000175	0.000042	
4	-0.000061	0.000023	0.000175	0.000171	-0.000045	-0.000175	-0.000016	
...	...	...	...	...	...	...	...	
3052	0.000200	0.000110	0.000029	0.000048	0.000015	-0.000098	-0.000125	
3053	0.000175	0.000144	0.000105	0.000116	0.000042	-0.000097	-0.000118	
3054	0.000130	0.000135	0.000136	0.000154	0.000062	-0.000086	-0.000100	

```

3055  0.000079  0.000084  0.000109  0.000148  0.000073 -0.000055 -0.000058
3056  0.000045  0.000011  0.000033  0.000099  0.000072 -0.000005  0.000010

```

```

          7          8          9  ...      3927      3928      3929
0      0.000142 -0.000217 -0.000439 ...  0.000089  0.000175  0.000142 \
1      0.000189 -0.000099 -0.000337 ...  0.000086  0.000151  0.000069
2      0.000229  0.000047 -0.000182 ...  0.000037  0.000068 -0.000053
3      0.000243  0.000179 -0.000016 ... -0.000032 -0.000033 -0.000165
4      0.000226  0.000265  0.000125 ... -0.000077 -0.000102 -0.000216
...      ...      ...      ...      ...      ...      ...
3052 -0.000057 -0.000054 -0.000123 ... -0.000016 -0.000029 -0.000098
3053 -0.000032 -0.000021 -0.000087 ... -0.000149 -0.000140 -0.000125
3054 -0.000011 -0.000006 -0.000072 ... -0.000274 -0.000242 -0.000142
3055  0.000014 -0.000010 -0.000085 ... -0.000356 -0.000303 -0.000137
3056  0.000050 -0.000022 -0.000116 ... -0.000370 -0.000302 -0.000101

          3930          3931          3932          3933          3934          3935          3936
0      0.000126  0.000144  0.000103  0.000020  0.000005  0.000063  0.000098
1      0.000019  0.000034  0.000012 -0.000041 -0.000045 -0.000018 -0.000011
2     -0.000116 -0.000079 -0.000053 -0.000048 -0.000029 -0.000046 -0.000092
3     -0.000210 -0.000141 -0.000071 -0.000009  0.000039 -0.000025 -0.000141
4     -0.000219 -0.000130 -0.000055  0.000036  0.000115  0.000019 -0.000167
...      ...      ...      ...      ...      ...      ...
3052 -0.000153 -0.000098  0.000037  0.000115  0.000080  0.000028  0.000061
3053 -0.000108 -0.000027  0.000109  0.000181  0.000120  0.000005 -0.000036
3054 -0.000065  0.000023  0.000146  0.000216  0.000149 -0.000005 -0.000114
3055 -0.000032  0.000033  0.000128  0.000206  0.000164  0.000007 -0.000142
3056 -0.000003  0.000008  0.000062  0.000156  0.000164  0.000040 -0.000111

```

[3057 rows x 3937 columns]

## KNN MODELING

```
[ ]: from sklearn.preprocessing import StandardScaler
```

```

x = image_data_df.values
y = np.where(x.mean(axis = 1) > 0, 1, 0)

scaler = StandardScaler()
X_normalized = scaler.fit_transform(x)

```

```
[ ]: from sklearn.model_selection import train_test_split
```

```

X_train, X_test, y_train, y_test = train_test_split(X_normalized, y,
    ↪ test_size=0.2, random_state=42)

from sklearn.neighbors import KNeighborsClassifier

```

```
knn = KNeighborsClassifier(n_neighbors=5) # Use 5 neighbors
knn.fit(X_train, y_train)
```

```
[ ]: KNeighborsClassifier()
```

```
[ ]: from sklearn.metrics import accuracy_score, classification_report

# Predict test labels
y_pred = knn.predict(X_test)

# Evaluate accuracy
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy * 100:.2f}%")

# Detailed metrics
print("Classification Report:")
print(classification_report(y_test, y_pred))
```

Accuracy: 90.85%

Classification Report:

	precision	recall	f1-score	support
0	0.81	0.80	0.80	143
1	0.94	0.94	0.94	469
accuracy			0.91	612
macro avg	0.87	0.87	0.87	612
weighted avg	0.91	0.91	0.91	612