Ioana Lupascu

Tufts University: DATA-0201a

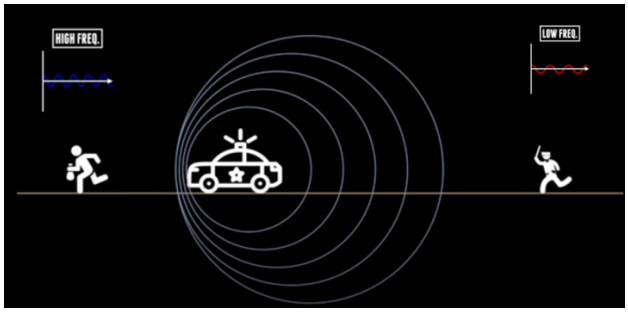
Introduction to Python for Data Science

Individual Project

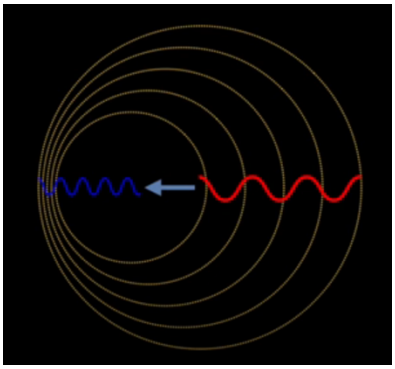
Table of Contents

Redshift Concept Explained

The universe is expanding, and as objects in space move away from us, they are accelerating. To measure this phenomenon, astronomers use **cosmic redshift** to explain how far away objects are. Cosmic redshift is a light effect comparable to the doppler effect in sound. Let's take the example of a police car siren. The waves of the siren sound travel through the air, into your ear. But if the police car is moving, the sound effect is different in front of the car compared to behind the car. The sound in front of the car is compressed, and creates a higher frequency sound. But the sound waves behind the car are stretched out, creating a lower frequency sound.



This effect is similar when measuring the distance to a galaxy. The concepts translates directly to light, radar, and any other EM signature used to measure distance. If the color is blue, you and the object are getting closer. Therefore, blueshift indicates the galaxy is close to us. When an object moves away from us, the light waves become stretched, therefore the frequency goes down into red wavelengths, and redshift indicates that the galaxy is further away.



Chadra Deep Field South is one image, taken by the Chandra X-ray Observatory satellite. It is noteworthy because the location gives a relatively clear “window” through the clouds of the Milky Way galaxy, which allows us to clearly see the rest of the universe in X-rays. COMBO-17 is a survey of the Chandra Deep Field South image, which contains a multi-color classification into the categories Star, Galaxy, and Quasar, and redshifts.

This analysis examines the COMBO-17 Dataset in detail. In order to study star formation in the galaxies which emerged after the Big Bang, astronomers examine objects in space which are far too faint for traditional spectroscopy. Instead, statistical analysis of multiband photometric datasets is performed. The COMBO-17 dataset shows the multiple photometric bands, as well

as associated corrections which are made for the effects of redshift. There are a total of 3462 rows and 65 columns in the data set.

Galaxy Redshift Magnitudes

Columns 2 – 3 in the COMBO-17 dataset show the galaxy redshift Magnitudes and associated error adjustments. Figure 1 shows the histograms for the Rmag magnitude values, which show the logarithmic measures of brightness. Figure 1 shows that e.Rmag, the error associated with the measurement process, increases and reaches an asymptote as the magnitudes increase.

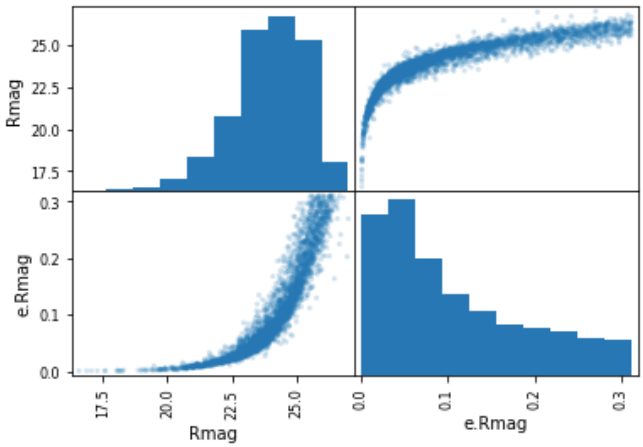


Figure 1. Redshift magnitudes and errors

Descriptive Statistics

	Nr	Rmag	e.Rmag
count	3462.000000	3462.000000	3462.000000
mean	5099.237435	23.939266	0.106156
std	2826.911998	1.435899	0.083147
min	6.000000	16.572000	0.001000
25%	2661.250000	23.099000	0.039000
50%	5158.500000	24.073500	0.080000
75%	7493.500000	25.029750	0.160000
max	9997.000000	27.000000	0.311000

Galaxy Sizes

Colums 4 – 5 of the dataset show the difference between the total aperture magnitude in the R band. This information is a rough measure of the sizes of the galaxies.

Descriptive Statistics. The table below shows the means, standard deviations, and distribution of mumax data relative to the mean, giving a fuller picture of the distribution of galaxy sizes.

	Nr	ApDRmag	mumax
count	3462.000000	3462.000000	3462.000000
mean	5099.237435	-0.190947	24.181846
std	2826.911998	0.452994	1.017225
min	6.000000	-4.700000	18.112000
25%	2661.250000	-0.425000	23.624250
50%	5158.500000	-0.141000	24.387500
75%	7493.500000	0.089000	24.957500
max	9997.000000	1.462000	25.833000

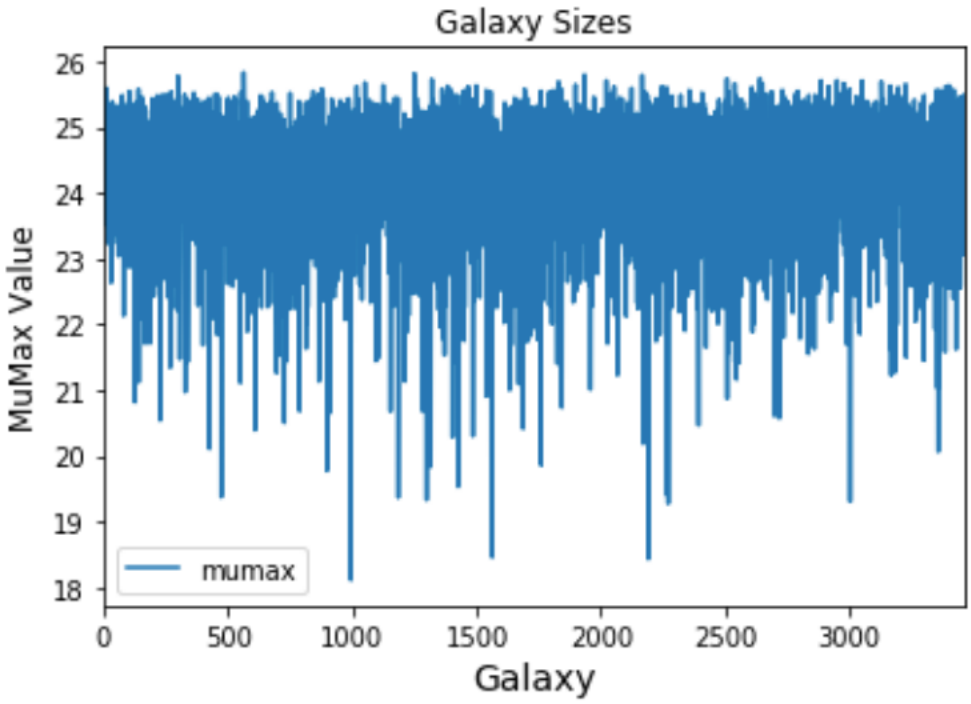


Figure 2. Galaxy sizes in COMBO-17 dataset

Smallest Galaxies

The smallest 10 galaxies in the COMBO-17 survey are identified below.

995	18.112
2191	18.433
1563	18.458
2271	19.276
3003	19.305
1299	19.330
1185	19.364
473	19.373
1559	19.381
2263	19.413

Largest Galaxies

The largest 10 galaxies in the COMBO-17 survey are identified below.

563	25.833
1250	25.816
1934	25.799
2166	25.791
299	25.780
2640	25.743
2507	25.737
1320	25.730
2950	25.711
2022	25.709

Redshift Estimates

Columns 6 – 9 of the COMBO-17 dataset are redshift estimates, which give more information about the reliability of the error measurements taken from the surface brightness approximations. Higher e.Mcz and chi2red values may indicate lower reliability of brightness interpretation. The table below shows the means, standard deviations, and distribution of redshift estimates.

	Nr	Mcz	e.Mcz	MCzml	chi2red
count	3462.000000	3462.000000	3462.000000	3462.000000	3462.000000
mean	5099.237435	0.728503	0.090381	0.770000	1.167392
std	2826.911998	0.319010	0.061970	0.375989	0.682852
min	6.000000	0.007000	0.000000	0.000000	0.140000
25%	2661.250000	0.519000	0.036000	0.502250	0.730000
50%	5158.500000	0.810000	0.076500	0.826500	1.020000
75%	7493.500000	0.981750	0.141000	1.024750	1.440000
max	9997.000000	1.379000	0.233000	1.400000	11.910000

Absolute Magnitudes

Columns 10 – 29 give the intrinsic luminosities in terms of the absolute magnitudes of the galaxies identified, and their associated measurement errors. The readings are based on the measured magnitudes and the calculated redshifts. The distributions of each band after redshift correction can be seen in Figure 3 below.

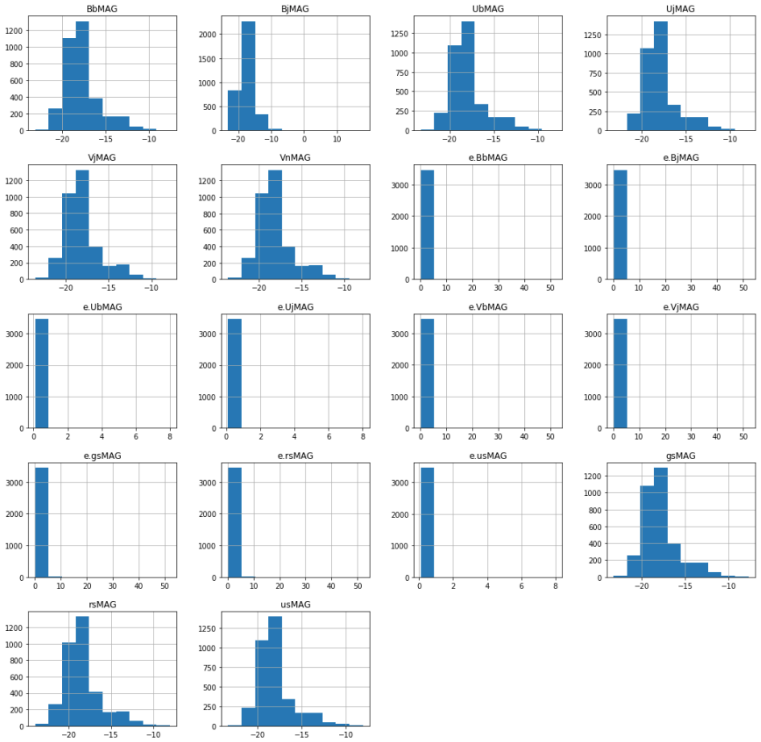


Figure 3. Intrinsic luminosities based on measured magnitudes and redshifts, with associated calculated measurement errors

Figure 4 below shows only the distributions of values for the readings in each band. Error measurements have been removed. All readings outside of 1 standard deviation away from the mean have been removed for analysis.

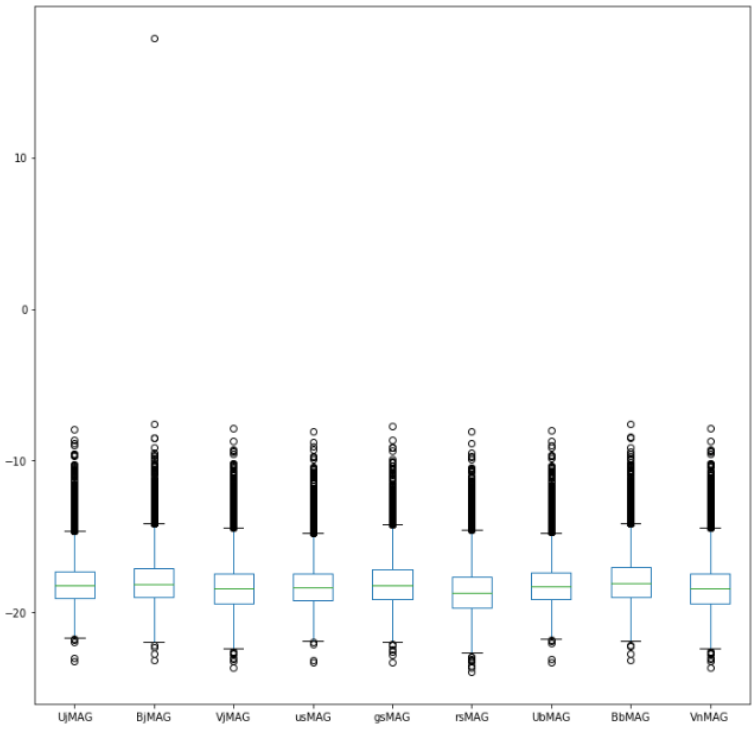


Figure 4. Distributions for absolute magnitude readings

Observed Brightness

Columns 30 – 55 show the observed brightness's in 13 bands in sequence from 420 nm in the ultraviolet to 915 nm in the far red. These are given in linear variables with units of photon flux densities, photons/m^2/s/nm. Each measurement also has a measurement error which can be used to distinguish measurement from intrinsic dispersions in the distributions.

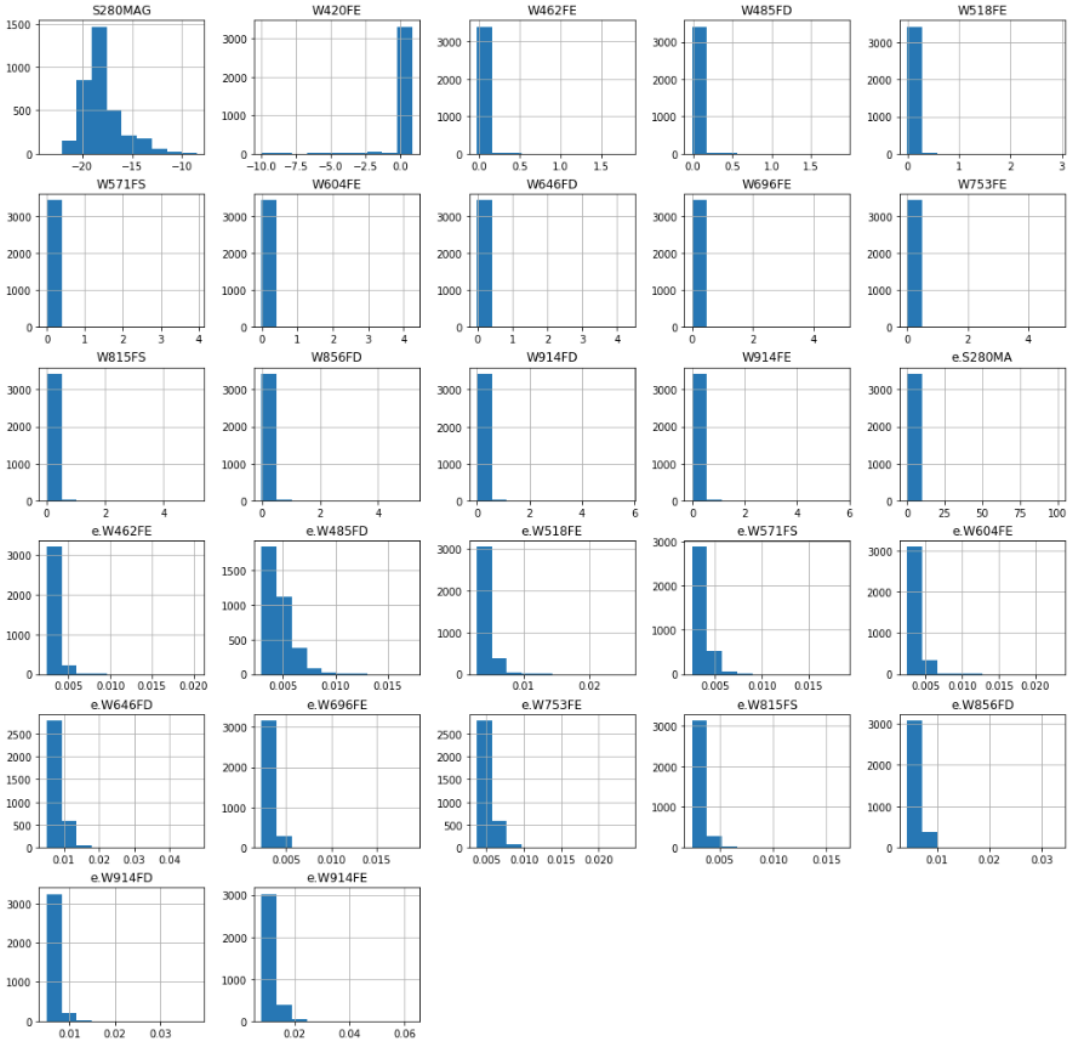


Figure 5. Observed brightnesses for galaxies in 13 bands from ultraviolet to far red range

Photometric Correlation

The analysis below shows a correlation of all features compared. The lighter parts of the image indicate a higher correlation factor between factors compared.

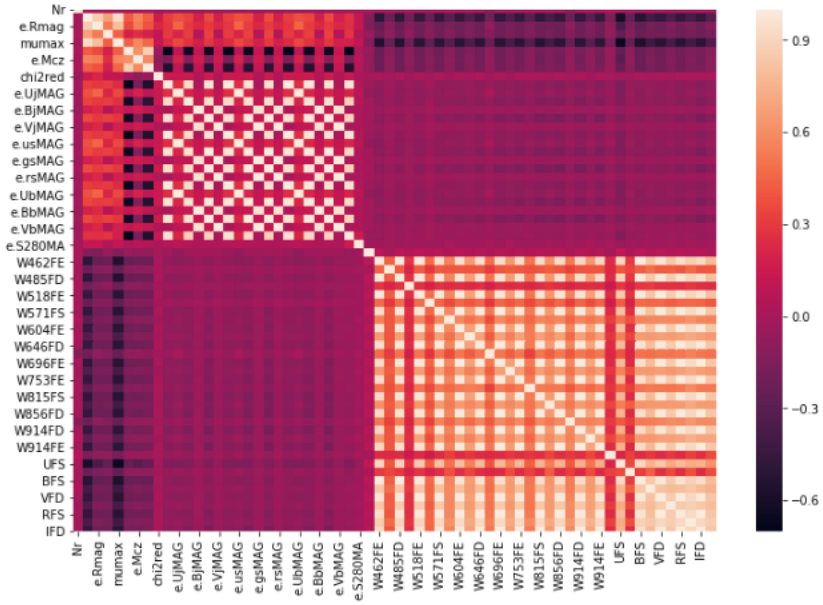


Figure 6. Correlation of all variables effecting intrinsic luminosity of galaxies.

The analysis below shows a correlation of features compared on intrinsic luminosity but excludes measurement errors. Lighter sections of the figure indicate a higher correlation between readings in the measurement bands assessed.

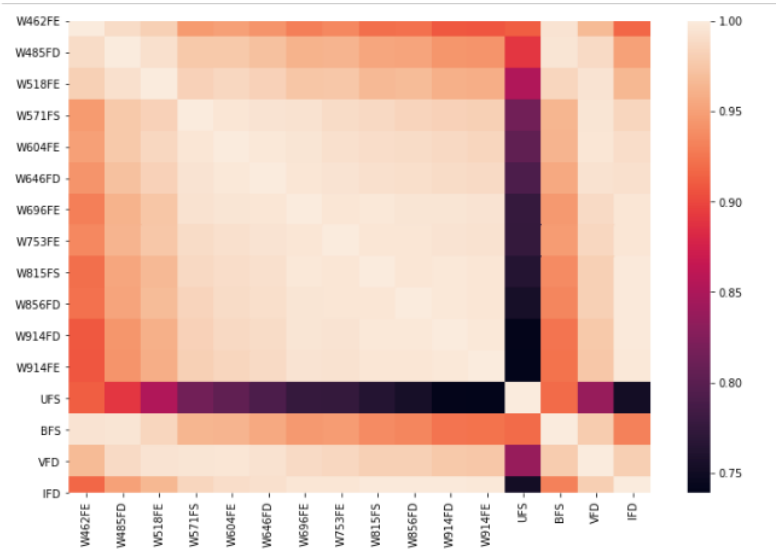


Figure 6. Correlation of all variables effecting observed brightness, excluding error readings.