Home: to the Combo-17 dataset below there is the link and the description

<https://www.kaggle.com/mrisdal/combo17-galaxy-dataset>

### Contents

[Wolf et al. (2004)](http://arxiv.org/abs/astro-ph/0403666) provide the first public catalog of a large dataset (63,501 objects) with brightness measurements in 17 bands in the visible band. (Note that the Sloan Digital Sky Survey provides a much larger dataset of 108 objects with measurements in 5 bands.) We provide here a subset of their catalog with 65 columns of information on 3,462 galaxies. These are objects in the Chandra Deep Field South field which Wolf and colleagues have classified as `Galaxies'. The column headings are formally described in their Table 3, and the columns we provide are summarized here with brief commentary:

**Col 1:** Nr, object number

**Col 2-3:** Total R (red band) magnitude and its error. This was the band at which the basic catalog was constructed. Magnitudes are inverted logarithmic measures of brightness. A galaxy with R=21 is 100-times brighter than one with R=26. The error is the standard deviation derived from detailed knowledge of the measurement process. This dataset is an excellent example of astronomical datasets where each variable is accompanied by heteroscedastic measurement errors of known variances.

**Col 4-5:** ApDRmag is the difference between the total and aperture magnitude in the R band. This is a rough measure of the size of the galaxy in the image where ApDRmag=0 corresponds to a point source. Negative values are not physically meaningful. mu*max is the central surface brightness of the object in the R band. The difference between Rmag and mu*max should also be an indicator of galaxy size.

**Col 6-9:** Mcz and MCzml are two redshift estimates. Mcz is the preferred value. e.Mcz is its estimated error, and chi2red is the reduced chi-squared value of the least-squares fit of the 17-band magnitudes to the best-fit template galaxy spectrum. Galaxies with large e.Mcz or chi2red might be omitted as unreliable.

**Col 10-29:** These give the absolute magnitudes (i.e. intrinsic luminosities) of the galaxy in 10 bands, with their measurement errors. They are based on the measured magnitudes and the redshifts, and represent the intrinsic luminosities of the galaxies; a galaxy with M=-15 is 100-times less luminous than one with M=-20. These magnitudes are not all independent of each others, but the are important for representing intrinsic properties of the galaxies. Below is one of several redshift-stratified plots of the B-band absolute magnitude (abscissa) against the difference of magnitude (i.e. ratio of luminosities) between the 2800A ultraviolet and blue band, which is a sensitive indicator of star formation. A redshift-dependent bimodal distribution is seen.

**Col 30-55:** Observed brightnesses 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/m2/s/nm. Again, each measurement is accompanied by a measurement error which can be used to distinguish measurement from intrinsic dispersions in the distributions.

**Col 56-65:** Observed brightnesses in 5 traditional broad spectral bands, UBVRI. These are largely redundant with the 13 bands in the previous columns.