Homework 1

---

1 Problem 1

Plotting every 50 points because the data set is so big!
Also, couldn't figure out how to get three panel with seaborn joinplot. So, plotted that seperately.

1.1 Stars

Figure 1 shows the distribution of r Vs g-i
Discussion:
The bimodal distribution in stars represents the distribution of stars on the H-R diagram. The X axis r corresponds to the magnitude in r band and the Y axis g-i corresponds to the color. Lower value of g-i means it is redder. The main sequence stars are the ones which are bluer and undergoing the hydrogen fusion process and the stars on the red sequence (lower g-r value) are the red giants (also have higher r luminosities).
The stars have with ~ 0 Proper motion(PM at lower RA (00 to 03 RA) while at higher RA, they do have some PM. Stars have PM along the DEC in both populations.

1.2 Galaxies

Figure 2 shows the distribution of r Vs g-i
Discussion:
In galaxies, we do not see a bimodal distribution like in stars. However, there are two populations of galaxies (not clearly separated on the H-R diagram of galaxies), spirals (blue) and ellipticals (red). From the RA and DEC PM, we see that majority of the galaxies have a PM in both directions (RA and DEC).

2 Problem 2

2.1 mag.outlier

Figure 3 shows the distribution of mag.outlier (Top) and the Q-Q plot (Bottom).
Sample Mean = 18.56 ; Sample Variance = 0.005
25th percentile = 18.51 ; 75th percentile = 18.61
Equation of line through these two points is Y = 0.07*X + 18.56
From the PDF< we see that the sample has outliers which is also seen in the Q-Q plot where the sample quantile deviates (not along 45 degree line) from that of theoretical value (i.e. that of normal distribution) with outliers seen at the tails. The red line passes through the 25th and 75th quantile values and the slope of this line is not 1 which again shows that it is not normal distribution.

2.2 mag.het

Figure 4 shows the distribution of mag.het (Top) and the Q-Q plot (Bottom).
Sample Mean = 18.55 ; Sample Variance = 0.002
25th percentile = 18.52 ; 75th percentile = 18.58
Equation of line through these two points is Y = 0.04*X + 18.55
This sample is biased towards higher magnitudes i.e. less luminous galaxies. The variance is lower than the outlier sample. The sample quantiles lie on a line with a slope not equal to 1 i.e. not normal distribution.

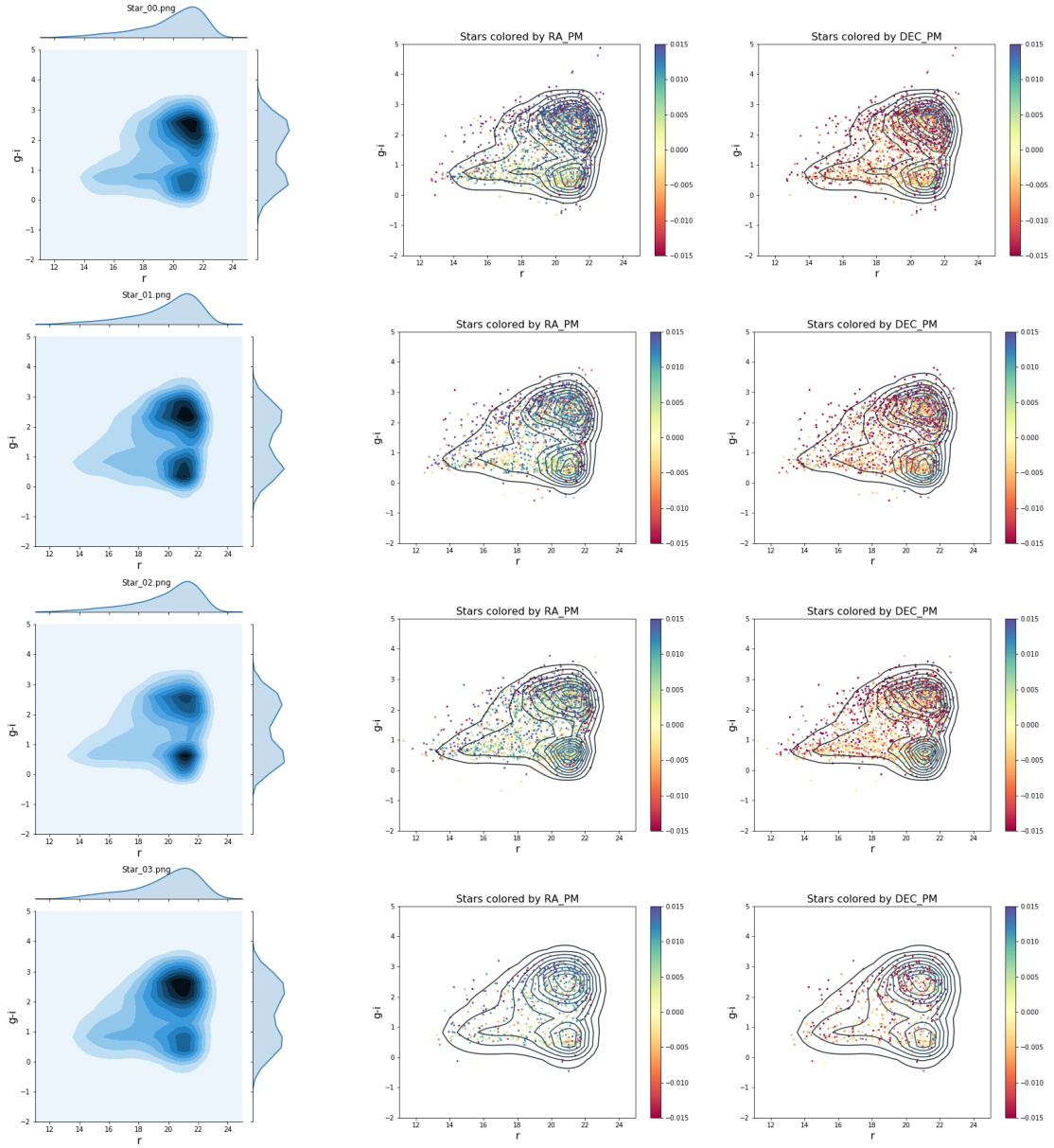
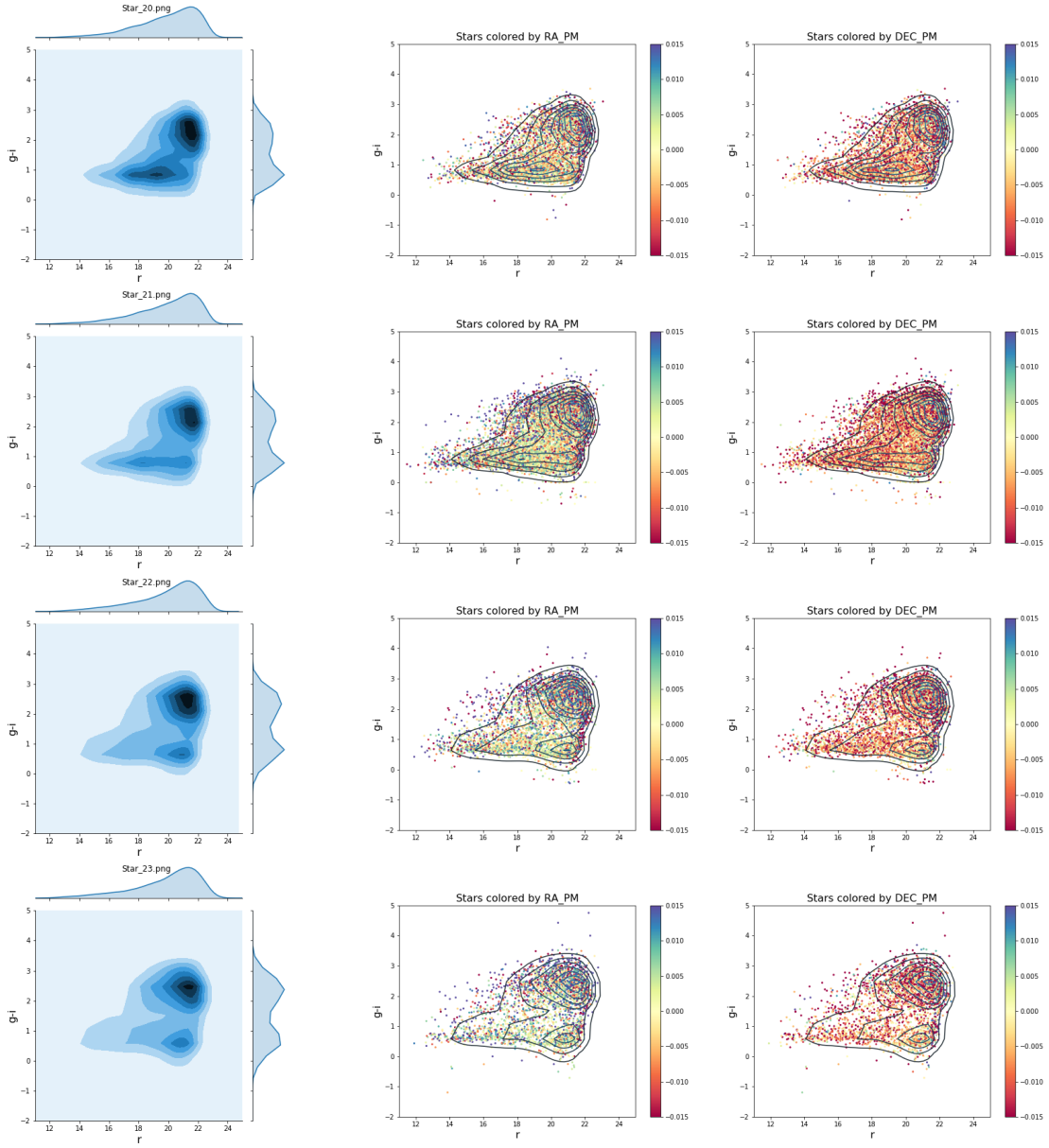


Figure 1: **Left:** Bivariate distribution of r Vs g-i in stars color coded by density. **Center:** Color coded by RA proper motion. **Right:** Color coded by DEC proper motion. Continued on next page

2.3 Normalized mag.het

Figure 5 shows the distribution of normalized mag.het (Top) i.e [mag.het-mena(mag.het)]/sample error and the Q-Q plot (Bottom).
As expected of a standard normal distribution, Sample Mean = 0.1 ($\sim$ 0) ; Sample Variance = 0.8 ($\sim$ 1)
25th percentile = -0.5 ; 75th percentile = 0.72
Equation of line through these two points is Y = 0.91*X + 0.1 i.e. almost lies on 45 degree line as expected of a standard normal distribution.



3 Problem 3

3.1 LMC

Figure 6 shows W Vs logP1 in mode F1 and 1. The residuals do not have a normal distribution (the normal distribution with sample mean and variance is shown in red). In Figure 7, we see that in LMC, the Cepheid variables have almost no proper motion across the galaxy. Figure 8 shows residuals Vs W and since the W parameter is insensitive to reddening, the scatter can not be because of extinction.

3.2 SMC

In Figure 10, we see that the residuals have an almost normal distribution in SMC compared to LMC i.e. larger scatter. From Figure 11, the PM across SMC is higher than that of LMC and we can see the rotation of SMC.

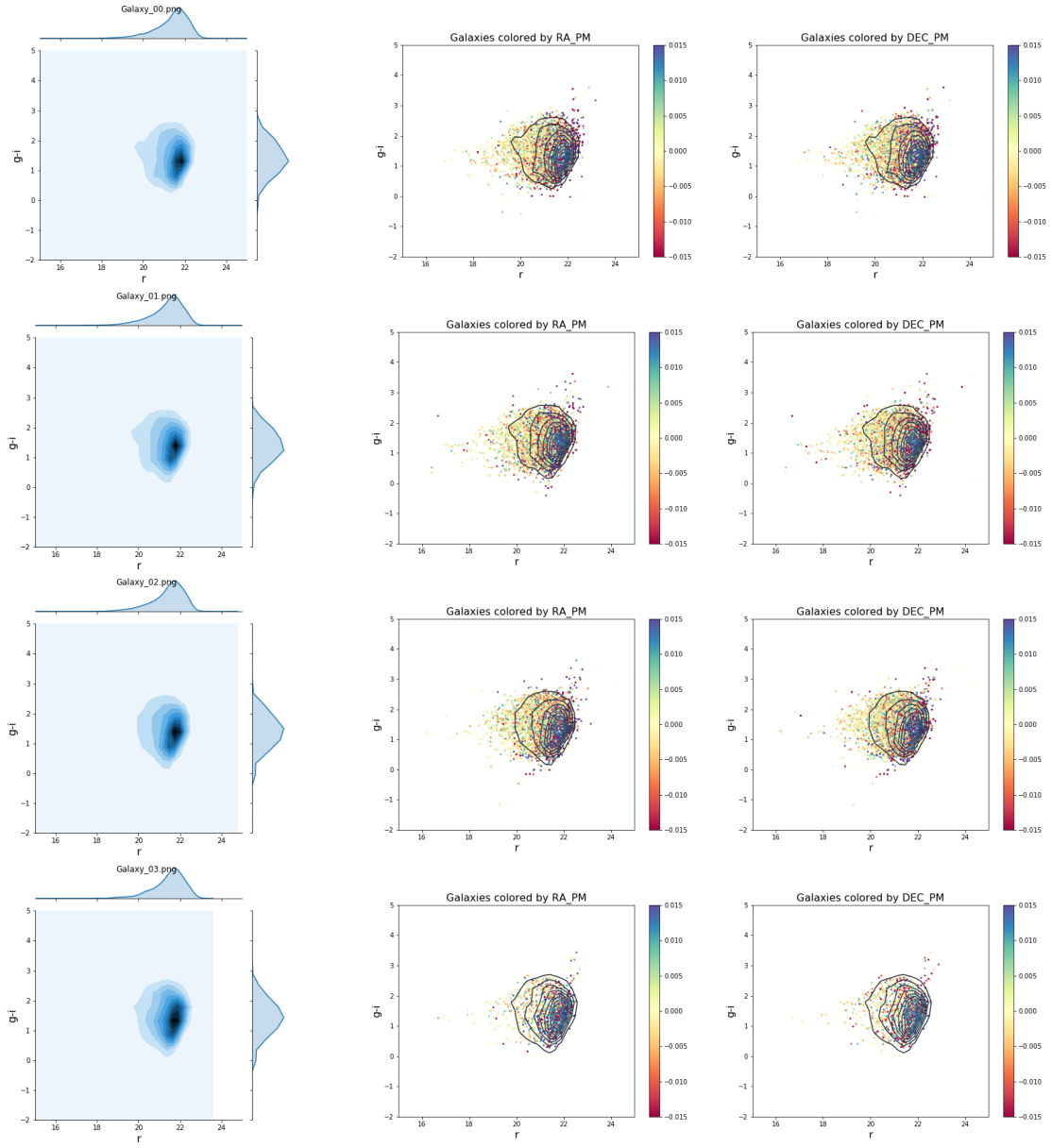
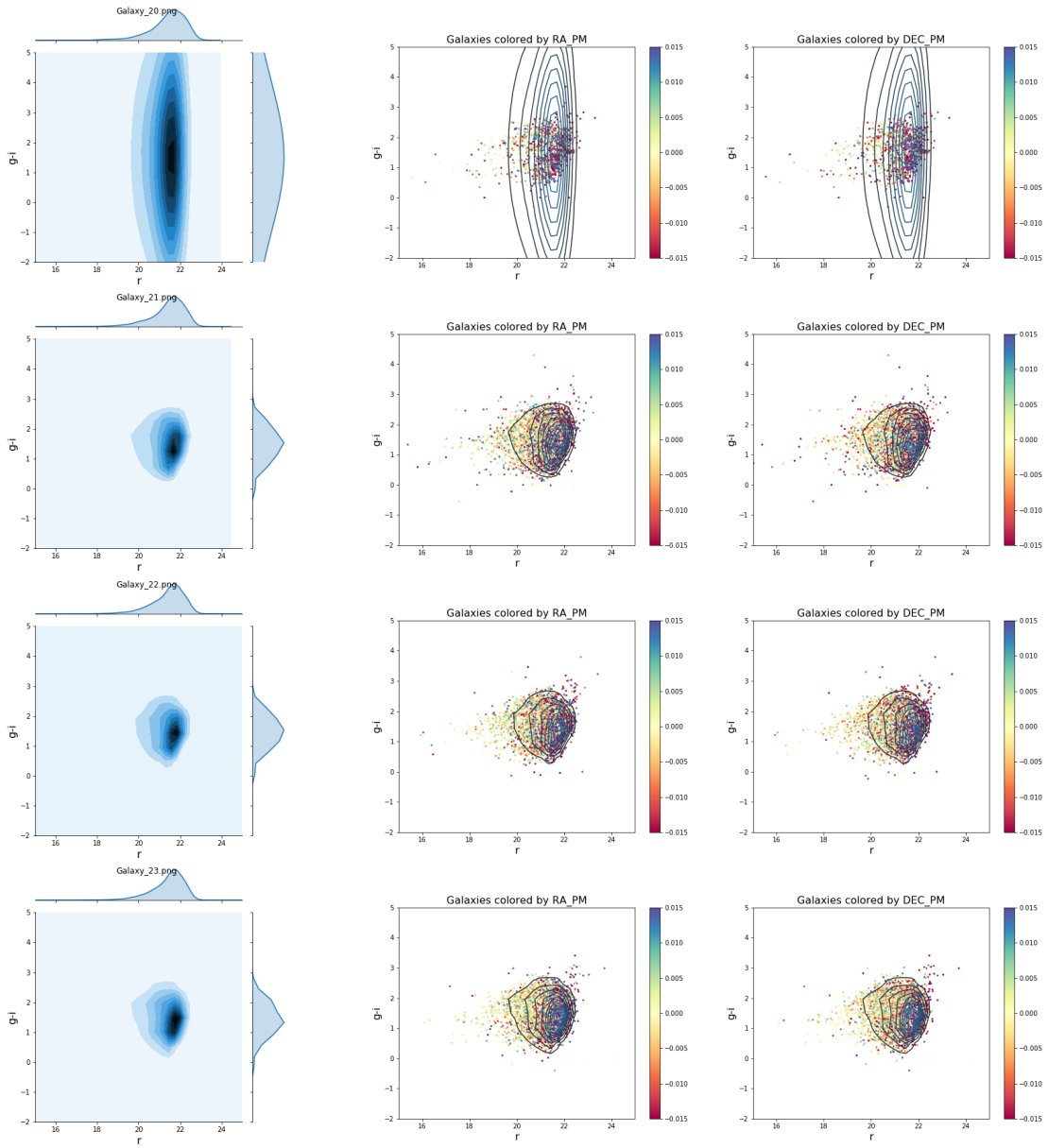


Figure 2: **Left:** Bivariate distribution of r Vs g-i in galaxies color coded by density. **Center:** Color coded by RA proper motion. **Right:** Color coded by DEC proper motion. Continued on next page


Galaxy_20.png
Galaxies colored by RA_PM
Galaxies colored by DEC_PM
Galaxy_21.png
Galaxies colored by RA_PM
Galaxies colored by DEC_PM
Galaxy_22.png
Galaxies colored by RA_PM
Galaxies colored by DEC_PM
Galaxy_23.png
Galaxies colored by RA_PM
Galaxies colored by DEC_PM
g-i
r

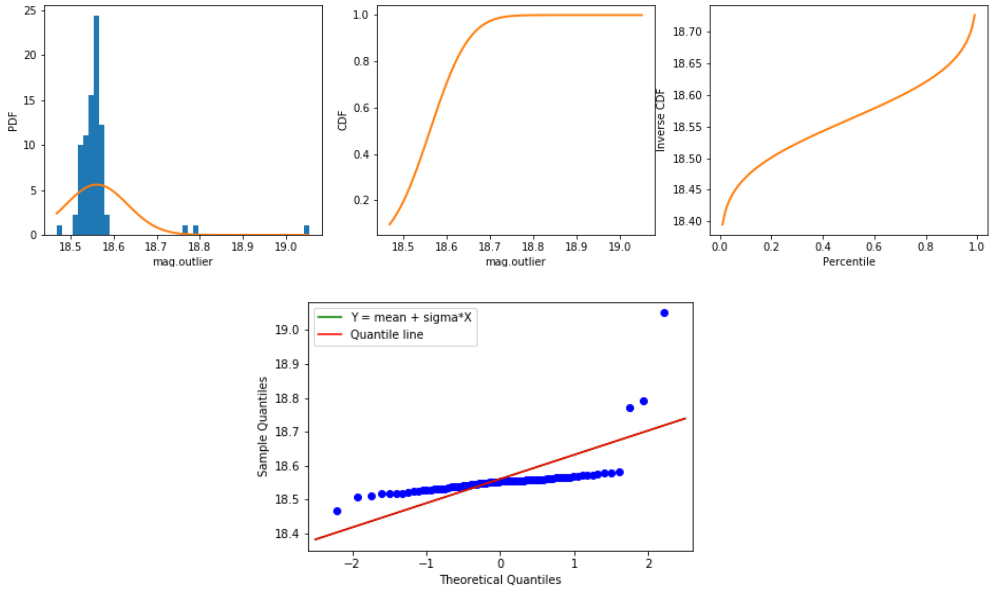


Figure 3: **Top:** Left: Probability distribution function(PDF) of the sample **mag.outlier**. Center: Cumulative distribution function (CDF). Right: Inverse CDF **Bottom:** Q-Q plot of the sample is shown in blue. Y = mean + sigma*X is plotted in green and the line passing through 25th and 75th percentile is shown in red.

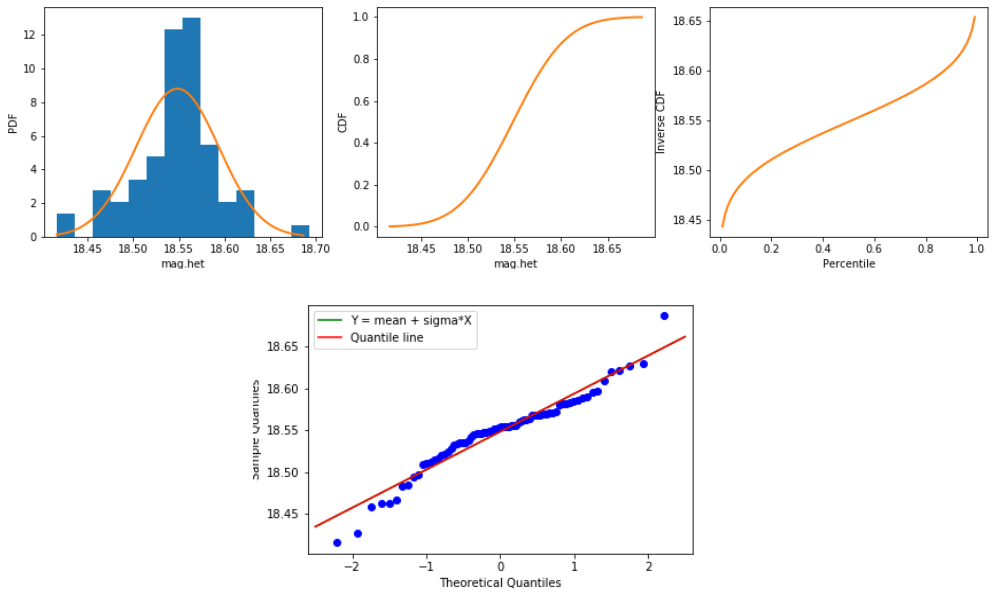


Figure 4: **mag.het**

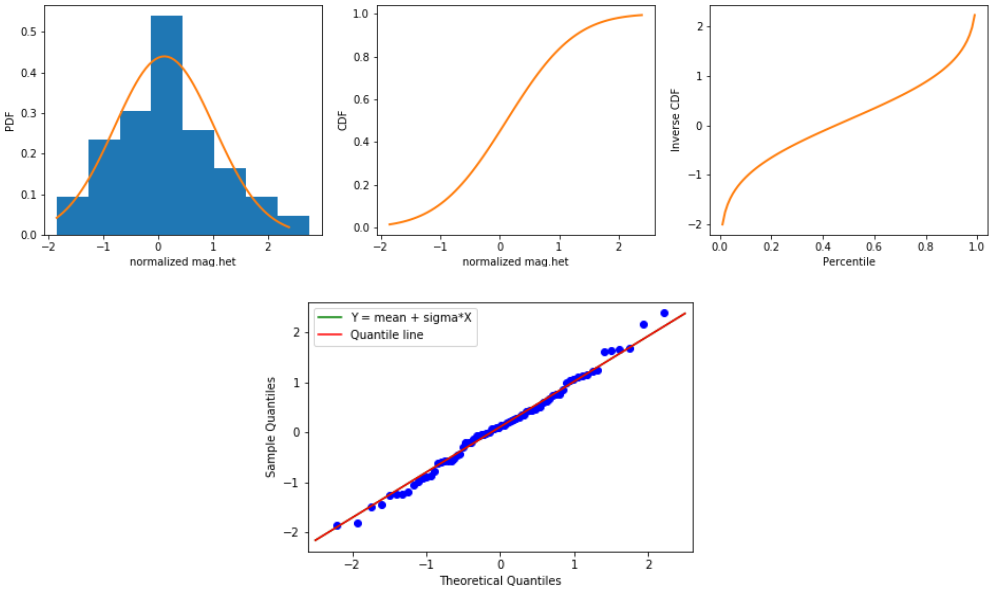


Figure 5: **Normalized mag.het**

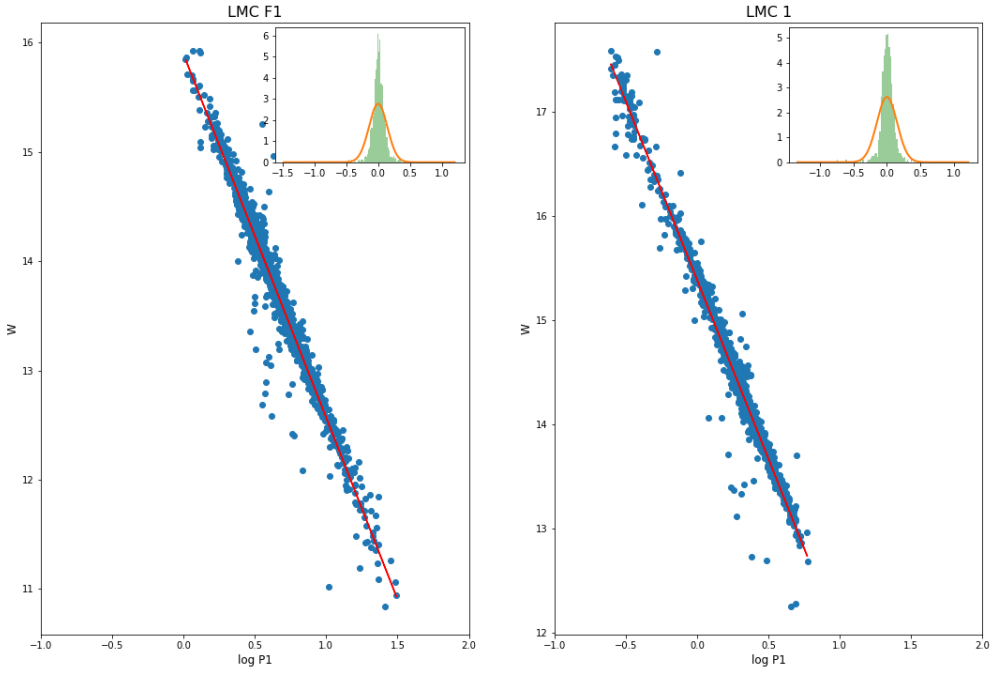


Figure 6: **Left:** Mode = F1, **Right:** Mode = 1. The residuals i.e. deviation of the data from the best fit line is shown in the inset with the normal distribution (with sample mean and variance) shown in red

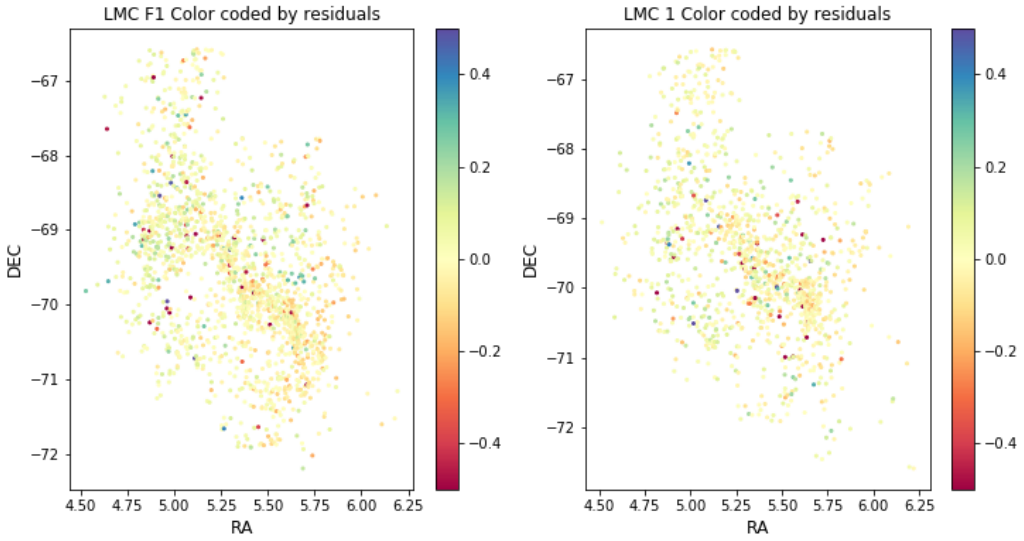


Figure 7

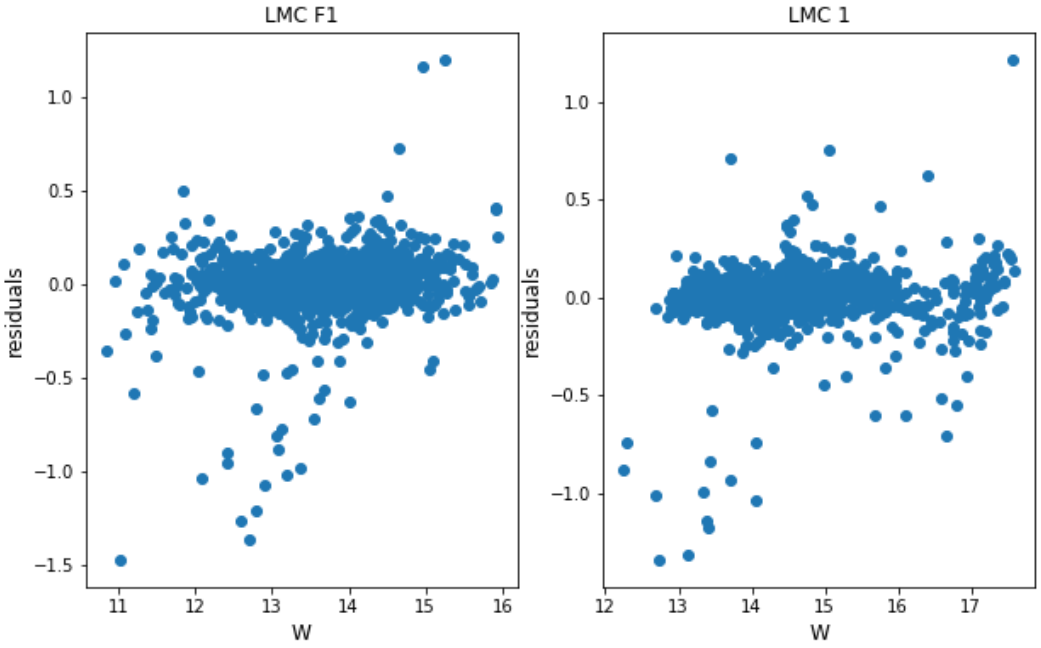


Figure 8

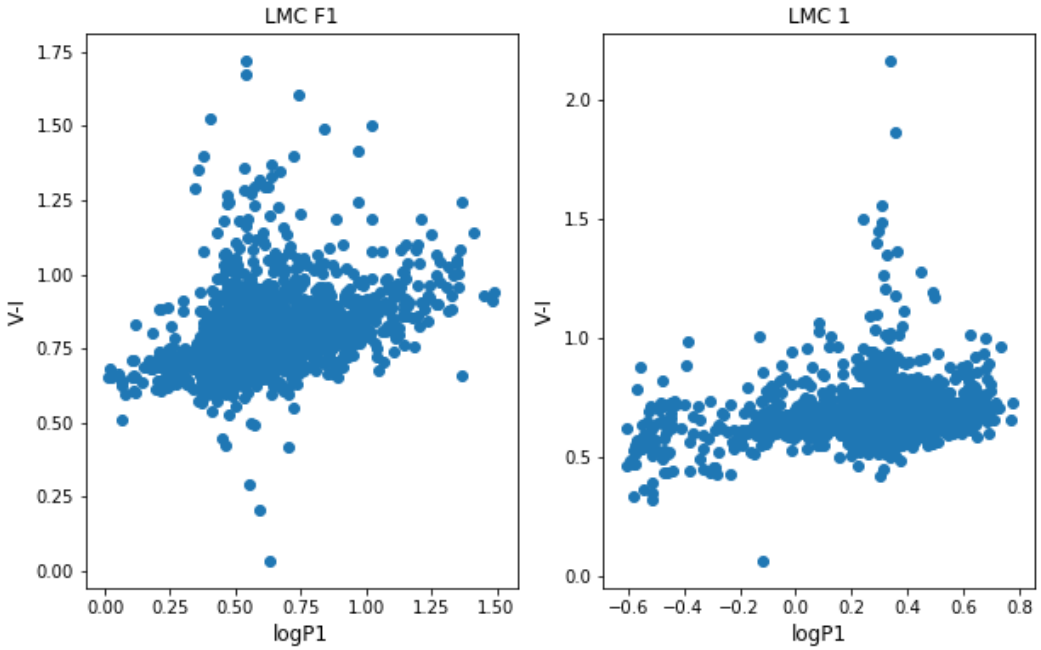


Figure 9

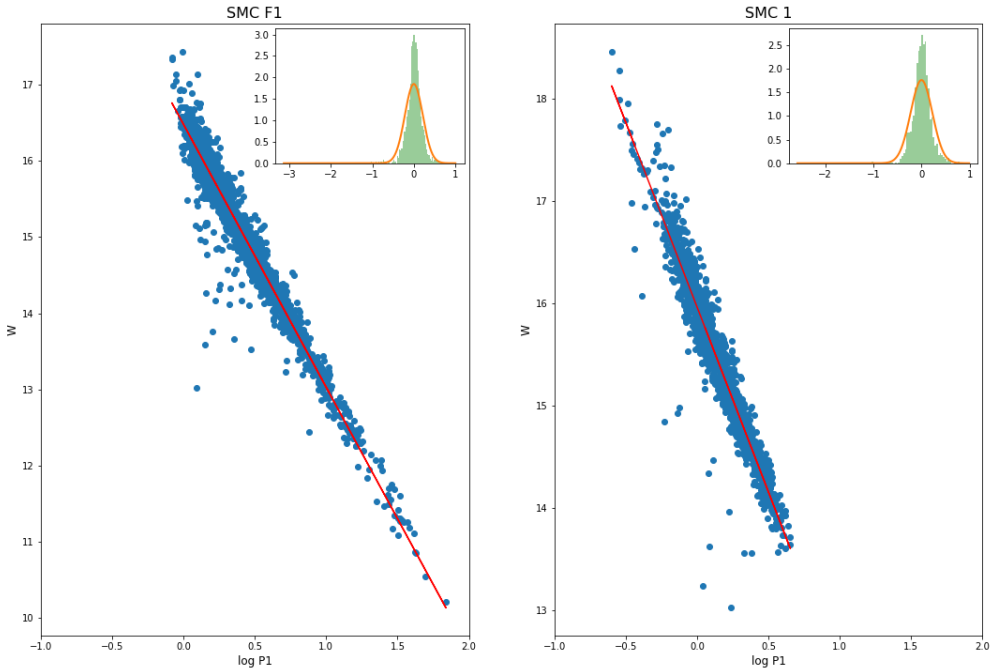


Figure 10

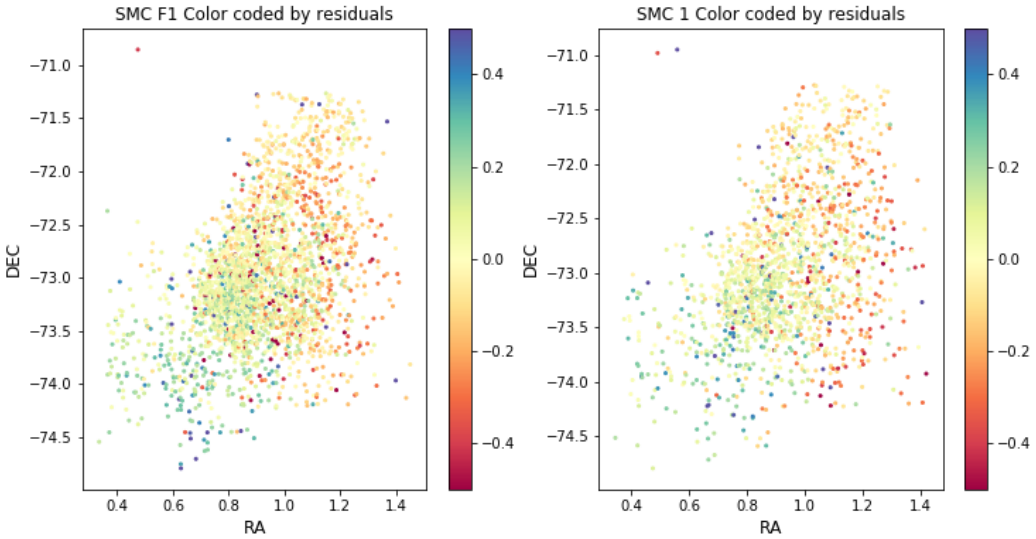


Figure 11

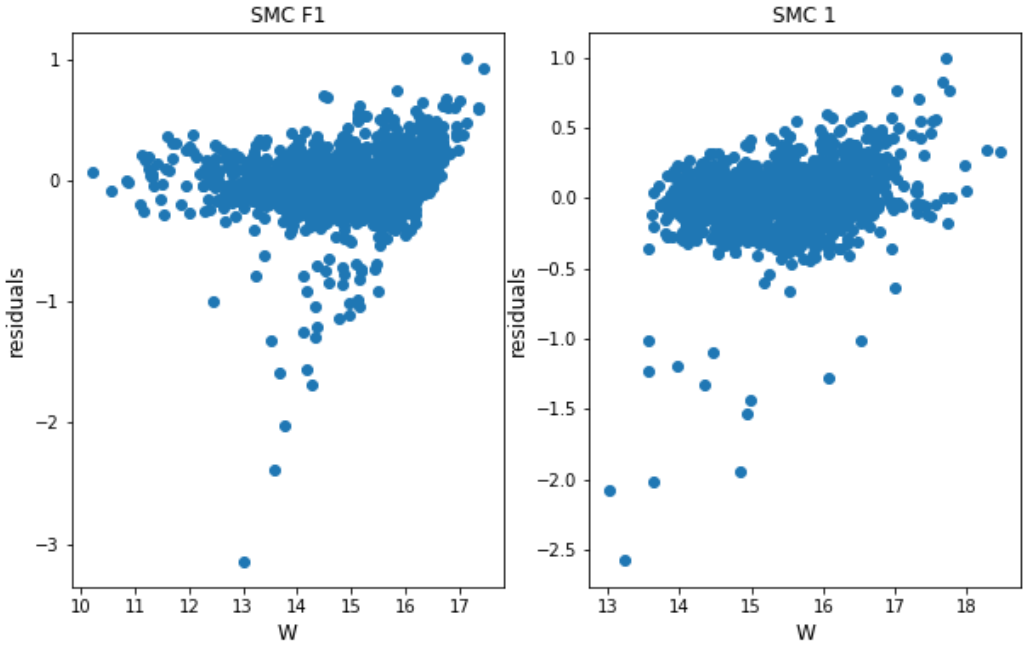


Figure 12

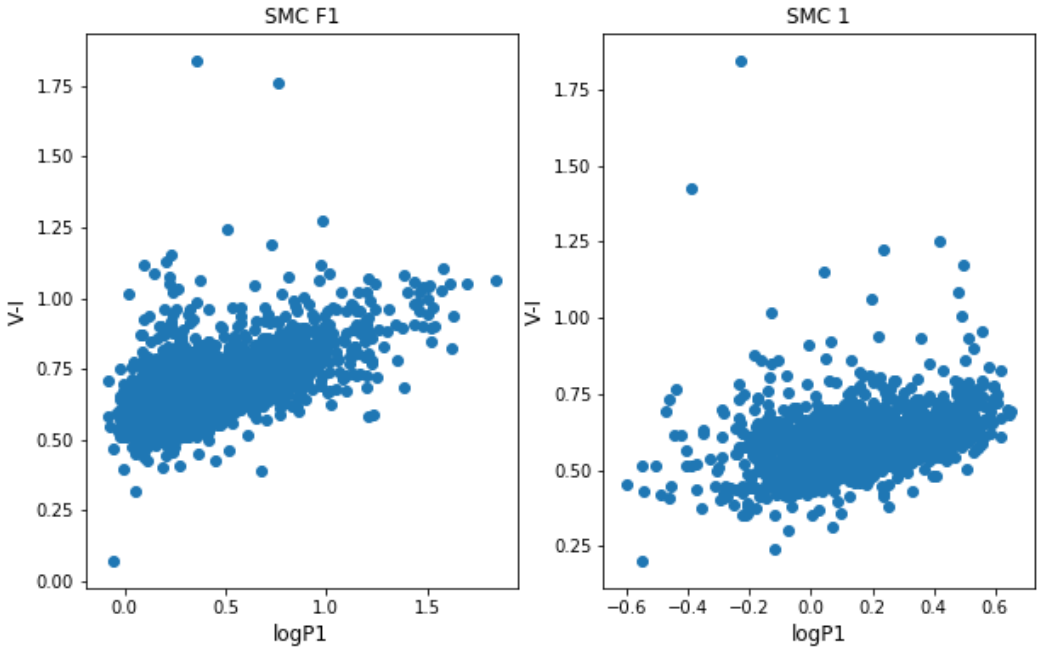

SMC F1
SMC 1
V-I
logP1
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
−0.6
−0.4
−0.2
0.2
0.4
0.6

Figure 13