ASTR596FDS_HW1_ChrisTandoi

February 6, 2020

Problem 1

[3]:
```
import matplotlib.pyplot as plt
import astropy
import astropy.io.fits as fits
import os
import seaborn as sns
import pandas as pd
import numpy as np
```

[4]:
```
HLC1 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_00_to_01.fits.gz'))
HLC2 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_01_to_02.fits.gz'))
HLC3 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_02_to_03.fits.gz'))
HLC4 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_03_to_04.fits.gz'))
HLC5 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_20_to_21.fits.gz'))
HLC6 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_21_to_22.fits.gz'))
HLC7 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_22_to_23.fits.gz'))
HLC8 = fits.open(os.path.join('/Users/christandoi/Google Drive/UIUC/Classes/
 ↪596fds/','HLC.RA_23_to_24.fits.gz'))
```

[5]:
```
def stats(file):
    MOT = file[1].data['MEAN_OBJECT_TYPE']
    PSF = file[1].data['MEAN_PSFMAG']
    RA = file[1].data['RA_PM']
    dec = file[1].data['DEC_PM']
    stars_r = [PSF[index][2] for index, value in enumerate(MOT) if value > 5]
    stars_gi = [(PSF[index][1] - PSF[index][3]) for index, value in␣
 ↪enumerate(MOT) if value > 5]
    galaxies_r = [PSF[index][2] for index, value in enumerate(MOT) if value <=␣
 ↪5]
```

```
    galaxies_gi = [(PSF[index][1] - PSF[index][3]) for index, value in␣
↪enumerate(MOT) if value <= 5]
    stars_ra = [RA[index] for index, value in enumerate(MOT) if value > 5]
    stars_dec = [dec[index] for index, value in enumerate(MOT) if value > 5]
    galaxies_ra = [RA[index] for index, value in enumerate(MOT) if value <= 5]
    galaxies_dec = [dec[index] for index, value in enumerate(MOT) if value <= 5]
    return stars_r, stars_gi, galaxies_r, galaxies_gi, stars_ra, stars_dec,␣
↪galaxies_ra, galaxies_dec
```

[6]:

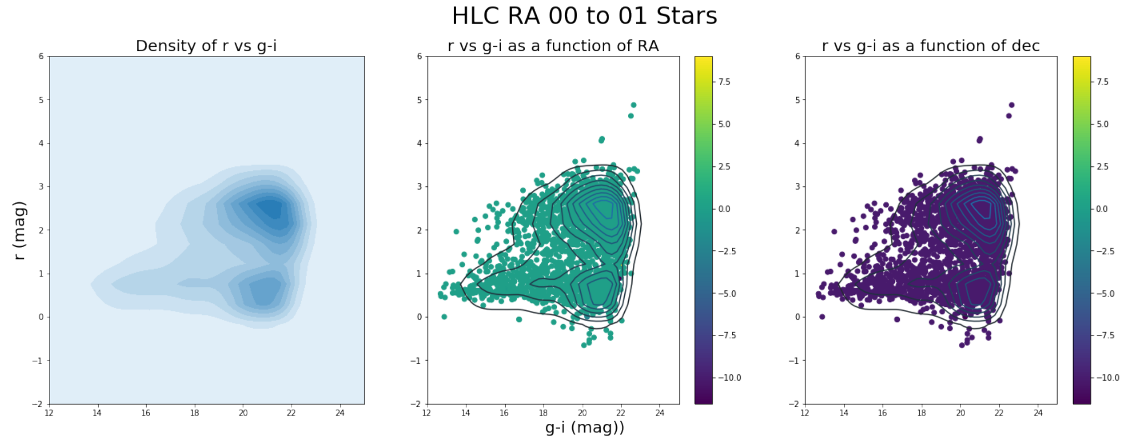
```
HLC1stats = stats(HLC1)
HLC2stats = stats(HLC2)
HLC3stats = stats(HLC3)
HLC4stats = stats(HLC4)
HLC5stats = stats(HLC5)
HLC6stats = stats(HLC6)
HLC7stats = stats(HLC7)
HLC8stats = stats(HLC8)
```

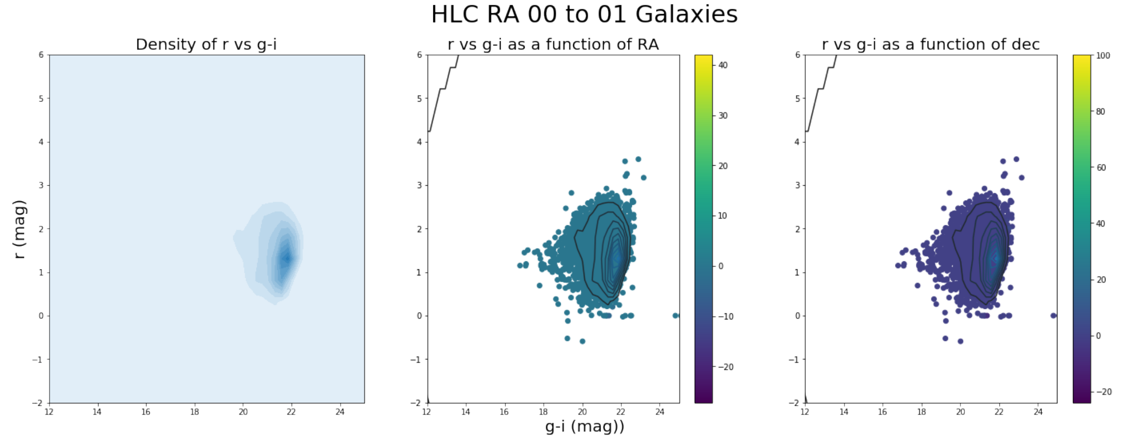
[10]:

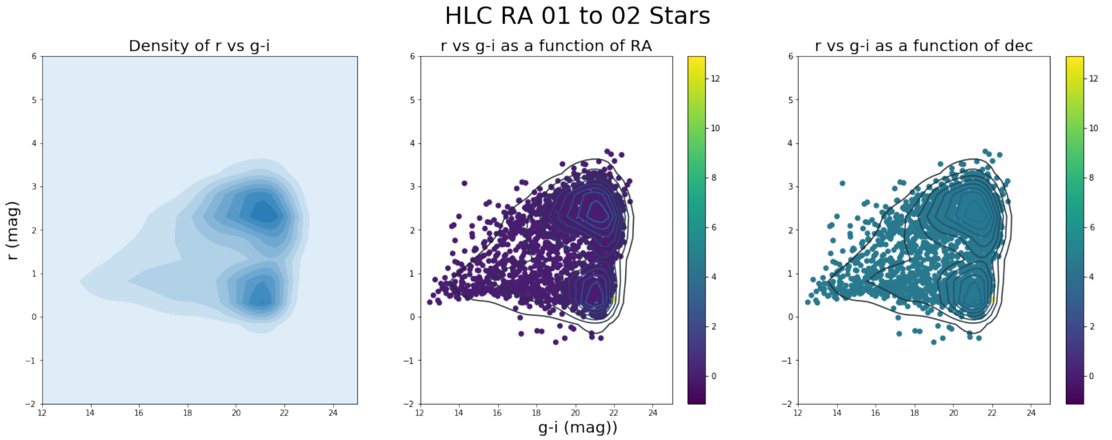
```
def plotting(file, name):
    bigtitle = {'fontsize':30}
    regtitle = {'fontsize':20}

    fig1 = plt.figure(figsize=(24,8))
    ax = fig1.add_subplot(1,1,1)
    ax.set_title(f'{name} Stars', pad=40, fontdict=bigtitle)
    ax.set_xlabel('g-i (mag))', fontdict=regtitle)
    ax.spines['top'].set_color('none')
    ax.spines['bottom'].set_color('none')
    ax.spines['left'].set_color('none')
    ax.spines['right'].set_color('none')
    ax.tick_params(labelcolor='w', top=False, bottom=False, left=False,␣
↪right=False)

    ax11 = fig1.add_subplot(1,3,1)
    ax11.set_xlim((12,25))
    ax11.set_ylim((-2,6))
    ax11.set_ylabel('r (mag)', fontdict=regtitle)
    ax11.set_title('Density of r vs g-i', fontdict=regtitle)
    sns.kdeplot(file[0][::50], file[1][::50], shade=True)

    ax12 = fig1.add_subplot(1,3,2)
    ax12.set_xlim((12,25))
    ax12.set_ylim((-2,6))
    ax12.set_title('r vs g-i as a function of RA', fontdict=regtitle)
    plot2 = ax12.scatter(file[0][::50], file[1][::50], c=file[4][::50])
    sns.kdeplot(file[0][::50], file[1][::50])
    fig1.colorbar(plot2, ax=ax12)
    ax13 = fig1.add_subplot(1,3,3)
    ax13.set_xlim((12,25))
    ax13.set_ylim((-2,6))
    ax13.set_title('r vs g-i as a function of dec', fontdict=regtitle)
    plot3 = ax13.scatter(file[0][::50], file[1][::50], c=file[5][::50])
    sns.kdeplot(file[0][::50], file[1][::50])
    fig1.colorbar(plot2, ax=ax13)

    fig2 = plt.figure(figsize=(24,8))
    ax = fig2.add_subplot(1,1,1)
    ax.set_title(f'{name} Galaxies', pad=40, fontdict=bigtitle)
    ax.set_xlabel('g-i (mag))', fontdict=regtitle)
    ax.spines['top'].set_color('none')
    ax.spines['bottom'].set_color('none')
    ax.spines['left'].set_color('none')
    ax.spines['right'].set_color('none')
    ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)

    ax21 = fig2.add_subplot(1,3,1)
    ax21.set_xlim((12,25))
    ax21.set_ylim((-2,6))
    ax21.set_ylabel('r (mag)', fontdict=regtitle)
    ax21.set_title('Density of r vs g-i', fontdict=regtitle)
    sns.kdeplot(file[2][::50], file[3][::50], shade=True)

    ax22 = fig2.add_subplot(1,3,2)
    ax22.set_xlim((12,25))
    ax22.set_ylim((-2,6))
    ax22.set_title('r vs g-i as a function of RA', fontdict=regtitle)
    plot22 = ax22.scatter(file[2][::50], file[3][::50], c=file[6][::50])
    sns.kdeplot(file[2][::50], file[3][::50])
    fig1.colorbar(plot22, ax=ax22)

    ax23 = fig2.add_subplot(1,3,3)
    ax23.set_xlim((12,25))
    ax23.set_ylim((-2,6))
    ax23.set_title('r vs g-i as a function of dec', fontdict=regtitle)
    plot23 = ax23.scatter(file[2][::50], file[3][::50], c=file[7][::50])
    sns.kdeplot(file[2][::50], file[3][::50])
    fig2.colorbar(plot23, ax=ax23)
```

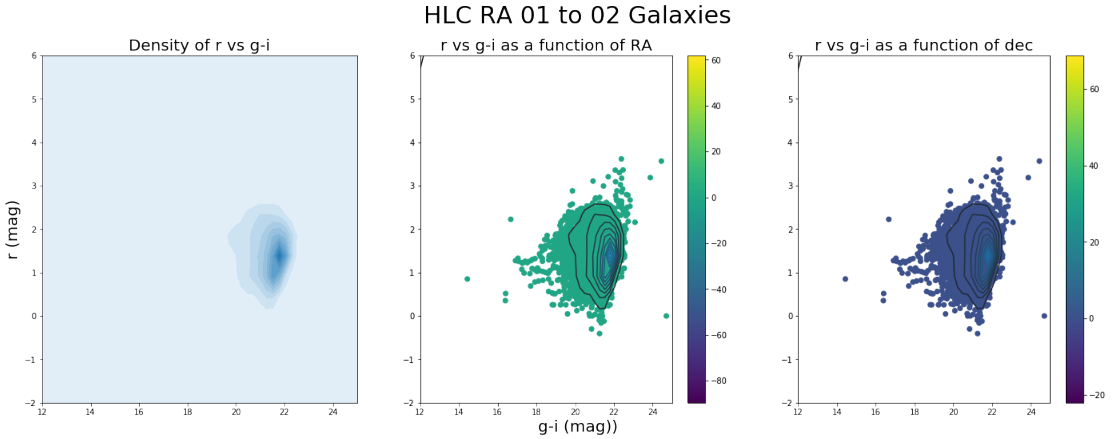
[11]:

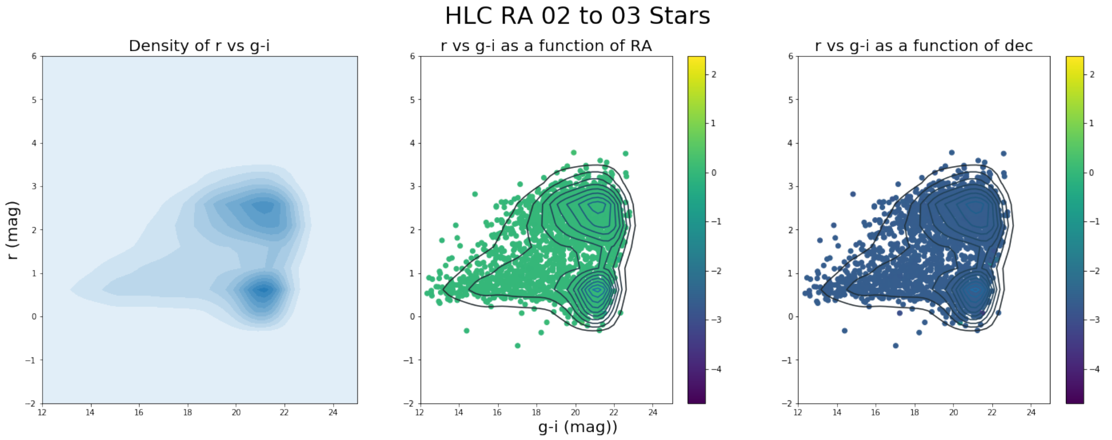
```
plotting(HLC1stats, 'HLC RA 00 to 01')
plotting(HLC2stats, 'HLC RA 01 to 02')
plotting(HLC3stats, 'HLC RA 02 to 03')
plotting(HLC4stats, 'HLC RA 03 to 04')
plotting(HLC5stats, 'HLC RA 20 to 21')
plotting(HLC6stats, 'HLC RA 21 to 22')
plotting(HLC7stats, 'HLC RA 22 to 23')
plotting(HLC8stats, 'HLC RA 23 to 24')
```



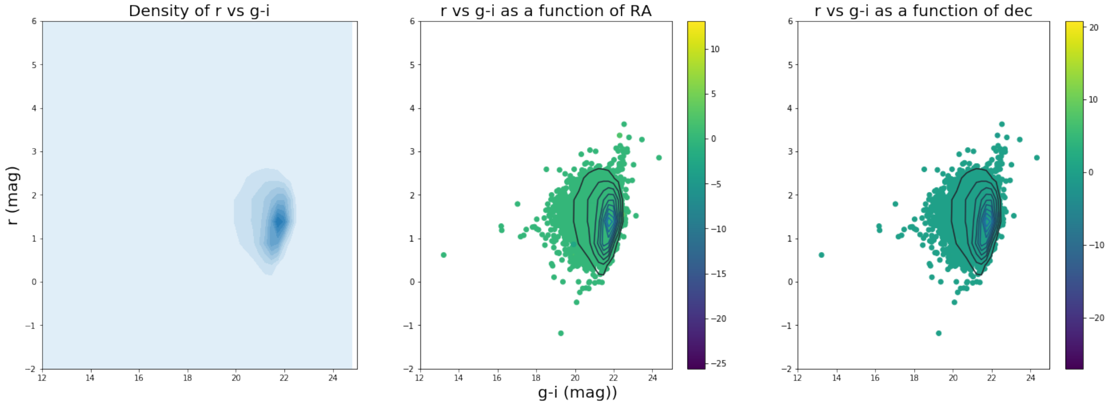



HLC RA 01 to 02 Stars


HLC RA 01 to 02 Galaxies


HLC RA 02 to 03 Stars

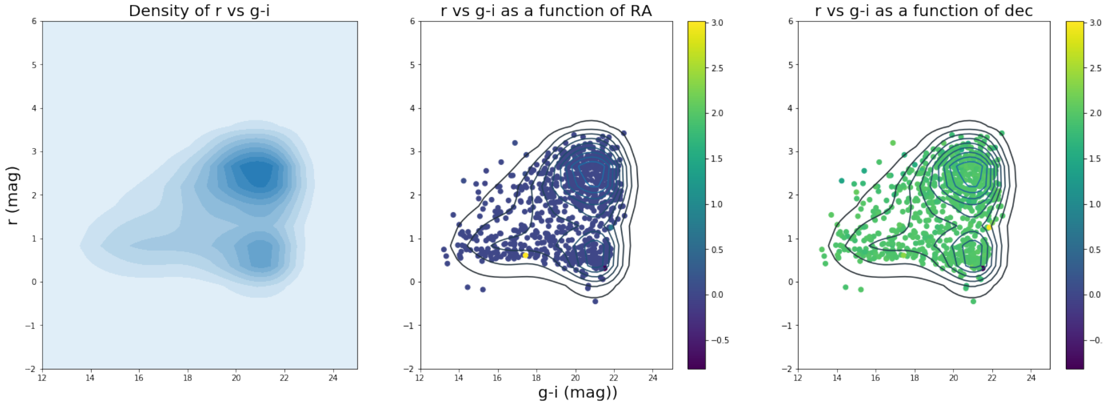
HLC RA 02 to 03 Galaxies


Density of r vs g-i

r vs g-i as a function of RA

r vs g-i as a function of dec

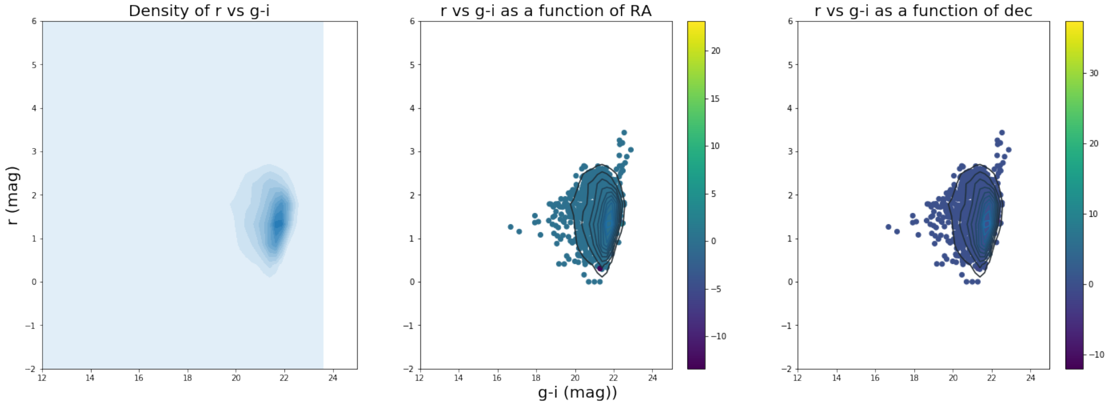
HLC RA 03 to 04 Stars


Density of r vs g-i

r vs g-i as a function of RA

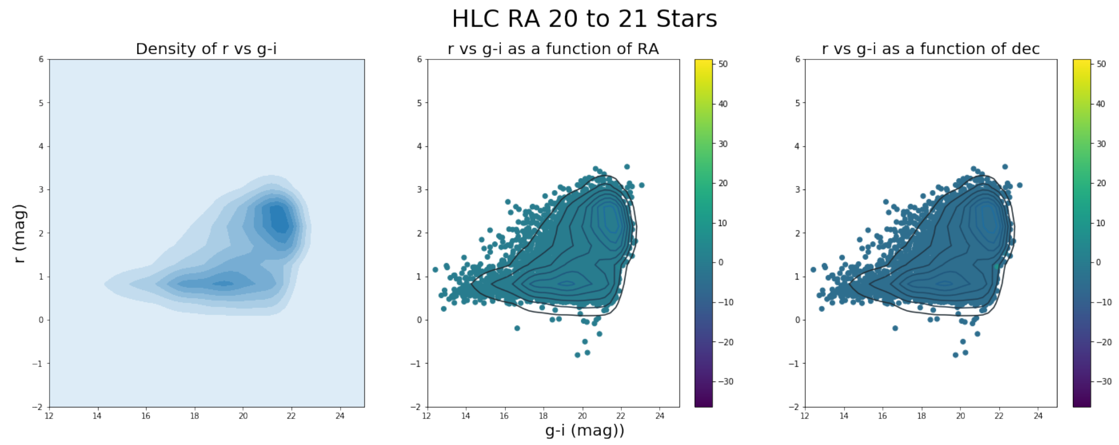
r vs g-i as a function of dec

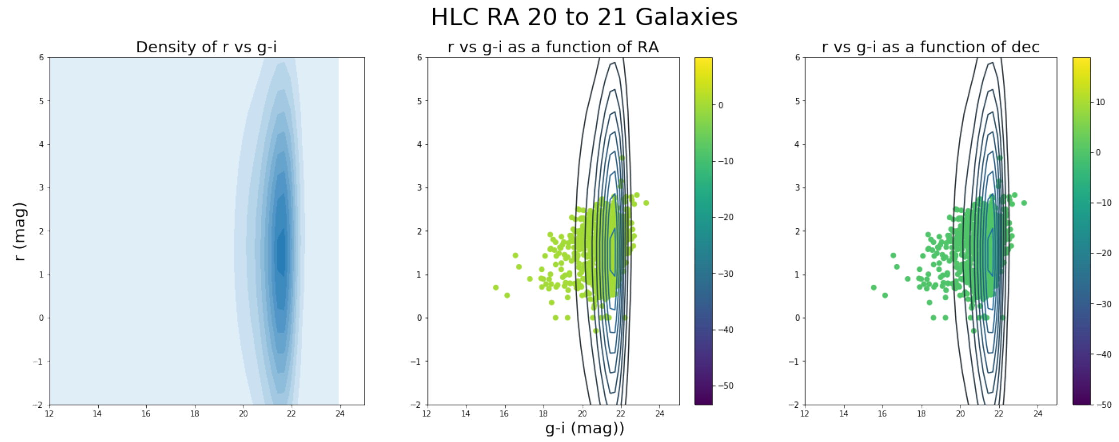
HLC RA 03 to 04 Galaxies

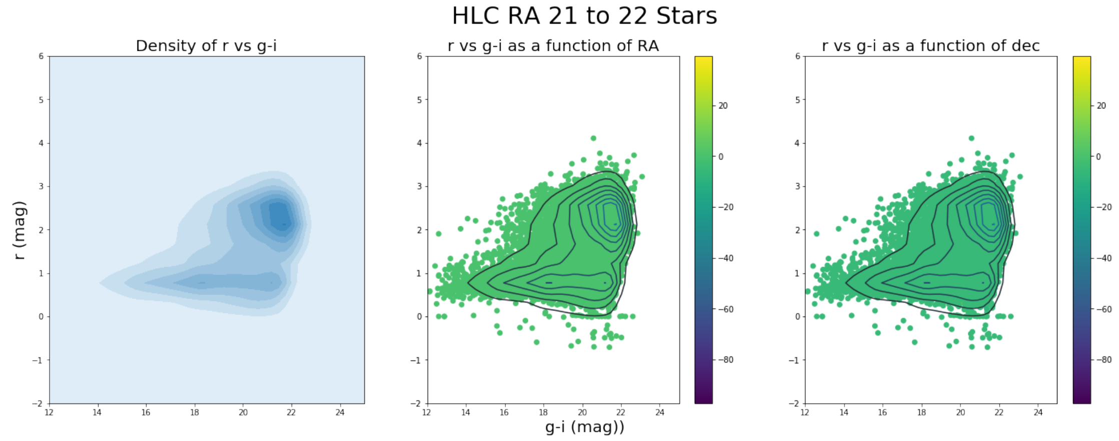

Density of r vs g-i

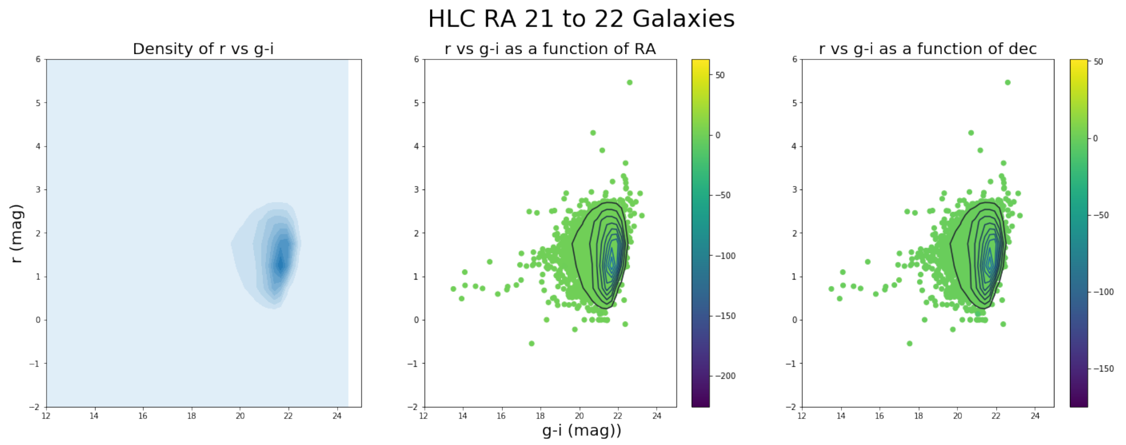
r vs g-i as a function of RA

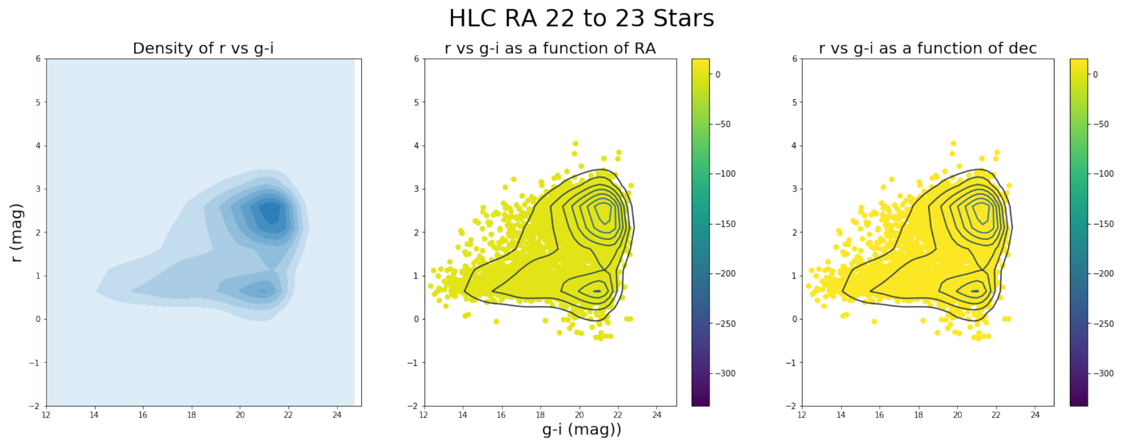
r vs g-i as a function of dec

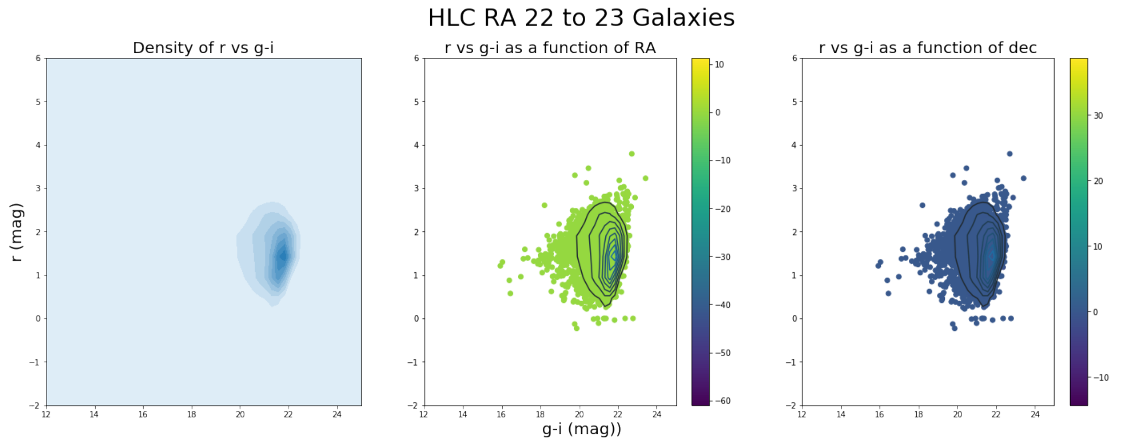

HLC RA 20 to 21 Stars
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
50
40
30
20
10
0
−10
−20
−30

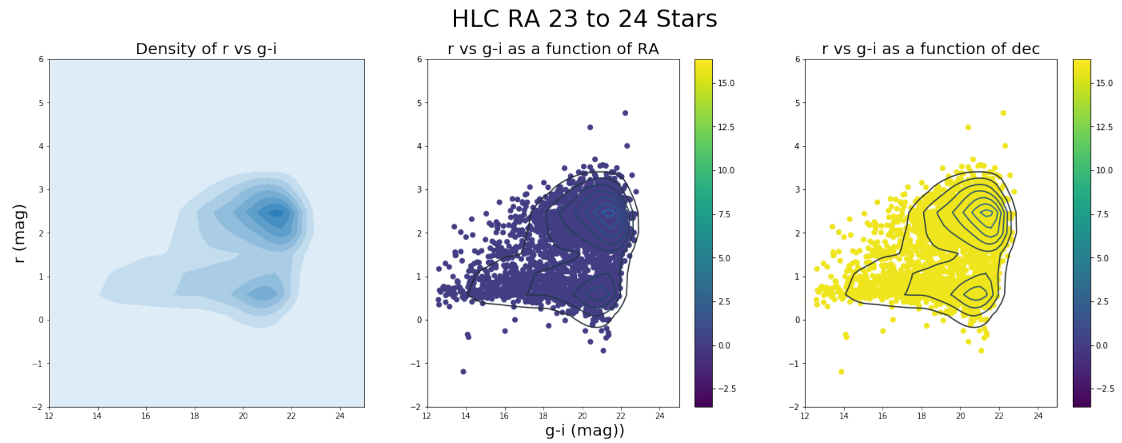

HLC RA 20 to 21 Galaxies
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
0
−10
−20
−30
−40
−50
10
0
−10
−20
−30
−40
−50

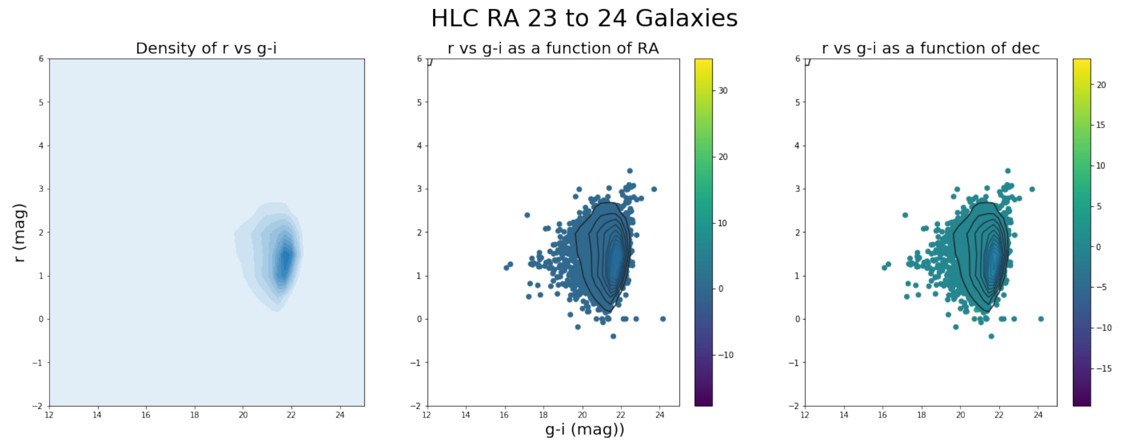

HLC RA 21 to 22 Stars
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
20
0
−20
−40
−60
−80


HLC RA 21 to 22 Galaxies
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
50
0
−50
−100
−150
−200


HLC RA 22 to 23 Stars
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
0
−50
−100
−150
−200
−250
−300


HLC RA 22 to 23 Galaxies
Density of r vs g-i
r vs g-i as a function of RA
r vs g-i as a function of dec
r (mag)
g-i (mag))
6
5
4
3
2
1
0
−1
−2
12
14
16
18
20
22
24
10
0
−10
−20
−30
−40
−50
−60
30
20
10
0
−10

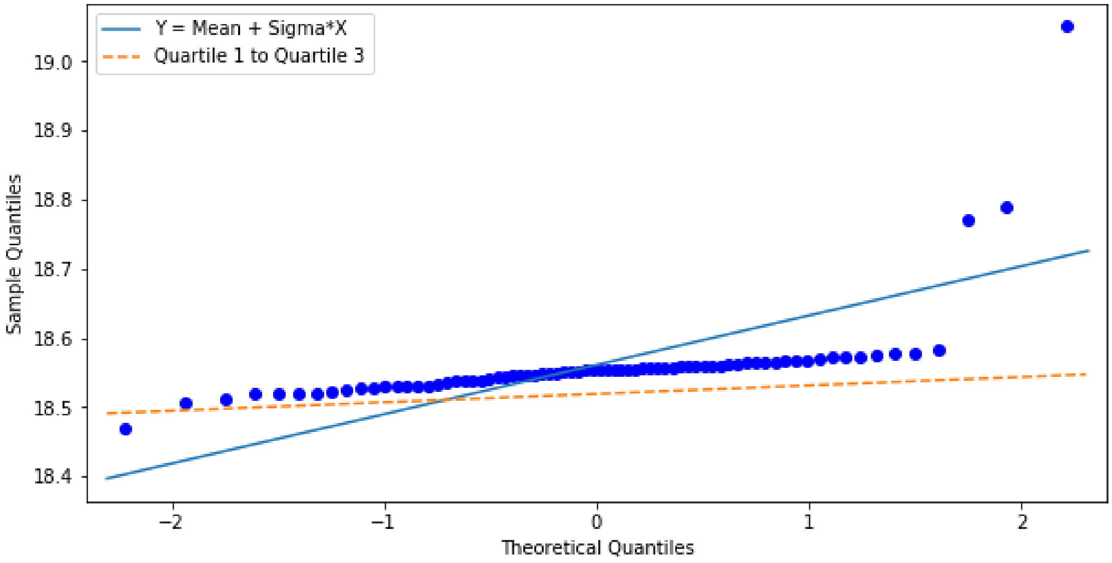




Problem 2

```
[12]: import statsmodels.api as sm

      introdemo = pd.read_csv(os.path.join('/Users/christandoi/projects/
       ↪ast596_2020_Spring/data/01/','IntroStat_demo.csv'))

      outliers = introdemo['mag.outlier']
      out_mean = np.mean(outliers)
      out_var = np.var(outliers)
      out_sigma = np.sqrt(out_var)

      fig = plt.figure(figsize=(10,5))
      ax = fig.add_subplot(1,1,1)

      sm.qqplot(outliers, ax=ax)
      nmin, nmax = ax.get_xlim()
```

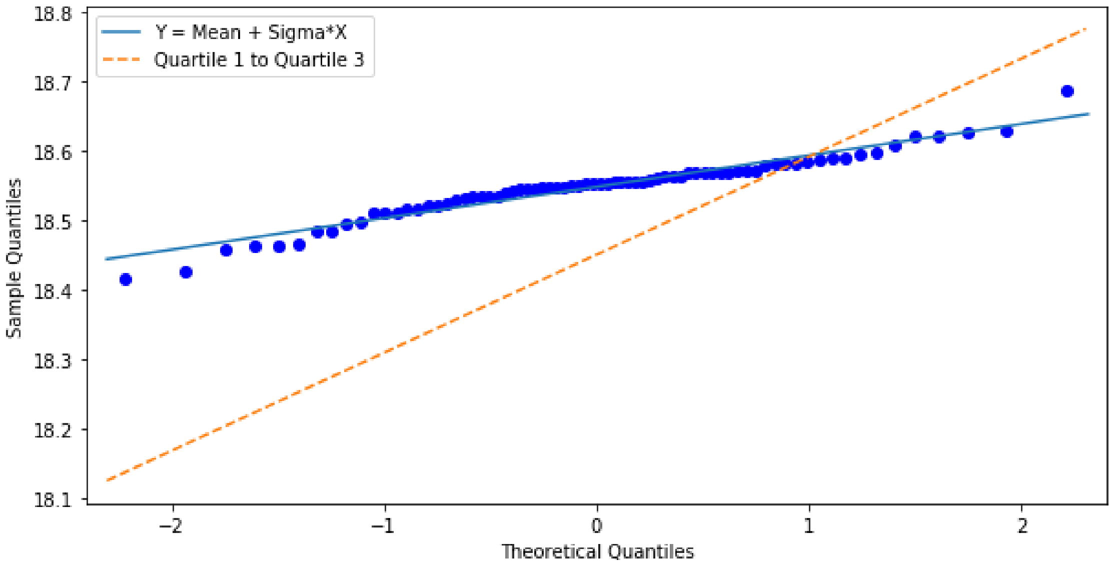
```
nvals = np.arange(nmin, nmax+0.01, 0.01)
y = out_mean + (out_sigma*nvals)
ax.plot(nvals, y, label='Y = Mean + Sigma*X')

quartile_first = np.quantile(outliers, 0.25)
quartile_third = np.quantile(outliers, 0.75)

sorted_outliers = list(outliers)
sorted_outliers.sort()

yi1 = np.where(np.isclose(sorted_outliers, quartile_first))[0][0]
yi2 = np.where(np.isclose(sorted_outliers, quartile_third))[0][0]

y1 = outliers[yi1]
y2 = outliers[yi2]

x1 = (yi1-.05)/len(sorted_outliers)
x2 = (yi2-.05)/len(sorted_outliers)

m = (y2-y1)/(x2-x1)
x = np.linspace(nmin,nmax,1000)
line = (m*x) - (m*x1) + y1
plt.plot(x, line, '--', label='Quartile 1 to Quartile 3')
plt.legend()

plt.show()
print(quartile_first, quartile_third)
```



```
18.5373952724961 18.56298225840465
```

[14]:

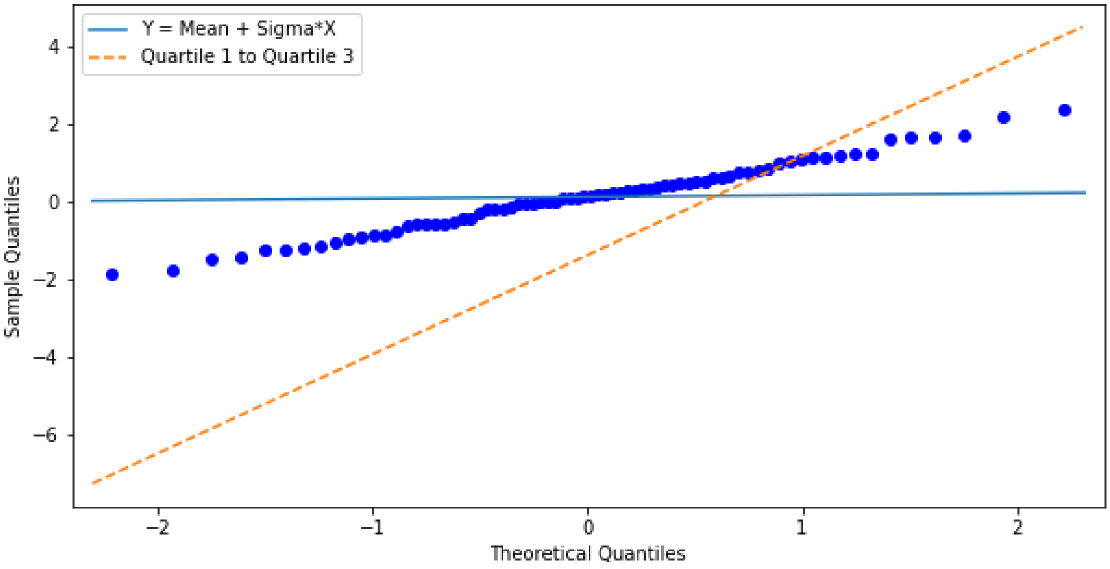
```python
het = introdemo['mag.het']
het_mean = np.mean(het)
het_var = np.var(het)
het_sigma = np.sqrt(het_var)

fig = plt.figure(figsize=(10,5))
ax = fig.add_subplot(1,1,1)

sm.qqplot(het, ax=ax)
nmin, nmax = ax.get_xlim()
nvals = np.arange(nmin, nmax+0.01, 0.01)
y_het = het_mean + (het_sigma*nvals)
ax.plot(nvals, y_het, label='Y = Mean + Sigma*X')

quartile_first = np.quantile(het, 0.25)
quartile_third = np.quantile(het, 0.75)

sorted_het = list(het)
sorted_het.sort()

#note: i have to change the absolute tolerance to score a hit with np.isclose
yi1 = np.where(np.isclose(sorted_het, quartile_first, atol=1e-3))[0][0]
yi2 = np.where(np.isclose(sorted_het, quartile_third, atol=5e-4))[0][0]

y1 = het[yi1]
y2 = het[yi2]

x1 = (yi1-.05)/len(sorted_het)
x2 = (yi2-.05)/len(sorted_het)

m = (y2-y1)/(x2-x1)
x = np.linspace(nmin,nmax,1000)
line = (m*x) - (m*x1) + y1
plt.plot(x, line, '--', label='Quartile 1 to Quartile 3')
plt.legend()

plt.show()
```



[15]:

```
het_error = introdemo['mag.het.error']
het_mean = np.mean(het)
adj_het = (het-np.mean(het))/het_error
adj_het_mean = np.mean(adj_het)
adj_het_var = np.var(adj_het)
adj_het_sigma = np.sqrt(het_var)

fig = plt.figure(figsize=(10,5))
ax = fig.add_subplot(1,1,1)

sm.qqplot(adj_het, ax=ax)
nmin, nmax = ax.get_xlim()
nvals = np.arange(nmin, nmax+0.01, 0.01)
y_adj_het = adj_het_mean + (adj_het_sigma*nvals)
ax.plot(nvals, y_adj_het, label='Y = Mean + Sigma*X')

quartile_first = np.quantile(adj_het, 0.25)
quartile_third = np.quantile(adj_het, 0.75)

sorted_adj_het = list(adj_het)
sorted_adj_het.sort()

#note: i have to change the absolute tolerance to score a hit with np.isclose
yi1 = np.where(np.isclose(sorted_adj_het, quartile_first, atol=1e-2))[0][0]
yi2 = np.where(np.isclose(sorted_adj_het, quartile_third, atol=1e-2))[0][0]

y1 = adj_het[yi1]
y2 = adj_het[yi2]
```

```
x1 = (yi1-.05)/len(sorted_adj_het)
x2 = (yi2-.05)/len(sorted_adj_het)

m = (y2-y1)/(x2-x1)
x = np.linspace(nmin,nmax,1000)
line = (m*x) - (m*x1) + y1
plt.plot(x, line, '--', label='Quartile 1 to Quartile 3')
plt.legend()
plt.show()
```



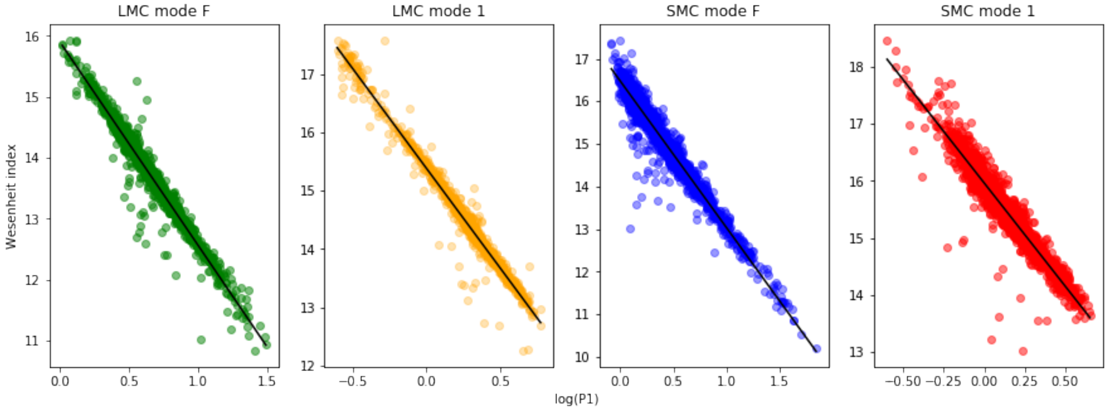
Problem 3

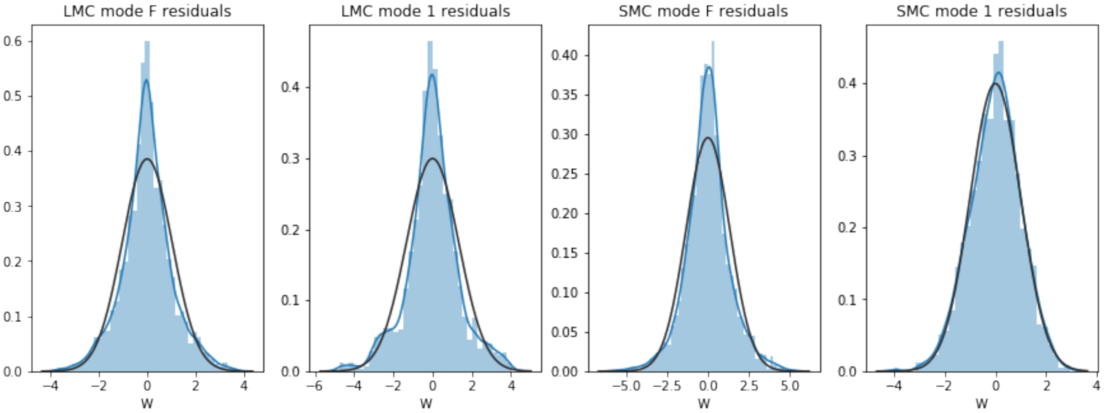
[83]:

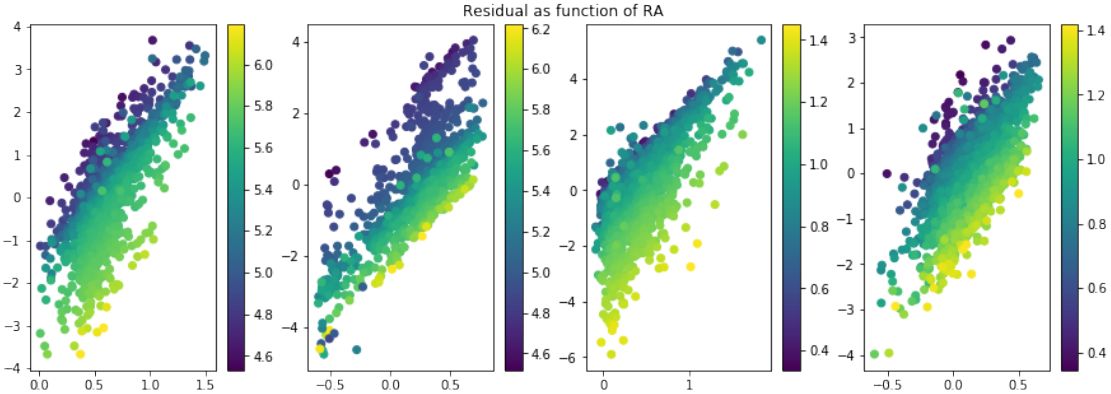
```
from scipy.stats import norm

cepheidsdata = pd.read_csv(os.path.join('/Users/christandoi/projects/
 ↪ast596_2020_Spring/data/01/','Cepheids.csv'))

LMC_F = cepheidsdata[(cepheidsdata.Cloud == 'LMC') & (cepheidsdata.Mode == 'F')]
LMC_1 = cepheidsdata[(cepheidsdata.Cloud == 'LMC') & (cepheidsdata.Mode == '1')]
SMC_F = cepheidsdata[(cepheidsdata.Cloud == 'SMC') & (cepheidsdata.Mode == 'F')]
SMC_1 = cepheidsdata[(cepheidsdata.Cloud == 'SMC') & (cepheidsdata.Mode == '1')]

fig = plt.figure(figsize=(15,5))
ax = fig.add_subplot(1,1,1)
ax.set_xlabel('log(P1)')
ax.spines['top'].set_color('none')
ax.spines['bottom'].set_color('none')
ax.spines['left'].set_color('none')
ax.spines['right'].set_color('none')
ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)

ax1 = fig.add_subplot(1,4,1)
ax1.set_title('LMC mode F')
ax1.set_ylabel('Wesenheit index')
ax1.scatter(LMC_F['logP1'], LMC_F['W'], color='g', alpha=0.5)
line1 = np.poly1d(np.polyfit(LMC_F['logP1'], LMC_F['W'], 1))(np.
 ↪unique(LMC_F['logP1']))
ax1.plot(np.unique(LMC_F['logP1']), line1, color='k')
residual1 = (line1 - LMC_F['W'])

ax2 = fig.add_subplot(1,4,2)
ax2.set_title('LMC mode 1')
ax2.scatter(LMC_1['logP1'], LMC_1['W'], color='orange', alpha=0.3)
line2 = np.poly1d(np.polyfit(LMC_1['logP1'], LMC_1['W'], 1))(np.
 ↪unique(LMC_1['logP1']))
ax2.plot(np.unique(LMC_1['logP1']), line2, color='k')
residual2 = (line2 - LMC_1['W'])

ax3 = fig.add_subplot(1,4,3)
ax3.set_title('SMC mode F')
ax3.scatter(SMC_F['logP1'], SMC_F['W'], color='blue', alpha=0.4)
line3 = np.poly1d(np.polyfit(SMC_F['logP1'], SMC_F['W'], 1))(np.
 ↪unique(SMC_F['logP1']))
ax3.plot(np.unique(SMC_F['logP1']), line3, color='k')
residual3 = (line3 - SMC_F['W'])

ax4 = fig.add_subplot(1,4,4)
ax4.set_title('SMC mode 1')
ax4.scatter(SMC_1['logP1'], SMC_1['W'], color='red', alpha=0.5)
line4 = np.poly1d(np.polyfit(SMC_1['logP1'], SMC_1['W'], 1))(np.
 ↪unique(SMC_1['logP1']))
ax4.plot(np.unique(SMC_1['logP1']), line4, color='k')
residual4 = (line4 - SMC_1['W'])

fig2 = plt.figure(figsize=(15,5))
ax = fig2.add_subplot(1,1,1)
ax.spines['top'].set_color('none')
ax.spines['bottom'].set_color('none')
ax.spines['left'].set_color('none')
ax.spines['right'].set_color('none')
ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)

ax21 = fig2.add_subplot(1,4,1)
ax21.set_title('LMC mode F residuals')
```

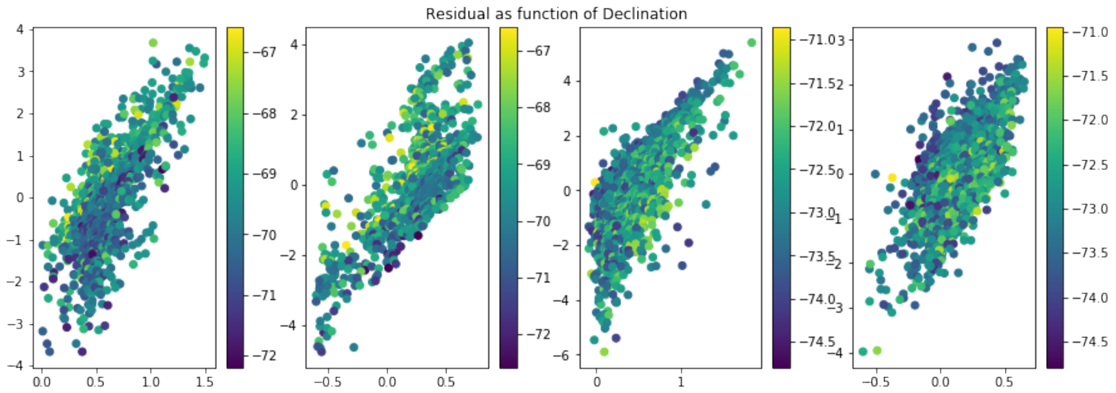
```
sns.distplot(residual1, fit=norm)

ax22 = fig2.add_subplot(1,4,2)
ax22.set_title('LMC mode 1 residuals')
sns.distplot(residual2, fit=norm)

ax23 = fig2.add_subplot(1,4,3)
ax23.set_title('SMC mode F residuals')
sns.distplot(residual3, fit=norm)

ax24 = fig2.add_subplot(1,4,4)
ax24.set_title('SMC mode 1 residuals')
sns.distplot(residual4, fit=norm)

fig3 = plt.figure(figsize=(15,5))
ax = fig3.add_subplot(1,1,1)
ax.set_title('Residual as function of RA')
ax.spines['top'].set_color('none')
ax.spines['bottom'].set_color('none')
ax.spines['left'].set_color('none')
ax.spines['right'].set_color('none')
ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)

ax31 = fig3.add_subplot(1,4,1)
plot31 = ax31.scatter(LMC_F['logP1'], residual1, c=LMC_F['RA0'])
fig3.colorbar(plot31, ax=ax31)

ax32 = fig3.add_subplot(1,4,2)
plot32 = ax32.scatter(LMC_1['logP1'], residual2, c=LMC_1['RA0'])
fig3.colorbar(plot32, ax=ax32)

ax33 = fig3.add_subplot(1,4,3)
plot33 = ax33.scatter(SMC_F['logP1'], residual3, c=SMC_F['RA0'])
fig3.colorbar(plot33, ax=ax33)

ax34 = fig3.add_subplot(1,4,4)
plot34 = ax34.scatter(SMC_1['logP1'], residual4, c=SMC_1['RA0'])
fig3.colorbar(plot34, ax=ax34)

fig4 = plt.figure(figsize=(15,5))
ax = fig4.add_subplot(1,1,1)
ax.set_title('Residual as function of Declination')
ax.spines['top'].set_color('none')
ax.spines['bottom'].set_color('none')
ax.spines['left'].set_color('none')
ax.spines['right'].set_color('none')
ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)
```

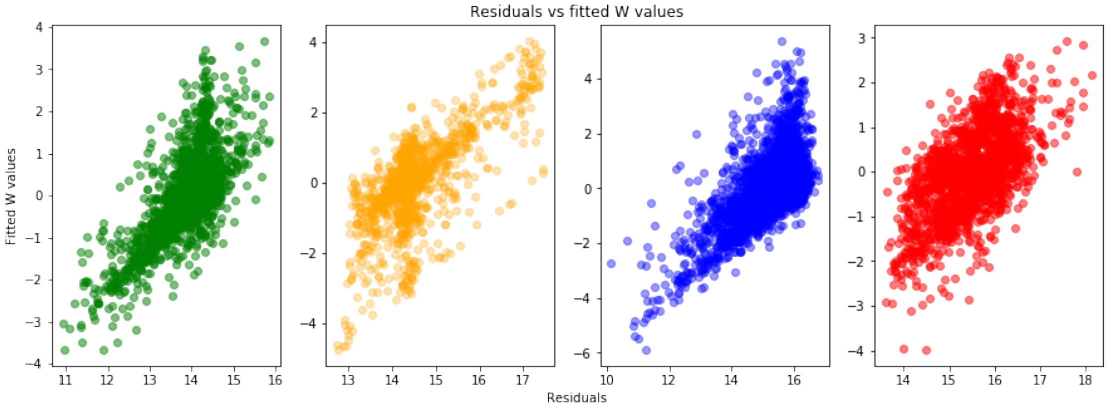
```
ax41 = fig4.add_subplot(1,4,1)
plot41 = ax41.scatter(LMC_F['logP1'], residual1, c=LMC_F['Decl0'])
fig4.colorbar(plot41, ax=ax41)

ax42 = fig4.add_subplot(1,4,2)
plot42 = ax42.scatter(LMC_1['logP1'], residual2, c=LMC_1['Decl0'])
fig4.colorbar(plot42, ax=ax42)

ax43 = fig4.add_subplot(1,4,3)
plot43 = ax43.scatter(SMC_F['logP1'], residual3, c=SMC_F['Decl0'])
fig4.colorbar(plot43, ax=ax43)

ax44 = fig4.add_subplot(1,4,4)
plot44 = ax44.scatter(SMC_1['logP1'], residual4, c=SMC_1['Decl0'])
fig4.colorbar(plot44, ax=ax44)

fig5 = plt.figure(figsize=(15,5))
ax = fig5.add_subplot(1,1,1)
ax.set_title('Residuals vs fitted W values')
ax.set_xlabel('Residuals')
ax.spines['top'].set_color('none')
ax.spines['bottom'].set_color('none')
ax.spines['left'].set_color('none')
ax.spines['right'].set_color('none')
ax.tick_params(labelcolor='w', top=False, bottom=False, left=False, right=False)

ax51 = fig5.add_subplot(1,4,1)
ax51.set_ylabel('Fitted W values')
plot51 = ax51.scatter(line1, residual1, color='g', alpha=0.5)

ax52 = fig5.add_subplot(1,4,2)
plot52 = ax52.scatter(line2, residual2, color='orange', alpha=0.3)

ax53 = fig5.add_subplot(1,4,3)
plot53 = ax53.scatter(line3, residual3, color='blue', alpha=0.4)

ax54 = fig5.add_subplot(1,4,4)
plot54 = ax54.scatter(line4, residual4, color='red', alpha=0.5)
```


LMC mode F
LMC mode 1
SMC mode F
SMC mode 1
Wesenheit index
log(P1)


LMC mode F residuals
LMC mode 1 residuals
SMC mode F residuals
SMC mode 1 residuals
W
W
W
W


Residual as function of RA


Residual as function of Declination
4
3
2
1
0
−1
−2
−3
−4
0.0
0.5
1.0
1.5
−67
−68
−69
−70
−71
−72
4
2
0
−2
−4
−0.5
0.0
0.5
−67
−68
−69
−70
−71
−72
4
2
0
−2
−4
−6
0
1
−71.0
−71.5
−72.0
−72.5
−73.0
−73.5
−74.0
−74.5
3
2
1
0
−1
−2
−3
−4
−0.5
0.0
0.5
−71.0
−71.5
−72.0
−72.5
−73.0
−73.5
−74.0
−74.5


Residuals vs fitted W values
Fitted W values
Residuals
4
3
2
1
0
−1
−2
−3
−4
11
12
13
14
15
16
4
2
0
−2
−4
13
14
15
16
17
4
2
0
−2
−4
−6
10
12
14
16
3
2
1
0
−1
−2
−3
−4
14
15
16
17
18