Data analysis in Astronomy {-}

Homework 1 {-}

Due Oct 11 11:59 pm

- Please submit your homework through NTU cool
- Please write your answers in English

Note: for this homework, please do not discuss it with your friends and classmates. Consider that you are the only person who has the dataset and wants to explore it. We will discuss it next week.

Answer these questions in the jupyter notebook and output as a pdf file which includes the code you wrote.

1. (Data exploration) Use the tools (SDSS navigate tools / topcats / Python) discussed in the lecture to explore this dataset (https://www.dropbox.com/s/xakm1krrpqrll8w/a_catalog_with_50_objects.fits?dl=0) and answer the following questions. Please describe what you observe and find and offer plots and evidence when needed: {-}

(45\%)

a). Make a color-magnitude plot with y-axis (SDSS_u-SDSS_z) and x-axis (SDSS_z). (10%)

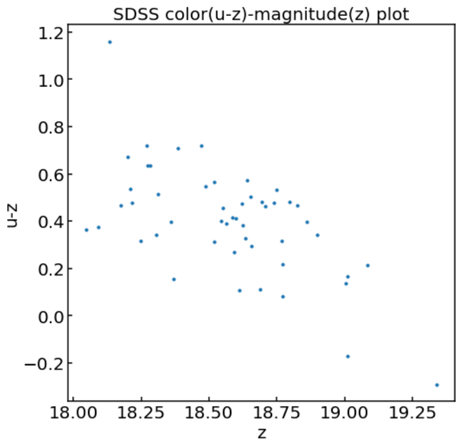
In [2]:
```python
# Answer a)
### List all the attributes names in the .fits file
import astropy.io.fits as pf

data = pf.open('a_catalog_with_50_objects.fits')
catalog = data[1].data
catalog.names
```

Out[2]:
```
['RA', 'DEC', 'SDSS_u', 'SDSS_g', 'SDSS_r', 'SDSS_i', 'SDSS_z']
```

In [3]:
```python
# Define plotting style
import numpy as np
import matplotlib.pyplot as plt
def my_plot_style():

    params = {'legend.fontsize': 20,
             'axes.labelsize': 20,
             'axes.titlesize':20,
             'xtick.labelsize':20,
             'ytick.labelsize':20,
             'xtick.major.size':5,
              'xtick.minor.size':2.5,
             'ytick.major.size':5,
              'ytick.minor.size':2.5,
             'figure.facecolor':'w',
             #'lines.linewidth' : 1.5,
              'xtick.major.width':1.5,
              'ytick.major.width':1.5,
              'xtick.minor.width':1.5,
              'ytick.minor.width':1.5,
              'axes.linewidth':1.5,
              'xtick.direction':'in',
              'ytick.direction':'in',
             'ytick.labelleft':True,
              'text.usetex' : False,
             'font.family': 'sans-serif'}

    plt.rcParams.update(params)
```

In [8]:
```python
# Plot the (u-z)-z scattering points
my_plot_style()

plt.figure(figsize=(8,8))
plt.title('SDSS color(u-z)-magnitude(z) plot')
plt.xlabel('z')
plt.ylabel('u-z')
plt.scatter(catalog['SDSS_z'],catalog['SDSS_u']-catalog['SDSS_z'],s=10)
```

Out[8]:
```
<matplotlib.collections.PathCollection at 0x7fe490a5a580>
```



3-5 sentences for each question.

b). What are their photometric properties? (Are they extended or point sources? Are they blue or red?) (5%)

[Describe]

c). What are the properties of their spectra? (Are there any emission lines and/or continuum?) (5%)

[Describe]

d). What kind of sources are we looking at? (stars, quasars, or galaxies) What are the evidence supporting your conclusion? (10%)

[Describe]

Answer:

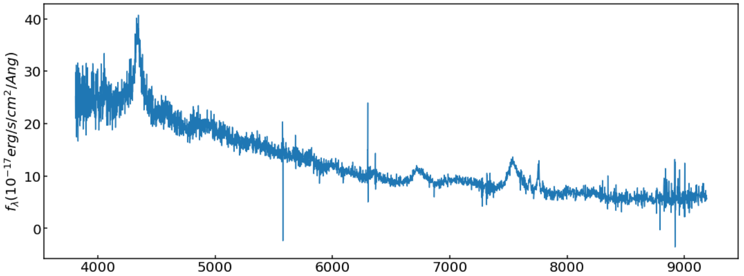
b) I used navigation tools in SDSS to search the spectra and data of these objects. They seem to be point sources, unlike the extended sources of galaxies. I look at the images of these objects, which are blue, and you can also see this from their spectra because their fluxes at blue light are greater than red light.

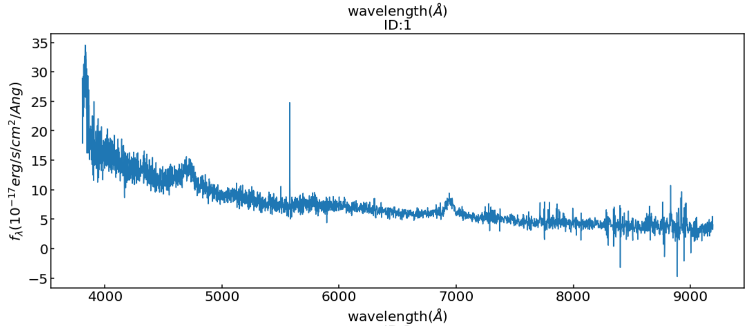
c) They have many emission lines, and have continuum at other places. They have very strong Ly-alpha lines and many ion lines. Their spectra tend to have higher intensity at short wavelengths around visible light area. Many of their emission lines are quite broad, more like a bulge than a thin line.

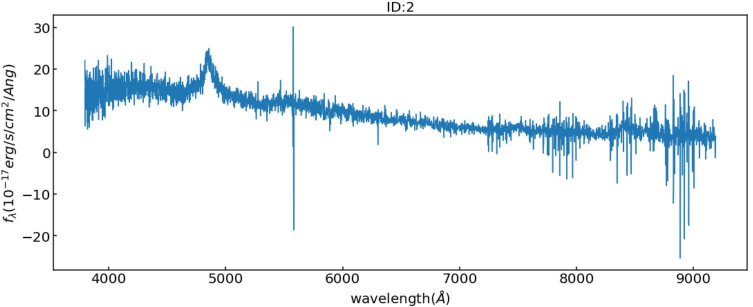
d) They are quasars according to SDSS's classification. Actually, we can judge them to be quasars from their spectra. For example, these their emission lines tend to have high redshifts, which are common characteristics of quasars. Also, unlike stars' spectra have many absorption line on black radiation spectra, these objects spectra have a lot of emission lines. Moreover, their emission lines are broad, unlike the thin lines of stars' spectra.

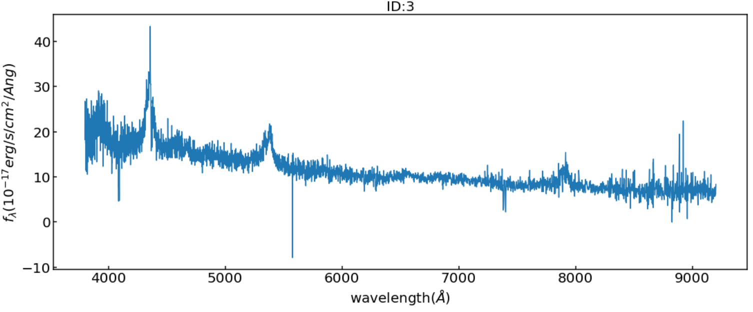
In [19]:
```python
# I used this code to plot the spectra of the first ten objects in the 50 object fits file.
# I cannot find some of their spectra probably due to SDSS's data problem.
from astropy import units as u
from astropy.coordinates import SkyCoord
from astropy.io import fits
from astroquery.sdss import SDSS
from matplotlib import pyplot as plt

objCoords = []
for i in range(len(catalog['ra'])):
    objCoords.append(SkyCoord(catalog['ra'][i], catalog['dec'][i], unit="deg"))
    #print(SkyCoord(catalog['ra'][i], catalog['dec'][i], unit="deg"))
spectrum_data = []
tot_num = 10
f = plt.figure(figsize=(18,8*tot_num))

for i in range(0,tot_num):
    try:
        objCoord = objCoords[i]
        xObj = SDSS.query_region(objCoord, spectro=True)
        imgObj = SDSS.get_spectra(matches=xObj)[0][1]
        spec_data = imgObj.data
        loglam=spec_data['loglam']  # Logaritmo de la longitud de onda
        flux=spec_data['flux'] # Flujo medido en unidades de Ergs/cm^2/s/AA
        spectrum_data.append([loglam,flux])

        ax = f.add_subplot(tot_num,1,i+1)
        plt.title("ID:{:d}".format(i))
        plt.xlabel(r'wavelength($\AA$)')
        plt.ylabel(r'$f_{\lambda}$($10^{-17} erg/s/cm^2/Ang$)')
        ax.plot(10**loglam,flux)


    except:
        i=i # do nothing
```

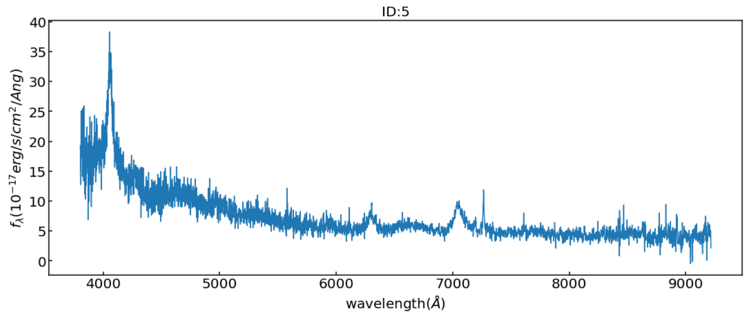
```
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
```

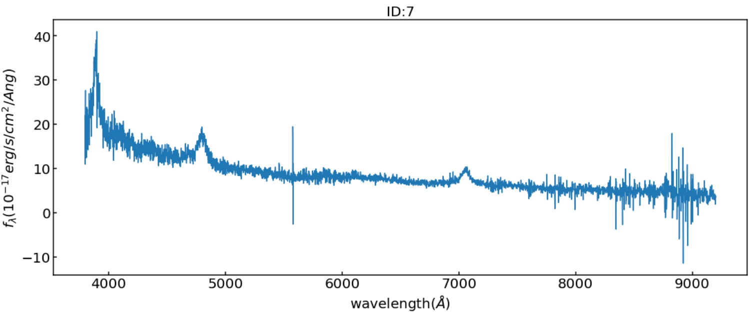


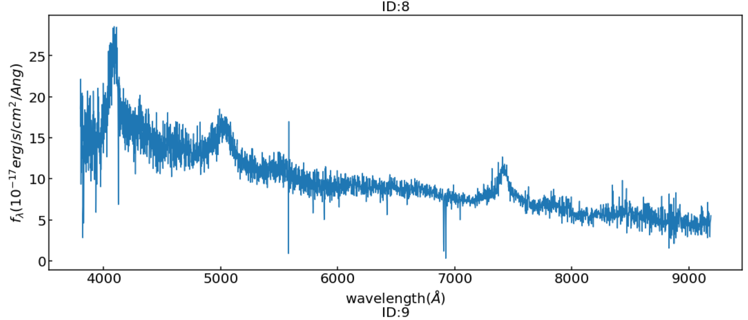


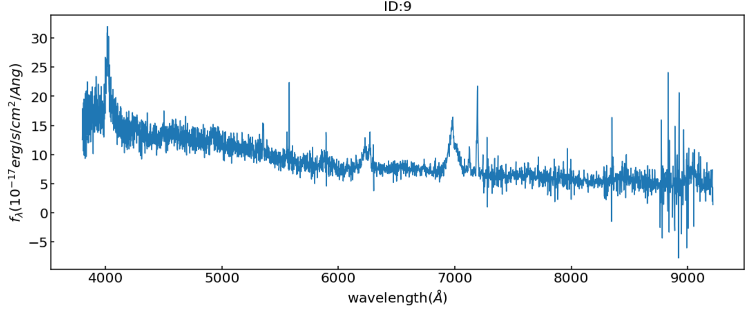












e). Do you find any interesting or unexpected objects? Download the spectra of the interesting / unexpected objects and plot it. (15%)

Describe why you think they are interesting. (hints: check the outliers in Figure 1)

Answer

e) Below I plot the spectrum of the object whose u-z value is above 1.0, probably due to severe redshift, and is an outlier in fig.1. You will notice a very high peak near 400nm, which is Ly-alpha line according to SDSS website's description. Normally, Ly-alpha line is at around 100nm, but here it is redshifted to 400nm.

In [27]:

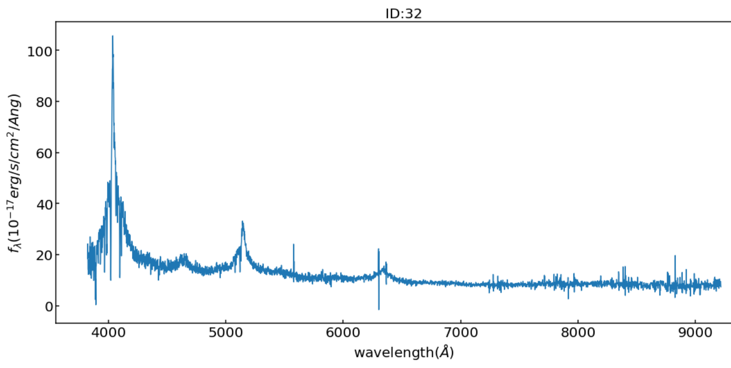
```
### I pick out the object which has u-z>1.0
u_z = catalog['SDSS_u']-catalog['SDSS_z']
search = np.where((u_z>1.0))
id = search[0][0]

print('Sky coordinates:'+str(SkyCoord(catalog['ra'][id], catalog['dec'][id], unit="deg")))
print('outlier (u-z>1.0) id:'+str(id))

from astropy import units as u
from astropy.coordinates import SkyCoord
from astropy.io import fits
from astroquery.sdss import SDSS
from matplotlib import pyplot as plt

objCoord = SkyCoord(catalog['ra'][id], catalog['dec'][id], unit="deg")
xObj = SDSS.query_region(objCoord, spectro=True)
xObj=xObj[0:1]
imgObj = SDSS.get_spectra(matches=xObj)[0][1]
spec_data = imgObj.data
loglam=spec_data['loglam']  # Logaritmo de la longitud de onda
flux=spec_data['flux'] # Flujo medido en unidades de Ergs/cm^2/s/AA
spectrum_data.append([loglam,flux])

f = plt.figure(figsize=(18,8))
ax = f.add_subplot(1,1,1)
plt.title("ID:{:d}".format(i))
plt.xlabel(r'wavelength($\AA$)')
plt.ylabel(r'$f_{\lambda}$($10^{-17} erg/s/cm^2/Ang$)')
ax.plot(10**loglam,flux)
```

```
Sky coordinates:<SkyCoord (ICRS): (ra, dec) in deg
    (166.85718909, 17.48409708)>
outlier (u-z>1.0) id:32
```

```
/home/albert/.local/lib/python3.8/site-packages/astroquery/sdss/core.py:862: VisibleDeprecationWarning: Reading
unicode strings without specifying the encoding argument is deprecated. Set the encoding, use None for the syst
em default.
  arr = np.atleast_1d(np.genfromtxt(io.BytesIO(response.content),
```

Out[27]:

```
[<matplotlib.lines.Line2D at 0x7fe48e1bf760>]
```



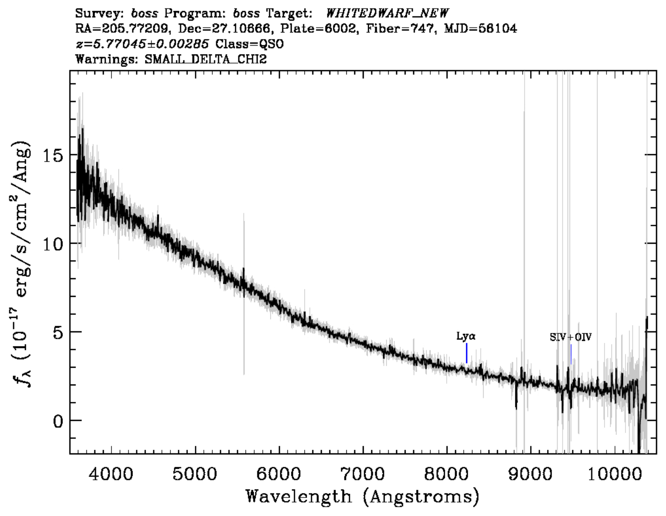
Below I found another spectrum of an object whose u-z value is smaller than 0.0.

In [29]:

```
### I pick out the object which has u-z<0.0
u_z = catalog['SDSS_u']-catalog['SDSS_z']
search = np.where((u_z<0.0))
id = search[0][0]

print('Sky coordinates:'+str(SkyCoord(catalog['ra'][id], catalog['dec'][id], unit="deg")))
print('outlier (u-z<0.0) id:'+str(id))

# This is one of the outlier (u-z<0.0), and we cannot find this outlier's spectra through python's code.
# So, I go to SDSS's official website to get its spectrum and put it at below text.
```

```
Sky coordinates:<SkyCoord (ICRS): (ra, dec) in deg
    (205.77201497, 27.10666546)>
outlier (u-z<0.0) id:20
```

You can see that this outlier's spectrum has no obvious emission lines and absorption lines. More details were discussed in this paper: https://arxiv.org/abs/1711.01122. Actually this object is not quasar, but a white dwarf. In the paper, the author has said that this object was falsely classified into quasar (you can see this point through the nearly disappearing Ly-\alpha line in the above spectrum) Its name is J134305.302+270623.98, and it may be a low temperature DB white dwarf. Its low temperature caused the undeveloped Helium lines.



2. (Visualization) Use Python to plot (this dataset https://www.dropbox.com/s/ev10vdr6tjb955s/Simple_mass_metallicity_210927.fits dl=0) mass (mass) in x-axis and metallicity (OH) in y-axis. The xscale should be [8,12] and the yscale should be around [8, 9.5]. Do your best to make it clean and informative. {-}

Add your plot here.

(25\%)

In [2]:

```
### List all the attributes names in the .fits file
import astropy.io.fits as pf

data = pf.open('Simple_mass_metallicity_210927.fits')
catalog = data[1].data
catalog.names
```

Out[2]:

```
['mass', 'OH']
```

In [3]:

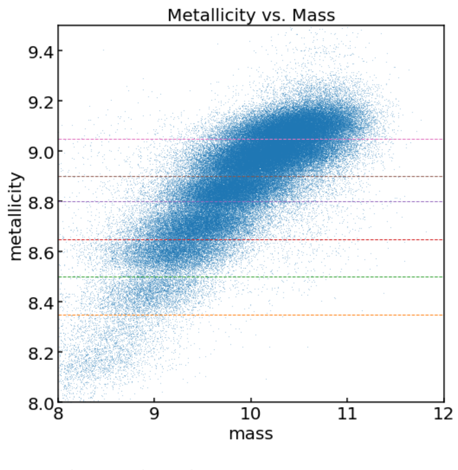
```
import numpy as np
import matplotlib.pyplot as plt
def my_plot_style():

    params = {'legend.fontsize': 20,
              'axes.labelsize': 20,
              'axes.titlesize':20,
              'xtick.labelsize':20,
              'ytick.labelsize':20,
              'xtick.major.size':5,
               'xtick.minor.size':2.5,
              'ytick.major.size':5,
               'ytick.minor.size':2.5,
              'figure.facecolor':'w',
              #'lines.linewidth' : 1.5,
               'xtick.major.width':1.5,
               'ytick.major.width':1.5,
               'xtick.minor.width':1.5,
               'ytick.minor.width':1.5,
               'axes.linewidth':1.5,
               'xtick.direction':'in',
               'ytick.direction':'in',
              'ytick.labelleft':True,
               'text.usetex' : False,
              'font.family': 'sans-serif'}

    plt.rcParams.update(params)
```

In [27]:

```
my_plot_style()

import matplotlib.cm as cm
plt.figure(figsize=(8,8))
plt.title('Metallicity vs. Mass')
plt.xlabel('mass')
plt.ylabel('metallicity')
plt.xlim([8,12])
plt.ylim([8,9.5])
plt.plot(catalog['mass'],catalog['OH'],'.',markeredgewidth=0.0,markersize=1.0)

# Plot lines seperating different populations
# You can see some gaps between the populations
x0=np.linspace(8,12,1000)
y0_value=8.35
y0=np.linspace(y0_value,y0_value,1000)

x1=np.linspace(8,12,1000)
y1_value=8.5
y1=np.linspace(y1_value,y1_value,1000)

x2=np.linspace(8,12,1000)
y2_value=8.65
y2=np.linspace(y2_value,y2_value,1000)

x3=np.linspace(8,12,1000)
y3_value=8.8
y3=np.linspace(y3_value,y3_value,1000)

x4=np.linspace(8,12,1000)
y4_value=8.9
y4=np.linspace(y4_value,y4_value,1000)

x5=np.linspace(8,12,1000)
y5_value=9.05
y5=np.linspace(y5_value,y5_value,1000)

plt.plot(x0,y0,linestyle='--',linewidth=1)
plt.plot(x1,y1,linestyle='--',linewidth=1)
plt.plot(x2,y2,linestyle='--',linewidth=1)
plt.plot(x3,y3,linestyle='--',linewidth=1)
plt.plot(x4,y4,linestyle='--',linewidth=1)
plt.plot(x5,y5,linestyle='--',linewidth=1)
```

Out[27]:

```
[<matplotlib.lines.Line2D at 0x7f2c60bbcbe0>]
```



3. Read Gunn (2020) https://ui.adsabs.harvard.edu/abs/2020ARA%26A..58....1G/abstract and answer the following questions: {-}

Please describe with your own words. (Note that direct copy and paste from the article is plagiarism.) {-}

(30%)

3-5 sentences for each question.

a). What makes SDSS so transformative according to Gunn? (10%) [Describe]

b). What are the unexpected key science results offered by SDSS? (10%) [Describe]

c). What did you learn after reading this article? (5%) [Describe]

d). What happened to the Hubble Space Telescope when it was just launched? [You might also need to google it.] (5%) [Describe]

Answer:

a) SDSS is transformative is that it has high quality data and open to everyone. The high data quality and homogenuity of the survy helped discovery of many scientific successes like weak lensing and high-redshifted quasars. And the basis of the high quality is that SDSS includes a large collaboration of people with many different skills in software, optics, and detectors, which is uncommon to most astronomers then, who often work in small groups. The openess of SDSS also enables astronomers worldwide to access large amount of high-quality data facilitating their numerical analysis in their research, creating unexpected scientific discoveries by different teams of researchers around globe.

b) Many important discoveries in cosmology and astronomy were discovered thanks to SDSS's data. For instance, weak lensing was found thanks to the high quality data (otherwise the noises and errors may deter you from observing it),and BAO matter clustering, which can be used as standard rulers for length scale in cosmology, was also not expected at first. Many other things like "green valley" in galaxy's properties' distribution, some results in asteroid populations, brown dwarfs, high-redshifted quasars(z~6, breaking record of z=4.9), large-scale structure results, and detection of stellar tidal streams in the Galaxy are all contributions thanks to SDSS.

c) What I have learned from this article is that one man cannot accomplish great things without other people's help. Gunn may be very intelligent and versatile, but he cannot do all the works in SDSS's project with his own strength. So, it's important for us to acquaint ourselves with many friends in different fields because they may be helpfuls someday. Another is his vision on future astronomy research. He sees the greater demand for astronomical data, and he seeks to solve the problem and thus facilitates astronomical research around world and even many generations to come.

d) Hubble Space Telescope(HST)'s mirror was accurately ground into wrong shape by Perkin-Elmer due to the null correctro's problem in this mirror's figuring process, and it caused severe spherical abberation. Although Kodah had ground a backup mirror, but since HST was already on the orbit around Earth, the problematic on HST could not be replaced. To avoid the entire project from abondoning, NASA sent astronomers to space to install optical correctors on HST to cancel the mirror's aberration.