Astro 376: Homework 4 (Python Assignment 2)

(Assigned on Tu. Apr 13, 2021. Due by 11:59pm CT on Th. Apr 22, 2021)

Overview of Assignment

- In this homework you will be applying what you have learned in the Python tutorial over the last few weeks, and going beyond the tasks you did in homework 3. For some questions, you will need to go beyond the Python tutorial and consult online help. We also assume that by now, you are comfortable with basic Mac OSX/Linux commands and some text editor of your choice (like Aquamacs or Vi). Recall that you can let Terminal choose a text editor for you if you type '`open file`'. If you need a refresher, please consult the Mac OSX/Linux tutorial that you practiced in class earlier. The number of points for each question is indicated in brackets, and the total score is 100 points.

- In this homework your assignment includes **identifying, statistically analyzing, and graphically illustrating the properties galaxies with the largest star formation rates and stellar masses at early epochs when the Universe up to half of its present age**. You will use a real scientific catalog from cutting-edge surveys of galaxies conducted with NASA's *Hubble Space Telescope* and Spitzer Space Telescope, as well as ground-based observations at many different wavelengths. Your main task is to write an Python program that performs the following operations:
  - Reads a very large formatted catalog containing properties of galaxies (e.g., their redshifts, stellar masses, and rates of star formation) out to early epochs.
  - Uses conditional operations on array indices to extract elements that satisfy specific criteria (e.g., specific redshift ranges, masses, etc).
  - Performs specific operations and basic statistical analyses on select elements.
  - Produces different graphical representations of the results, including scatter plots, and histograms.
  - Produces formatted output catalogs

Detailed Instructions

(a) Create a sub-directory called 'hwk4' in your python directory and change (cd) to this directory. Retrieve the file *'hwk4-catalog.txt'* by typing in a **single** line

```
curl -u cosmos:ast376web www.as.utexas.edu/~sj/a376-sp21/secure/hwk4/
hwk4-catalog.txt -o hkw4-catalog.txt
```

(b) Put all the files that you create for this homework into the sub-directory '∼/python/hwk4'.

(c) Please write a single Python program to do all the questions in this homework and put in *short comments within the program to indicate which part of the code addresses which question.* For instance, the comments might read 'Code for question 1', 'Code for question 2', etc. *Please name your Python program 'hwk4-myname.py', where 'myname' is your last name (e.g., hwk4-jogee.py.)*

(d) For full credit, **you must use the correct way to program in Python.** Getting to the right answer using a bad method will only get you partial credit. Note that copying the programs of your fellow students is considered as **cheating** and will lead to a "Fail" grade for both the person who copies the program and the person who allow their program to be copied. We reserve the right to ask students to discuss their programs and methodology after submission.

(e) **After working out questions 1 to 8 below, please submit the following on Canvas:**

- Your python (.py) script file (named "hwk4-myname.py".)
- A compressed .tar.gz file ('named "hwk4-myname-tar.gz") containing the PDF version of the figures you made for questions 1, 3, and 5.
- Two separate output catalogs (named "hwk4-myname-massive.txt' and "hwk4-myname-sfr.txt") for questions 7 and 8, respectively.
- A text file (named as "'hwk4-myname-answers.txt") containing your answers to questions 2, 4, and 6. The format of this text file is not important, we simply need to see the answers written in it. Note that while we will review your python (.py) script file when we grade, you should not expect us to run the code to obtain the answers to questions 2, 4, and 6. One simple way for you to make a text file with the answers is to have your python script print the answers and then you can direct (pipe) this printed output into a text file using the command:

```
python hwk4_myname.py > hwk4-myname-answers.txt
```

1. The catalog 'hwk4-catalog.txt' has information on a subset of galaxies, which have been observed with the optical Advanced Camera for Surveys (ACS) on the *Hubble Space Telescope* and the mid-infrared MIPS instrument at a wavelength of 24 microns aboard the *Spitzer Space Telescope*. In addition, observations from ground-based telescopes from the ultra-violet (UV) to near-infrared (NIR) have also been taken. Read through the header information in the catalog to understand the different quantities that have been recorded for each galaxy.

   In this question, 'log' denotes the logarithm to the base 10, M is the stellar mass in unit of $M_\odot$, and $\mathrm{SFR_{UV}}$ is the unobscured UV-based star formation rate in units of $M_\odot$ yr^{-1}, determined from UV data. Write a Python program, which makes a plot of log($\mathrm{SFR_{UV}}$) versus log(M), for those galaxies with photometric redshift z in the range $0.47 < z \leq 0.62$, and stellar masses $M \geq 5 \times 10^9$ $M_\odot$. This redshift range corresponds to lookback times of 4 to 5 billion years, an epoch when the Universe was 9.7 to 8.7 billion years old or about two-thirds of its present age. Plot the galaxies as black triangles or black stars. Label the x-axis with 'log($M/M_\odot$)', and the y-axis with 'log($\mathrm{SFR_{UV}}/M_\odot$ yr^{-1})'. Add a legend labeled 'z =0.47–0.62' in one corner of the plot. Save a copy of the figure in PDF format and turn it in. [**25 points**]

2. For the galaxies selected in part (1), use Python to determine the minimum, maximum, and median UV-based star formation rate in unit of $M_\odot$ yr^{-1}. Submit a copy of your answers in a text file. [**10 points**].

3. The UV data traces the unobscured star formation, namely star formation that is not obscured by dust. Thus, the UV-based star formation rate is a lower limit to the true total

star formation rate. The *Spitzer* 24 micron observations trace the obscured part of the star formation and yield the infrared-based obscured star formation rate, $\mathrm{SFR_{IR}}$. Unfortunately, only galaxies with fairly large infrared luminosities are detected by the *Spitzer* 24 micron observations, and $\mathrm{SFR_{IR}}$ are only available for those systems. In these cases, the total star formation rate, $\mathrm{SFR_{tot}}$, can be calculated as ($\mathrm{SFR_{IR}}$ + $\mathrm{SFR_{UV}}$). The value of the 'whmips' parameter in the catalog indicates whether or not there was a successful *Spitzer* 24 micron detection.

For galaxies with a successful *Spitzer* 24 micron detection, a photometric redshift z in the range $0.47 < z \leq 0.62$, and stellar masses $M \geq 5 \times 10^9\ M_\odot$, write a Python program, which makes a plot of $\log(\mathrm{SFR_{tot}})$ versus $\log(M)$. Plot galaxies with $SFR_{\mathrm{tot}} \geq 3.0\ M_\odot$ yr^{-1} (i.e., with star formation rates higher than that of our Milky Way) as red stars, and the remaining galaxies as black stars. Label the x-axis with '$\log(M/M_\odot)$', and the y-axis with '$\log(\mathrm{SFR_{tot}}/M_\odot\ \mathrm{yr}^{-1})$'. Add a legend label for both data sets. Label the red stars as: '$\mathrm{SFR_{tot}} \geq 3.0 M_\odot\ \mathrm{yr}^{-1}$' and label the black stars as '$\mathrm{SFR_{tot}} < 3.0 M_\odot\ \mathrm{yr}^{-1}$. Save a copy of the figure in PDF format and turn it in. [**10 points**].

4. For the galaxies selected in part (3), use Python to determine the minimum, maximum, and median value of the total star formation rate ($\mathrm{SFR_{tot}}$), as well as the median value of the UV-based star formation rate ($\mathrm{SFR_{UV}}$), in units of $M_\odot\ \mathrm{yr}^{-1}$. For these galaxies, what is the ratio R of (median $\mathrm{SFR_{tot}}$/median $\mathrm{SFR_{UV}}$)? This ratio R is a measure of how obscured the star formation sites are. Submit your answer in a text file. [**10 points**].

5. In addition to inspecting the mean or median value of the star formation rates, one can glean important conclusions from the distribution of star formation rates. For the galaxies selected in part (3), use Python to make a plot that overlays two normalized histograms: a histogram of $\mathrm{SFR_{tot}}$ in solid lines, and a histogram of $\mathrm{SFR_{UV}}$ in dotted lines. Label the y-axis with 'Fraction of Galaxies' and the x-axis with 'SFR ($M_\odot\ \mathrm{yr}^{-1}$)'. In one corner of the plot, add a legend to show that the solid line represents $\mathrm{SFR_{tot}}$, while the dotted line represents $\mathrm{SFR_{UV}}$. Save a copy of the figure in PDF format. [**15 points**].

6. Our Galaxy, the Milky Way currently has a stellar mass of $4.3 \times 10^{10}\ M_\odot$ and a total star formation rate of $\sim 3 M_\odot yr^{-1}$. Among the galaxies selected in part (3), how many systems are already more massive than the Milky way, 4 to 5 billion years ago? Among the galaxies selected in part (3), how many systems had a $\mathrm{SFR_{tot}}$ exceeding that of the Milky Way, 4 to 5 billion years? Submit your answer in a text file. [**10 points**].

7. Among the galaxies selected in part (3), extract the 30 galaxies with the largest stellar mass, and write out their properties (listed in the original catalog 'hwk4-catalog.txt') into a file called "hwk4-myname-massive.txt") where 'myname' is your last name (e.g., "hwk4-jogee-massive.txt"). Be sure to save the data as strings if using numpy. Submit a copy of this file and of your Python program. [**10 points**].

8. Repeat step (7), but extract the 30 galaxies with the largest total star formation rate and call the output catalog "hwk4-myname-sfr.txt") where 'myname' is your last name (e.g., "hwk4-jogee-sfr.txt"). Submit a copy of this file and of your Python program. [**10 points**].

END OF ASSIGNMENT