Exercises Sheet 2

This set of exercises follows on from the topics discussed in the 'Python Introduction Part 2', and focus on the different visualisations we encountered with matplotlib.

These exercises will need to be completed in script. You should create a script for each exercise. At the start of each script you will need to import matplotlib.pyplot *and* numpy.

Remember - for any function you haven't met before, you can use ? to find out information, e.g. ?print

For each of these exercises you can also play around with the matplotlib themes and also the style of the plot.

Exercises

1. **Line plots**

   We will be working the following table of star data:

Class	Radius (R_o)	Mass (M_o)	Luminosity (L_o)	Temp (K)	Name
O6	18	40	500,000	38,000	Theta1 Orionis C
B0	7.4	18	20,000	30,000	Phi1 Orionis
B5	3.8	6.5	800	16,400	Pi Andromedae A
A0	2.5	3.2	80	10,800	Alpha Coronae Borealis A
A5	1.7	2.1	20	8,620	Beta Pictoris
F0	1.3	1.7	6	7,240	Gamma Virginis
F5	1.2	1.3	2.5	6,540	Eta Arietis

   a. Create two separate numpy arrays for mass and luminosity. Make a line plot of mass versus luminosity. Display with plt.show().
   b. Add labels to the axis.
   c. Add a title to the plot.
   d. Make the y-axis logarithmic using plt.yscale.
   e. Change the style and colour of the line to something of your choosing.
   f. Create a new array for star temperature. Create a multi-panel plot that contains your first plot, and a fully labelled plot of mass vs temperature.

For the following exercises we are going to use a dataset from the Sloan Sky Digital Survey (SDSS), which has taken observations of 10,000 objects in space. I obtained the data from https://www.kaggle.com/lucidlenn/sloan-digital-sky-survey/downloads/sloan-digital-sky-survey.zip/2 and the information on the Kaggle website contains further information about the dataset. Or see here: http://www.sdss3.org/dr9/imaging/imaging_basics.php

The data we are using can be obtained by with the following code:

```
> import pandas as pd

> url='https://docs.google.com/spreadsheets/d/1xc1YpcP9sHKnV3l9tpkc156-
lEjms6D5Xflb9Yw9JiU/export?format=csv&gid=1442705777'

> data=pd.read_csv(url)
```

The SDSS basically takes images across the sky in 5 filter bands *u, g, r, i, z*. We will look at the relationship between the magnitudes of objects in the u and g filter band-passes.

Add the following lines of code to your script:

```
> u_mag=data['u']
> g_mag=data['g']
> object=data['class']
```

We can continue with the exercises.

2. Scatter plots

a. Create a scatter plot of *u* magnitudes vs *g* magnitudes.
b. Label the plot.
c. The given magnitudes are from objects made up of stars, galaxies, and QSO's. Using relational operators, create new arrays of magnitudes for stars, and new arrays for galaxies. Create a single scatter plot showing the two different data sets.
d. Add labels and a legend to the scatter plot.
e. Is there a difference between stars and galaxies?

3. Histograms

a. Plot a histogram of the *u* magnitudes for all objects. Label the plot appropriately.

b. Change the bins size to 20, then to 40. How does this influence the shape of the histogram?
c. Let's compare the magnitudes of different deep space objects by plotting multi-panel histograms, one for each of stars, galaxies, and QSO's.
d. Instead of plotting them side by side. Create a single plot with each histogram over-plotted. You may have to change from counts to density to properly compare each distribution (look at the docs for density keyword). You will also have to play with the opacity to show each distribution clearly.