Problem 2

For this problem, I ploted the Hubble data, created 9 groups of it and ploted that as well. For both data sets I generated a line of best fit using the least squares algorithm.

```
import numpy as np
import math
from HubbleFit import least_squares
from matplotlib import pyplot as plt
from read_hubble import read_hubble
from read_hubbleGroups import read_hubbleGroups
```

```
The history saving thread hit an unexpected error (DatabaseError('database disk image is malformed')).H
```

Below is a function to create a line of the slope and intercept.

```
def abline(slope, intercept):
    axes = plt.gca()
    x_vals = np.array(axes.get_xlim())
    y_vals = intercept + slope * x_vals
    return(x_vals, y_vals)
```

Below, I read in both the full data and the 9 Groups data set, and run a least squares fit on each.

```
distance, velocity = read_hubble("Hubble.txt")
distanceGroups, velocityGroups = read_hubbleGroups("HubbleGroups.txt")
```

```
z = least_squares(distance,velocity)
zGroups = least_squares(distanceGroups,velocityGroups)
```

Below, I plot the full data and its least squares line as well as the 9 grouped data and its best fit line.

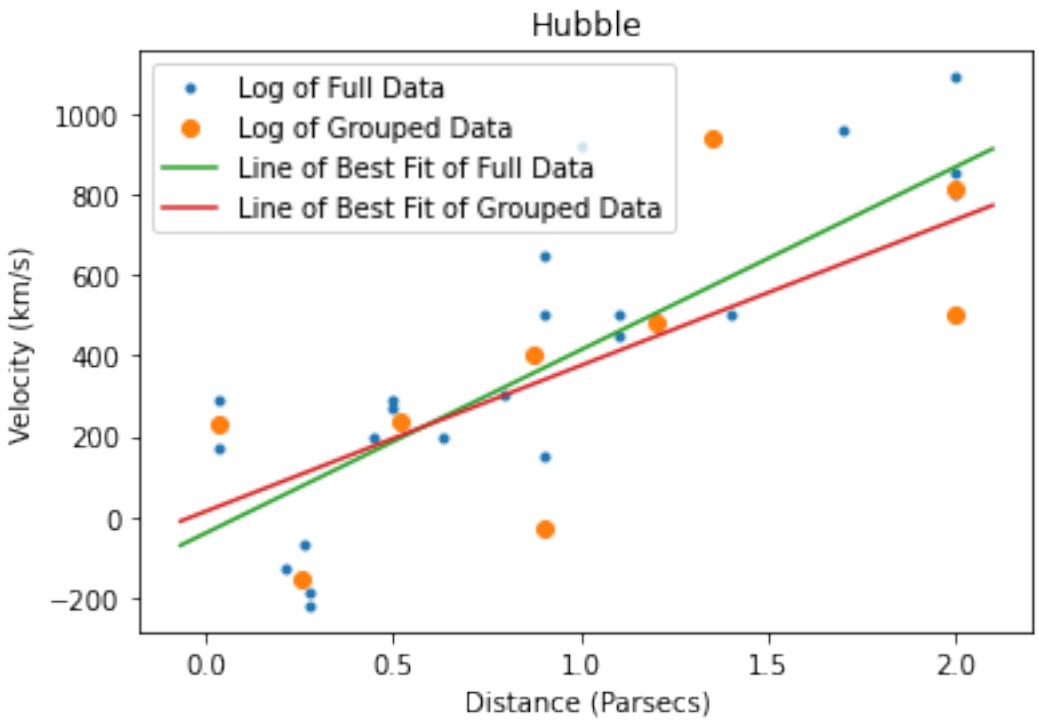
```
plt.plot(distance,velocity, '.',  label='Log of Full Data')
plt.plot(distanceGroups,velocityGroups, 'o',  label='Log of Grouped Data')

xLine, yLine = abline(z[1],z[0])
xLineGroups, yLineGroups = abline(zGroups[1],zGroups[0])

plt.plot(xLine,yLine, label = 'Line of Best Fit of Full Data')
plt.plot(xLineGroups,yLineGroups, label = 'Line of Best Fit of Grouped Data')

plt.title("Hubble")
plt.xlabel("Distance (Parsecs)")
plt.ylabel("Velocity (km/s)")
plt.legend()


plt.show()
```



Comparing the slope of the fitted straight line to Hubble's value of K.

```
print("The full data best fit line slope is: ", z[1])
print("The groups best fit line slope is: ", zGroups[1])
print("Hubble constant k was found to be around 500.")
```

```
The full data best fit line slope is:  454.1584409226284
The groups best fit line slope is:  361.73946381986593
Hubble constant k was found to be around 500.
```