Data Science – Learning Management Questions

**Week Five – Question #1**

from pandas import \*

from ggplot import \*

import pandas

def lineplot(hr\_year\_csv):

# A csv file will be passed in as an argument which

# contains two columns -- 'HR' (the number of homerun hits)

# and 'yearID' (the year in which the homeruns were hit).

#

# Fill out the body of this function, lineplot, to use the

# passed-in csv file, hr\_year.csv, and create a

# chart with points connected by lines, both colored 'red',

# showing the number of HR by year.

#

# You will want to first load the csv file into a pandas dataframe

# and use the pandas dataframe along with ggplot to create your visualization

#

# **You can check out the data in the csv file at the link below:**

**# https://s3.amazonaws.com/content.udacity-data.com/courses/ud359/hr\_year.csv**

#

# You can read more about ggplot at the following link:

# https://github.com/yhat/ggplot/

**gg = #YOUR CODE GOES HERE**

return gg

**Week Five – Question #2**

import pandas

from ggplot import \*

def lineplot\_compare(hr\_by\_team\_year\_sf\_la\_csv):

# Write a function, lineplot\_compare, that will read a csv file

# called hr\_by\_team\_year\_sf\_la.csv and plot it using pandas and ggplot.

#

# This csv file has three columns: yearID, HR, and teamID. The data in the

# file gives the total number of home runs hit each year by the SF Giants

# (teamID == 'SFN') and the LA Dodgers (teamID == "LAN"). Produce a

# visualization comparing the total home runs by year of the two teams.

#

# You can see the data in hr\_by\_team\_year\_sf\_la\_csv

# at the link below:

# https://s3.amazonaws.com/content.udacity-data.com/courses/ud359/hr\_by\_team\_year\_sf\_la.csv

#

# Note that to differentiate between multiple categories on the

# same plot in ggplot, we can pass color in with the other arguments

# to aes, rather than in our geometry functions. For example,

# ggplot(data, aes(xvar, yvar, color=category\_var)). This might help you

# in this exercise.

**gg = #YOUR CODE GOES HERE**

return gg

**Week Five – Question #3**

**Week Five – Question #4**

**Week Five – Question #5**