Anamitra Bhattacharyya
Predict 420-DL, Section 55
Assignment 1 (April 10, 2016)

1) Python code:

```
import pandas as pd #import panda
#all files are in the current working directory

#Read in Airlines labels data
airlinelabs=pd.read_csv("airlinelabs.csv", sep=',', header=None)
# get the airlines data into a pandas data frame
airlinesDF=pd.read_csv("airlines.dat", sep=',', header=None)
airlinesDF.columns=airlinelabs.loc[0] # assign airlinelabs to airlinesDF columns
print(airlinelabs) # print variable names

#Read in Airports labels data
airportlabs=pd.read_csv("airportlabs.csv", sep=',', header=None)
# get the airports data into a pandas data frame
airportsDF=pd.read_csv("airports.dat", sep=',', header=None)
# Assign airportlabs names to airportsDF column dataframe
airportsDF.columns=airportlabs.loc[0]
print(airportlabs) #print airportlabs variable names

#Read in Routes labels data
routelabs=pd.read_csv("routelabs.csv", sep=',', header=None) # get a DataFrame with one
row of data from a file w/o a header
# get the routes data into a pandas data frame
routesDF=pd.read_csv("routes.dat", sep=',', header=None)
routesDF.columns=routelabs.loc[0] # Assign routes names to routesDF data frame
print(routelabs) # print routelabs variable names


#Pickling Airlines, Airports and Routes data frames
import cPickle as pickle
pickle.dump(airlinesDF,open('airlineslist.p', 'wb'))
pickle.dump(airportsDF,open('airportslist.p', 'wb'))
pickle.dump(routesDF,open('routeslist.p', 'wb'))

#Number of departing routes from Chicago O'Hare Airport (ORD)
routesDF[routesDF.srcAirport=="ORD"].shape
#Number of incoming routes from Airport EGO
routesDF[routesDF.destAp=="EGO"].shape
```

Output:

a) airlineslabs:

```
    ...: print(airlinelabs)
          0         1         2     3     4         5        6       7
0  airlineID  airName  airAlias  iata  icao  callSign  country  active
```

b) airportlabs:

```
    ...: print(airportlabs)
     0       1       2          3       4       5           6            7   \
0  apID  apName  apCity  apCountry  apIata  apIcao  apLatitude  apLongitude

          8           9      10    11
0  apAltitude  apTimezone  apDST  apTz
```

c) routelabs:

```
    ...: print(routelabs)
         0          1           2        3       4         5          6  \
0  airline  airlineID  srcAirport  srcApID  destAp  destApID  codeshare

       7          8
0  stops  equipment
```

d) Departing and incoming flights

```
In [173]: routesDF[routesDF.srcAirport=="ORD"].shape
Out[173]: (558, 9)

In [174]: routesDF[routesDF.destAp=="EGO"].shape
Out[174]: (11, 9)
```

Number of departing flights from ORD = 558
Number of incoming flights from EGO = 11