**US Airline data from 1999 – 2009**

**Psudo Code**

**Fly Date (Data Capture)**

* Import Libs
  + Pandas
  + Matplotlib
  + Datetime
  + Scipy
  + Statsmodels
  + numpy
* Load airline route data into DataFrame - define datatype for integers
* Recast 'Fly Date' from YYYYMM format to separate columns of 'Month' and 'Year'
* Create an empty list to store the chunks of 1,000,000 rows at a time
* Iterate over each chunk of data
  + Drop rows with null values
  + Recast 'Fly Date' from YYYYMM format to columns for 'Month' and 'Year
  + Append the modified chunk to the list
* Concatenate all of the chunks into one DataFrame
* Return the number of rows and columns in the DataFrame
* Reorder column order so 'Month' and 'Year' are repositioned to come after 'Fly Date
* Return updated DataFrame

**Top 10 Route (Data Capture)**

* Reorganizing and renaming Columns as a new Data frame
  + Grabing route and destination Data
  + Reorganizing passanger count by route
  + #merge approprate columns with newly calc data

**Data Cleanup**

* Reordering and Naming columns
  + Reshape the Data
  + Remove all Rows without info or NaN info
  + Get amount of unique rows in column (airport)
* Split date in to two columns
  + Convert Column to string
  + Assign new column to date info split by month and year

**Calculate Percentage of seats used on Flight**

* Access data Frame
  + Read data frame
  + Round math by 2 decimals places
  + Reordering Columns
  + Export data frame

**Chart Generation**

* Reads CSV data
  + Count of unique routes, summing the value in the Flights column (number of flights in the given month)
  + Total number of unique routes
    - Get count of each occurrence of a destination code
    - Creates variables for Number of entries
    - stores 25 rows of data in new DataFrame (variable)
      * create a pie chart that shows the percentage of flights to each destination airport in the shortened dataframe
      * Calc the sum of data in flights column
  + Group data in to one column and calc Sub of rows in 2 columns
* Creates varibles for Number of entries
  + create a pie chart that shows the percentage of flights to each destination airport in the shortened dataframe
  + Calc the sum of data in flights column

**Percent Full vs Route Frequency**

* Group data by 2 columns and calc the mean
* Merge data
* assigns it the length of the full Frame data frame after dropping the missing

values.

* + Remove all nan Data
  + Assign it the length of the full frame after dropping missing values
* Create scatter plot