Solutions to Assignment 1

Problem 1

The object most appropriate would be data frame as it's easy to access.

The missing values can be dealt with na.rm = T

Problem 2

In this problem we first filter the data for carrier delays and then make all the negative values as 0. Then we get the data from column 6 to 14 separately which has different types of arrays.

We only take the arrival and departure array into consideration as the other delays are assumed to be caused because of these arrays. Thus we find the sum of those arrays.

first we create the function called TotalNumDelays(Carrier)

then we filter the all data related to carrier

now we get all the negetive values and make them 0

we create required.data in which we calculate the sum of all the columns from 6 to 14 without the NA values

from the required.data we find the sum of the first two columns

to get different types of arrays

Results:

```
The Total Numbers of different types Delays of your Carrier " AA " are

          DEP_DELAY           ARR_DELAY           CANCELLED       CARRIER_DEL
AY       WEATHER_DELAY
               19035               17848                1574                49
00               1433
          NAS_DELAY      SECURITY_DELAY LATE_AIRCRAFT_DELAY
X
               4767                   6                4973
0

As all the delays are included in Arrival and departure, we just calculate th
e sum of these two

 so The total number of delays of a Carrier " AA " are 36883
```

Solutions to Assignment 1

Problem 3

In this problem, we approach the same way as we did in the previous problem. We first separate the data for a particular origin and filter all the negative value as zero and find the columns for delays from origin. Then we add the departure delay as we only need to find the delays from a single origin

#First we create a function called TotalNumDelays(ORIGIN)

to separate data for a particular Origin

now we make all the negative values to zero

#to count elements of column

#to get different types of delays from your Origin

#to get the total number of delays

```
The Total Numbers of  Delays from your Origin from different types " JFK " ar
e

           DEP_DELAY           ARR_DELAY           CANCELLED       CARRIER_DEL
AY       WEATHER_DELAY
                3694                3708                 761                10
87                800
           NAS_DELAY      SECURITY_DELAY LATE_AIRCRAFT_DELAY
X
                1640                  15                 796
0

The total number of delays are sum of Departure Delay and Arrival Delay as th
e other delays are included in the Departure Delay and Arrival Delay,

The total number of delays from the Origin " JFK " are 3694
```

Problem 4

#check if delays.in.flights exists

First create function called AvgDelay(Carrier,Dest)

create a data.frame with columns for CARRIER, ARR_DELAY and DEST

Solutions to Assignment 1

remove all the NA values

Create a Subset that reads data from DEST and CARRIER

calculate average arrival delay for a carrier flying into the specified destination airport.

gives Avg delays of the Destination Airport code and Carrier Code

Result:

```
> AvgDelay ('AA','DFW')
[1] 33.8675
```

Problem 5

To improve a function so that we don't have to load the data frame repeatedly we can just do this

TotalNumDelays("AA")

TotalDelaysByOrigin("JFK")

AvgDelay ('AA','DFW')

```
> # No need to rerun the Program
> TotalNumDelays("AA")
The Total Numbers of different types Delays of your Carrier " AA " are

           DEP_DELAY           ARR_DELAY           CANCELLED       CARRIER_DEL
AY       WEATHER_DELAY
               19035               17848                1574                49
00               1433
            NAS_DELAY      SECURITY_DELAY LATE_AIRCRAFT_DELAY
X
                4767                   6                4973
0

As all the delays are included in Arrival and departure, we just calculate th
e sum of these two

 so The total number of delays of a Carrier " AA " are 36883
```

Solutions to Assignment 1

```
> TotalDelaysByOrigin("JFK")
The Total Numbers of  Delays from your Origin from different types " JFK " ar
e

           DEP_DELAY            ARR_DELAY            CANCELLED
                3694                 3708                  761
       CARRIER_DELAY        WEATHER_DELAY            NAS_DELAY
                1087                  800                 1640
      SECURITY_DELAY LATE_AIRCRAFT_DELAY                    X
                  15                  796                    0

The total number of delays are sum of Departure Delay  as we need the delays
from a particular origin so ,

The total number of delays from the Origin " JFK " are 3694

> AvgDelay ('AA','DFW')
[1] 33.8675
```