Visualization of Flight Delay Causes with Shiny app

Carl Edem Dekpor

5/3/2022

INTRODUCTION

My project is about using shiny, shinydashboard and some R packages to visualize data and create a web application. The dataset used is the 2008 fight delay causes in America. The data was provided by Professor Rege for a project in my Data Analytics and visualization class. Professor Abe mentioned the shiny package in class, did some reading on it and I thought it will be interesting to visualize and analyze the same dataset with it.

What:

This project will focus on developing a multi-shiny app web dashboard in R using the "Shinydashboard" packages to visualize flights delay and cancellation data from 2008 in the US.

Why:

Seems like a fun project to undertake, doing a similar project in my Data Analytics and visualization (using tableau) class. It will be good to compare the outcomes and to develop my R coding skills.

How:

Shiny app – is an interactive web app built on HTML. I will be exploring the reactive and interactive functions of shiny library. Interactively on how an element reacts to user actions and reactive changes that affect the output. UI objects will be used i.e., control widgets, I will be exploring the server function of shiny app. The graphical interface will help users quickly visualize important metrics of my data set. (Flight delay data and cancellation data in excel format)

BODY

We have all been affected by the impact of flight delays and cancellations. This project will look at the major causing factors of flight delays and cancellations among some carriers in the USA for year 2008.

The causes of flight delays have been identified as;

Air Carrier: The cause of the cancellation or delay was due to circumstances within the airline's control (e.g., maintenance or crew problems, aircraft cleaning, baggage loading, fueling, etc.).

Extreme Weather: Significant meteorological conditions (actual or forecasted) that, in the judgment of the carrier, delays or prevents the operation of a flight such as tornado, blizzard or hurricane. National Aviation System (NAS): Delays and cancellations attributable to the national aviation system that refer to a broad set of conditions, such as non-extreme weather conditions, airport operations, heavy traffic volume, and air traffic control.

Late-arriving aircraft: A previous flight with same aircraft arrived late, causing the present flight to depart late.

Security: Delays or cancellations caused by evacuation of a terminal or concourse, re-boarding of aircraft because of security breach, inoperative screening equipment and/or long lines in excess of 29 minutes at screening areas.

The tool used for my analysis is "R" with R packages like – (Shiny, shinydashboard, dpplyr, plotly, DT) R is a programming language for statistical computing and graphics supported by the R Core Team and the R Foundation for Statistical Computing. Created by statisticians Ross Ihaka and Robert Gentleman, R is used among data miners and statisticians for data analysis and developing statistical software. Users have created packages to augment the functions of the R language. R shiny allows you to access a complete web application framework in the R environment which can be customized.

Shiny app has two main components – *UI (user interface)* and *Server function* which defines how the app works. Functions like fluidpage() , selectInput(), tableOutput(), verbatimTextOutput().

R shiny allows you to access a complete web application framework in the R environment which can be customized. Shiny app has two main components – UI (user interface) and Server function which defines how the app works. Functions like fluidpage() , selectInput(), tableOutput(), verbatimTextOutput().

These are the structures for building the shiny web app and the dashboard. Histograms and boxplots will be used to analyze the various attributes recorded to detail how they affect the various carriers. This can be extended to analyze or compare the impacts of the delay causes and cancellations in different years.

Structure for creating a shiny app

Developing the UI (user interphase)

Code for the UI

```
library(shiny) # shiny features
library(shinydashboard) # shinydashboard functions

dashboardPage(
  dashboardHeader(title = "2008 FLIGHT DATA ANALYSIS WITH R Shiny",
                  titleWidth =650),
  dashboardSidebar(
    sidebarMenu(
      id = "sidebar",
menuItem("Flight Dataset", tabName = "data", icon =icon("database")),
menuItem(text= "visualization", tabName = "viz", icon = icon("chart-line")),
selectInput(inputId = "var1", label= "Select Attributes", choices = c1,
selected = "TotalOperations"),
selectInput(inputId = "var3", label= "Select the x variable", choices = c1,
selected = "CarrierDelays"),
selectInput(inputId = "var4", label= "Select the y variable",
choices = c1, selected = "LateAircraftDelays"),
      menuItem(text = "Carrier Location", tabName = "map", icon = icon("map"))
    )
```

```
  ),
  dashboardBody(
    tabItems(
      #tab item
      tabItem(tabName = "data",
              tabBox(id="t1", width =12,
     tabPanel("About", icon=icon("address-card"), fluidRow(
     column(width = 8, tags$img(src="airplane.jpg", width =600 , height = 300),
                               tags$br() ,
                               tags$a("Image from google"), align = "center"),
                          column(width = 4, tags$br() ,
tags$p("A flight delay is when an airline flight takes off and/or lands later
than its scheduled time. The Federal Aviation Administration (FAA) considers a
flight to be delayed when it is 15 minutes later than its scheduled time.
A cancellation occurs when the airline does not operate the flight at all for
a certain reason.
source:wikipidia")
                          )
                        )),
tabPanel("Data", icon= icon("table"), dataTableOutput("dataT")),
tabPanel("Structure", icon= icon("address-book"),
verbatimTextOutput("structure")),
tabPanel("summary stats", icon= icon("address-card"),
verbatimTextOutput("summary")),
                        )
                 )),
    tabItem(tabName = "Viz",
            #tab box
             tabBox(id="t2", width =8,
tabPanel(tiitle ="Distribution", value = "distro", plotlyOutput("histplot"))

      ),
    #tab item 3
    tabItem(tabName = "map",
            box(h1("placeholder UI"))
            )
    )))
```

Code for the server

```
library(DT)

function(input, output, session){
  #structure
  output$structure <- renderPrint(
    # structure of the data
    us_flights %>%
      str()
  )
  #summary
  output$structure <- renderPrint(
    us_flights %>%
      summary()
```

```
  )

  #DataTable
  output$dataT <-renderDataTable(
    us_flights
  )

  # stack histogram and boxplot
 output$histplot <- renderPlotly({
   p1 = us_flights %>%
    plot_ly() %>%
    add_histogram(~TotalOperations) %>%
  layout(xaxis = list(title = "TotalOperations"))

  #box plot
  p2 = us_flights %>%
    plot_ly() %>%
    add_boxplot(~TotalOperations) %>%
    layout(yaxis = list(showticklabels = F))

  # stack plot
  subplot(p2, p1, nrows = 2) %>%
    hide_legend() %>%
    layout(title ="Distribution chart - Histogram and Boxplot",
           yaxis =list(title = "Frequency"))
})


}
```

Snapshots of the Run app

Introduction page - A breif summary on flight delays

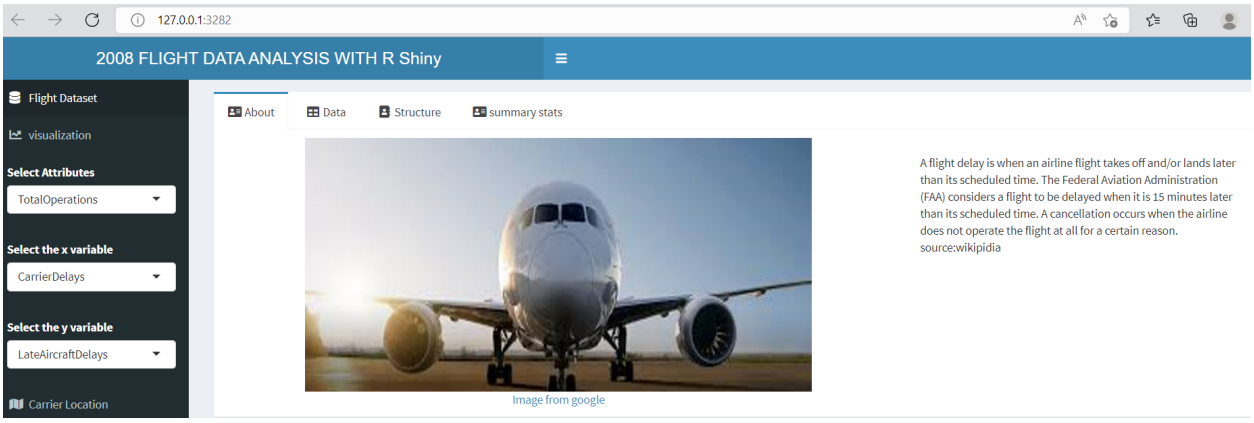

Figure 1: Introduction page

Statistical summary of the attributes

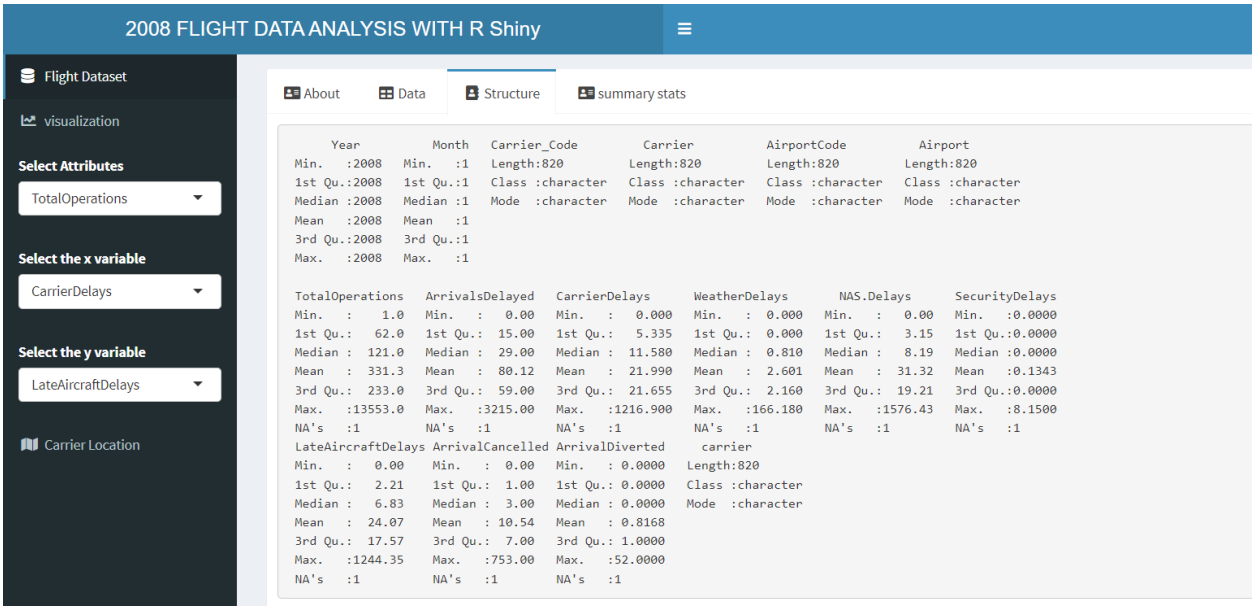

Figure 2: Summary of Attributes

Visualization of the arrivals delayed and total operations

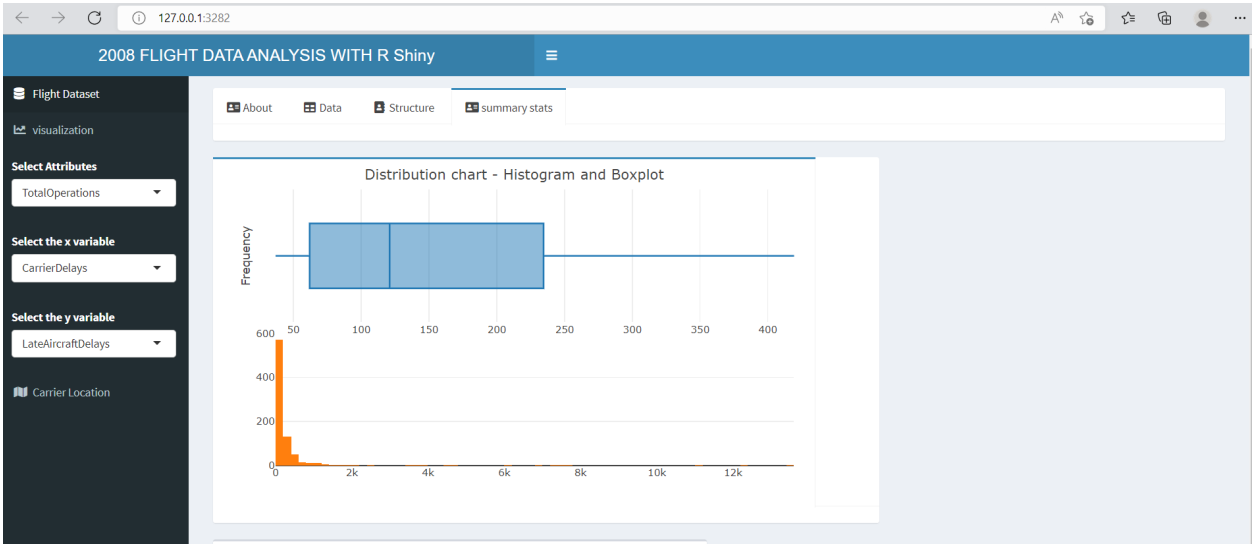

Figure 3: Histogram and Boxplot

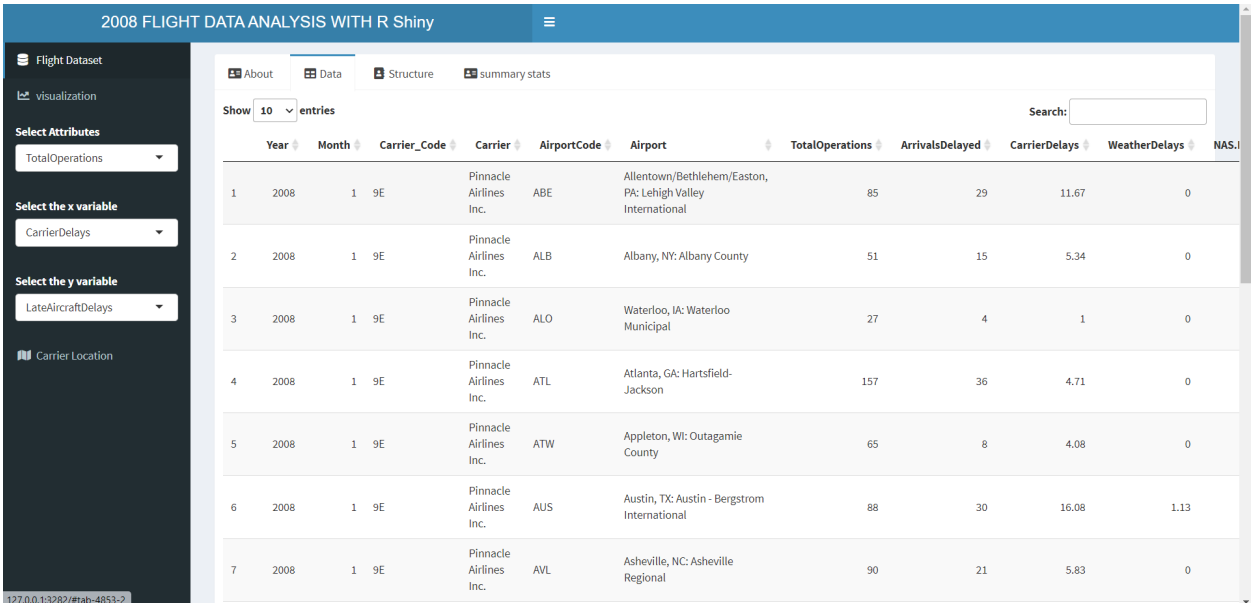

Figure 4: Loaded Dataset app

Loaded dataset in the app

How to run the application

To run the application you will need to copy UI.R and Server.R files into one file. ie Flight.R with the dataset and images in your working folder.

TOPICS FROM CLASS

Installing R packages and Some basics R commands

I applied various topics treated in class like – Installing packages in "R", Listing data, working with data frames, loading files in different formats in "R", Showing files and setting the working directory, assignment of variables, importing images, getting help with functions, and invoking commands

Exposure to reading academic-type paper.

In writing and understanding some of the codes and resources used on the internet. The exposure to R packages and documentation in class was very helpful.

Github

Creating and use a repository. Making changes to a file and push them to GitHub as commits.

R Markdowns

Provides a quick and reproducible reporting from R with a nice layout.

Statistics

Statistical analysis

CONCLUSION

Putting all the codes together has been really a good learning experience. I have had to go over a lot of documentations and similar ideas on the internet to be able to develop a functioning application. With more time I could have used a larger dataset and compare more years which will involve more scripts. I'm proud what I have been able to achieve with this short period of time. I hope to continue after the semester. Overall, my R coding skills have improved immensely.

References

[https://en.wikipedia.org/wiki/R_(programming_language)] [https://www.bts.gov/explore-topics-and-geography/topics/airline-time-performance-and-causes-flight-delays] [https://www.youtube.com/watch?v=tlOBVZx8Hy0]

Shinyapp Link

[https://carldekpor.shinyapps.io/flightdelay-Carldekpor/]