REPORTS ON MISHANDLED BAGGAGE OF MAJOR CARRIERS IN U.S - DOMESTIC SYSTEM ONLY

JINU KINGCY SEBASTIN

GEORGE MASON UNIVERSITY

STAT 515: APPLIED STATISTICS AND VISUALIZATION FOR ANALYTICS

Abstract

Air travel in the United States, has witnessed a remarkable transformation over the recent years and has dawned as the preferred means of transportation between cities or states involving long distances. In spite of it being advantageous and time-saving, air travel has also turned out to be more chaotic and undesirable for some travelers who deal with issues on lost or mishandled bags. Major carriers in the U.S are required to report monthly statistics on flight delays, mishandled baggage and other air travel related data to the Department of Transportation's[DOT] Bureau of Transportation Statistics[BTS]. From the data published by DOT, it is contained that complaints on lost bags by customers have been elevated since last decade. By analyzing the data on mishandled baggage, it is inferred that there has been an increase in the number of passengers flying major carriers over the last few years whereas the reports on lost bags have seen a declination along the scale. Heightened competition in the airline sector to deliver quality and outstanding travel experience to its customers is considered a factor in the overall reduction of baggage mishandling.

REPORTS ON MISHANDLED BAGGAGE OF MAJOR CARRIERS IN U.S - DOMESTIC SYSTEM ONLY

Despite advancements made in the airline industries in recent years, baggage handling remains one of the biggest bottlenecks amongst major air carriers. Statistics represent that the main cause for mishandled baggage were delayed, lost or damaged bags with the predominant reason being delayed bags. Analyzing the data is the crucial step for any statistical analysis and choosing the right visualization for data aids the users to easily comprehend and interpret the information.

Data on lost baggage

The article on analyzing data about lost baggage from the Minitab blog emphasizes the veracity about major air carriers and the complaints by their customers on mishandled bags. The author tries to point out the fact that major carriers handling the largest number of passengers tend to have the highest rate of mishandled baggage. A bar chart comparison of total mishandled bags and rate of mishandled bags per 1000 customers for each airline is portrayed. However, the article lacks a sense of depth in comparing the mishandled bags over the years, also the charts are monotonous and are not visually appealing. Further, it is missing information whether there has been an improvement among the airline industries with respect to baggage handling.

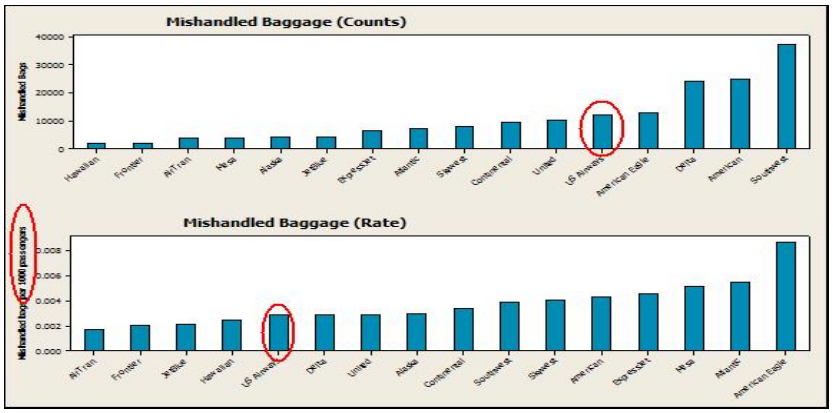

Figure 1. Bar chart representation of mishandled baggage per airline

Implementing redesigned graph

The data was extracted from U.S. Department of Transportation and was examined to develop an improved graph by adding factors that are considered critical in making comparisons and sound judgement about the major carriers. The software used for Data visualization and analysis is R Studio.

Total reports vs. reports per 1000 passengers. The data provides information on baggage report from 2013-2017. It is observed that the total reports on lost bags by each airline provides a raw ranking of the major carriers, while the actual comparison is to be made between the number of reports per 1000 passengers for an airline and their total reports on baggage mishandling. The rate of reports per 1000 passengers is calculated by dividing the total baggage reports by total number of passengers flying the airline. This representation delivers a vivid understanding on the basis of airlines being ranked.

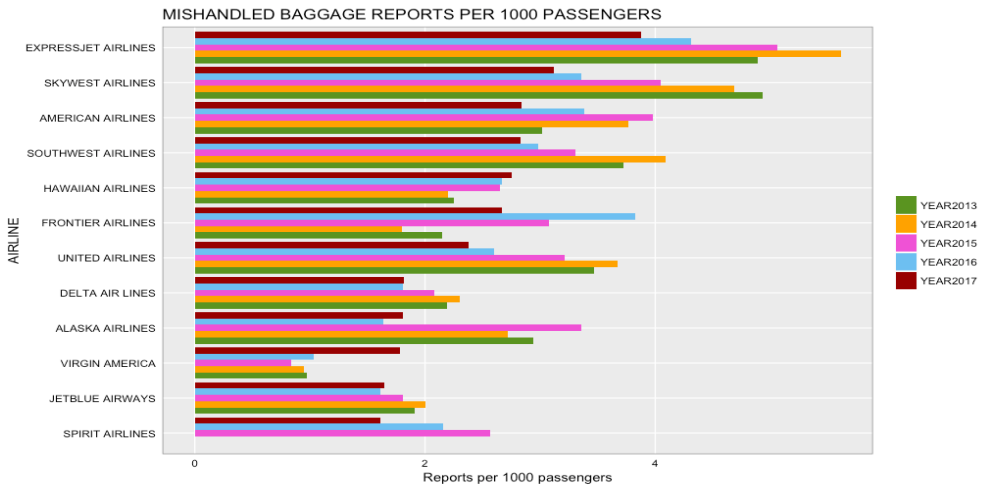

Figure 2. Horizontal bar chart representation of mishandled bags per 1000 passengers

Trend of lost bags and customers flying. A significant comparison of trends was drwan between mishandled bags and customers flying the U.S based carriers. A positive trend interpreting the fact that over the last five years, customers flying the U.S airlines has increased

in millions and a rising trend was observed until the year 2014 which has witnessed the highest reports on baggage loss, then declining gradually over the years, thus proving the reduction in baggage mishandling.

Summary. The redesigned graph is made more comprehensive with main factors such as comparisons based on the last five years has been added to the existing graph. Conclusions from the bar charts supports the fact that competition among major air carriers has led to the reduction in mishandled bags over the years.

References

Runkel, P. (2011). The Minitab Blog. *Analyzing data about lost baggage: It's all relative(frequency)*.

Kheel, J. M. (2018). The Points Guy. *The best and worst US airlines in 2018*

SITA Baggage report (2016). *Air transport industry insights: The Baggage report*. Retrieved from https://www.sita.aero/baggage-report-2016

Trefis Team (2016) The State of Air travel in the U.S has changed over the years. Retrieved from https://www.forbes.com/sites/greatspeculations/2016/10/11/the-state-of-air-travel-in-the-u-s-has-changed-over-the-years/#1f96e3129b6b

Appendix 1

R Script

```
library(ggplot2)
library(reshape2)
library(cowplot)
library(ggthemes)
report_bag <- read.csv("reportsper1000.csv")
report_bag$AIRLINE <- factor(report_bag$AIRLINE, levels = report_bag$AIRLINE)
report.m <- melt(report_bag, id.vars='AIRLINE')
report.m
reports_plot <- ggplot(report.m,aes(AIRLINE,value,fill= variable))+
geom_bar(stat="identity", position = "dodge")+labs(x="AIRLINE",y="Reports per 1000
passengers",title="MISHANDLED BAGGAGE REPORTS PER 1000
PASSENGERS")+theme(legend.position="top")+scale_fill_manual(values =
c("olivedrab","orange","orchid","skyblue2","darkred"))+hw+
theme(legend.title=element_blank())
reports_plot +coord_flip()
total_reports <- read.csv("totalreports.csv")
total_reports$AIRLINE <- factor(total_reports$AIRLINE, levels =
total_reports$AIRLINE)
totalrep.m <- melt(total_reports, id.vars='AIRLINE')
totalrep.m
totalplot <- ggplot(totalrep.m,aes(AIRLINE,value,fill= variable))+
```

```
 geom_bar(stat="identity", position = "dodge")+labs(x="AIRLINE",y="Total reports
(Thousands)",title="TOTAL REPORTS ON MISHANDLED
BAGGAGES")+theme(legend.position="top")+scale_fill_manual(values =
c("olivedrab","orange","orchid","skyblue2","darkred"))+hw+
theme(legend.title=element_blank())
totalplot+coord_flip()
comparison_report <- read.csv("combinedata.csv")
comparison_report$AIRLINE <- factor(comparison_report$AIRLINE, levels =
comparison_report$AIRLINE)
compreport.m <- melt(comparison_report, id.vars='AIRLINE')
compreport.m
comparison_plot <- ggplot(compreport.m,aes(AIRLINE,value,fill= variable))+
geom_bar(stat="identity", position = "dodge")+labs(x="AIRLINE",y="Number of
Reports (Thousands)",title="TOTAL REPORTS AND REPORTS PER 1000
PASSENGERS \nON MISHANDLED
BAGGAGES")+theme(legend.position="top")+scale_fill_manual(values =
c("olivedrab","orange","orchid","skyblue2","darkred"))+hw+
theme(legend.title=element_blank())
comparison_plot+coord_flip()
enplane <- read.csv("enplanedata.csv")
enplane$AIRLINE <- factor(enplane$AIRLINE, levels = enplane$AIRLINE)
enplane.m <- melt(enplane, id.vars='AIRLINE')
enplane.m
```

```
enplane_plot <- ggplot(enplane.m,aes(AIRLINE,value,fill= variable))+
 geom_bar(stat="identity", position = "dodge")+labs(y= "No. of passengers
(Millions)",x= "AIRLINE",title="ENPLANED PASSENGERS(2013 -
2017)")+theme(legend.position="top")+scale_fill_manual(values =
c("olivedrab","orange","orchid","skyblue2","darkred"))+hw+theme(legend.title=element
_blank())
enplane_plot +coord_flip()
data1 <- read.csv("trenddata1.csv")
data1.m <- melt(data1, id.vars='YEAR')
data1.m
data1_plot <- ggplot(data1.m,aes(x=YEAR,y=value))+
  geom_line(aes(colour=variable,group=variable))+
geom_point(aes(colour=variable,shape=variable,group=variable),size=3)+ylab(label =
"No. of Passengers flying(in millions)")+
  xlab("Year") + labs(title="Trend showing No. of passengers flying US
Airlines\nDomestic only")+hw+
theme(legend.title=element_blank())+theme(legend.position = 'bottom')
data1_plot
data2 <- read.csv("trenddata2.csv")
data2.m <- melt(data2, id.vars='YEAR')
data2.m
data2_plot <- ggplot(data2.m,aes(x=YEAR,y=value))+
```

```
 geom_line(aes(colour=variable,group=variable))+
geom_point(aes(colour=variable,shape=variable,group=variable),size=3)+scale_color_m
anual(values = "blue")+hw+ylab(label = "Mishandled baggage reports(in millions)")+
  xlab("Year") + labs(title="Trend showing mishandled baggage reports US
Airlines\nDomestic only")+theme(legend.title=element_blank())+ theme(legend.position
= 'bottom')
data2_plot
plot_grid(data1_plot,data2_plot,labels = "AUTO")
```

Appendix 2

Barplots

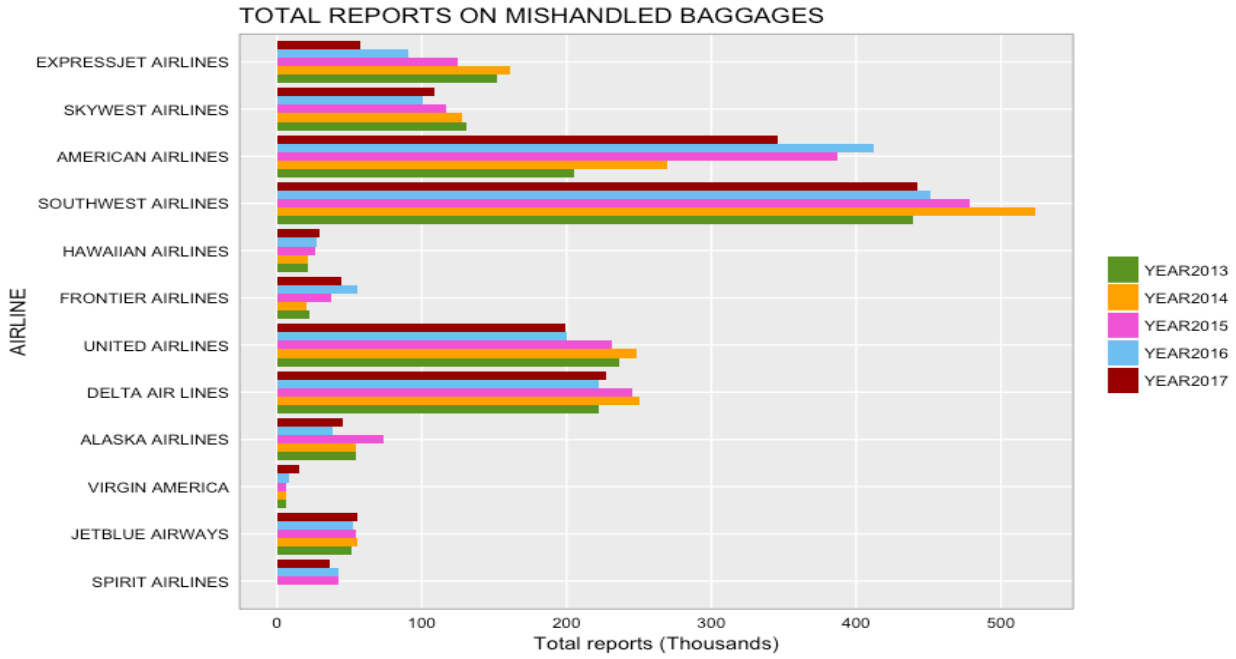

Figure 3. Bar chart representation of total reports on mishandled baggage

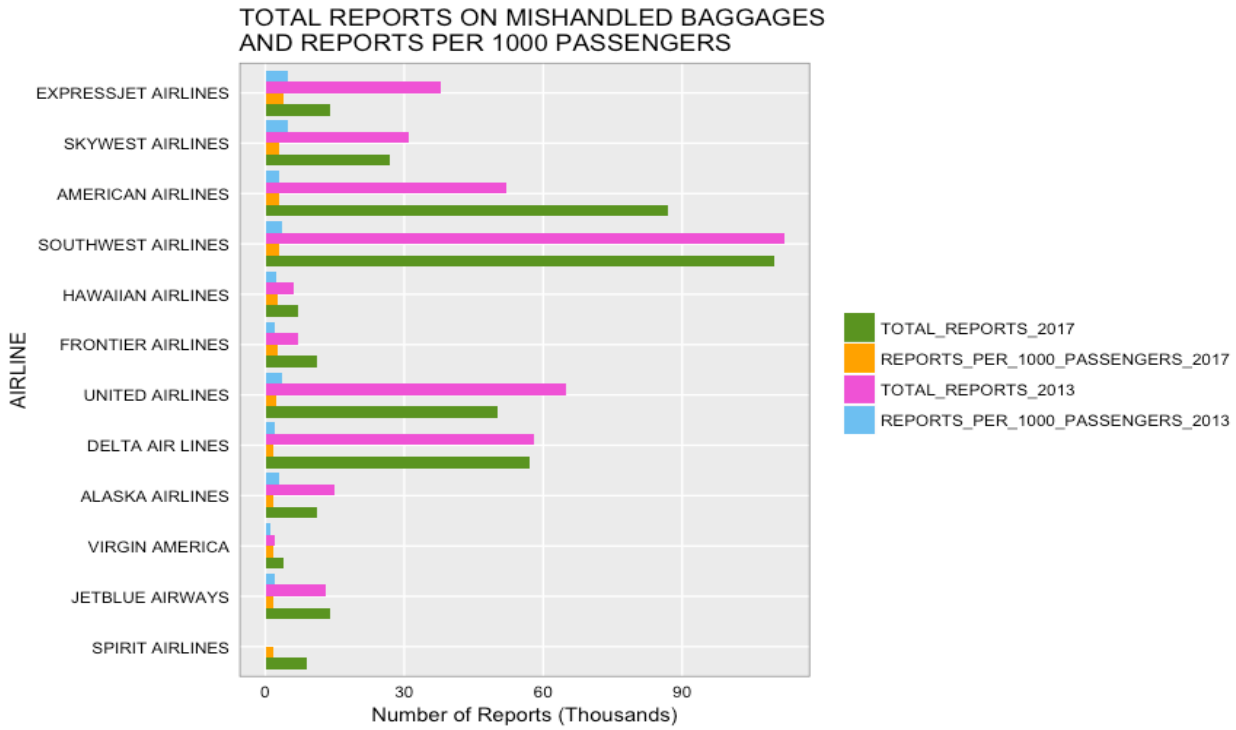

Figure 4. Bar chart depicts total reports on mishandled bags and reports per 1000 customers

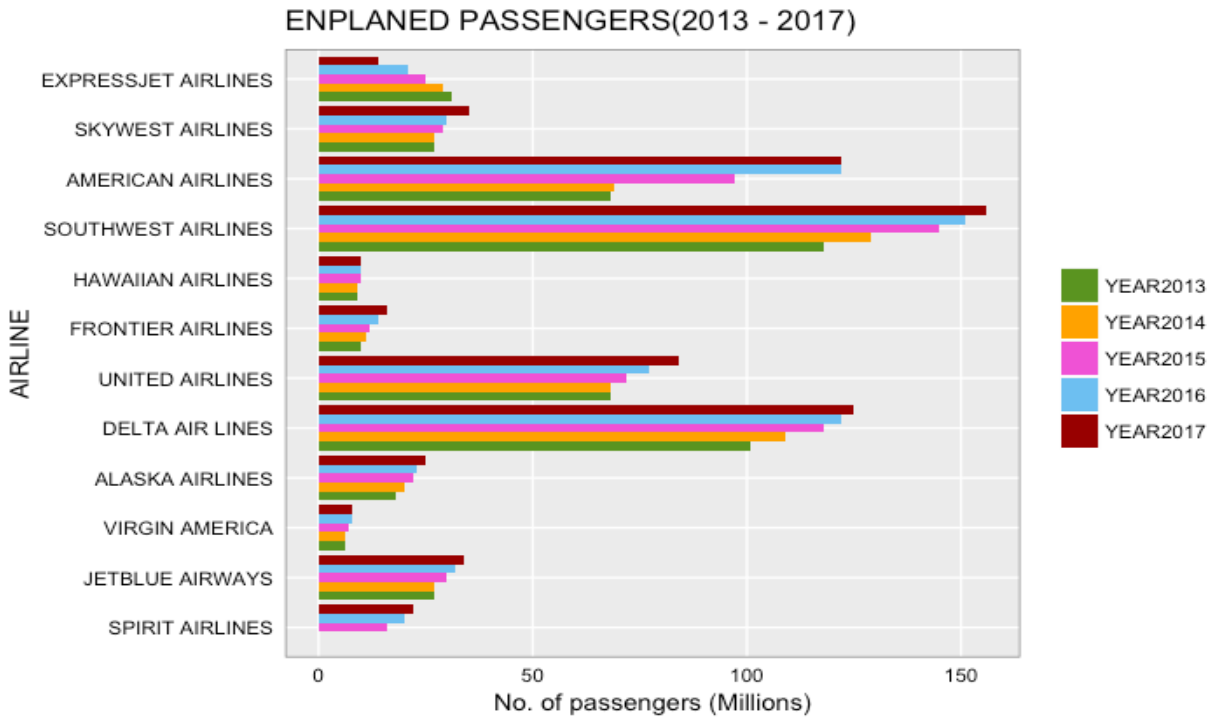

Figure 5. Bar chart representation of passengers flying each airline from 2013-2017

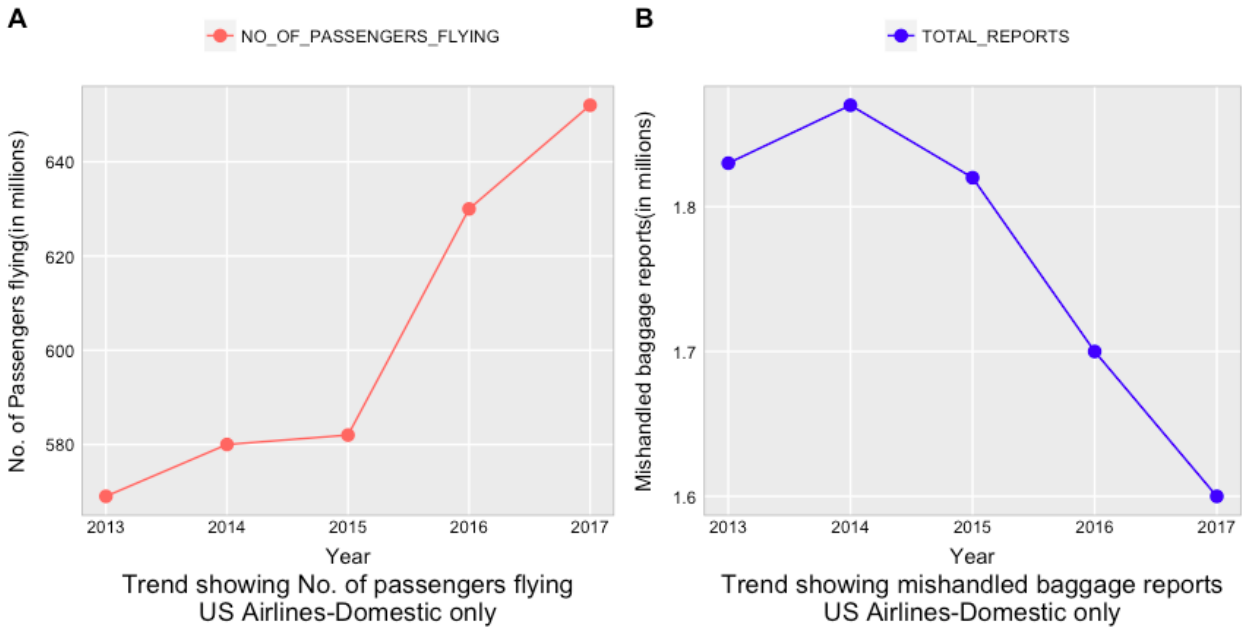

Figure 6. Line chart representing trend of passengers flying and the number of baggage mishandled over the years