Abstract

In this project our main focus is to study the trends of airline accidents that have occured between 1985-2014 and create a functioning dashboard that will help to find the faults and decrease the number of fatilities in future. We picked up a dataset from github. We interviewed the target audience and observed their adoption rates, later we did task abstraction and data abstarction (the most important part of vizualisation) and carried out the field study and collected the anecdotal evidences of utility. We later finalised the effective idioms and encoding attributes and did the quantitative and qualitative result image analysis. We later tried to run our algorithm and checked the computational complexity. We focused on each aspect properly so that not even one fault gets cascaded to the down level.

Introduction

1. Objective

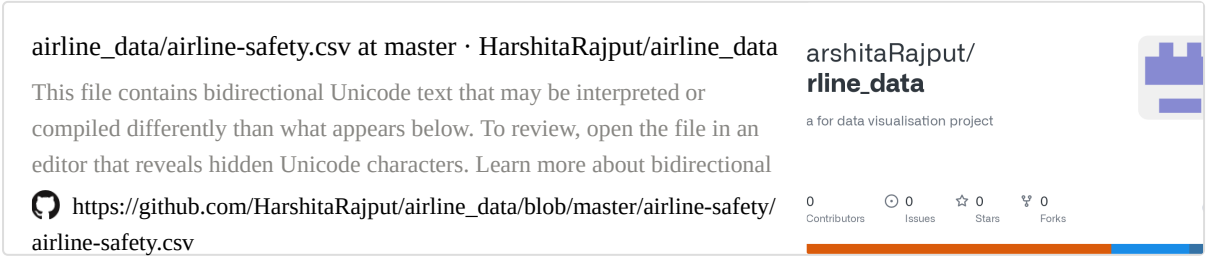
To visualize the statistics of fatal accidents and fatalities of different airlines during the time span of 1985-1999 and 2000-2014 to predict the success of these airlines in the coming years.

2. Problem statement:

Although airlines are one of the safest modes of transport, a countable number of accidents still occour on an yearly basis. Even after taking all the safety precautions, plane crashes become unavoidable in certain situations.

3. Functional Requirements

Dataset link:



Dashboard used: flexdashboard

Libraries used:

- ggplot2
- flexdashboard
- plotly

Software version: R 4.1.2

Data Abstraction

Dataset:

Name	Definition
Airline	name of airline(asterisk: regional subsidiaries are included)
Avail_seat_km_per week	Kilometres travelled in one week
Incidents_85_99	Number of incidences such as emergency landings during 1985- 1999
Fatal_accidents_85_99	Number of fatal accidents during 1985 to 1999
Fatalities_85_99	Number of fatalities during 1985 to 1999
Incidents_00_14	Number of incidences such as emergency landings during 2000- 2014
Fatal_accidents_00_14	Number of fatal accidents during 1985 to 1999
Fatalities_00_14	Number of fatalities during 1985 to 1999

Data Abstraction:

Name	Type
Airline	Categorical
Avail_seat_km_per week	Numeric (Discrete)
Incidents_85_99	Numeric (Discrete)
Fatal_accidents_85_99	Numeric (Discrete)
Fatalities_85_99	Numeric (Discrete)
Incidents_00_14	Numeric (Discrete)
Fatal_accidents_00_14	Numeric (Discrete)

Name	Type
Fatalities_00_14	Numeric (Discrete)

- this data is a quantitative dataset with numeric attributes and one categorical attribute serving as the primary key.
- dataset has been split into two parts for analysis by column,i.e 85_99 and 00_14 representing two timelines.

Task abstraction

```
#reading the dataset from local files
library(ggplot2)
library(flexdashboard)
library(plotly)
airline<-read.csv("C:/Users/Harshita/WinterSem22/Data Vis/Lab/airlinedata.csv")
```

Task: Which of the airlines had faced more than 400 fatalities during 1985-1999?

A1) Attributes required: airline, fatalities_85_95

Idiom:

Targets: airline, fatalities_85_95

Action: analyze (record the corresponding attributes from the data)

```
plot<ggplot(airline,aes(x=airline,y=fatalities_85_99))+geom_bar(stat="i dentity")
plot
```

Q2) Task: Which of the airlines had faced more than 400 fatalities during 2000-2014?

A2) Targets: airline, fatalities_00_14

Actions: analyse(record, derive)

```
plot<-ggplot(airline,aes(x=airline,y=fatalities_00_14))+geom_bar(stat="identity")
 plot
```

Q3) Task: How many fatal accidents led to fatalities of these airlines during 1985-1999?

A3) Attributes required: fatal_accidents_85_99, fatalitites_85_99

Action: analyze (record the corresponding attributes from the data)

```
ggplot(airline,aes(x=fatal_accidents_85_99,y=fatalities_85_99))+
geom_bar(stat="identity",width=0.4)
p
```

Q4) How many fatal accidents led to fatalities of these airlines during 2000-2014?

A4) Attributes required: fatal_accidents_00_14, fatalities_00_14

Action: analyze (record the corresponding attributes from the data)

```
ggplot(airline,aes(x=fatal_accidents_00_14,y=fatalities_00_14))+
geom_bar(stat="identity",width=0.1)
plot
```

Q5) Comparing the fatalities of 1985-1999 and 2000-2014, has the number of fatalities decreased or increased?

A5) Attributes required: fatalities_85_99, fatalitites_00_14

The decrease in fatalities comparing fatalities of 1985-1999 and 2000-2014

Action: analyze (record the corresponding attributes from the data)

```
plot(airline$fatalities_85_99,type="l",col="blue", xlab="Airlines",ylab= "Fatalities")
lines(airline$fatalities_00_14,type="l",col="red")
legend(39,530,legend<-c("1985-1999","2000-2014",col=c("blue","red")))
```

Q6) How many incidents have led to fatal accidents of these airlines during 1985- 1999?

A6) Attributes required: incidents_85_99, fatalities_85_99

Action: analyze (record the corresponding attributes from the data)

```
plot<-ggplot(airline,aes(x=incidents_85_99,y=fatalities_85_99),xlim=c(0,
40))+geom_bar(stat="identity")+xlim(0,30)
plot
```

Q7) How many incidents have led to fatal accidents of these airlines during 2000- 2014?

A7) Attributes required: incidents_00_14, fatalities_00_14
Action: analyze (record the corresponding attributes from the data)

```
plot<-ggplot(airline,aes(x=incidents_00_14,y=fatalities_00_14))+geom_bar
(stat="identity")
plot
```

Dashboard Implementation:

Dashboard 1: Task Abstraction plots

code:

````
---
title: "Task abstraction"
output:
  flexdashboard::flex_dashboard:
    orientation: columns
    vertical_layout: fill
---

```{r setup, include=FALSE}
library(flexdashboard)
library(ggplot2)
airline<-read.csv("C:/Users/Harshita/WinterSem22/Data Vis/Lab/airlinedata.csv")
```

Column {data-width=350}
-----------------------------------------------------------------------

### fatalities of different airlines during 1985-1999

```{r}
````
````

````
plot<-ggplot(airline,aes(x=airline,y=fatalities_85_99))+geom_bar(stat="identity")
plot
```



### decrease in fatalities comparing fatalities of 1985-1999 and 2000-2014
```{r}
plot(airline$fatalities_85_99,type="l",col="blue", xlab="Airlines",ylab= "Fatalities")

lines(airline$fatalities_00_14,type="l",col="red")

legend(39,530,legend<-c("1985-1999","2000-2014",col=c("blue","red")))

```

Column {data-width=350}
-----------------------------------------------------------------------
### fatal accidents during 2000-2014 that led to fatalities
```{r}
ggplot(airline,aes(x=fatal_accidents_00_14,y=fatalities_00_14))+geom_bar(stat="identit
y",width=0.1)


```

### fatalities of different airlines during 2000-2014

```{r}
ggplot(airline,aes(x=airline,y=fatalities_00_14))+geom_bar(stat="identity")


```

### fatal accidents during 1985-1999 that led to fatalities

```{r}
ggplot(airline,aes(x=fatal_accidents_85_99,y=fatalities_85_99))+geom_bar(stat="identit
y",width=0.5)


```
````

Dashboard 2: Interactive plots

code:

````
---
title: "Interactive plots"
output:
  flexdashboard::flex_dashboard:
````
````

````
    orientation: columns
    vertical_layout: fill
---

```{r setup, include=FALSE}
library(flexdashboard)
library(plotly)
airline<-read.csv("C:/Users/Harshita/WinterSem22/Data Vis/Lab/airlinedata.csv")
```

Column {data-width=600}
-----------------------------------------------------------------------

###  Horizontal Stacked Bar Graph showing Fatalities of different airlines.


```{r}


p <- plot_ly(airline, x = airline$fatalities_85_99, y = airline$airline, type = 'bar',
 orientation = 'h', name = 'Fatalities between 1985 to 1999',
                  marker = list(color = 'rgba(246, 78, 139, 0.6)',
line = list(color = 'rgba(246, 78, 139, 1.0)'
                                                ,
width = 3))) %>%
add_trace(x = airline$fatalities_00_14, name = 'Fatalities between 2 000 to 2014',
marker = list(color = 'rgba(58, 71, 80, 0.6)',
line = list(color = 'rgba(58, 71, 80, 1.0)',
width = 3))) %>%
layout(barmode = 'stack',
xaxis = list(title = ""),
yaxis = list(title =""))
 p
```




### Line plot showing the fatalities of airline during 1985-1999


```{r}
p <- plot_ly(airline, x = airline$airline, y = airline$fatalities_85_99, type = 'scatte
r', mode = 'lines')
p

```



Column {data-width=350}
-----------------------------------------------------------------------

### fatal accidents which led to fatalities of airlines during 2000-2014


```{r}
p <- plot_ly(airline, x = airline$airline, y = airline$fatalities_00_14, type = 'scatte
````
````

```
r', mode = 'lines', color='red')
p
```

incidents which have led to fatal accidents of these airlines

```{r}
p <- plot_ly(airline, x = airline$airline, y = airline$fatalities_85_99, name = 'Fatali
ties
1985-1999', type = 'scatter', mode = 'lines',
line = list(color = 'rgb(205, 12, 24)')) %>%
add_trace(y = airline$fatalities_00_14, name = 'Fatalities 2000-2014 ', line = list(col
or = 'rgb(22, 96, 167)'))
p
```

Result Analysis

- *the horizontal bar graph shows the fatalities of different airlines during two time periods stacked on top of each other. Overall, it can be concluded that China Airlines has had the most number of fatalities*
- *The line graph shows the number of fatalities of different airlines during the year 1985-1999. China Airlines has shown the most number of fatalities.*
- *The above line graph shows the number of fatalities of different airlines during the year 2000-2014. Malaysian Airlines has shown the most number of fatalities.*
- *Using multiple line plots in the same graph, we can conclude that the number of fatalities for most of the airlines have decreased. But some airlines have had more fatalities in 2000-2014 than during 1985-1999.*

Conclusion

The first dashboard shows the fatalities caused by different airlines during 1985-99 and 2000-2014, and the second dashboard shows a few interactive plots regarding the same. The airline which caused the most fatalities during this time in China Airlines. The total fatalities count is 6,295. The airline which caused the most fatalities during this time is

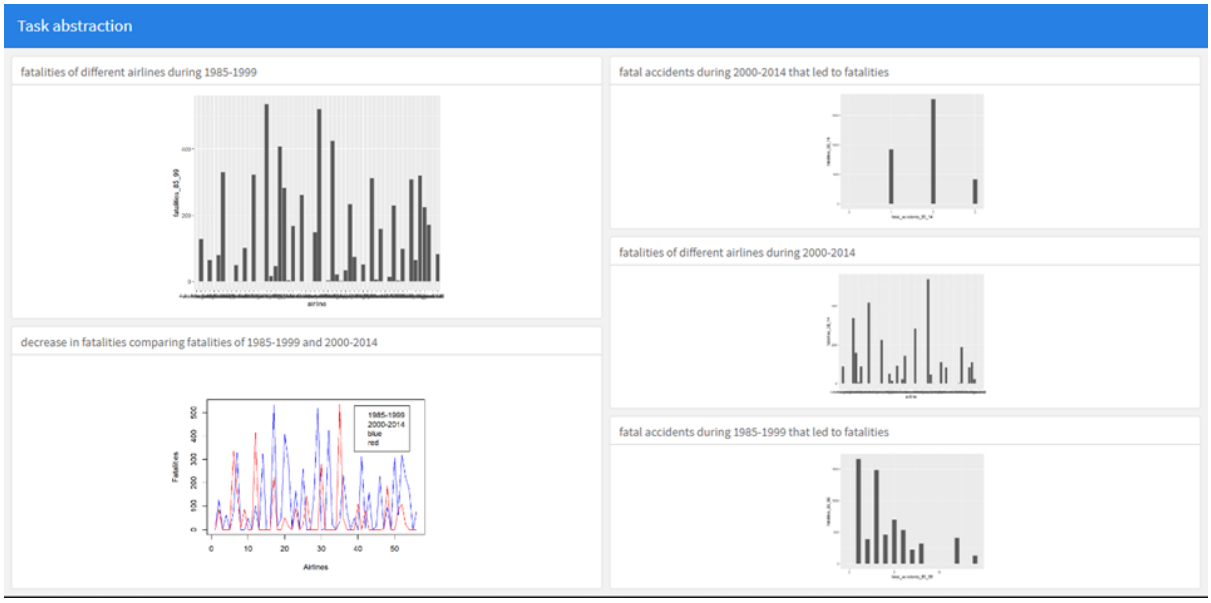
Malaysian Airlines. The total fatalities is 3,109 and plots showing the different plane sizes and types that had fatalities are shown.

It can be seen there has a decrease in the number of fatalities from 6,295 to 3,109. So airlines have become more reliable. But some airlines still cause a lot of fatalities and their issues much be looked into to make travel feel safe for its customers.

Appendix

1. Dashboard Screenshots:

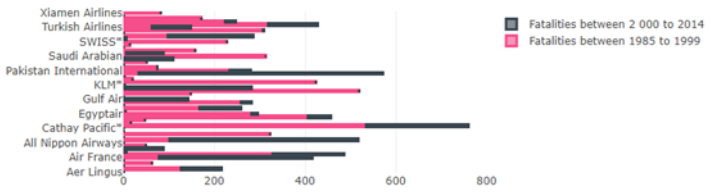
a)Task abstraction:



b) Interactive plots:

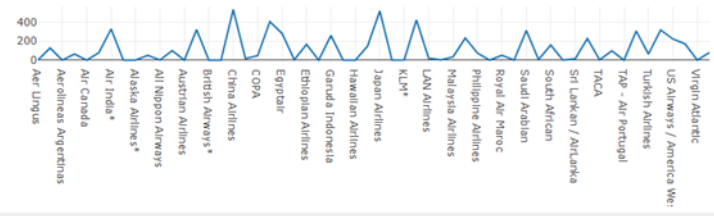
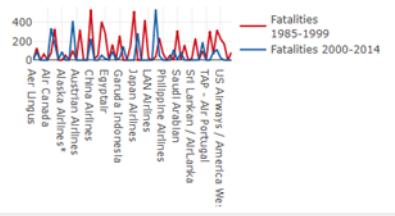
Interactive plots

Horizontal Stacked Bar Graph showing Fatalities of different airlines.



fatal accidents which led to fatalities of airlines during 2000-2014

Line plot showing the fatalities of airline during 1985-1999



incidents which have led to fatal accidents of these airlines