First Pair Assignment for Introduction to Collaborative Social Science Data Analysis (MPP-E1180)

Paulo Kalkhake & Dan Murphy
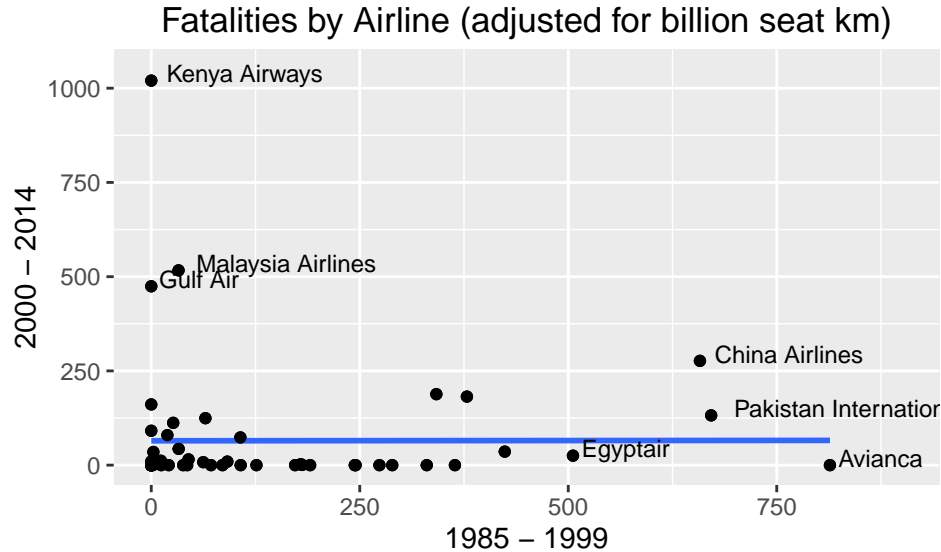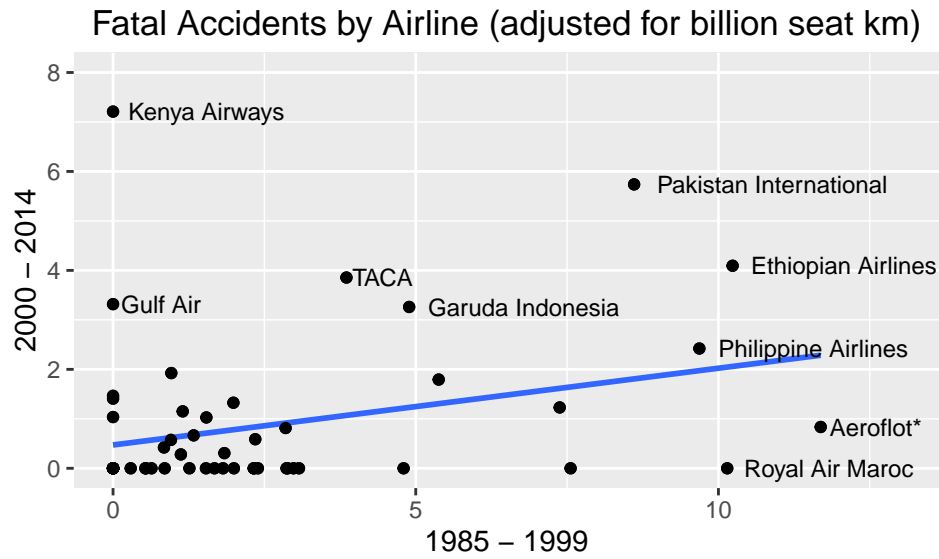
6 October 2016

Description

This Paper contains some of the results from the first pair assignment in Introduction to Collaborative Social Science Data Analysis (MPP-E1180). We performed basic transformations on the data, and illustrated the data's distributions using a variety relevant of descriptive statistics.

Data Set 1: Airline Safety

The first data set is from fivethirtyeight.com and is based on the Aviation Safety Network's database. It breaks down the number of incidents, fatal accidents, and fatalities in 56 airlines. These incidents, accidents, and fatalities are separated into two groups, those that occured in between 1985 to 1999 and those that occured between 2000 and 2014. The data set also includes the available seat kilometers per week for each airline, which fivethirtyeight defines as the number of seats multiplied by the number of kilometres the airline flies.

Our analysis confirms the author's results: There seems to be no clear relationship between rate of fatal accidents in period one (1985 - 1999) and period two (2000 - 2014).

Fatalities by Airline (adjusted for billion seat km)

Fatal Accidents by Airline (adjusted for billion seat km)

However, there seems to be some correlation, if one looks at the number of incidents, i. e. whether they resulted in a fatality or not, in both periods.

Incidents by Airline (adjusted for billion seat km)
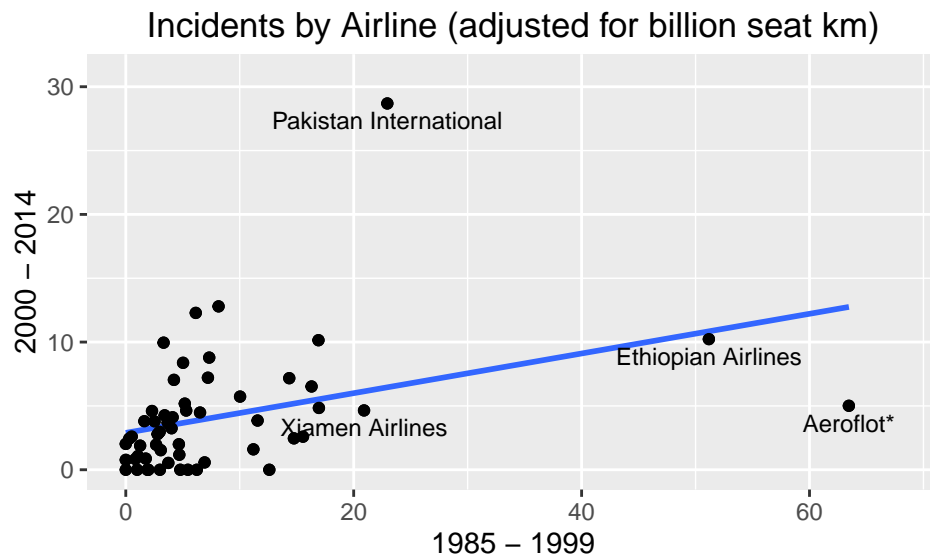
Data Set 2: Worldwide Alcohol Consumption in 2010

The second data set also comes from fivethirtyeight.com, and deals with global alcohol consumption. Specifically, it provides the World Health Organization's (WHO) data for the total litres of alcohol consumed per person in 193 countries around the world in 2010. It also includes adjusted data showing the number of servings of beer, wine, and spirits consumed by each person.

Our analysis shows the strongest relationship between beer drinking and total litres of alcool consumed.
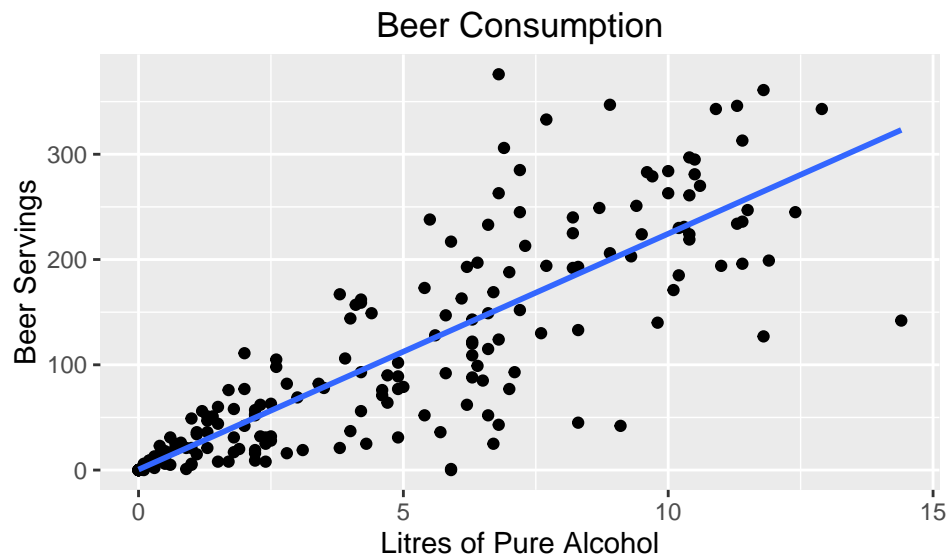
Yet, our analysis also shows diverse drinking habits among the top four total alcohol consumers. Andorrans have the most variety in their consumption, while Belarus and Grenada show strong preferences for spirits. Of the top four, only Lithuania favors beer.