

# BA Homework #3

*Business Analytics - Spring 2020, NYU*

---

## INSTRUCTIONS & SUBMISSION

Complete the problems below and submit via NYUClasses. Submit a PDF file with your answers, graphs, and R code.

### Problem 1: CitiBike anomaly detection & neighborhood usage

CitiBike data reveals how users go about using bikes for travel needs such as commuting, entertainment, and tourism. However, the data is highly criticized for anomalies.

#### **Link to dataset:**

<https://github.com/jcbonilla/BusinessAnalytics/blob/master/BADData/Citibike%20Data.csv>

#### **Questions:**

Explore this CitiBike dataset and answer the following questions:

1. What anomalies detectable with tripduration and the age of the user
2. Based on your anomalies analysis in (1) , what recommendation would you provide to the CitiBike operators

### Problem 2: Aviation Accidents

Aviation accidents involve human injuries, fatalities, and aircraft damages in multiple levels.

#### **Link to dataset:**

<https://github.com/jcbonilla/BusinessAnalytics/blob/master/BADData/aviation.csv>

#### **Questions:**

Using the aviation data provide, develop an analysis that does the following:

A visualization of fatal vs. non-fatal crashes in the US from the 1940s through 2013.

### Problem 3: Retail Targets

Assume you are a Northeaster (NE) regional manager and have been asked to benchmark your region against the performance of other regions in the USA. Analyze and visualize the 10-year percent change in demographics.

The NE Region includes the following states:

- Maine(ME), New York(NY), New Jersey(NJ), Vermont(VT), Massachusetts(MA), Rhode Island(RI), Connecticut(CT), New Hampshire(NH), and Pennsylvania(PA)

For each 10-year percent change. Demographics include:

- Population, Income, Density, Ownhome
- Percentage of U18, college, white, black

***Link to dataset:***

<https://github.com/jcbonilla/BusinessAnalytics/blob/master/BAData/HDLData.csv>