
Application User Guide
IS428 Visual Analytics for Business Intelligence
AY 2018/19 T1
The Defenders
Tan Xinyi, Lim Li Xuan, Yeo Qi Xun

# Introduction

With the rise in cybersecurity threats in the world today, our goal is to empower respective policy makers, government or cybersecurity departments of companies to identify the target victims of data breaches and establish correlations between the types of breaches and the sources of breaches.

The Defenders has created an interactive visual analytics web application that aims to assist the users of the application to visually discover interesting patterns from the data that can guide their policymaking to reduce the occurence of these data breaches.

The dataset that we are visualising is the global data breaches reported to Gemalto from 2013 to 2017. It may not be totally representative of all the data breaches that are happening around the world as the velocity of data breaches is far too high and not all data breaches are reported. However, it is still useful to assist in the analyst's discovery of plausible patterns of data breaches.

Users should be able to execute the following actions using our application:

1. Explore type, source and impact of breaches globally
2. Compare the nature and impact of reported breaches between countries
3. Identify trends in breaches for each industry over time

# Prerequisites for Executing our Application

You should have the following installed in your local machine:

- Web Browser (Preferably Firefox)
- R (latest version should be fine)
- RStudio (latest version should work)

1. Unzip the packaged zip files into any location on your local machine and remember the location
2. Start up your RStudio and load the server.R and ui.R into it
3. Ensure that you have the required libraries at the top of the files installed
    a. If not installed
        i. Click on Tools -> Install packages
        ii. Type in the names of required packages shown at the top of both server.R and ui.R (each module is denoted by library(<name of module>) so type in the <name of module> at the textbox)
        iii. Click install and install them
    b. If installed
        i. Skip to Step 4
4. Click on the right green arrow with Run App beside it
5. Open in browser for the best viewing experience

# Basic Navigation

There are tabs located conveniently at the top of the page that the analyst can toggle to view the different tabs on display. There are a total of four tabs that can be seen: Overview, Risk of Attacks, Timeline of Attacks and Treemap Comparison.

## Overview

The overview page features a scatter plot of type of breach against source of breach. Hover on the individual points to see finer details about each breach

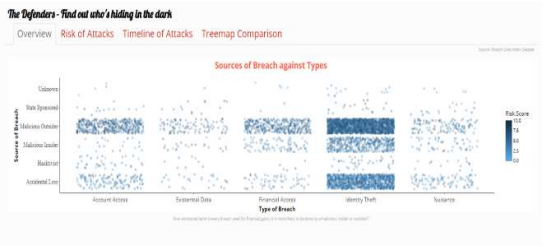


Fig 1: Overview of Breaches

## Risk of Attacks

The risk of attack page features multiple boxplots of data breaches along with its original data points. Hover on the individual points to see finer details about each breach

1. **Selection of Type of Analysis -** The analyst can toggle between having type of breach and source of breach on the x-axis
2. **Selection of Measure of Severity -** The analyst can toggle between the view of individual points being compared by risk score or by number of records breached.
3. **Filter by Year-** The analyst can move the slider filter points in a particular year between 2013 and 2018
4. **Filter by Industry-** The analyst can click on the radio button to filter breaches found in the chosen industry.

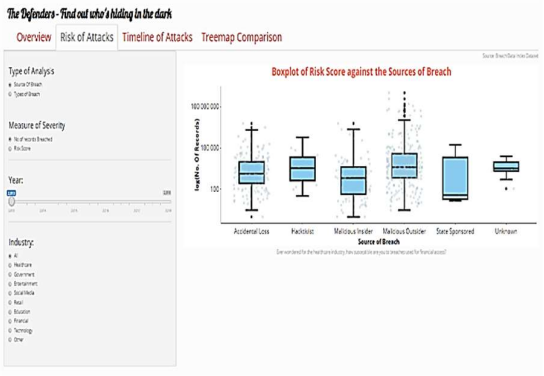


Fig 2: Boxplot of Risk of Attacks

## Timeline of Attacks

The timeline of attacks page features a calendar heatmap of breaches happening in various industries from 2013 to 2018. Hover on the individual cells to see finer details about number of records breached for that specific industry on the specific day.

1. **Filter by Industry-** The analyst can click on the dropdown list to filter breaches found in the chosen industry.


Fig 3: Calendar Heatmap of Breaches

## Treemap Comparison

The treemap comparison page features two treemaps visualising finer grain details of the data breaches positioned side-by-side. Hover on the individual squares to see finer details about the characteristics of breaches from the segment.

1. **Selection of Type of Analysis-** The analyst can toggle between having type of breach or source of breach as the first layer of the treemap
2. **Filter by Year-** The analyst can move the slider filter points in a particular year between 2013 and 2018
3. **Filter by Country-**The analyst can click on the dropdown list of continents and the list of countries will dynamically change to be from the selection of continent. Upon choosing the country to analyse, the respective treemaps will reshape
4. **Drilldown by Clicking-** The analyst can click on the area demarcated by the different components of the type of analysis (either type of breach or source of breach) to gain a better visualisation of the finer details of industry breakdown in the respective selections.

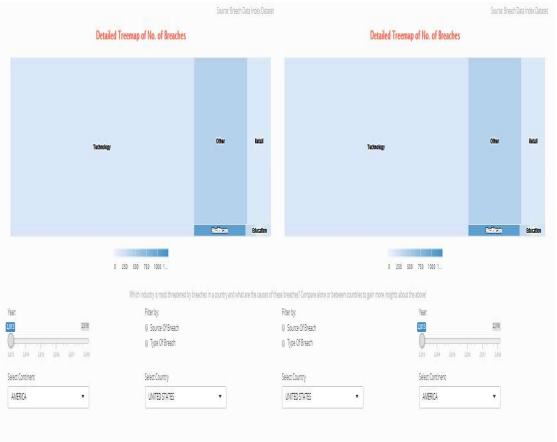


Fig 4: Treemap Comparison of Breaches

# Appendix

These are a glossary of important terms used in the Application

| Term | Definition |
|---|---|
| Source of Breach | The place or person from which the data breaches originates or is obtained from |
| Type of Breach | The category from which the data breach falls under |
| Risk Score | A metric to calculate the damage of the breach to the victims.<br>The formula is given by:<br>*Log 10 (Type of Data * Source of Breach * Action)* |

Type of Data

| Value | Type of Data |
|---|---|
| 1 | Nuisance (email addresses, affliation, etc) |
| 2 | Account access (username/passwords to social media, websites, etc) |
| 3 | Financial access (bank account credentials, credit card data) |
| 4 | Identity theft (information that can be used to masquerade as someone) |
| 5 | Existential data (information of national security value or threatens business survival) |

Source of Breach

| Value | Source of Breach |
|---|---|
| 1 | Lost device such as laptop, DVD, or USB thumb drive |
| 2 | Stolen device |
| 3 | Malicious insider |
| 4 | Malicious outsider |
| 5 | State espionage |

Action (whether or not the stolen data has been used to cause harm)

| Value | Action |
|---|---|
| 1 | No action |
| 5 | Publication of embarassing or harmful information (Wikileaks, hacker logs, etc) |
| 10 | Use of financial identity to obtain funds or apply for loans |

Risk Score

| Breach Level Index Score | Characterization |
|---|---|
| 9 – 10 | Breach with immense long term impact on breached organisation, customers and/or partners. Very large amount of highly sensitive information lost (at least 10 – 100 million records). Massice notification process costs. Potentially existential financial loss for breached organisation in costs. Use of lost sensitive information for damaging purposes |
| 7 – 8.9 | Breach with significant exposure to business with large legal and/or regulatory impact. Large amounts of sensitive information lost. Significant notification process cost and public image impact |
| 5 – 6.9 | Breach with short to midterm exposure to business with some legal and/or regulatory impact. Significant amounts of moderately sensitive information lost. Some notification process cost and financial loss |
| 3 – 4.9 | Breach with low long-term business impact. Usually involves the loss several thousands of records of semi sensitive information. Limited breach notification and financial exposure |
| 1 – 2.9 | Breach with no material effect. Less than one thousand records. Little damage is done |