Special Issue: Teaching Cases

In this issue:

4. **Black Box Thinking: Analysis of a Service Outsourcing Case in Insurance**  
   Paul D. Witman, California Lutheran University  
   Christopher Njunge, California Lutheran University

14. **Wooden Peg Game: Implementations as Both a Web App and as an Android App**  
    Cynthia J. Martincic, Saint Vincent College

19. **Advancing Student Productivity: An Introduction to Evernote**  
    Melinda Korzaan, Middle Tennessee State University  
    Cameron Lawrence, The University of Montana

27. **Single Sourcing, Boilerplates, and Re-Purposing: Plagiarism and Technical Writing**  
    Michelle O’Brien Louch, Duquesne University

34. **Too Much of a Good Thing: User Leadership at TPAC**  
    Brett Connelly, Miami University  
    Tashia Dalton, Miami University  
    Derrick Murphy, Miami University  
    Daniel Rosales, Miami University  
    Daniel Sudlow, Miami University  
    Douglas Havelka, Miami University

43. **Analyzing Security Breaches in the U.S.: A Business Analytics Case-Study**  
    Rachida F. Parks, University of Arkansas at Little Rock  
    Lascelles Adams, Bethune-Cookman University

49. **Stalled ERP at Random Textiles**  
    Robert Brumberg, Miami University  
    Eric Kops, Miami University  
    Elizabeth Little, Miami University  
    George Gamble, Miami University  
    Jesse Underbakke, Miami University  
    Douglas Havelka, Miami University

58. **Front-End and Back-End Database Design and Development: Scholar’s Academy Case Study**  
    Rachida F. Parks, University of Arkansas at Little Rock  
    Chelsea A. Hall, University of Arkansas at Little Rock
The Information Systems Education Journal (ISEDJ) is a double-blind peer-reviewed academic journal published by ISCAP, Information Systems and Computing Academic Professionals. The first year of publication was 2003.

ISEDJ is published online (http://isedj.org). Our sister publication, the Proceedings of EDSIGCon (http://www.edsigcon.org) features all papers, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the ISEDJ journal. Currently the target acceptance rate for the journal is under 40%.

Information Systems Education Journal is pleased to be listed in the 1st Edition of Cabell's Directory of Publishing Opportunities in Educational Technology and Library Science, in both the electronic and printed editions. Questions should be addressed to the editor at editor@isedj.org or the publisher at publisher@isedj.org. Special thanks to members of AITP-EDSIG who perform the editorial and review processes for ISEDJ.

2016 AITP Education Special Interest Group (EDSIG) Board of Directors

Scott Hunsinger  
Appalachian State Univ  
President

Leslie J. Waguespack Jr  
Bentley University  
Vice President

Wendy Ceccucci  
Quinnipiac University  
President – 2013-2014

Nita Brooks  
Middle Tennessee State Univ  
Director

Meg Fryling  
Siena College  
Director

Tom Janicki  
U North Carolina Wilmington  
Director

Muhammed Miah  
Southern Univ New Orleans  
Director

James Pomykalisky  
Susquehanna University  
Director

Anthony Serapiglia  
St. Vincent College  
Director

Jason Sharp  
Tarleton State University  
Director

Peter Wu  
Robert Morris University  
Director

Lee Freeman  
Univ. of Michigan - Dearborn  
JISE Editor

Copyright © 2016 by the Information Systems and Computing Academic Professionals (ISCAP). Permission to make digital or hard copies of all or part of this journal for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial use. All copies must bear this notice and full citation. Permission from the Editor is required to post to servers, redistribute to lists, or utilize in a for-profit or commercial use. Permission requests should be sent to Jeffry Babbs, Editor, editor@isedj.org.
Analyzing Security Breaches in the U.S.: A Business Analytics Case-Study

Rachida F. Parks
rfparks@ualr.edu
University of Arkansas at Little Rock
Little Rock, AR

Lascelles Adams
adamsl@cookman.edu
Bethune-Cookman University
Daytona Beach, FL, USA

Abstract

This is a real-world applicable case-study and includes background information, functional organization requirements, and real data. Business analytics has been defined as the technologies, skills, and practices needed to iteratively investigate historical performance to gain insight or spot trends. You are asked to utilize/apply critical thinking skills that will produce measurable insights from historical performance data that can be transformed into actionable insights. By critical analysis, reporting and visualization, you will engage with the three major analytic activities: (1) extract, transform and load (ETL) the data; (2) create reports and visualization graphs using business intelligence tools (e.g., IBM Cognos Insight, Tableau, Excel, SQL Server Reporting Services (SSRS)); and (3) engage in critical thinking to identify actionable items that will assist with decision making or recommendations. The Privacy Rights Clearinghouse (PRC) Chronology of Data Breaches reported more than 800 million records breached from over 4500 data breaches since 2005. Security breaches continue to increase: Therefore, there is an urgent need to analyze their patterns and provide meaningful insights to support better decision making.

Keywords: Information security, Business intelligence, Business analytics, Teaching case, Critical thinking

1. CASE SUMMARY

Data breaches have gone mainstream. Whether it is Sony PlayStation Network, Target, or Anthem Health attack, more and more consumers are receiving notifications from companies stating that sensitive, personally identifiable information has been exposed and possibly compromised. Across all industries, as the number of data breaches has increased, eliminating data breaches and protecting business-critical data remains a top priority as well as government’s interest in protecting its citizens.

The Bureau of Consumer Protection, Information Security Commission (ISC) Office recently organized a task force seeking to understand the nature of data breaches across the government, and selected industries including: retail, financial and insurance, and healthcare. The task force is seeking your help to collect and analyze the data obtained from the Privacy Rights Clearinghouse Chronology of Data Breaches.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSO</td>
<td>Businesses Other</td>
</tr>
</tbody>
</table>
Data breaches are endangering the privacy and confidentiality of consumers and resulting in dire organizational consequences, such as reputation damage, monetary penalties, and civil and criminal liabilities. Organization leaders recognize the importance of keeping track of breaches trends and their impacts.

Therefore, there is an urgent need to have tools that communicate patterns difficult to see. As part of the task force, your team will produce a report to help the bureau understand the patterns of data breaches in the government and the selected industries identified above.

The case study includes a set of functional requirements (technical, business, and critical thinking), the data set, and the business intelligence tools to be used. Below are the detailed technical and business functional requirements.

## 2. TECHNICAL REQUIREMENTS

**Extract – Transform – Load Process**

Clean, meaningful and useful data play a very important role in getting better analysis and more insightful results. The ETL process consists of three steps; extract, transform, and load. First, the extract step consists of pulling data from a source. Next, the transform step involves converting the raw data into a form that can be loaded into a target system for further processing. Lastly, the load step entails putting the transformed data into the target system. In the end the ETL process take source data, cleans it, and integrates it into a target system.

Before beginning the analysis, you will perform an ETL process set using PRC data.

- **Extraction**: extract data from the PRC website. The dataset can be downloaded as a CSV file from http://www.privacyrights.org/data.
- **Transformation**: transform data into a suitable format as required by the target system. The transformation process takes place in Excel by saving the CSV file into an Excel file. You can delete the columns for street address, the name of the business, postal code, city, description, and a few unnamed columns. You can remove either the “total records” or “records breached” column to eliminate redundancy. In addition, a few columns need to have their headers changed. For example "location" to "city" and the “state” column to be changed from ",,.

Additional columns can be derived from the “Date” column such as the year and

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>How it occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISC</td>
<td>Unintended Disclosure</td>
<td>Information made public via web</td>
</tr>
<tr>
<td>HACK</td>
<td>Hacking or Malware</td>
<td>Electronic entry by outside party</td>
</tr>
<tr>
<td>CARD</td>
<td>Payment Card Fraud</td>
<td>Fraud via debit or credit cards accomplished by means other than hacking</td>
</tr>
<tr>
<td>INSD</td>
<td>Insider</td>
<td>Someone with legitimate access intentionally breaches non electronic records lost, discarded, or stolen</td>
</tr>
<tr>
<td>PHYS</td>
<td>Physical Loss</td>
<td>Data lost, stolen or discarded through electronic devices</td>
</tr>
<tr>
<td>STAT</td>
<td>Stationary device</td>
<td>Lost, stolen or discarded stationary device</td>
</tr>
<tr>
<td>UNKN</td>
<td>Unknown or other</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Data breaches are endangered via http://www.isedj.org 2014 (2) Month 2016

Privacy Rights Clearinghouse (PRC), accessible via http://www.privacyrights.org/data-breach, is a nonprofit corporation in California that was established in 1992. PRC keeps up-to-date information of data breaches across all industries and the government within the US. PRC aims to provide timely and historical information on data breaches and to educate stakeholders (e.g., consumers, businesses, and policymakers) of currents trends in data breaches. PRC reported more than 800 million records breached from over 4500 data breaches since 2005. These breaches span across financial institutions, retail, educational institutions, government and military, healthcare, non-profit and other businesses (see table 1). PRC also provides data about different types of breaches as outlined in table 2.

Table 2. Type of Breaches

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>How it occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISC</td>
<td>Unintended Disclosure</td>
<td>Information made public via web</td>
</tr>
<tr>
<td>HACK</td>
<td>Hacking or Malware</td>
<td>Electronic entry by outside party</td>
</tr>
<tr>
<td>CARD</td>
<td>Payment Card Fraud</td>
<td>Fraud via debit or credit cards accomplished by means other than hacking</td>
</tr>
<tr>
<td>INSD</td>
<td>Insider</td>
<td>Someone with legitimate access intentionally breaches</td>
</tr>
<tr>
<td>PHYS</td>
<td>Physical Loss</td>
<td>Non electronic records lost, discarded, or stolen</td>
</tr>
<tr>
<td>PORT</td>
<td>Portable device</td>
<td>Data lost, stolen or discarded through electronic devices</td>
</tr>
<tr>
<td>STAT</td>
<td>Stationary device</td>
<td>Lost, stolen or discarded stationary electronic device</td>
</tr>
<tr>
<td>UNKN</td>
<td>Unknown or other</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Table 2. Type of Breaches

Data breaches are endangering the privacy and confidentiality of consumers and resulting in dire organizational consequences, such as reputation damage, monetary penalties, and civil and criminal liabilities. Organization leaders recognize the importance of keeping track of breaches trends and their impacts.

Therefore, there is an urgent need to have tools that communicate patterns difficult to see. As part of the task force, your team will produce a report to help the bureau understand the patterns of data breaches in the government and the selected industries identified above.

The case study includes a set of functional requirements (technical, business, and critical thinking), the data set, and the business intelligence tools to be used. Below are the detailed technical and business functional requirements.

## 2. TECHNICAL REQUIREMENTS

**Extract – Transform – Load Process**

Clean, meaningful and useful data play a very important role in getting better analysis and more insightful results. The ETL process consists of three steps; extract, transform, and load. First, the extract step consists of pulling data from a source. Next, the transform step involves converting the raw data into a form that can be loaded into a target system for further processing. Lastly, the load step entails putting the transformed data into the target system. In the end the ETL process take source data, cleans it, and integrates it into a target system.

Before beginning the analysis, you will perform an ETL process set using PRC data.

- **Extraction**: extract data from the PRC website. The dataset can be downloaded as a CSV file from http://www.privacyrights.org/data.
- **Transformation**: transform data into a suitable format as required by the target system. The transformation process takes place in Excel by saving the CSV file into an Excel file. You can delete the columns for street address, the name of the business, postal code, city, description, and a few unnamed columns. You can remove either the “total records” or “records breached” column to eliminate redundancy. In addition, a few columns need to have their headers changed. For example "location" to "city" and the “state” column to be changed from ",,.

Additional columns can be derived from the “Date” column such as the year and
the quarter. An example of a finalized and transformed dataset is provided in Appendix A.

- Loading: load the data into the business intelligence (BI) application. This starts by loading or importing the transformed file into your BI tool. During the loading, you have to identify dimensions and measures among the available attributes. You can choose several dimensions such as entity, type, state, year, and quarter. An example of measures is the number of breached records and the count of breaches

**Business Intelligence Tools**
The ISC Office and organization leaders want a tool that allows them to know the state of information breaches across different industries, impacted geographical areas (States with higher breaches) and type of breaches by which they are threatened.

BI tools are used to retrieve, analyze, and report data to support decision making by providing meaningful insights. Several BI tools are available and some providers offer trial versions or free student versions. Some of the BI tools currently available are EXCEL, SSRS, Tableau, and IBM Cognos Insights.

You will need to select the appropriate BI software to use and install it; keeping in mind that some BI tools are only windows compatible.

### 3. BUSINESS REQUIREMENTS

**Reporting or Descriptive analytics**
Descriptive analysis answers the “What” questions (See Appendix B) and provides a view of both current and historical results. Descriptive analytics tells the business how they are performing and help identify key issues in their current performances. Using the dataset provided by PRC, you need to address the requirements outlined below:

1. **Total Number of Breaches:** This report should outline the number of breaches by year; broken down by government and the selected industries (e.g., retail, financial and healthcare). The visual representation should allow the changes and trends throughout the years among the above-mentioned industries and government. Advanced report options include exploring any seasonal trends (e.g., quarterly analysis).

2. **Type of Data Breaches Report:** This report compares the different types of data breaches reported. This report can be broken into two sub-reports. The first sub-report will show different breach trends across all different years provided, while the second sub-report will represent different type of breaches broken down by industry to be able to see the most prevalent type of breaches within different industries. Advanced report options should combine the findings from the previous report (total number of breaches) to the type of data breaches. This will allow comparison of the highest occurrences of breaches to the type of breaches reported.

3. **Geographic Location of Breaches Report:** This report describes the geographic locations (States) associated with the breaches in the United States. Advanced report options can integrate this geographic representation to be interactively displayed on a U.S. map.

4. **Citizens Impacted by Breaches Report:** This report provides a comparison of the total individuals impacted by the breaches in comparison to the total number of breach occurrences. Advanced report options should consider adding the type of breaches and their location to this report for more meaningful insights.

5. **Cost Analysis of Breaches Report:** This is an advanced report options where you can include a new measure called “Estimated Cost of a Breach”. This measure is calculated based on the average cost per breach published by the Ponemon Institute. By multiplying the records breached by the average cost per record you will be able to determine an estimated cost for an entire breach for an entity. In your reports, you can analyze (1) the total cost of breaches per industry, (2) the cost per type of breach and, (3) cost of breaches over the years.

**Visual Communication**
Organization leaders need to be able to make quick and accurate decisions. Therefore, a need for simple graphs that stand out and aid in their decision making is very important. These graphs provide insights into the data and address the key issues of the problem identified in descriptive analytics reports.
You should create visual representation of all the above reports along with a short analysis of the reports.

4. ACTIONABLE INSIGHTS: CRITICAL AND ANALYTICAL THINKING

While having descriptive and historical reports is very important, business leaders are looking for the needle in the haystack. Meaning they want you to provide them with recommendations that are actionable. Your recommendations should be based on the insights gathered through the reports and supplemented by recommendations to safeguards against data breaches. These preventative recommendations can be:

- Pertinent to specific types of breaches.
- Specific to a particular industry or government or span across all the industries and government.
- Either technical recommendations (e.g., encryption), human recommendations (e.g., education/training, awareness, social engineering), or policies.

5. CONCLUSION

As data breaches continue to increase, countermeasures can only be effective if they align with what is causing the breaches. Therefore, a thorough analysis of breaches and their trends is crucial. A real and up-to-date data set of security breaches is provided by the privacy rights clearinghouse. The Information Security Commission Office is seeking your help to clean the data, build and analyze reports, and ultimately provide insights into the state of data breaches in the U.S. for the purpose of better preventative safeguards.

Note: Teaching Notes and Case Supplements are available by contacting the authors
Appendix A – Sample of a Transformed Dataset

<table>
<thead>
<tr>
<th>Date Made Public</th>
<th>Name</th>
<th>Entity Type</th>
<th>City</th>
<th>State</th>
<th># Records Breached</th>
<th>Year</th>
<th>Quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>26-Jan-09</td>
<td>Heartland Paym BSF</td>
<td>HACK</td>
<td>Princeton</td>
<td>New Jersey</td>
<td>190,000,000</td>
<td>2009</td>
<td>Q1</td>
</tr>
<tr>
<td>17-Jan-07</td>
<td>T: stones (TIX), iBSF</td>
<td>HACK</td>
<td>Framingham</td>
<td>Massachusetts</td>
<td>100,000,000</td>
<td>2007</td>
<td>Q1</td>
</tr>
<tr>
<td>5-Feb-15</td>
<td>Anthem BSF</td>
<td>HACK</td>
<td>Indianapolis</td>
<td>Indiana</td>
<td>80,000,000 Media</td>
<td>2015</td>
<td>Q1</td>
</tr>
<tr>
<td>2-Oct-09</td>
<td>U.S. Military Gov</td>
<td>PORT</td>
<td>Washington</td>
<td>District Of Columbia</td>
<td>76,000,000</td>
<td>2009</td>
<td>Q4</td>
</tr>
<tr>
<td>2-Sep-14</td>
<td>The Home Depot BSF</td>
<td>HACK</td>
<td>Atlanta</td>
<td>Georgia</td>
<td>56,000,000 Media</td>
<td>2014</td>
<td>Q3</td>
</tr>
<tr>
<td>16-Jun-05</td>
<td>CardSystems BSF</td>
<td>HACK</td>
<td>Tucson</td>
<td>Arizona</td>
<td>40,000,000</td>
<td>2005</td>
<td>Q2</td>
</tr>
<tr>
<td>13-Dec-13</td>
<td>Target Corp. BSR</td>
<td>HACK</td>
<td>Minneapolis</td>
<td>Minnesota</td>
<td>40,000,000 Media</td>
<td>2013</td>
<td>Q4</td>
</tr>
<tr>
<td>10-Nov-11</td>
<td>Steam (The Val BSR)</td>
<td>HACK</td>
<td>Bellevue</td>
<td>Washington</td>
<td>35,000,000 Databreaches.net</td>
<td>2011</td>
<td>Q4</td>
</tr>
<tr>
<td>22-May-06</td>
<td>U.S. Department Gov</td>
<td>PORT</td>
<td>Washington</td>
<td>District Of Columbia</td>
<td>26,500,000</td>
<td>2006</td>
<td>Q2</td>
</tr>
<tr>
<td>2-Aug-08</td>
<td>Countrywide Fr BSF</td>
<td>INSD</td>
<td>Calabasas</td>
<td>California</td>
<td>17,000,000</td>
<td>2008</td>
<td>Q3</td>
</tr>
<tr>
<td>26-Mar-08</td>
<td>Bank of New York BSF</td>
<td>PORT</td>
<td>Pittsburgh</td>
<td>Pennsylvania</td>
<td>12,500,000 Media</td>
<td>2008</td>
<td>Q1</td>
</tr>
<tr>
<td>27-Apr-11</td>
<td>Sony, Playstation BSR</td>
<td>HACK</td>
<td>New York</td>
<td>New York</td>
<td>12,000,000</td>
<td>2011</td>
<td>Q2</td>
</tr>
<tr>
<td>3-Jul-07</td>
<td>Fidelity Natl BSF</td>
<td>INSD</td>
<td>Jacksonville</td>
<td>Florida</td>
<td>8,500,000</td>
<td>2007</td>
<td>Q3</td>
</tr>
<tr>
<td>30-Mar-12</td>
<td>Global Payment BSF</td>
<td>CARD</td>
<td>Atlanta</td>
<td>Georgia</td>
<td>7,000,000</td>
<td>2012</td>
<td>Q1</td>
</tr>
<tr>
<td>27-Apr-12</td>
<td>Office of the Tre Gov</td>
<td>DISC</td>
<td>Austin</td>
<td>Texas</td>
<td>6,500,000 Media</td>
<td>2012</td>
<td>Q2</td>
</tr>
<tr>
<td>26-Oct-12</td>
<td>South Carolina Gov</td>
<td>HACK</td>
<td>Columbia</td>
<td>South Carolina</td>
<td>6,400,000 Media</td>
<td>2012</td>
<td>Q4</td>
</tr>
<tr>
<td>14-Sep-07</td>
<td>TD Ameritrade BSF</td>
<td>HACK</td>
<td>Omaha</td>
<td>Nebraska</td>
<td>6,300,000</td>
<td>2007</td>
<td>Q3</td>
</tr>
<tr>
<td>6-Jan-09</td>
<td>CheckFree Corp BSF</td>
<td>HACK</td>
<td>Atlanta</td>
<td>Georgia</td>
<td>5,000,000</td>
<td>2009</td>
<td>Q1</td>
</tr>
<tr>
<td>18-Aug-14</td>
<td>Community Hea MED</td>
<td>HACK</td>
<td>Franklin</td>
<td>Tennessee</td>
<td>4,500,000</td>
<td>2014</td>
<td>Q3</td>
</tr>
<tr>
<td>17-Mar-08</td>
<td>Hannaford Bros BSF</td>
<td>HACK</td>
<td>Portland</td>
<td>Maine</td>
<td>4,200,000</td>
<td>2008</td>
<td>Q1</td>
</tr>
<tr>
<td>28-Aug-13</td>
<td>Advocate Medi MED</td>
<td>STAT</td>
<td>Park Ridge</td>
<td>Illinois</td>
<td>4,000,000</td>
<td>2013</td>
<td>Q3</td>
</tr>
<tr>
<td>6-Jun-05</td>
<td>Citigroup, UPS BSF</td>
<td>PORT</td>
<td>New York</td>
<td>New York</td>
<td>3,900,000</td>
<td>2005</td>
<td>Q2</td>
</tr>
</tbody>
</table>
Appendix B -- Questions to consider in your analysis

<table>
<thead>
<tr>
<th>Type of question</th>
<th>Sample question pertaining to the data set provided</th>
</tr>
</thead>
</table>
| WHAT questions            | • What type of breaches?  
                           | • What is the breakdown of breaches per year?  
                           | • What the breakdown of breaches per geographic location?  
                           | • What entities/type of entities are impacted the most?  
                           | • What are the seasonal trends if any (quarterly analysis)?  
                           | • What is the total number of breaches in comparison to the total of individuals impacted by the breaches? |
| Descriptive Analytics     |                                                                                                                       |
| WHY Questions             | These types of questions allow you to further analyze your findings. It involves more depth in your analysis/discussion |
| Getting to the meat       | getting and may require the use of other data sets.  
                           | • Why is a particular breach type higher than others?  
                           | • Why are breaches the highest in certain geographical areas (States)?  
                           | • Why is it important to focus on the number of breaches rather than the number of individuals impacted or vice versa?  
                           | • Why are breaches lower in certain States? (may have to check the stringency of their privacy/security policies) |
| HOW Questions             | • How do the above findings and discussion provide a competitive advantage or improve decision making?  
                           | Prescriptive Analytics  | • How can one achieve the best outcome using the following?  
                           | (check figure below)    |   ○ Technical countermeasures  
                           |                                                                 |   ○ Physical countermeasures  
                           |                                                                 |   ○ Policies and regulations  
                           |                                                                 |   ○ Training and education  

![Diagram showing the relationship between types of analytics and degree of complexity.](http://www.isedj.org)