Data visualization: an exploratory study into the software tools used by businesses

Michael Diamond
Jacksonville University

Angela Mattia
Jacksonville University

Abstract

Data visualization is a key component to business and data analytics, allowing analysts in businesses to create tools such as dashboards for business executives. Various software packages allow businesses to create these tools in order to manipulate data for making informed business decisions. The focus is to examine what skills employers are looking for in potential job candidates, and compare with the ability to include those technological skills in a business school curriculum. The researchers explored a variety of software tools, and reported their initial results and conclusions.

Keywords: data visualization, business analytics, dashboards

Introduction

The visualization and interpretation of data is becoming an important skill in today's business world. Indeed, the inclusion of data visualizations and dashboards allows executives to make decisions in a short amount of time given the information which is needed (Eigner, 2013). This is a critical part of the evolution of business intelligence to what is now known as the area of business analytics (Negash, 2004).

Furthermore, the growing importance of 'big data' has created a shortage of employees with data and business analytics skills, with the expectations of a need of close to 1.8 million jobs in these areas (Manyika et al., 2012; Lohr, 2012). The data scientist has been named the 'Sexiest job of the 21st Century' according to Harvard Business Review October 2012 edition (Davenport & Patil, 2012).

A variety of data visualization software packages are important to those who plan to work in the current business environment. Students enrolled in business school are often exposed to some form of statistical and data analysis, due to the growing number of employment opportunities which require these specialized skills. These software packages are also highly researched and maintain a high amount of relevance in academia, as well as corporate America.

We examined the number of hits on employment websites, Indeed and SimplyHired, a search on Google Scholar for related articles and an advanced search on Google with keywords including the software, along with the words "college" and "university." The results can be found below in Table 1.

Table 1: Number of hits on websites by software package

Software	SimplyHired	Indeed	Google Scholar	Advance Google
IBM	244,986	11,561	23,501	27,055,100
Microsoft	1,114,897	411,478	42,262	23,231,000
Pentaho	1,157	703	19	341,000
QlikView	233,639	2,082	39	427,000
R	30,322	55,288	58,700	254,000
SAS	414,994	17,094	206,000	97,300,000
Stata	1,006	1,414	28,800	87,400
Tableau	294,600	7,773	850	33,000,000

Based on our research these software packages are some of the leaders or those who are up-in coming in terms of relevance in academia and the business world. The researcher plans to explore the differences in skill levels for business students to create data dashboards for executives, much like that of an analyst. This is an exploratory study of the various software packages available to explore the visualization of data when doing business analytics.

Methodology

The researcher explored four different software packages to visual data given a data set of a bike supply chain store. It was explored to examine the bike sales by territory (United States, Canada, Germany, United Kingdom, France) and the percentage of bike sales by United States

market (Northeast, Northwest, Central, Southeast, Southwest). Furthermore, the other two charts displayed the lowest sales amount in a United States market by color of bike and another shows the trend of bike sales by quarter. The dashboards were created using Microsoft Excel, Power BI, Tableau, and Watson Analytics.

The goal of the study is to explore the functionality of the various software packages for data visualization purposes, as well as compare the skill level for various levels of business school students in terms of influencing instruction and curricula.

Results

Each dashboard was created and compared to how each software package was viewed in terms of the researchers experience with creation of the dashboards. The results of the creation and final products can be found below.

Microsoft Excel

When completing the first dashboard, the researcher created it using Microsoft Excel (Figure 1). While creating this dashboard, creation was relatively straightforward. The obstacles included labeling the sale amount in terms of dollars, guaranteeing the units are correct and easily accessible for an executive to make quick decisions, as well as grouping the bike sales by quarter. To separate the data, it required utilizing the use of PowerPivot extension to manipulate the data, add a slicer and group it into quarters based on sales date. The slicer provides the executive to be able to examine sales based on quarter.
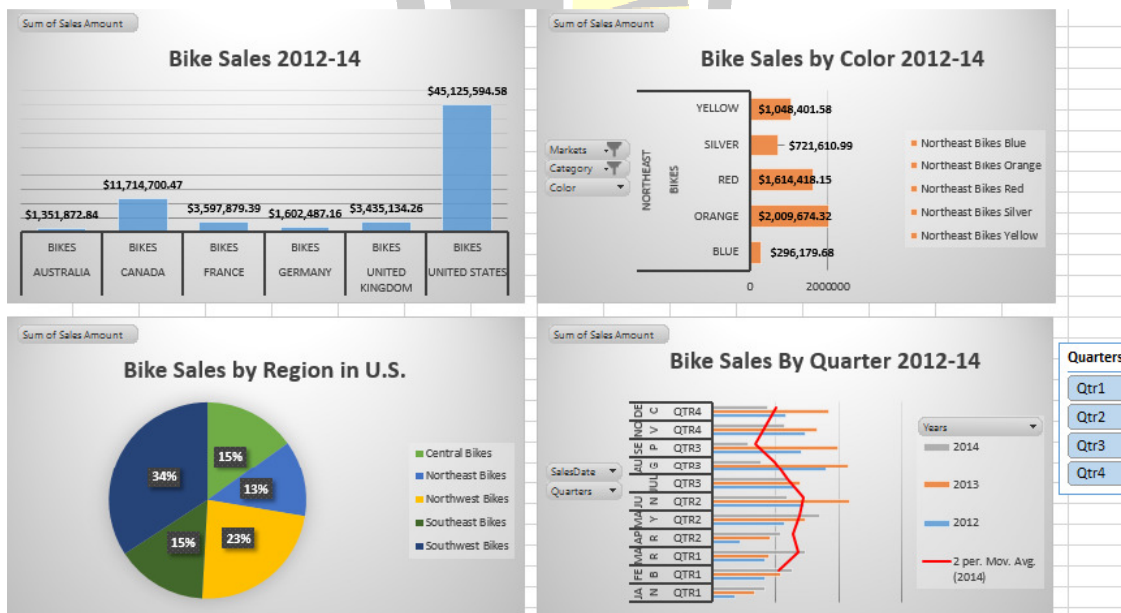

Figure 1. Excel Dashboard

Microsoft Power BI

The next dashboard was created by utilizing another Microsoft product, Power BI (Figure 2). Power BI was released on July 24, 2015 as a way for Microsoft to venture into higher level business intelligence tools (Announcing, 2015).

When creating this dashboard, two large obstacles came into play, one of which was the ease of including data labels for any of the graphs. The other issue was that Power BI currently does not have easy abilities to include trend lines for dashboards, therefore the moving average trend line was not easily created. The same graph did not allow the data to be broken up into quarters as it takes a higher level of technical knowledge to create. The other difficulty with the dashboard was the difficulty to include data labels on the dashboard.
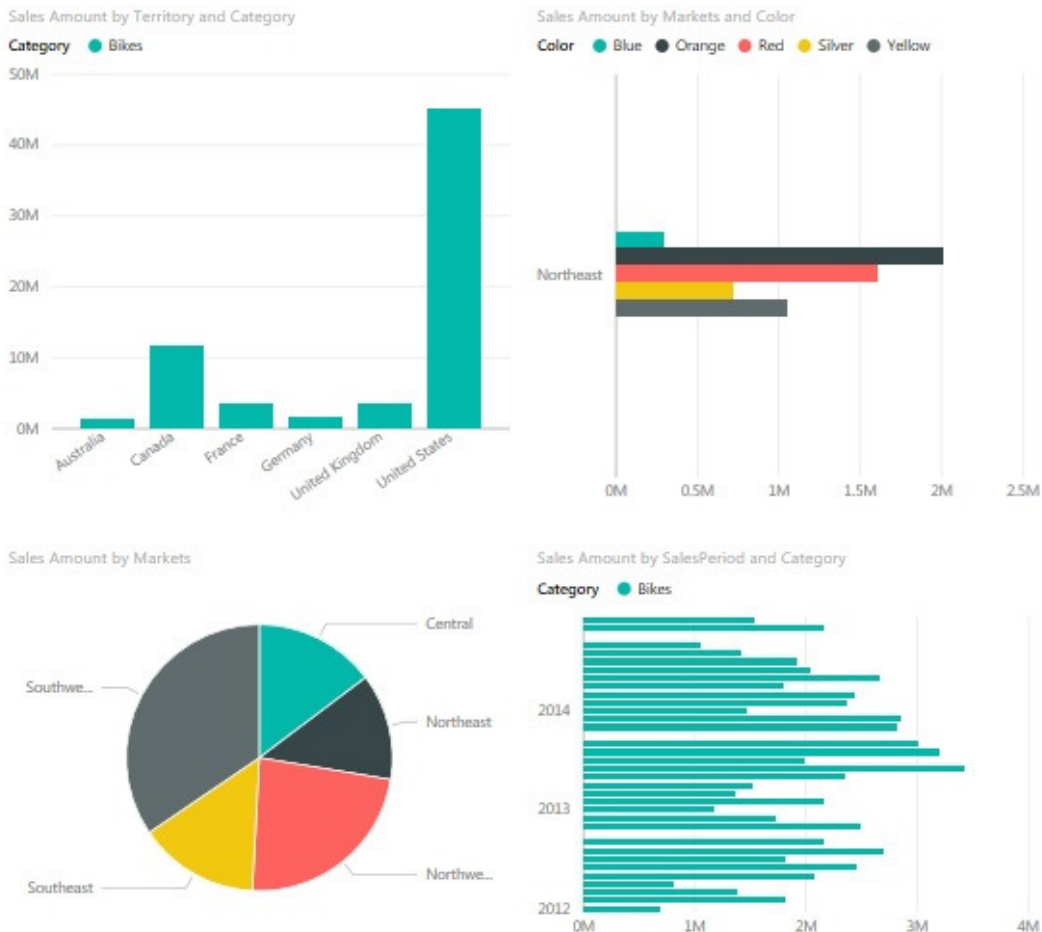
Figure 2. Power BI Dashboard

Tableau

When completing the next dashboard, the researchers created it using Tableau (Figure 3). When creating a dashboard using Tableau, the researcher found it to be quite similar to the functionality of Excel, but was easier to create graphs. The drag and drop functions allowed for a quick creation of the dashboard. The trend lines were not able to be a connected moving average, but did show moving average by month. Any different lines would require more advanced

programming skills. All labels were able to be created and a slicer type tool was utilized for executives to view the trend chart by quarter if needed.
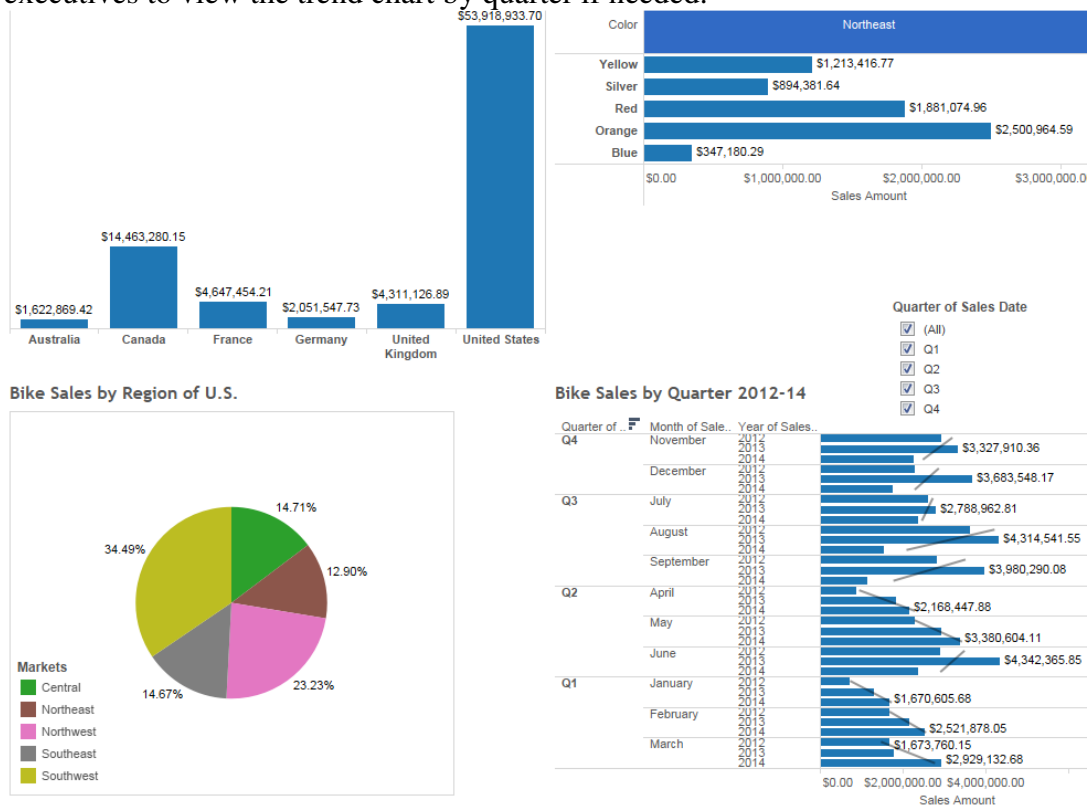

Figure 3. Tableau Dashboard

IBM Watson Analytics

Creating this dashboard with Watson Analytics (Figure 4) began with a wizard, which asked you to type the research question from the data which you upload. The difficulty came when the researcher typed in the question, the labels from the spreadsheet needed to be utilized to create the correct chart. When not using the wizard, the user interface is relatively straightforward. The third chart about bike color was unable to be placed sideways like the other dashboards. Also the fourth chart was not able to be split up into quarters, as the chart cannot hold the data. As you can see, the data is unable to fit on the chart without scrolling to view each month's data. The other observation was that creation of the charts and other functions were relatively slow, which was another issue with the software.
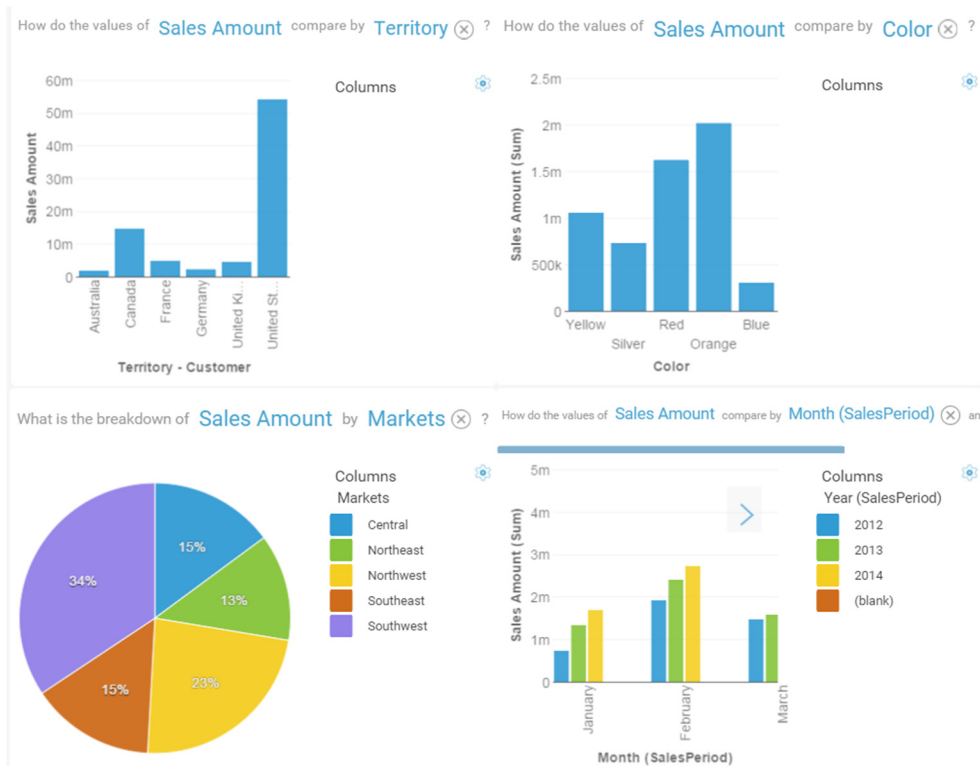
Figure 4. Watson Analytics Dashboard

Conclusion

Exploring various visualization tools provide an insight for business schools across the world to be able to determine which software packages are worth investing the time and capital in to best prepare their students for employment opportunities. Meeting the market demand and preparing students with skills to land meaningful employment is very important to a business schools' overall value. We explored some of these tools and compared what software is out there based on academic execution, as well as the place each organization has within the data visualization space currently.

Overall, we feel that it is imperative for business schools to explore and analyze their curriculum so they are best preparing their students in the area of analytics and data analysis to allow for the skills needed by many businesses today. The use of the dashboard has grown significantly over the past couple of years and it is not showing any signs of slowing down, therefore a focus on data visualization for strategic business decisions is imperative in today's business education curriculum.

References

"Announcing Power BI General Availability Coming July 24th." PowerBI. N.p., 10 July 2015.
<http://blogs.msdn.com/b/powerbi/archive/2015/07/10/announcing-power-bi-general-availability-coming-july-24th.aspx>
Eigner, W. "Current Work Practice and Users' Perspectives on Visualization and Interactivity in Business Intelligence." 2013 17th International Conference on Information Visualisation. 2013.

Lohr, S. "The Age of Big Data." *The New York Times*. 11 Feb. 2012. <
        http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-
        world.html?_r=0>

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A.H. "Big
        Data: The next Frontier for Innovation, Competition, and Productivity." Big Data: The
        next Frontier for Innovation, Competition, and Productivity. May
        2011.<http://www.mckinsey.com/insights/business_technology/big_data_the_next_fronti
        er_for_innovation>

Negash, S. "Business Intelligence." *Communications of the Association for Information Systems*
        13:177-195, 2004.