Data Analytics Using an Integrated Microsoft Platform

Frederick J. Riggins
College of Business
North Dakota State University
NDSU Dept 2410, PO Box 6050
Fargo, ND 58108-6050
701-231-5102
fred.riggins@ndsu.edu

Kelly T. Slaughter
The Naveen Jindal School of Management
The University of Texas at Dallas
800 West Campbell Road, JSOM 2.703
Richardson, TX
972-883-4755
kts130030@utdallas.edu

Abstract: This paper provides an introduction to performing data analytics and business intelligence using the full suite of Microsoft tools including extracting data from Dynamics 365 using Power Query and SQL Server Integration Services, data model development in Power Pivot and SQL Server Analysis Services using DAX, visualization and report generation and dissemination using advanced features in Excel 2016, Power BI and SQL Server Reporting Services, and introductory data mining analysis using SQL Server Analysis Services and SQL Server R Services. An accompanying set of tutorials that is in development provides students with a hands-on introduction to these tools.

INTRODUCTION

In recent years the number of tools available to perform data analysis and business intelligence has increased dramatically. For the professional data scientist, tools such as R, Python, Hadoop, Spark, SAS, SPSS, etc. form a powerful toolkit for data analysis and manipulation. However, an even bigger trend is the proliferation of self-service business intelligence tools that allow the typical business professional, without a background in statistical analysis, to engage in rather sophisticated data analysis and visualization. This list of tools includes the advanced data analysis features found in Microsoft Excel 2016, Microsoft Power BI, Tableau, SAP BusinessObjects tools, and IBM Watson Analytics just to name a few.

As these tools continue to be adopted by business managers and analysts seeking to make sense of the growing volume of available data, there is growing interest in university-level business education programs for a straight-forward, easy to integrate hands-on curriculum to train business students on the use of these tools. This paper provides an introduction to a hands-on curriculum that introduces students to the full suite of Microsoft tools in a series of tutorials that can be implemented in a single course to a wide range of business students. The curriculum seeks to show students how they can

1) extract data from enterprise systems such as Microsoft Dynamics 365 Operations (ERP), Dynamics 365 Financials (Accounting), and Dynamics 365 Sales (CRM), as well as other source files using the Microsoft Power Query editor or SQL Server Integration Services,
2) build rich tabular data models in Microsoft Power Pivot or SQL Server Analysis Services using the Data Analysis Expressions (DAX) programming language,
3) begin building effective visualizations in Microsoft Excel 2016, Power View, Power Maps, and SQL Server Reporting Services,
4) build and share interactive dashboards and reports in Microsoft Power BI, and
5) perform data mining using Microsoft SQL Server Analysis Services and leverage Microsoft's alignment of its solutions with the open source tool R to execute R-supported analytics.

An accompanying set of student tutorials for each of these objectives is currently in development with some tutorials having already been used in the classroom.

A DATA ANALYTICS FRAMEWORK

The data analytics process begins with identification of the goals and objectives of the analysis project as shown in Figure 1 [Kale and Jones (2015; Sharda, Delen and Turban (2014)]. Data can be gathered from a number of sources and input into a data staging area for data cleansing before being loaded into the data model using the extract-

transform-load (ETL) process. The data model, in our example a tabular data model, is then developed to build relationships between the data tables, define key performance indicators, and construct useful hierarchies. Data analysis can then occur by slicing and dicing data in pivot tables for descriptive analytics or data mining and forecasting models can be applied to the data for more in-depth predictive and prescriptive analytics. Visualization, spreadsheet, and report building tools are then used to create meaningful presentations to display the results of the analysis. Finally, the analysis can be disseminated to appropriate stakeholders throughout the organization. After getting feedback of the results, the process can continue in an iterative fashion. This process can support self-service business intelligence for the typical business user or more advanced data analysis depending upon the need of the project and the sophistication of the user. As described in Riggins and Klamm (2017) data governance is a critical issue with self-service business intelligence, therefore the process must include traceability and repeatability of the steps taken to gather, cleanse, import, and analyze the data.

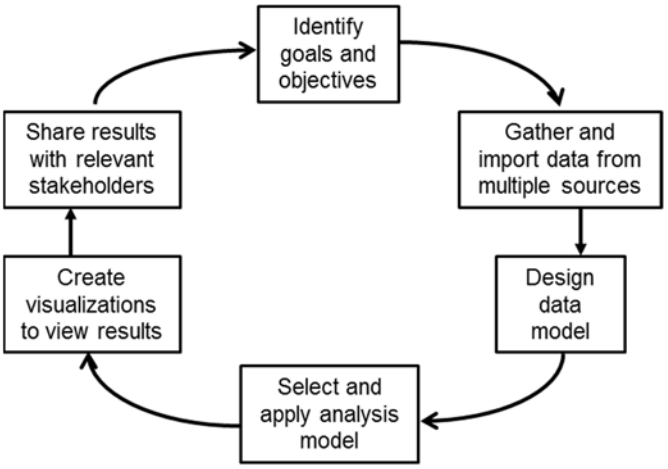

Figure 1: Six Phases in the Data Analytics Process

There are a number of self-service business intelligence tools that could be used to accomplish the goals of the data analytics framework. The broad acceptance and use of Microsoft tools across the business community makes the integrated Microsoft suite an attractive educational platform for achieving the objectives of the framework. The entire process can be illustrated in Figure 2.

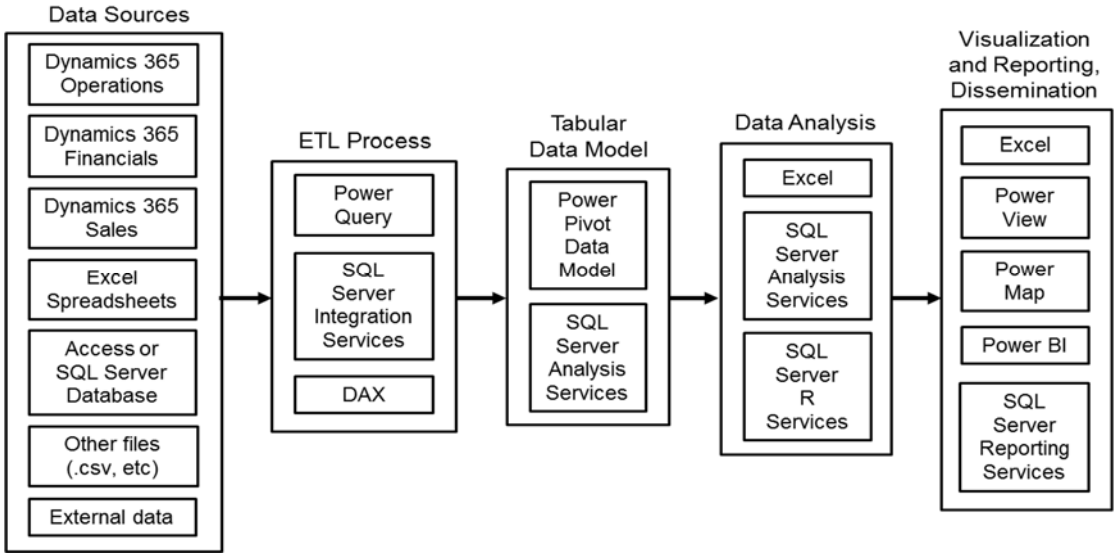

Figure 2: Data Analytics across the Microsoft Suite

Data can come from a number of sources including Dynamics 365, Excel spreadsheets, other database systems or files, and external sources. Depending upon the sophistication of the user and the anticipated level of analysis, the data extraction and cleansing process takes place using Power Query or SQL Server Integration Services with data being manipulated using visual tools or DAX [Larson (2017a); Russo and Ferrari (2015)]. The tabular data model can then be constructed in either Power Pivot or SQL Server Analysis Services [Larson (2017a); Russo and Ferrari (2013); Russo and Ferrari (2015)]. Once the data model is sufficient basic data analysis can take place in Excel 2016, with advanced data mining occurring using SQL Server Analysis Services or through R as integrated into SQL Server 2016 as supported through the Microsoft R Server [Larson (2017a)]. Finally, visualizations and reports can be created in Excel 2016, Power View, Power Maps, Power BI, or SQL Server Reporting Services, with the results shared using Power BI [Russo and Ferrari (2017); Larson (2017b)].

AN INTEGRATED INTRODUCTORY TRAINING CURRICULUM

The goal is to provide an integrated, introductory training curriculum that can be completed by the undergraduate or graduate business student within one course. The curriculum must be easy to manage by the instructor in terms of hands-on delivery to the students. Five assignments with accompanying student tutorials are either in use or in the process of being developed as follows:

1) **Data Extraction and Cleansing** – makes use of the Power Query editor to gather data from a number of sources as shown in Figure 2 to load data into a preliminary data model in Power Pivot. Learning objectives include understanding data sources, the ETL process, data model development, and hands-on use of Power Query and Power Pivot.

2) **Data Manipulation and Visualization** – manipulate the data model in Power Pivot using DAX and build visualizations in Excel 2016 using pivot tables, Power View and Power Map. Learning objectives include data hierarchies, key performance indictors (KPIs), slicing and dicing data, basic visualizations, and hands-on use of DAX, Power Pivot, Excel 2016, Power View, and Power Map.

3) **Advanced Visualization and Reporting** – create advanced visualizations and reports in Power BI and share results with others. Learning objectives include best-practices for dashboard development, choosing the right visualization, and hands-on use of DAX and Power BI.

4) **Advanced Descriptive Analytics** – perform more sophisticated data manipulations using SQL Server Integration Services and SQL Server Analysis Services tabular model, with advanced visualizations using SQL Server Reporting Services.

5) **Data Mining for Predictive and Prescriptive Analytics** – use SQL Server Analysis Services data model to illustrate data mining concepts and techniques, and introduce more advanced analytic models using R as integrated in SQL Server R Services and through the Microsoft stand-alone R Server.

REFERENCES

Kale, N. and N. Jones, *Practical Analytics – Applied Analytics Concepts using Market-leading Software Tools*, 2015, Epistemy Press.

Larson, B., *Delivery Business Intelligence with Microsoft SQL Server 2016, Fourth Edition*, 2017a, McGraw Hill Education, New York, NY.

Larson, B., *Microsoft SQL Server 2016 Reporting Services, Fifth Edition*, 2017b, McGraw Hill Education, New York, NY.

Riggins, F.J. and B.K. Klamm, "Data Governance Case at KrauseMcMahon LLP in an Era of Self-Service BI and Big Data," *Journal of Accounting Education*, vol. 38, March 2017, pp. 23-36.

Russo, M. and A. Ferrari, *Microsoft Excel 2013 – Building Data Models with PowerPivot*, 2013, O'Reilly Media, Sebastopol, CA.

Russo, M. and A. Ferrari, *The Definitive Guide to DAX – Business Intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI*, 2015, Microsoft Press, Redmond, WA.

Russo, M. and A. Ferrari, *Analyzing Data with Microsoft Power BI and Power Pivot for Excel*, 2017, Microsoft Press, Redmond, WA.

Sharda, R., Delen, D., and E. Turban, *Business Intelligence – A Managerial Perspective on Analytics, 3rd Edition*, 2014, Pearson, Boston, MA.