

# Data Mining with Excel 2007

Marc Holliday

Tom Higginbotham

2010

A large, 3D green arrow points diagonally upwards from the bottom left towards the top right. Below the arrow, the year '2010' is displayed in large, 3D green numbers. The background is a solid blue color.

# Agenda



- What is Data Mining?
- What is the Excel Data Mining Add-in?
- System Requirements
- Data Mining with Excel Demonstration
- What is PowerPivot for Excel?
- PowerPivot Demonstration
- Questions

# Performance Management



# Data Mining Defined



The process of extracting patterns from data to solve a problem.

Four Key Activities:

1. Clustering
2. Classification
3. Regression
4. Association Rule Learning

# A Real World Example

Problem Statement : Student retention rates are falling

Hypothesis : We can improve and predict student retention rates if we better understand student activities and their relationship to course completion



## Data Mined:

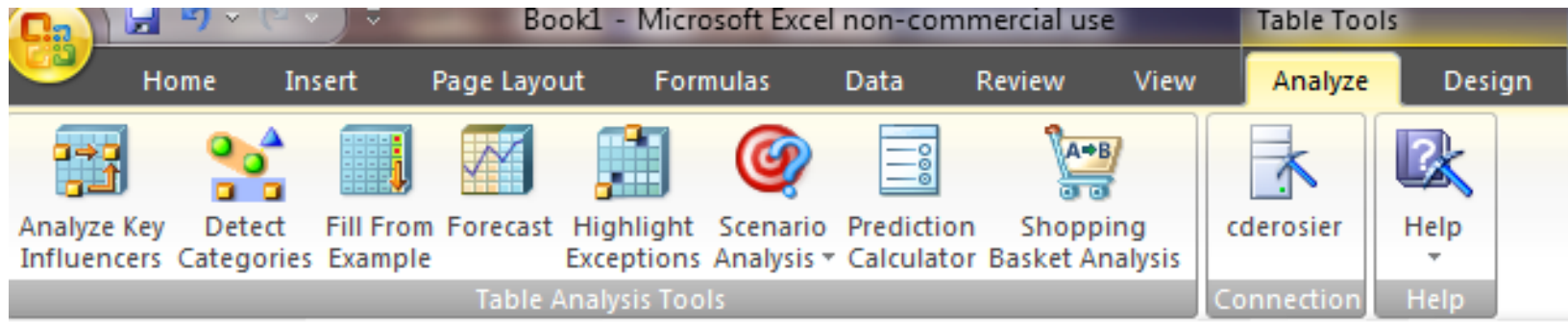
- First Log In
- Syllabus download
- Book purchase & date
- Student enrollment history
- Contact with Instructor
- Contact with other students
- Time to Graduation
- Full or part time student
- Assignment completion before add/drop date

# Office Data Mining Components

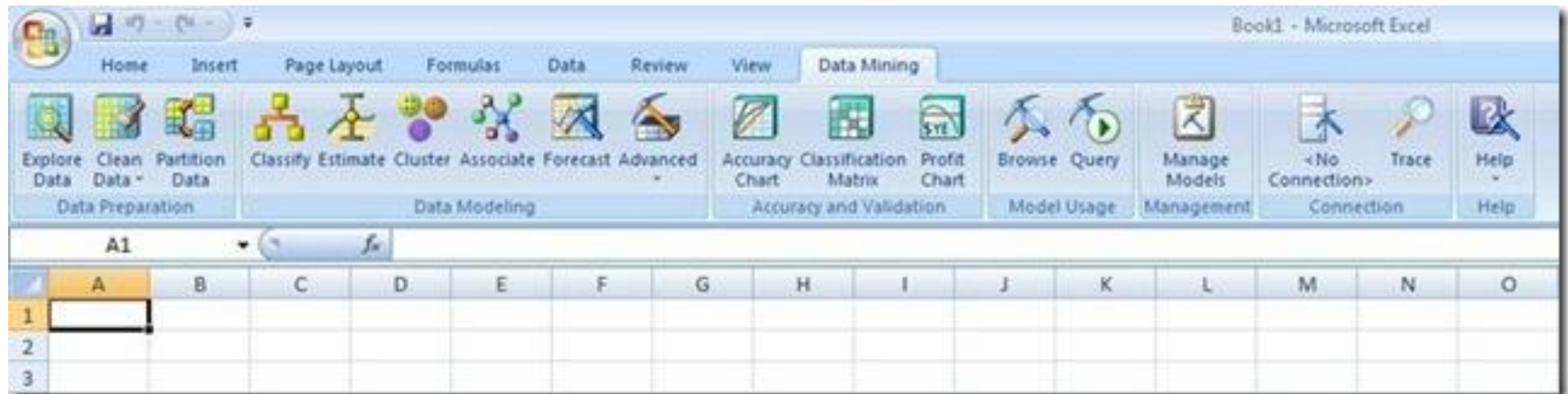


- Table Analysis Tools for Excel
- Data Mining Client for Excel
- Data Mining Templates for Visio

# Table Analysis Tools



# Data Mining in Excel





# **EXCEL DATA MINING DEMO**

# System Requirements



- 32-bit environment only
- Office 2007 Professional or higher
- A connection to SQL Server 2005 or 2008 Analysis Services
- Microsoft .NET Framework 2.0

# Download Location

## Microsoft SQL Server 2008 Data Mining Add-ins for Microsoft Office 2007

### Brief Description

Download SQL Server 2008 Data Mining Add-ins for Office 2007. This package includes two add-ins for Microsoft Office Excel 2007 (Table Analysis Tools and Data Mining Client) and one add-in for Microsoft Office Visio 2007 (Data Mining Templates).



### On This Page

- ↓ [Quick Details](#)
- ↓ [System Requirements](#)
- ↓ [Additional Information](#)
- ↓ [What Others Are Downloading](#)
- ↓ [Overview](#)
- ↓ [Instructions](#)
- ↓ [Related Resources](#)

[Download](#)

SQL Server 2008:

<http://www.microsoft.com/downloads/details.aspx?FamilyId=896A493A-2502-4795-94AE-E00632BA6DE7&displaylang=en>

SQL Server 2005:

<http://www.microsoft.com/downloads/details.aspx?FamilyID=7c76e8df-8674-4c3b-a99b-55b17f3c4c51&displayLang=en>

# Introducing PowerPivot

Microsoft®  
Excel 2010



2010



# Why Use PowerPivot?



- Create relationships between a variety of data sources
  - ERP and CRM System
  - SQL Server and/or any ODBC source
- Process massive amounts of data in seconds
  - Demos of 100+ million rows
- Take advantage of familiar Excel tools and features
  - Pivot Tables and Slicers

# PowerPivot Versions

- **Stand Alone Installer:**
  - [PowerPivot for Microsoft® Excel X86 Executable \(32 bit\)](#)
  - [PowerPivot for Microsoft® Excel X64 Executable \(64 bit\)](#)
- **PowerPivot for SharePoint 2010 Enterprise – Additional configuration required:**
  - How to: Install PowerPivot for SharePoint on an Existing SharePoint Server - <http://technet.microsoft.com/en-us/library/ee210616.aspx>
  - How to : Install PowerPivot from Command Prompt - <http://technet.microsoft.com/en-us/library/ee210645.aspx>
  - PowerShell Configuration Script - <http://technet.microsoft.com/en-us/library/ee210697.aspx>

# Self Service BI with PowerPivot - Goals

## Self-service BI platform

- The system should enable business users to pull data from a variety of sources without requiring those sources to have strong relationships defined between them.
- The system should enable business users to identify and define relationship among the source data.
- The system should enable business users to use Excel to design reports based on the consolidated data.
- The system should enable business users to share their analysis across the enterprise.

## Managed infrastructure

- The system should enable IT to monitor user activities.
- The system should enable IT to secure business analysis models (PowerPivot workbooks)
- The system should enable IT to support high availability (HA) and disaster recovery (provided with standard SharePoint techniques).



# **POWERPIVOT DEMO**

# PowerPivot Resources



- **Websites**
  - <http://www.powerpivot.com>
  - <http://powerpivotgeek.com>
  - <http://powerpivotpro.com>
  - <http://powerpivot-info.com>
- **Excel Add-in Download (32 or 64 bit)**
  - <http://www.microsoft.com/downloads/details.aspx?displaylang=en&FamilyID=e081c894-e4ab-42df-8c87-4b99c1f3c49b>
- **Sample Data Download**
  - <http://powerpivotsampledata.codeplex.com/>
  - <http://www.microsoft.com/downloads/details.aspx?displaylang=en&FamilyID=eac83429-c6e5-48a6-87cf-00a4141e5441>

# Questions?



Marc Holliday

[mholliday@socius1.com](mailto:mholliday@socius1.com)

614.280.9880 x2024