Data Mining: Using e-book data to visualize and analyze a collection

Yvonne Lee, MS, Laura Abate, MLS, Julie Silverman, MLS, MBA
Himmelfarb Health Sciences Library, George Washington University

Methods
Source data compiled in MS Access
- SerialsSolutions report with broad subject categories for department assignments
- E-Book usage data
- Faculty counts by department
- Data cleaned up & merged by ISBN/E-ISBN
- Exported merged data into Excel

Visualization Charts & Graphs
Data analyzed and displayed in Excel
- Excel PivotTables/Charts with Slicers filter
- Schools and Departments
- Interactive dashboard with hyperlinks increases flexibility and ease of navigation
- Designed to simulate a webpage experience
- Detailed title lists allow staff to move between charts/graphs and underlying title lists

Results
Analysis examples
- Size of collection and usage by Department & School
- Relationship of faculty size, collection size and usage
- Currency of collection by title count by publication years
- Usage comparison by consecutive years by Department
- E-book usage by Vendor by School

Conclusion
Effective data mining and visualization
- Provides broad and detailed collection data to support decision making
- Creates innovative ways of viewing and analyzing e-book collection
- Identifies areas which need additional data
- Promotes efficient fund management and better collection support for Schools and Departments
- Demonstrates Excel’s powerful tools and advanced features

Excel Hyperlink
PivotTable Slicer