Earth Science Data Analytics/Science Skills Needed: Overall Experiences /Operational Needs

**Data Analytics / Data Science**
- Need skills in: mathematics, numerical modeling, statistics, software engineering and the ability to integrate data across multiple domains.
- Need expertise in tools and techniques: rule learning, classification, cluster analysis, data fusion, machine learning, neural networks, anomaly detection, modeling, time series analysis, visualization.
- Need knowledge in particular science domains where data analytics can advance our understanding of science.
- The role is a hybrid one... skills to support domain scientists with data and computational needs to communicate across domains.

**Operational Needs**
- Need to facilitate making data more useful.
- Knowledge of the data lifecycle.
- Understand the data and how to understand the data.
- Know how to apply the techniques to the discipline.
- Additional Requirements include:
  - Understanding data, metadata, and data integration.
  - Knowledge of the importance of the data lifecycle.
  - Training in Data Visualization.

**General Experiences**
- Need to facilitate making data more useful.
- Knowledge of the data lifecycle.
- Understanding the data and how to understand the data.
- Know how to apply the techniques to the discipline.
- Additional Requirements include:
  - Understanding data, metadata, and data integration.
  - Knowledge of the importance of the data lifecycle.
  - Training in Data Visualization.

**Question:** What do university level Data Analytics/Science Programs focus on?

**Methodology:**
1. Surveyed the curriculums of 167 universities offering 267 Data Analytics/Science degree programs.
2. Categorized program based on focus area, by degree and on-line/on-campus.
3. Surveyed all curriculums to determine course topics providing specific training, relevant to Earth science.
4. Charted course topics against program focus area to see what is emphasized per program.
5. Compare changes to 2013 survey.
6. Compare to ‘real world’ needs.

<table>
<thead>
<tr>
<th>Program Focus Areas</th>
<th>B on line</th>
<th>B on campus</th>
<th>M on line</th>
<th>M on campus</th>
<th>PhD on line</th>
<th>PhD on campus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business analytics</td>
<td>10</td>
<td>17</td>
<td>22</td>
<td>22</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Data/Information science</td>
<td>15</td>
<td>19</td>
<td>21</td>
<td>21</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Data analytics</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Quantitative Analysis/Applied Statistics</td>
<td>4</td>
<td>5</td>
<td>17</td>
<td>17</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Information technology/systems management</td>
<td>4</td>
<td>4</td>
<td>19</td>
<td>19</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Computer science</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Health/Bio analytics</td>
<td>19</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Business management</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Environmental Informatics</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Program Pertaining to Data Science/Data Analytics: Course Topics Most Offered**
- **Data Mining, Mathematics, Statistics, Machine Learning, Data Visualization**
- **Data Fusion, Data Warehousing, Pattern Recognition, GIS, Remote Sensing, Text Mining, Information/Knowledge Management**
- **Predictive Modeling, Data Fusion, Data Warehousing, Pattern Recognition, GIS, Remote Sensing, Text Mining, Information/Knowledge Management**
- **Artificial Intelligence, Clustering, Time Series, Data Warehousing, Pattern Recognition, GIS, Remote Sensing, Text Mining, Information/Knowledge Management**

**Data Analytics/Data Science Techniques Practiced**
- **Anomaly Detection, Artificial Intelligence, Classification, Cluster analysis, Data Compression, Data Engineering, Data Fusion, Data Mining, Data Warehousing, Database Management, Machine Learning, Mathematics, Modeling, Neural networks, Pattern Recognition, Rule learning, Signal Processing, Statistics, Time Series Visualization**
- **Ability to integrate data across multiple domains**
- **Support domain scientists with data and computational needs to communicate across domains (be interdisciplinary)**
- **Knowledge of data life cycle**
- **Software engineering - Programming**

**Skills Practiced**
- **Business intelligence, data science, data analytics, and computer science**
- **Mathematical modeling, data lifecycle, and data management**
- **Anomaly detection, artificial intelligence, classification, cluster analysis, data compression, data engineering, data fusion, data mining, data warehousing, database management, machine learning, mathematics, modeling, neural networks, pattern recognition, rule learning, signal processing, statistics, time series visualization**

**Summary Table of Best Practices**
- **Repeated Exposure**
- **Sharing of Vocabulary**
- **How, Where, When and Who to Find/Ask for Resources & Help**
- **Sharing of Community:** "both ways" Communication

**What Else Universities Should Consider Offering**
- **Increased in general Data Analytics/Science programs**
- **Comprehensive systems (physical, biological, social, engineered) may be studied, modeled and understood using similar approach, regardless of the application domain.**

**What Universitiess Offer (Findings)**
- **Decreased for Business programs.**
- **Increased in general Data Analytics/Science programs.**
- **Decreased in Information technology/systems and Computer Science programs.**
- **MBA and Master of Information programs show up.**
- **Some programs are interdisciplinary with other departments.**
- **A few universities listed below, are very in tune with professional needs.**
- **Many programs are introductory, offering ‘generic’ courses, e.g., ‘Introduction to Data Science’.**
- **Some really good Bachelor degree programs have appeared.**
- **PhD programs are obviously more research than course work in nature.**

In 3 years, programs have increased in number, some more interdisciplinary and specialized, and much more refined in providing a broader range of relevant courses.