### General Experiences

- **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
- Should be interdisciplinary from the start.
- Learn your math and statistics.
- Know the importance of the data lifecycle.
- Understand what the data says and how to understand the data.
- Know the territory: What information is available. Where to get it. How it is generated. How to use it. How it can be used.
- Understand data, metadata, and data integration.
- Know how to apply the techniques to the discipline.
- Learn through internships.

### References

- Needs against program focus area

#### Charted course topics

- Earth science specific training, relevant to

#### 3. Campus Science degree programs


Specialized, much more refined in providing a broader range of relevant skills practiced:

- General Experiences
- Operational Needs

#### Data Science, Data Analytics, Information Systems

- Database Management/Analysis
- Quantitative Analysis:
- Data Mining, Mathematics, Statistics
- Other Relevant Courses Offered:
  - Programming, Neural Networks, Data Analysis, Artificial Intelligence, Clustering, Time Series, Data Warehousing, Pattern Recognition, GIS, Remote Sensing, Text Mining, Information/ Knowledge Management

#### Data Analytics/Data Science Techniques Practiced

- 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

#### What Else Universities Should Consider Offering

- Every Earth science program should contain training in Data analytics/science and Programming (data analysis)

---

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

<table>
<thead>
<tr>
<th>Data Analytics / Data Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Data scientist studies methods of analyzing data, ways of storing it, and ways of presenting it</td>
</tr>
<tr>
<td>• Data analytics is performed by the practitioners who applies tools and techniques to co-analyze data.</td>
</tr>
<tr>
<td>• Both, data science and data analytics require very similar skill sets.</td>
</tr>
<tr>
<td>• Once acquired, it becomes up to the individual to determine how best to use these skills, based on their interest and aptitude</td>
</tr>
</tbody>
</table>

---

**Earth Science Data Analytics**

- Lots of interdisciplinary, heterogeneous data!
  - Scientists spend 30% of their time programming, but 90% are self-taught!
  - Unlike laboratory and field equipment, software is often not carefully validated.
  - Computing errors can have disproportionate impacts on scientific process.

**Summary Table of Best Practices**

- Only programs that really understand how to apply data science are making the biggest impact.
- Too many programs are not computing part of the process.
- There is a need to move beyond traditional data science.

---

**What the Universities Offer (July, 2016 study and comparison with 2013 Study)**

<table>
<thead>
<tr>
<th>Program Focus Areas</th>
<th>B+ online</th>
<th>B online</th>
<th>M online</th>
<th>M cae online</th>
<th>PhD online</th>
<th>PhD campus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business analytics</td>
<td>10 22 41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data/analytics</td>
<td>1 4 12 34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantitative Analysis/Management</td>
<td>5 17 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer science</td>
<td>1 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health/Bio analytics</td>
<td>15 15 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business management</td>
<td>1 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental Informatics</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**What Else Universities Should Consider Offering**

- In 3 years, the percentage of degree programs that claim to pertain to Data Analytics/Data Science:
  - 1. Decreased for Business programs
  - 2. Increased in general Data Analytics/Science programs
  - 3. Decreased in information technology/systems and Computer Science programs
  - PhD...program called ‘Transformation of information’ shows up

**Program Pertaining to Data Science/Data Analytics: Course Topics Most Offered**

- **Overall:** Statistics, Data Mining, Database Management/Analysis
- **Data Science, Data Analytics, and Computer Science:**
  - Data Mining, Mathematics, Statistics, Machine Learning, Data Visualization
- **Data Science, Data Analytics, Information Systems:**
  - Database Management/Analysis
- **Quantitative Analysis:**
  - Data Mining, Mathematics, Statistics
- **Other Relevant Courses Offered:**
  - Programming, Neural Networks, Data Analysis, 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**

---

**Question:**

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

**Methodology:**

1. Surveyed the curriculums of 167 universities offering 267 'Data Analytics' or Data Science degree programs
2. Categorized program based on focus area, by degree and on-line/orcampus
3. Surveyed all curriculums to determine course topics providing specific training, relevant to Earth science
4. Chatted 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