We will evaluate the various quantitative data collection and analysis methods in standard industry practice. These methods are what will be used throughout this program, so you should become familiar with the terminology.

The second part presents a variety of career paths for data analysts and an overview of how several industries are currently using data analytics. Pay special attention to the intersection of skills necessary for a data analyst to possess, and think of the steps you can take to gain or improve on these in your own skill set. This may give you an idea of the career path and industry you would like to pursue, or enhance your understanding of a career path and industry you have already chosen.

**Typical Quantitative Techniques Used in Advanced Analytics**

Several quantitative techniques apply to analytics projects, including:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation</td>
<td>Randomized repetitions of a set of discrete events in order to model real-world systems and phenomena (e.g., queues)</td>
</tr>
<tr>
<td>Optimization</td>
<td>Algorithm selects the best possible outcome, subject to satisfying constraints</td>
</tr>
</tbody>
</table>
# Analytics Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix Algebra</td>
<td>Calculations involving matrices solve multidimensional problems</td>
</tr>
<tr>
<td>Fitting Functions to Data</td>
<td>Also called “curve fitting,” using numerical methods to interpolate data</td>
</tr>
<tr>
<td>Survival Analysis</td>
<td>Originally used by life scientists, but adopted by marketers and actuaries</td>
</tr>
<tr>
<td>Time Series</td>
<td>When data are “auto-correlated,” such as time-dependent data (also called “Box-Jenkins”)</td>
</tr>
</tbody>
</table>
| Predictive Analytics and Machine Learning | > Descriptive: calculates metrics to characterize the distribution of values of data (mean, standard deviation, range, etc.)  
  > Predictive: estimates parameters using historical data and making predictions of future outcomes (multivariate regression, generalized linear regression, etc.)|
| Classical Statistics        | > Unsupervised learning: characterizes the data to establish classes without using explicit metrics, e.g., k-means clustering  
  > Supervised learning: Classify and describe the data with pre-defined ‘labels,’ e.g., decision trees                                                   |
| Bayesian                    | Used to augment classical analysis when there is prior knowledge about how the data was generated                                          |

## Typical Challenges and Pitfalls in An Analytics Project

1. Poorly defined problem
   - Unclear goal of problem-solving
   - Scope is unclear, e.g., how many SKUs to analyze
   - Mixed objectives, e.g., economic analysis of a product category promotion for retailer versus CPG mixed

2. Limited IT resources
   - Cloud data can’t be acquired off-line within a reasonable time
   - Can’t run the complete model due to computation limitation
   - Too slow to generate results in real time
• Can’t share the data and results with network limitation

3. Less-best approach
• Selected less effective modeling method
• Incremental accuracy doesn’t offset the extra complexity
• Inadequate or incorrect performance monitoring criteria

4. Incomplete or incorrect data
• Primary dataset unavailable
• Complementary data unavailable, e.g., missing competitor pricing data
• Coarse data or aggregated data
• Very sparse data with missing values

5. Insufficient communication
• Insufficient data dashboard to communicate the analysis result
• Lack of soft skills to sell the results and insights
• Long feedback cycle to make the results less relevant
• Isolated org structure to stifle collaboration

Careers in Analytics
Many skill sets come to play during the course of a data analysis project workflow. These include hacking skills, math and statistics knowledge, and substantive expertise. Whereas traditional research relies primarily on math/statistics and domain expertise, modern data science typically draws from all three sets. The hacking skills reflect themselves in the understanding of available tools and technologies.

Full Suite of Data Scientist Skill Set
### Technical area
- Data Mining
- Predictive Modeling
- Machine Learning
- NLP
- Text Analytics

### Technology
- Data storage and processing
- Computing environment
- Computer programming
- Visualization
- BI / reporting

### Academic discipline
- Probability Theory
- Statistics
- Computer Science
- Operations Research
- Economics

### Domain knowledge
- Healthcare
- Retail marketing
- Financial services
- Manufacturing
- Telecommunication

---

**Rapid Growth Projected in Big Data Market**

Wikibon projects the Big Data market will top $84B in 2026, attaining a 17% Compound Annual Growth Rate (CAGR) for the forecast period 2011 to 2026.

![Big Data Market Forecast, 2011-2026 ($US B)](chart.png)

The Big Data market reached $27.36B in 2014, up from $19.6B in 2013. These and other insights are from Wikibon’s excellent research of Big Data market adoption and growth. The graphic above provides an overview of their Big Data Market Forecast.

Note–Amazon’s annual revenue is about $100B
**Significant Gap in Data Analytics Demand and Talent**

The demand for deep analytical talent in the United States could be 50 to 60 percent greater than its projected supply by 2018.

The demand and supply are projected to be noticeably mismatched in the 2018 forecast. This would be a striking change from 2008 when 150,000 slots were filled by ‘Data Analytics’ graduates with an exceeded supply of 30,000. In 2018, the projected demand (est. over 440,000) is expected to exceed the projected supply (est. 300,000) by 140,000.

---

**Summary | Careers in Analytics**

We presented a terminology overview of quantitative data collection and analysis methods that you will become more familiar with as you progress through this program. These methods are used in common practice throughout most industries and are enhancing the data collection and analysis trends of today.

We also discussed a variety of career paths for data scientists and an overview of how several industries are currently using data analytics. We learned about the growing demand for data analytics talent and the talent gap that exists in the field. To be a part of that talent demand, there is a vast skill set that data analysts need to possess, and throughout this course, we will examine each of them in more detail, while giving you the tools you need to gain or enhance these skills for your future career in data analytics.