EMC Education Services

Education Solutions for Data Science & Big Data Analytics

EMC Proven Professional Data Science & Big Data Analytics training

Joseph Kambourakis, Lead Data Science Instructor
Data Scientist – An Emerging Career

SPOTLIGHT ON BIG DATA

Data Scientist: The Sexiest Job of the 21st Century
Profile of a Data Scientist

- Curious & Creative
- Technical
- Quantitative
- Skeptical
- Communicative & Collaborative

© Copyright 2012 EMC Corporation. All rights reserved.
Roles by Technical and Quantitative Skills

- **Quantitative Skills**
  - Quantitative Analysts, Statisticians, Business and data analysts
  - Data Scientists
  - Recent STEM Grads
  - Business Intelligence Professionals, IT

- **Technical Ability**
# Data Science and Big Data Analytics Course and EMCDSA Certification

## Intended Audience
- **Business and data analysts** looking to add big data analytics skills
- **Database professionals** looking to exploit their analytic skills
- **Recent college graduates** looking to move into data science and big data
- **Managers** of business intelligence, analytics, or big data groups

## Details
- “Open” curriculum
- Practitioner’s approach
- Enables immediate participation on analytics projects
- Prepares for EMC Proven Professional Data Science Associate (EMCDSA) **Certification**
# Data Science and Big Data Analytics

## Hands-on practitioner’s approach to analyzing Big Data

<table>
<thead>
<tr>
<th>Introduction</th>
<th>Big Data Overview</th>
<th>State of the practice in analytics</th>
<th>The role of the Data Scientist</th>
<th>Big Data Analytics in industry verticals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Introduction to Big Data Analytics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytics Lifecycle</td>
<td>Key roles for a successful analytic project</td>
<td>Main phases of the lifecycle</td>
<td>Developing core deliverables for stakeholders</td>
<td>End-to-end data analytics lifecycle</td>
</tr>
<tr>
<td>Basic Methods</td>
<td>Introduction to R</td>
<td>Analyzing and exploring data with R</td>
<td>Statistics for model building and evaluation</td>
<td>Using R to execute basic analytic methods</td>
</tr>
<tr>
<td>Tools</td>
<td>Using MapReduce/Hadoop for analyzing unstructured data</td>
<td>Hadoop ecosystem of tools</td>
<td>In-database Analytics, MADlib and Advanced SQL Techniques</td>
<td>Advanced analytics and statistical modeling for Big Data – Technology and Tools</td>
</tr>
<tr>
<td>Lab</td>
<td>How to operationalize an analytics project</td>
<td>Creating the Final Deliverables</td>
<td>Data Visualization Techniques</td>
<td>Hands-on Application of Analytics Lifecycle to a Big Data Analytics Problem</td>
</tr>
</tbody>
</table>

Endgame, or Putting it all together
Video-ILTs : A Powerful Mode of Learning

- Multi-media capture of actual class delivered by one or more subject matter experts
- Near-classroom experience
- Includes walk-through labs, student discussions, slides, student handout, lab guides
- Questions/discussions via online community
- DVD based, easy navigation
- Cost efficient and scalable
- Selective learning – skip, repeat
- Modular, Self-paced, Portable
- Useful as a reference
Questions?

For more information:

http://education.emc.com/datascience

Joseph Kambourakis
joseph.kambourakis@emc.com

David Dietrich
@imdaviddietrich