Syllabus

Course Number: MSDS 600
Course Title: Introduction to Data Science

Course Description:
Introduces foundational topics of data science including data manipulation, data analysis using statistics and machine learning, techniques for working with Big Data, communication of analysis using information visualization, and ethical use of data analyses.

Prerequisite Courses:
None

Course Outcomes:
Upon completion of this course, learners should be able to:
1. Discuss the emerging field of Data Science.
2. Discuss applications of Data Science.
3. Describe the Data Science project lifecycle.
4. Distinguish the role of a Data scientist in engineering driven, data-product driven, or commerce driven domains.
5. Characterize Big Data using volume, variety, and velocity.
6. Discuss ethical concerns associated with Big Data and Data Science.
7. Use statistical analysis and machine learning techniques to characterize and model data.
8. Create a recommender system using the Hadoop software ecosystem.

Course Materials:

Required Texts:
None

Required Resources:
“From the Expert” presentations linked within each week and provided in Course Resources folder.

Pre-Assignment:
See Course Assignments and Activities table below.
## Course Assignments and Activities:

<table>
<thead>
<tr>
<th></th>
<th>Topics</th>
<th>Readings</th>
<th>Activities Assignments and Associated Points</th>
</tr>
</thead>
</table>
| 1 | Introduction to Data Science         | *From the Expert | Introductions – initial post required by Wednesday of Week 1  
Discussion Questions/threads (1.25%)  
Install Python (not graded) |
| 2 | Understanding Data                   | *From the Expert | Discussion Questions/threads (1.25%)  
**Project** – Sentiment Analysis using Python (15%) |
| 3 | Exploratory Data Analysis            | *From the Expert | Discussion Questions/threads (1.25%)  
Exercise – Install R and RStudio (9%) |
| 4 | Statistical Modeling Experimental Design | *From the Expert | Discussion Questions/threads (1.25%)  
Exercise – Data Analysis using R (9%) |
| 5 | Machine Learning                     | *From the Expert | Discussion Questions/threads (1.25%)  
**Project** – Machine Learning with R (15%) |
| 6 | Recommender Systems Introduction     | *From the Expert | Discussion Questions/threads (1.25%)  
Exercise – Remote access of the Hadoop lab system (9%) |
| 7 | Introduction to Hadoop (ecosystem)   | *From the Expert | Discussion Questions/threads (1.25%)  
Exercise – Using Hadoop (9%) |
| 8 | Implementing a Recommender System    | *From the Expert | Discussion Questions/threads (1.25%)  
**Project** – Implementing a Recommender System using the Hadoop ecosystem (15%) |
Summary of Assignments and Percentage Weight:

<table>
<thead>
<tr>
<th>Assignments</th>
<th>Weighted Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discussion Questions (8 at 1.25% each)</td>
<td>10%</td>
</tr>
<tr>
<td>Exercises (5 exercises at 9% each)</td>
<td>45%</td>
</tr>
<tr>
<td>Projects (3 projects at 15% each)</td>
<td>45%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

CC&IS Grading Scale

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Percentage</th>
<th>Grade Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93 to 100</td>
<td>4.00</td>
</tr>
<tr>
<td>A–</td>
<td>90 to less than 93</td>
<td>3.67</td>
</tr>
<tr>
<td>B+</td>
<td>87 to less than 90</td>
<td>3.33</td>
</tr>
<tr>
<td>B</td>
<td>83 to less than 87</td>
<td>3.00</td>
</tr>
<tr>
<td>B–</td>
<td>80 to less than 83</td>
<td>2.67</td>
</tr>
<tr>
<td>C+</td>
<td>77 to less than 80</td>
<td>2.33</td>
</tr>
<tr>
<td>C</td>
<td>73 to less than 77</td>
<td>2.00</td>
</tr>
<tr>
<td>C–</td>
<td>70 to less than 73</td>
<td>1.67</td>
</tr>
<tr>
<td>D+</td>
<td>67 to less than 70</td>
<td>1.33</td>
</tr>
<tr>
<td>D</td>
<td>63 to less than 67</td>
<td>1.00</td>
</tr>
<tr>
<td>D–</td>
<td>60 to less than 63</td>
<td>0.67</td>
</tr>
<tr>
<td>F</td>
<td>Less than 60</td>
<td>0</td>
</tr>
</tbody>
</table>

Additional information about grading can be found in the latest edition of the University Catalog, available at [http://www.regis.edu/Academics/Course%20Catalog.aspx](http://www.regis.edu/Academics/Course%20Catalog.aspx).

CC&IS Policies and Procedures

Each of the following CC&IS Policies & Procedures is incorporated here by reference. Students are expected to review this information each term, and agree to the policies and procedures as identified here and specified in the latest edition of the University Catalog, available at [http://www.regis.edu/Academics/Course%20Catalog.aspx](http://www.regis.edu/Academics/Course%20Catalog.aspx) or at the link provided.

- The CC&IS Academic Integrity Policy.
- The Student Honor Code and Student Standards of Conduct.
- Incomplete Grade Policy, Pass / No Pass Grades, Grade Reports.
• The Information Privacy policy and FERPA. For more information regarding FERPA, visit the U.S. Department of Education.

• The HIPPA policies for protected health information. The complete Regis University HIPAA Privacy & Security policy can be found here: http://www.regis.edu/About-Regis-University/University-Offices-and-Services/Auxiliary-Business/HIPAA.aspx.


The CC&IS Policies & Procedures Syllabus Addendum summarizes additional important policies including, Diversity, Equal Access, Disability Services, and Attendance & Participation that apply to every course offered by the College of Computer & Information Sciences at Regis University. A copy of the CC&IS Policies & Procedures Syllabus Addendum can be found here: https://in2.regis.edu/sites/ccis/policies/Repository/CCIS%20Syllabus%20Addendum.docx.