



**FLORIDA INTERNATIONAL UNIVERSITY  
UNIVERSITY CURRICULUM COMMITTEE**  
*Proposal for a New Course*

**DO NOT TYPE IN THIS BOX**

Bulletin # : 6

Academic Year : 2015/16

1. School/College Engineering and Computing  
Div./Dept. in Which Taught School of Computing and Information Sciences

2. IDC 5 XXX 3 CIP Code (Leave this blank): \_\_\_\_\_  
Alpha Prefix 1st Digit Last 3 Digits "C"-lec-lab "L"-Lab Cr. Hrs.

**CAP 5768 Keyed PS & CAT**

3. Grading Method (select one):  Graded  Pass/Fail

4a. Course Title Introduction to Data Science

b. Abbreviated course Title (for computer class schedules, transcripts) Intro to Data Science  
LIMITED TO 25 Characters (including spaces)

5. Statewide Course Numbering Subject Matter Area Interdisciplinary Computing

6. Catalog Description/Major Topics (not to exceed 200 characters including spaces)  
*College of Medicine and College of Law: Attach description not exceeding 1,000 characters including spaces.*

Foundations of databases, analytics, visualization and management of data. Practical data analysis with applications. Introduction to Python, SQL, R, and other specialized data analysis toolkits.

7. Attach detailed syllabus course outline and course justification on separate page(s).

8. Prerequisite(s): STA 3164 or equivalent

9. Corequisite(s): None

10. Objective(s) of Course:  
Introduction and exposure to fundamental tools in Data Science -- to read, clean, manage, manipulate and visualize data, and to apply statistical and computation tools to extract knowledge.

11. Does this course duplicate/overlap other courses at FIU?  No  Yes  
If yes, please explain: \_\_\_\_\_

12. What other closely related department(s) have been consulted about this course?  
Statistics, MIS, Biostatistics

**PROPOSAL REQUESTED BY:**

Faculty Contact <u>Giri Narasimhan</u>	 (Nagarajan Prabakar)	<u>2 / 23 / 2016</u>
(Type name)	(Signature)	
<u>giri@cs.fiu.edu</u>	(305) 348-3748	
(Email address)	(Phone number)	
Chairperson (Dept./Div.) <u>Mark Weiss</u>	 (Signature)	<u>2 / 23 / 2016</u>
(Type name)	(Signature)	
Chairperson (Curr. Comm.) <u>Bilal El-Zahab</u>	 (Signature)	<u>3 / 01 / 2016</u>
(Type name)	(Signature)	
College/School Dean <u>Ranu Jung</u>	 (Signature)	<u> / / 2016</u>
(Type name)	(Signature)	

Submit one original form. Attach one copy of the course syllabus containing: course description, objectives, learning outcomes, major topics and textbooks.

# Introduction to Data Science

---

## **Course Justification**

Davenport and Patil (Harvard Business Review, Oct 2012) called Data Scientist the “sexiest job of the 21st Century”. Big industry is being transformed by the power of analytics informing them how to adapt to changes. Many insights and discoveries in science are attributable to analytics. Data from [www.indeed.com](http://www.indeed.com) suggests that job postings for data scientist positions has increased tenfold since 2012. A growing number of universities are offering such programs in the US. Within the SUS, a related program is underway at UCF, and the New College, UWF, and Miami-Dade College appear to be headed in that direction.

Recognizing that the nation is facing economic competition on an unprecedented scale, the Florida Board of Governors has recommended that the State University System prepare graduates to “excel in the global society and marketplace”. Today’s marketplace is more data driven and more adaptive than ever before. Retailers arrange their shelves as well as their product websites based on customer purchase patterns. Personalized medicine is becoming possible because of increased availability of data on individuals. The entire finance industry is built around analysis of trends.

The new course suggested here is a core course in the proposed Masters degree program in Data Science (MS-DS) and will prepare students for the other courses in that Masters program. This new course will introduce students to the theory and practice of Data Science and expose them to a variety of application areas where it is needed. It will introduce them to the programming languages of Python and R and will expose them to tools used by Data Scientists.

# New Course Syllabus

---

**Course:** IDC- 5XXX

**Course Title:** Introduction to Data Science

**Credits:** 3

## Catalog Description

Foundations of databases, analytics, visualization and management of data. Practical data analysis with applications. Introduction to Python, SQL, R, and other specialized data analysis toolkits.

## Prerequisites

Undergraduate course in statistics such as STA 3164 or equivalent

## Type

This is a core course for MS – Data Science

## Objectives

The goal of this course is to provide students with an overview of the entire data analysis process, while providing them with basic tools (programming languages and toolkits) to navigate through the program. Students will also be exposed to many different applications of the data science approach.

## Topics

1. Introduction
  - What is Data Science? Examples, History, Application areas
2. Databases
  - Databases and relational approach
  - Introduction to SQL query language and basic analytics
  - Database algorithms, Hadoop, MapReduce
3. Data Analytics
  - Statistical foundations, exploratory data analysis
  - Introduction to R programming language
  - Basics of Data Mining and Machine Learning
4. Visualization
  - Visual data analytics and tools
5. Data management
  - Provenance, security, privacy, data cleaning, and data curation
6. Miscellaneous Topics
  - Big Data and its impact on analytics
  - Case studies from different application areas
  - Programming in Python, Tools for data analysis

## Textbook(s)

- *Mining of Massive Datasets*, by Leskovec, Rajaraman, and Ullman