**Course 5: Data Analytics and Data Visualization (2 credits)**

### By the time you finish this course you should: Use scraping techniques to download data, process data for analysis, analyze data and visualize data. You should also scrape data from IMDB and create a report of ratings of movies by genre as a media firm analyst. As trade analyst you should scrape data from BSE and use historical data to find the best performing stocks, and the best predictors

### Resource: <https://drive.google.com/file/d/0B0XhJYAEOB8mRzhXRGxXM0xDN3RMV3Y2OVhwbjhJcl9jcGk0/view?usp=sharing>

### Module 1: Data Repositories and resources

Learning Objectives:

* Identify different data repositories, data resources
* Identify different formats of data
* Discover different methods to acquire data from data resources

### Module 2: Data Scraping Tools

Learning Objectives:

* Practice Data Scraping Tools Beautiful Soup, Scrapy or Selenium to download static and dynamic web data

### Module 3: Data Parsing and Transformation

Learning Objectives:

* Download IMDB movie rating data
* Parse and transform data to find top rated movies
* Download BSE historical stock prices data
* Parse and transform it to find top gainers for the last one year

### Module 4: Data Analysis: Summarization by Segment

Learning Objectives:

* Summarize the data by segment
* Discover groups performing far better and far below

### Module 5: Data Analysis: Prediction using Linear Regression

Learning Objectives:

* Analyze relationships between items in the dataset
* Predict top rated items for future

### Module 6: Data Visualization: Vector Graphics SVG

Learning Objectives:

* Explain image files suffer from what problems
* Explain how SVG address problems image files face
* Create SVG graphs

### Module 7: Data Visualization: Design

Learning Objectives:

* Use code to generate visuals from data
* Use a template that will convert data into a HTML file -- typically embedding SVG within that. (Tornado’s template)
* Convert image files into PDF without using a browser

### Module 8: Data Visualization: Templates

Learning Objectives:

* Illustrate how aesthetic sense is key to creating visualisations
* Discover how Color, Typography and Layout can be used to great effect in increasing the impact the design

### Module 9: Interactive Visualizations

Learning Objectives:

* Use JavaScript, to create visualizations dynamic (changing over time) as well as interactive
* Structure templates that produce interactive visualizations, and design them to bring out the full meaning of the data -- without overwhelming the user.

### Module 10: Data Visualization Tools: Sci2, Tablo

Learning Objectives:

### Create model visualizations by using tools Sci2 and Tablo