



Trường ĐH Khoa Học Tự Nhiên Tp. Hồ Chí Minh  
**TRUNG TÂM TIN HỌC**

# **R PROGRAMMING LANGUAGE FOR DATA SCIENCE**

**Giảng viên: Khuất Thùy Phương**

**2020**



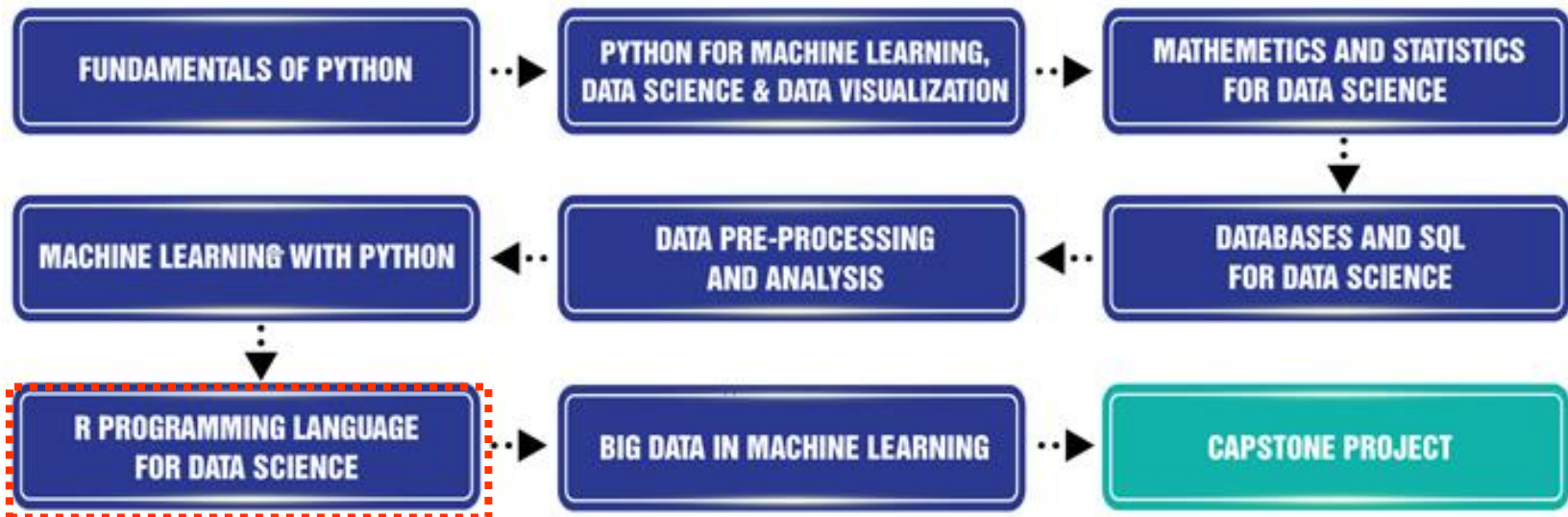
# R PROGRAMMING LANGUAGE FOR DATA SCIENCE

- **Hình thức học:** học OFFLINE 2 buổi/tuần (cho đến khi có thông báo mới)
  - **Thời gian:** 22/08 – 20/09 (5 tuần)
  - **Lịch học :** **Sáng Thứ Bảy & Chủ Nhật** (7h30 – 12h00)
- **Đánh giá môn học:**
  - **50%:** hoàn thành bài tập hàng tuần – tạo và share thư mục LDS7\_HoTen (Google Drive) – Nộp kết quả vào thứ Năm hàng tuần.
  - **50%:** Thi tại lớp hoặc làm bài thi ở nhà và nộp (Ngày thi 20/09/2020)
- **Phụ trách:**
  - **Ms. Phương**
  - Email: tubirona@gmail.com

# R PROGRAMMING LANGUAGE FOR DATA SCIENCE



## DATA SCIENCE





## □ Part 1: Basic

- Tổng quan lập trình R
- R cơ bản
- Cấu trúc điều kiện
- Cấu trúc lặp
- Function



## □ Part 2: Data Structure

- Vector
- Factor
- Array
- Matrix
- List
- Data Frame
- Tidyverse



## ❑ Part 3: Làm việc với dữ liệu tập tin

- Txt
- CSV
- Excel
- XML
- JSON



## ❑ Part 4: Graph và Chart

- Pie Chart
- Bar Chart
- Histogram
- Line Graph
- Scatterplot
- ggplot2



## ❑ Part 5: Thống kê

- Descriptive Statistic
- Normal Distribution
- Binomial Distribution
- Time Series Analysis





## ❑ Part 6: Machine Learning

- Giới thiệu, phân loại
- Linear Regression
- Logistic Regression
- Decision tree
- K-Means

