

Fall 2021 Syllabus 9.1.21 version

Duke Masters of Engineering Management

Fundamentals of Data Science

Fall 2021

Syllabus

(7.20.21 version)

Wednesdays. 7:00-9:45 pm

Teer 115

EGRMGMT 585-01, 02, 03

(6770, 6771, 8846)

Course materials for all sections will
be found on the EGRMGMT 585-01 Sakai site

Professor: Daniel Egger

Daniel.egger@duke.edu

Professor Office Hours

Zoom, Fridays 9-10 am

<https://duke.zoom.us/j/93179766560>

In-person CIEMAS 3575 Fridays 10-11 am

or by appointment

Teaching Assistant: Jiajun Walker Wang

jjajun.wang@duke.edu

TA Office Hours

Tuesday 5-6 pm

<https://duke.zoom.us/j/99222510314>

This course has no prerequisites

No programming knowledge is assumed, and all assignments can be done using MS Excel

Recommended Reading

Machine Learning: An Applied Mathematics Introduction

by Paul Wilmott (2019)

Fundamentals of Data Science is designed to prepare students to take an introductory Machine Learning (with Python) Course in their Spring Semester if they wish. The Wilmott text is an excellent general introduction as it explains many of the core math topics we will cover in the Fall Semester, presented from the point of view of their contributions to Machine Learning.

Class 1 – Aug 25

Overview of Course Topics

CRISP-DM and Data Science Project Best Practices

Marketing Case Study

Homework “0” Given

[Note: if students were switching classes or otherwise not aware the Homework 0 was due Sept 1, they will not be penalized].

Class 2 – Sept 1

Homework 0 Due

Supervised Learning Methods - Part 1

Classification Methods, the Confusion Matrix, Classification Cost Functions

Case Studies

Recommended Reading before class: Wilmott Chapters 1-2

Homework 1 given

Class 3 - Sept 8

Homework 1 due

Supervised Learning Methods - Part 2

Introduction to Regression & Logistic Regression

Algorithms and Cost Functions

Case Studies

Recommended reading before class: Wilmott Chapters 3 & 6

Homework 2 given

Class 4 - Sept 15

Homework 2 due

Time Series Analysis and Financial Data Overview

Homework 3 given

Class 5- Sept 22

Homework 3 due

Unsupervised Learning Methods

K Means Clustering and Scree Plots

Review of Basic Probability

Recommended Reading before class: Wilmott Chapter 4

Homework 4 Given

Read "Handout: Review of Basic Probability" and complete practice problems before next class

Class 6 – Sept 29

Homework 4 Due

Forecasting Probability Distributions

Bayes' Theorem and Maximum-Likelihood Methods

Kernel Density Estimation and Histograms

Recommended Reading before class: Wilmott Chapter 5

Homework 5 given

Class 7 – Oct 6

Homework 5 due

Network Analysis, Contagion, and “Going Viral”

Review of Course Material to Date

Class 8 - Oct 13

Midterm Exam

Class 9 – Oct 20

Experimental Design and A/B Testing

p-values and other basic statistics concepts

Introduction to Python, part 1

Homework 6 give

Class 10 – Oct 27

Homework 6 due

Scrubbing Messy & Missing Data

Introduction to Python, part 2

Homework 7 given

Class 11 – Nov 3

Homework 7 due

Looking Forward to Machine Learning

Overview of Spring 2022 Machine Learning course & topics

Recommended Reading before class: Wilmott Chapters 7, 9-11

Mini-projects Assigned

Class 12 – Nov 10

Project Outlines Due

Class time reserved for mini-project feedback

Class 13 – Nov 17

Mini-Project Presentations

no class Nov 24

no final exam

Grading:

Class Participation	10%
Homework Assignments (7)	30%
Midterm Exam	30%
Final Projects	30%

Class participation is: being prepared, and volunteering to speak in class. Being prepared means you read any assigned reading carefully enough to be able to express your own opinion on any unresolved issues or matters of judgement in the text, and can give reasons for your opinion if asked.

*Please check Sakai for assignment and schedule changes as **this syllabus is subject to change.***