MONTGOMERY COLLEGE - Rockville Campus

Business, Science, Math, and Technology Division

Course Syllabus

I. Instructor Information

Abdirisak Mohamed
Email: Abdirisak.Mohamed@MontgomeryCollege.edu
Phone: (240) 672-7529
Office: SC/454C
Mail box: SC/250C
Office Hours:

Tuesday	Thursday

These hours are in the office SC/454C or in the classroom.

II. General Course Information

DATA 201 – STAT METHODS IN DATA SCIENCE (3 Credits)

CRN

Mondays:

Room: SC 360

Course Description

Statistical concepts and applications related to data science including advanced exploratory data analysis, nonparametric inference and simulation for larger datasets, logistic regression modeling, statistical programming, and basics of machine learning. PREREQUISITE(S): A grade of C or better in DATA 101 or consent of department. Three hours each week.

Format: This course is taught in a lecture/discussion format. Attendance and participation are expected.

Homework: Complete the following DataCamp courses:

1. https://www.datacamp.com/courses/intro-to-python-for-data-science

2. https://www.datacamp.com/courses/python-for-r-users

Projects:

There will be four projects to be submitted as Jupyter notebooks. You can work in groups of 2 or 3 but it is not required. However, you are required to be compliant with the academic code of conduct.

1. The first project will be on exploratory data analysis using Matplotlib.

Work on a dataset of your choice.

(1a) Read the data into a dataframe.

(1b) Clean the data using Pandas to handle missing data, subset the dataframe, select features.
(1c) Show graphs (e.g., scatterplot, histogram etc.) using Matplotlib.

2. The second project will be on Linear Regression:

You can work on the dataset from project 1 or choose another dataset. The outcome/target variable must be continuous quantitative variable. The main goal of this project is to use the Scikitlearn Linear Regression, split the data into training and testing data, check the score (variance explained by the model) and plot the residual.

3. The third project is on Logistic Regression

The target/label variable is discrete/qualitative and could be binary or multi-class. Choose a suitable dataset. The main goal of this project is to use the Scikitlearn Logisitic Regression, split the data into training and testing data, check the accuracy (confusion matrix / AuC).

4. The fourth project will be on a machine learning algorithm of your choice:

Decision Trees, K-Means, or Neural Networks

You can work on datasets of your choice.

Here are some resources:

https://archive.ics.uci.edu/ml/index.php

https://www.kaggle.com/datasets

https://data.worldbank.org/data-catalog/

https://www.data.gov/

Books:

These books are all free

1. https://github.com/jakevdp/PythonDataScienceHandbook/
2. http://www.greenteapress.com/thinkstats/
3. http://www.holehouse.org/mlclass/
4. http://www-bcf.usc.edu/~gareth/ISL/

III. Course Outcomes

Upon course completion, a student will be able to:

- Select appropriate existing analytical and presentational tools for specific analyses of large databases.
- Develop new and appropriate analytical and presentational tools for specific analyses of large databases through programming.
- Demonstrate a competency with data science practices that allows for reproducible results.
- Summarize findings based on complex analyses in a concise way for a general audience using multivariate graphics and statistical measures.

IV. Grading

A. Requirements

Homework assignments, two tests, and four projects
Attendance is required.

B. Course Grade

The course grade is distributed as follows:

Project 1	15%
Project 2	15%
Project 3	20%
Project 4	20%
Homework	10%
Test 1	10%
Test 2	10%

C. Standards

Letter grades will be determined as follows:

A	90 – 100%
B	80 – 89%
C	70 – 79%
D	60 – 69%
F	< 60%

In the final grade calculation, all partial points are rounded up to whole points.

D. Late Policy

See attendance policy under Section A. Requirements.
No late homework will be accepted.

E. Audit Policy

Students auditing the class are expected to attend regularly. Those who do not may be dropped. "Audits" may take exams, but are not required to do so. Audit students do not earn a grade or credit for taking the course and cannot change to become a credit student in the middle of the semester.

V. Classroom Policies

A. Attendance and Withdrawal from Class

Attendance is required for this class, see Section A. Requirements.
It is the responsibility of the student to officially withdraw from a class or classes. If a student neglects to follow the official withdrawal procedure and merely ceases to attend classes, the grade of "F" will be recorded at the end of the semester.

B. Academic Honesty

Students are encouraged to work together on homework and project assignments.
No form of academic dishonesty will be tolerated, examples of which can be found in the Student Code of Conduct.

C. Classroom Conduct

Cell phones must be turned off or silenced during class time. Anyone receiving or making a call or a text message will be asked to leave the classroom.

Students are expected to adhere to the Student Code of Conduct. If a student persists in disruptive behavior, he or she may be asked to leave the room.

In addition to course requirements and objectives that are in this syllabus, Montgomery College has information on its web site (see link below) to assist you in having a successful experience both inside and outside of the classroom. It is important that you read and understand this information. The link below provides information and other resources to areas that pertain to Student Success such as: Student Behavior (Student Code of Conduct); Student e-mail, College Tobacco Free Policy; Course Withdrawal and Refund Information; Resources for Military Service Members, Veterans and Dependents; how to access information on delayed openings and closings; how to register for Montgomery College's Alert System and how closings and delays can impact your classes.

http://cms.montgomerycollege.edu/mcsyllabus/

Finally, any student who may need accommodations due to a disability, please contact the Disability Support Services office on your campus as soon as possible (R-CB122; G-SA172; or TP/SS-ST233). If you have an accommodation letter from DSS, please contact me to discuss arrangements for your accommodations. Any student who may need assistance in the event of an emergency evacuation must identify to the Disability Support Services Office; guidelines for emergency evacuations are at:
https://cms.montgomerycollege.edu/EDU/Plain2.aspx?id=4162

If you have any questions please bring them to your professor. As rules and regulations change they will be updated and you will be able to access them through the link. If any student would like a written copy of these policies and procedures, the professor would be happy to provide them. By registering for this class and staying in this class, you are indicating that you acknowledge and accept these policies.

.

D. Cancellation of Classes

If class is cancelled, read the corresponding sections as indicated in the attached schedule. Any quiz or test scheduled for class that is cancelled will be given when the class next meets.

- If the College opens or closes at a time when more than 50% of a class period will be missed, that class will be cancelled for the day.
- If less than 50% of a class will be missed, that class will meet for the remaining portion of its regularly scheduled time.

- For the most up-to-date information regarding College openings, closings, or emergencies, all students, faculty, and staff are encouraged to sign up for email and text alerts via the Montgomery County MC ALERT at https://alert.montgomerycountymd.gov/index.php?CCheck=1

Important Dates:
FEB-03-2019– Refund-6% Drop Deadline
FEB-17-2019– No Grade,Change Audit/Credi Deadline
APR-21-2019– W Grade Drop Deadline

E. Tentative course schedule (subject to change)

Class #	*Date*	*Material*	*Due Dates*
1	*Monday, January 28, 2019*	*Introduction and Course Overview*	
2	*Monday, February 4, 2019*	*Python Basics; Numpy*	
3	*Monday, February 11, 2019*	*Pandas*	*Homework Assignment 1: 2/10/2019*
4	*Monday, February 18, 2019*	*Matplotlib*	
5	*Monday, February 25, 2019*	*nonparametric Statistical simulations: Bootstrapping*	
6	*Monday, March 4, 2019*	*Introduction to Machine Learning*	*Homework Assignment 2: 3/4/2019*
7	*Monday, March 11, 2019*	*Spring Break*	
8	*Monday, March 18, 2019*	*Linear Regression*	*Project 1: 3/18/2019*

9	*Monday, March 25, 2019*	*Logistic Regression*	
10	*Monday, April 1, 2019*	*ScikitLearn: Model hyperparameter, GridSearch, Cross-validation*	
11	*Monday, April 8, 2019*	*Causal Inference*	
12	*Monday, April 15, 2019*	*Decision Trees, Randomforest, XGboost*	*Project 2: 4/15/2019*
13	*Monday, April 22, 2019*	***Neural Networks, Tensorflow***	
14	*Monday, April 29, 2019*	*Unsupervised Learning: K-Means, EM-algorithm*	
	Monday, May 6, 2019	*Project 3 Presentation.*	*Project 3: 5/6/2019*
15	*Monday, May 13, 2019*	*Project 4 Presentation*	*Project 4: 5/13/2019*

The professor reserves the right to make changes to this syllabus. All changes will be made in writing.

Last updated April 1, 2019