Course Syllabus Part II
DSC 530 Data Exploration and Analysis

Course Resources

Course Text

- *Hands-on Data Analysis with Pandas*, 2nd Edition, by Stefanie Molin, ©2021, Packt Publishing, ISBN: 978-1-80056-345-2

- *Practical Statistics for Data Scientists*, 2nd Edition, by Peter Bruce, Andrew Bruce, and Peter Gedeck, © 2020, O'Reilly, ISBN: 78-1-492-07294-2

Required Resources

In this course, you will need to be able to:

- Access the Internet through an internet browser from a Windows or Mac computer.
- Access the electronic textbooks and Python codes in the textbook.
- Collaborate online via Microsoft Office Suites and Blackboard.

- Utilize Python, Jupyter, and Miniconda environments on a Windows or Mac computer. • Download and install necessary software and packages through command line interface.

Supplemental (not required) Resources

Check supplemental resources in weekly content in Blackboard.

Course Schedule

Week	Topic	Required Reading
1 & 2	Preparing for Exploratory Data Analysis, Data Manipulation, and Processing	**Hands-On Data Analysis with Pandas** (Molin, 2021) Chapter 1: Introduction to Data Analysis (Only the 1st section between Page 3-11) Chapter 2: Working with Panda Data Frame (All sections) Chapter 3: Data Wrangling with Pandas (All sections)
3 & 4	Sampling Methods, Descriptive Statistics, and Data Distribution	**Practical Statistics for Data Scientists (Bruce, Bruce, & Gedeck, 2020)** Chapter 1. Exploratory Data Analysis (All sections) Chapter 2: Data and Sampling Distributions (All sections) **Hands-On Data Analysis with Pandas (Molin, 2021)** Chapter 1: Introduction to Data Analysis (Statistical Foundations section)
5 & 6	Data Aggregation, Exploration, and Test Statistics	**Practical Statistics for Data Scientists (Bruce, Bruce, & Gedeck, 2020)** Chapter 3. Statistical Experiments and Significance Testing (All sections) **Hands-On Data Analysis with Pandas (Molin, 2021)** Chapter 4: Aggregating Pandas DataFrames (All sections except Time Series section)
7 & 8	Data Visualization and Customization Techniques	**Hands-On Data Analysis with Pandas (Molin, 2021)** Chapter 5 Visualizing Data with Pandas and Matplotlib (All sections) Chapter 6 Plotting with Seaborn and Customization Techniques (All sections)

9 & 10	Regression, Classification, and Clustering Analysis	**Practical Statistics for Data Scientists (Bruce, Bruce, & Gedeck, 2020)** Chapter 4 Regression and Prediction (All sections except Polynomial and Spline Regression) Chapter 5 Classification (All sections) Chapter 7 Unsupervised Learning (All sections) **Hands-On Data Analysis with Pandas (Molin, 2021)** Chapter 9 Getting Started with Machine Learning, Clustering, Regression, and Classification (All sections)
11	Final Project	N/A

Course Activities

There are 3 main types of activities in this course: Online discussion/participation, required readings and exercises, and final course project. Details are provided in the course LMS Blackboard along with a detailed rubric of grading criteria.

Grade and Point Breakdown

This is the points and percentage breakdown for the 11-week course format.

Component	Percentage	Point Value Each Week	Number of Times	Total
Discussion/ Participation	20%	40 Points	6 Submissions per course, 20 posts per 2 weeks, in the final week, 10 posts required	240
Coding Assignment	30%	72 Points	5 Submissions per course	360
Final Project	50%	600 Points	1 Submission per course	600
			Total Points	1200

Late Work

Late work is not accepted unless arrangements are made with the instructor for very special, unavoidable circumstances. If in doubt, reach out to the instructor as soon as possible.

Online Participation/Discussion

Participation through online discussion is required as part of the course besides reading and trying out Python codes. The use of generative AI tools is encouraged for researching topics and understanding

complex concepts. The Microsoft Teams channel posts should be about individual thoughts and personal experiences. The post should not be completed by pasting ChatGPT or AI responses from your input prompts. Repeated instances of this will result in a 0 for the entire participation grade as it is considered plagiarism and unethical.

Post Requirements

The DSC 530 11-week course runs in a 2-week block format. You must have 20 posts every 2 weeks and be active at least 4 days within this time. Being active means you have posts on different days of the week over the two weeks and do not make all the posts in 1 day to receive full credit.
There is no minimum length requirement, and posts do not need to be written in a formal style. You should communicate how you would with a peer/colleague. The more the post sounds like you in a natural way, the better.

Expectations for Students

- Students should expect to spend approximately 10-12 hours per week to complete the activities and assignments in this course.
- Students will log in as often as needed to complete their assignments and progress through the course.
- Students will treat their classmates and the instructor with respect and courtesy.
- Students are responsible for keeping current with the reading assignments and coming to class prepared to discuss the work assigned.
- Students are responsible for knowing what assignments are due and when.
- Students will submit only their own work and will not commit plagiarism or other acts of academic dishonesty.
- Students will contact the instructor as soon as personal problems arise that may affect the student's ability to complete assignments on time.

Expectations for Faculty

- The instructor will treat all students with respect and courtesy.
- The instructor will make grading criteria clear and follow the criteria scrupulously in evaluating student work.
- The instructor will provide feedback about student work within 6 days of due dates (or 24 hours prior to the next due date)—feedback that helps the student learn and improve.

- The instructor will respond to all student phone and email messages within 48 hours.