MSIS 502: Business Data Analysis Summer 2019

*“For every leader in the company, not just for me, there are decisions that can be made by analysis. These are the best kinds of decisions. They’re fact-based decisions.”* – Jeff Bezos, Amazon CEO.

*“Computers are able to see, hear and learn. Welcome to the future.”* – Dave Waters

In today’s business world, data is the most valuable resource. Many organizations have become data-rich yet still knowledge-poor. How to extract business insights and mine business value from data becomes a fundamental problem for every company. What can be learned from such “big data” to improve decisions? How can an organization leverage its massive data for strategic advantage?

Automation, Robotic and complex analytics together are shaping the future of the industries, making them more sophisticated and smarter. The applications of machine learning are shining like a halo facilitating less equipment-failure, better on-time deliveries, advancement in quality, faster training cycles, advanced automation of design and production process. All this adds up to form endless possibilities for the industries opening the gates to a brand new world of highly efficient, customer-driven factories that seamlessly connects to the surrounding smart ecosystem.

The primary objective of this course is to introduce you to the various techniques used to extract business intelligence, and to learn and reason by machines without the intervention of humans. We will use Python, the most widely used programing language in academia and industry for data analytics. The course will cover general concepts and techniques including Python basics, data visualization, business decision-making, and regression analysis.

The key objectives of this course are two-fold: (1) to provide you with the foundation of identifying a problem, setting a data analysis plan, performing data analyses, and interpreting the results; and (2) to provide you with hands-on experience in applying statistical and data analysis techniques to practical real-word business problems using Python, especially with data science packages including numpy, matplotlib, sklearn, etc.

.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Time** |  | Monday/Wednesday night & Saturday |
| **Location** |  | PCAR 292 |
| **Instructor** |  | Yingfei Wang |
|  |  | PCAR 527, yingfei@uw.edu |
| **Office Hours** |  | Monday 3:00PM – 4:30PM @ PCAR 527 |
|  |  | Or by appointment |
| **Software** |  | Python, Jupyter Notebook |

**EXPECTATIONS, GUIDELINES AND POLICIES**

**Prerequisites**

Students are expected to have taken an intro-level Statistics class before, and have basic problem analyzing and solving skills such as mathematical calculation and algebra. Knowing at least one programming language (preferably Python) is a plus.

# Deliverables

You are responsible for completing each deliverable by its deadline. Deadlines are outlined in the syllabus or in supplementary documents accessible through Canvas. Assignments are due **before class** on their respective due dates. Unless under extreme circumstances, **make-up exam and quizzes will not be arranged.**

|  |  |
| --- | --- |
| **Grading and Workload** |  |
| Individual Assignments | 20% (10% each) |
| Individual Quizzes | 20% (10% each, open-book, open-note, in-class) |
| Final Exam | 50% (Open-book, open-note, in-class) |
| Participation | 10% |

This is a hands-on class. You will learn new skills by practicing in class. I expect many of you will volunteer to share your solutions but if we don’t have any volunteers, I will cold call *randomly*. Active participation to class discussions will constitute the participation portion of your grade.

\*Solutions for in-class exercises will be discussed in class. The solutions will not be distributed.

# Attendance Policy

Class attendance is very important. The exams will be based on the actual class coverage. The students are expected to attend classes and are responsible for obtaining information about missed classes from other students such as changes to the due dates and contents of exams and assignments. They are also expected to actively participate in all class discussions. Arriving late to class will also likely detract from your participation grade, especially when disruptive.

# Late Policy

Late individual assignments will be penalized 15% of the overall mark for each 24-hour period. (For instance, a homework turned in two days late will receive only 70% credit).

Exceptions to these rules will of course be made for serious illness or other emergency circumstances of a similar nature; in these cases, please contact me as soon as you are aware of the problem.

# Technology Courtesy Policy

Cellphones, tablets, e-readers, and etc. are expected to be turned off during class, unless instructed otherwise. Violation of the technology courtesy policy is likely to lower the participation portion of your grade.

# Foster Code of Conduct

By being a student in this course you acknowledge that you are a part of a learning community at the Foster School of Business that is committed to the highest academic standards. As a part of this community, you pledge to uphold the fundamental standards of honesty, respect, and integrity, and accept the responsibility to encourage others to adhere to these standards.

# Course Support

Because material in the course is cumulative, avoid falling behind. It will be difficult to do well in the course if you do not understand material as it is presented. If you do not understand material, please seek assistance as soon as possible.



# Tentative Course Schedule

**Bring your laptop to all classes**

|  |  |  |  |
| --- | --- | --- | --- |
| **Session** | **Date/Time** | **Topic** | **Due** |
| 1 |  | Introduction to Pandas for Data Analysis |  |
| 2 |  | Data Munging in Pandas |  |
| 3 |  | Exploratory Data Analysis and visualization |  |
| 4 |  | Linear Regression and visualization |  |
| 5 |  | Reporting Analysis Results |  |

## Disclaimer

I hope to cover all of the material on this schedule. If I find that the pace of the class is too quick, or that students are having difficulty with particular sections, I reserve the right to spend more time on specific topics and push subsequent topics to later dates, or skip them completely.