DSCI 5340-001, 003: Predictive Analytics and Business Forecasting
Spring 2021 syllabus

CLASS (DAY/TIME): Thursday 6:30pm-9:20pm (Remote)
INSTRUCTOR: Dr. Mahdi Fathi
OFFICE HRS: Tuesday 2:00 – 3:00 PM (online), Thursday 2:00 – 3:00 PM (online),
Friday 8:00 – 9:00 PM (online) or by appointment
CONTACT INFO: OFFICE PHONE: (940) 565-3111
E-MAIL (preferred): mahdi.fathi@unt.edu

Textbooks
- Forecasting, Time Series, and Regression, 4th edition, 2005, Bowerman,
- Brocklebank & Dickey, SAS for Forecasting Time Series, 2nd Ed., Wiley 2009,
- In addition, a number of reference texts will be available in PDF (free of charge)

Software
SAS (main environment). Also, R, Excel, Minitab, and IBM SPSS (some exposure).
(NOTE: All packages are available in our COB virtual lab)

Canvas Learning Management System
Materials for the DSCI 5340 course will be posted on Canvas LMS.

Course Description
This semester we will be working together to learn about how to create Predictive Models
which have many applications for Social Goods. Areas of interest include, but are not
limited to:

- Health and healthcare
- Humanitarian operations
- Disaster relief
- Education
- Social services
- Environment
- Sustainability
- Sharing economy
- Transportation
- Urban planning
- Fraud, collusion, and corruption
- Government policy
- Poverty
- Privacy
- Cyber security
- Crime and terrorism

This course has 12 Modules as follow:

<table>
<thead>
<tr>
<th>Module</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>An Introduction to Forecasting</td>
</tr>
<tr>
<td>2</td>
<td>Basic Statistical Concepts</td>
</tr>
<tr>
<td>3</td>
<td>Simple Linear Regression</td>
</tr>
<tr>
<td>4</td>
<td>Multiple Linear Regression</td>
</tr>
<tr>
<td>5</td>
<td>Model Building and Residual Analysis</td>
</tr>
<tr>
<td>6</td>
<td>Time Series Regression</td>
</tr>
<tr>
<td>7</td>
<td>Decomposition Methods</td>
</tr>
<tr>
<td>8</td>
<td>Exponential Smoothing</td>
</tr>
<tr>
<td>9</td>
<td>Nonseasonal Box-Jenkins Models and Their Tentative Identification</td>
</tr>
<tr>
<td>10</td>
<td>Estimation, Diagnostic Checking, and Forecasting for Nonseasonal Box-Jenkins Models</td>
</tr>
<tr>
<td>11</td>
<td>Box-Jenkins Seasonal Modeling</td>
</tr>
<tr>
<td>12</td>
<td>Advanced topics in Time Series</td>
</tr>
</tbody>
</table>

Also, we have two sessions on real case studies including regression models and time series models.

**Course Objectives**

This course introduces the theory and practice of forecasting, time series, and regressions with an emphasis on practical skills. Having completed DSCI 5340, you will be able to apply forecasting techniques in behavioral decision and management problems, and model and forecast a time series as well as read papers from the literature for writing high-quality journal papers in predictive modeling. More generally, you will obtain an appreciation for the role of dependence in statistical modeling.

**Course Activities**

- Weekly Module Activities
- Discussion forums
- Researching, and Understanding of Journal Papers from "International Journal of Forecasting"
- Following the events and topics in "International Institute of Forecasters (IIF)"
- Searching for Dataset and Novel data science projects in
  - Github
  - Kaggle
  - United States Department of Transportation
  - Challenges for City of Frisco, Dallas, Irving data
○ GUN VIOLENCE ARCHIVE
○ UC Irvine Machine Learning Repository
○ Google data set
○ Dallas city data set
○ Data Science Central
○ COVID-19 Open Research Dataset Challenge

Class Attendance
Regular class attendance and informed participation are expected.

Course Prerequisites
Graduate status and some introductory graduate course in Business Statistics such as DSCI 5010, or DSCI 5180, or consent of the ITDS department, are required. While a high degree of mathematical skill is not necessary in an “applied” course such as this, there are certain insights into the course that are gained through the analytics involved. Statistical software such as SAS, Excel, R, Minitab, and IBM SPSS will be used to demonstrate specification applications. Information on the use of these software packages will be provided in the course and students are not required to have prior experience with the software.

Course Data Sets
Several course datasets required for the Mini-Case assignments, as well as examples covered in class, will be posted on Canvas.

Point Allocation
Quizzes (12@5 pts, 4 dropped) 40 pts
Homework (12@5 pts) 60 pts
Project 1 (Regression Data) 40 pts
Project 2 (Time series Data) 60 pts
Journal Paper (Extra Credit) up to 100 pts

TOTAL 100 pts

An extra credit (up to 100 pts) will be considered for contributing in writing a high quality journal paper.

Letter Grades
180+ pts (=90%) = A
160+ pts (=80%) = B
140+ pts (=70%) = C
120+ pts (=60%) = D
Below 60 pts = F

**Quizzes**
There will be 12 Quizzes, worth 5 points each. These will be typically closed books, at the instructor’s discretion. The Quizzes will be multiple-choice or SAS coding. They will cover the material presented in class on the day of the Quiz. Make-up Quizzes will not be allowed, but the 4 lowest grades among the 12 Quizzes will be dropped.

**Homework**
There will be 12 HWs that will require individual work. Related handouts will be distributed in class and related datasets will be posted on Canvas.

**Group Project**
There will be two projects (regression data and time series data) that will require team work for developing predictive models. Related handouts will be distributed in class and related datasets will be posted on Canvas. You will be asked to form teams of 2 members. You will have to select a business problem that interests you (the instructor will suggest a problem in case you run out of ideas.) Problem data/facts can be real, obtained from published sources, or made up by your team in a way that they correspond to realistic situations. Your team will prepare a written report and a PowerPoint presentation, to be presented in class.

**Miscellaneous Policies**
IMPORTANT DATES: Dates of drop deadlines, exams, final exams, etc., are published in the university catalog and schedule of classes. It is your responsibility to be informed with regard to these dates. Unawareness is no excuse. Do not wait until the last day to drop the course if you are not making satisfactory progress in this class. Your instructor may not be available at that time.

**Campus Closures**
Should UNT close campus, it is your responsibility to keep checking your official UNT e-mail account (EagleConnect), the UNT Web site, and Canvas, to learn if your instructor plans to modify class activities, and how. This may include changing assignment due dates, rescheduling quizzes and exams, etc.

**Student Perceptions of Teaching (SPOT)**
Student Perceptions of Teaching (SPOT) utilizes IASystem® and is a requirement for all organized classes at UNT. This short Web-based survey will be available to you at the end of the semester, providing you with a chance to comment on how this class is taught. I am very interested in this feedback from my students, as I work to continually improve my teaching. I consider SPOT to be an important part of your class participation.

**Use of Cell Phones**
As a courtesy to your instructor and to your fellow classmates, you are asked to set your cell phone to *vibrate*, or switch it off. In case of a personal emergency, if you must use your cell phone, you are asked to step out of the classroom.

**Students with Disabilities**
UNT and the College of Business comply with the Americans with Disabilities Act in making reasonable accommodations. If you have an established disability you should register with the Office for Disability Accommodation and receive further instructions. Please see your instructor as soon as possible if you have any questions.

**Academic Integrity**
This course adheres to the UNT policy on academic integrity. The policy can be found at [https://vpaa.unt.edu/fs/resources/academic/integrity](https://vpaa.unt.edu/fs/resources/academic/integrity). Practices that violate academic integrity, such as “cheating” or “plagiarism”, are strongly discouraged. If you engage in academic dishonesty related to this class, you may receive a failing grade on the test or assignment, or a failing grade in the course. In addition, the case may be reported to the UNT Dean of Students/Academic Integrity Office, which maintains a database of related violations.

**Class Schedule (Subject to change; Effective 8/24/2020)**
I assume each week starts on Monday so Week 1, for example, refers to Jan 11th, Mon – Jan 17th, Sun.

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
<th>Readings</th>
<th>HW Due</th>
<th>Quiz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>An Introduction to Forecasting</td>
<td>Module 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Basic Statistical Concepts</td>
<td>Module 2</td>
<td>HW1</td>
<td>Quiz 1</td>
</tr>
<tr>
<td>3</td>
<td>Simple Linear Regression</td>
<td>Module 3</td>
<td>HW2</td>
<td>Quiz 2</td>
</tr>
<tr>
<td>4</td>
<td>Multiple Linear Regression</td>
<td>Module 4</td>
<td>HW3</td>
<td>Quiz 3</td>
</tr>
<tr>
<td>5</td>
<td>Model Building and Residual Analysis</td>
<td>Module 5</td>
<td>HW4</td>
<td>Quiz 4</td>
</tr>
<tr>
<td>6</td>
<td>Time Series Regression</td>
<td>Module 6</td>
<td>HW5</td>
<td>Quiz 5</td>
</tr>
<tr>
<td>7</td>
<td>Case Study Review-Regression Modeling</td>
<td>Case study Module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Project 1 presentation-Regression Modeling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>UNT Spring Break</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Decomposition Methods</td>
<td>Module 7</td>
<td>HW6</td>
<td>Quiz 6</td>
</tr>
<tr>
<td>11</td>
<td>Exponential Smoothing</td>
<td>Module 8</td>
<td>HW7</td>
<td>Quiz 7</td>
</tr>
<tr>
<td>12</td>
<td>Nonseasonal Box-Jenkins Models and Their Tentative Identification</td>
<td>Module 9</td>
<td>HW8</td>
<td>Quiz 8</td>
</tr>
<tr>
<td>13</td>
<td>Estimation, Diagnostic Checking, and Forecasting for Nonseasonal Box-Jenkins Models</td>
<td>Module 10</td>
<td>HW9</td>
<td>Quiz 9</td>
</tr>
<tr>
<td>14</td>
<td>Box-Jenkins Seasonal Modeling</td>
<td>Module 11</td>
<td>HW10</td>
<td>Quiz 10</td>
</tr>
<tr>
<td>15</td>
<td>Advanced topics in Time Series</td>
<td>Module 12</td>
<td>HW11</td>
<td>Quiz 11</td>
</tr>
<tr>
<td>16</td>
<td>Case studies Review-Time series modeling</td>
<td>Case study Module</td>
<td>HW12</td>
<td>Quiz 12</td>
</tr>
<tr>
<td>17</td>
<td>Project 2 presentation- Time series modeling</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>