

## STA3XX Intermediate Data Analytics

### Catalog Description

This is a non-technical and introductory course of applied data analytics using a programming language such as Python as the primary computational tool. Topics include data preparation, exploratory and visual data analysis, statistical models, and machine learning algorithms for predictive analytics. Students will gain hands-on analytic experience through real-world projects.

### Pre-requisites

STA319 or permission of instructor.

### Syllabus Components

- **Course Objectives:** Upon successful completion of this course, students will be equipped with basic programming and data analytics skills to complete real-world projects from raw data.
- **Learning Outcomes:** By the end of the class, students should be able to
  - query relational databases.
  - perform data management and prepare analytic data sets to use software tools.
  - conduct basic exploratory data analysis.
  - perform basic visual analysis with appropriate software tools.
  - conduct thorough analysis using appropriate analytic methods and correctly interpret the results.
  - draw reasonable conclusions with sufficient explanation and elaboration.
  - use a computational notebook to draft a well-organized analytic report with thoughtful and convincing details.
- **Textbooks (optional)**
  - will provide a list of optional books
- **Class Notes:** Class notes will be provided
- **Coverage:** See the list of tentative topics
- **Assessment:** There will be assignments and projects
  - Regular coding and analysis assignments (data preparation, small data analysis, project proposal, etc) (40%)

- 2 projects (One on data integration and one on analytics ) (40%, 20% each project)
- Class attendance and participation (10%)
- Project presentation (10%)

## **Suggested Weekly Topics**

### **Week 1: Introduction and Logistics**

- Scope of analytics: roadmap and tools
- Platforms and languages for this course
- Computational notebook

### **Week 2: Data Types, Data Frame, Input/Output**

- Data types and structures
- Data input/output
- Accessing data frames

### **Week 3: Syntax of Functions and Control Flow**

- Syntax of functions and user-defined functions
- Control flow and iterations

### **Week 4: Basic Statistical Graphics**

- Basics of statistical graphics: design elements and taxonomy
- Drawing with graphics libraries
- Computational graphics

### **Week 5: Preparing Analytic Data Sets: Merge, Transformation, Imputations**

- Organize data for different analyses
- Working with multiple data tables
- Transformation and standardization

### **Week 6: Visual and Exploratory Data Analysis**

- The role of EDA in data analytics
- Methods and Types of EDA
- Tools for visual analytics
- Special visualizations

### **Week 7: Statistics for Data Analytics**

- Sampling Designs - Probabilistic Sampling
- Fundamental distributions for data analytics
- Role of regression modeling in data analytics

**Week 8: Prediction with Regression Models**

- Linear regression as a predictive model
- logistic predictive modeling
- Performance evaluation

**Week 9: Rule-based Learning Algorithms**

- Supervised and unsupervised machine learning
- Survey of classification algorithms
- Survey of Methods of Clustering

**Week 10: Basic of Financial Risk Analytics**

- Types of financial risks
- Analytics in financial risk management
- Term project guidelines:

**Week 11: Case Study: Credit Risk Analytics**

- Credit Risk types, scoring, and modeling
- Should this loan be approved - predicting loan default

**Week 12: Case Study: Fraud risk Analytics**

- Types and nature of financial fraud
- Fraud Analytics: expert rules and models
- Case study: Credit card fraud modeling

**Week 13: Project Preparation**

- Project topics:
- Project SAP and Analysis

**Week 14: Project Report**

- Finalizing the term project report.
- Prepare project presentation