1 Information

- **Contacting Lecturer**: Email is the only way to contact the lecturer outside of class. Email of the lecturer arunku@wharton.upenn.edu. Please start the subject line of the email ALWAYS with “STAT 111:”.

- **Office Hours**: Lecturer’s office hours are open. Do not hesitate to make an appointment for any time 9-5 Mon-Fri via the email address above. Please indicate two times that are convenient to you and one will be confirmed.

- **Lectures**: There is only one section for this class. Lectures are scheduled five days each week Monday through Friday 11:00 AM – 12:45 PM from May 21 - June 27, 2018.

- **Textbook**: The most relevant material for the course are class notes. Depending on the topic covered in class, printed notes will be handed out in class and will be available on course webpage (see below). *It is highly recommended that you take detailed notes during the class.* There is no required textbook. Some suggested references are as follows:

- **Course Webpage**: This course uses the online system CANVAS. Canvas is available to all Penn students at [https://canvas.upenn.edu](https://canvas.upenn.edu) using pennkey authentication. Canvas page will have all the class materials and grades throughout the course period. For technical questions about using canvas you can contact the Wharton student computing support office at 215 898 8600 or at [https://computing.wharton.upenn.edu/](https://computing.wharton.upenn.edu/)

- **Homeworks**: The lectures are five days every week. To keep up with the pace of the class, there will be two homeworks each week handed out on Monday and Friday and additionally, there will be weekly in-class quiz on Wednesday. Quizzes will mostly concentrate on concept understanding. Each homework will be due on the third day. The homework handed out on Monday is due on Thursday and the one handed out on Friday is due on the Monday following. No late homeworks will be considered.

- **Assessment**: The final grade will be evaluated as follows:
  1. Homeworks: 25%
  2. Quizzes: 20%
  3. Class participation: 10%
  4. Final Exam: 45%

Final exam will be on June 27th (the last day of classes) and is in-class.
• **Classroom Expectations:** Classes will start and end on time every day. You are expected to arrive on time for the class. All cell phones must be turned off. It is highly recommended that you take detailed notes and participate during the class. All classes will be interactive in nature and you are expected to engage in the class discussions.

• **Software – JMP:** The course will in part be given in association with use of the statistical package JMP. You should either buy and then install this package on your computer or (a better option, since buying JMP is expensive) use the Wharton computers that have it installed. If you decide to choose the second option, it might be helpful to check out the Wharton Virtual Lab feature which allows you to use the Wharton computing environment remotely from your laptop: [http://supportcenteronline.com/link/portal/632/655/Article/5657/5a-Virtual-Lab-for-Laptops](http://supportcenteronline.com/link/portal/632/655/Article/5657/5a-Virtual-Lab-for-Laptops). Alternatively Penn students can get a JMP license through e-academy at [http://www.onthehub.com/jmp/](http://www.onthehub.com/jmp/) for 30 for a 6 month license or 50 for a year license.

You will not be able to use Wharton computers until you have created a Wharton account. If you are a non-Wharton student please create a class account at: [https://whartonstudentsupport.zendesk.com/hc/en-us/articles/202127736-Creating-a-Wharton-CLASS-Account](https://whartonstudentsupport.zendesk.com/hc/en-us/articles/202127736-Creating-a-Wharton-CLASS-Account).

2 Course Description and Syllabus

The content of this course falls into two broad categories, namely Probability theory and Statistics. The reason why we discuss probability theory will be given in the first lecture. A more detailed list of the topics covered within these two categories is given in the syllabus below. The aim of the course is to give you an introduction to the concepts in probability and provide you with a basic idea of statistical inference.

The effective syllabus of the course will be whatever is covered in the class. The following is the general guideline for the course. We may not be able to cover some of the topics due to time constraints.

1. Introduction: Statistics and Probability Theory
   - What is Statistics?
   - What related Statistics and Probability
   - Some Uses and Misuses of Statistics

2. Elementary Set Theory and Events
   - What are Events?
   - Notation
   - Unions, Intersections and Complements of Events
   - Venn diagrams

3. Probability of Events
   - The three Axioms of Probability
   - Mutually Exclusive Events
   - Independence of Events
   - Conditional Probability

4. Random Variable and its Probability Distribution
   - What is a Random Variable?
   - Definition of Discrete Random Variable
   - Definition of Probability Distribution
   - Measures of Central Tendency and Dispersion.
5. The Binomial Distribution
   - Definition
   - Binomial Formula and its Characteristics.

6. Many Random Variables
   - Independent and Identically Distributed Random Variables
   - Mean and Variance of Sum of Random Variables

7. Continuous Random Variables
   - Definition
   - Probability Density Function
   - Mean and Variance of a Continuous Random Variable

8. The Normal Distribution
   - Definition and The Bell Curve
   - Central Limit Theorem
   - The Chi-square Distribution

9. Data and Graphical Representation
   - Histogram and Summary Statistics
   - Scatter Plots

10. Introduction to Statistical Inference
    - What is Statistical Inference?
    - Three types of Inference Problems
    - Examples

11. Estimation of a Parameter
    - General Principles of Estimation of a Parameter
    - Estimation in Binomial Distribution
    - Estimation of Mean and Variance
    - Examples

12. Confidence Intervals
    - What is a Confidence Interval?
    - Confidence Interval in Binomial Distribution
    - Confidence Interval for Mean
    - Confidence Interval for Difference of Means

13. Statistical Hypothesis Testing
    - What is Hypothesis Testing?
    - Null and Alternative Hypothesis
    - Errors in Hypothesis Testing
    - Two Approaches to Hypothesis Testing
    - Critical Value and P-value
14. Simple Linear Regression
   - Assumptions of Simple Linear Regression
   - Estimation of Parameters
   - Testing about Parameters

15. Robust Statistics
   - What is Robustness?
   - Median and Weighted Means.