**How to Better Understand the World**

An Introduction to Surveying and Statistics

****

**1. Course Description**



The purpose of this course is to ignite the innate curiosity within each student, empowering them to comprehend and interpret the world through the lens of statistics. We will explore fundamental techniques that statisticians employ, while also delving into the underlying principles of human decision-making. Ultimately, this course aims to nurture students' confidence in their ability to harness the power of statistics, enabling them to tackle even the most complex challenges with informed and insightful perspectives. By the end of this course, students will have gained a solid foundation in statistical thinking, data analysis, and interpretation, which will be valuable in their academic and professional lives. The wide range of applications and examples will help engage students and demonstrate the practical relevance of statistics in various fields.

**2. Introduction**



**Global-Side Mentor**

| **INSERT A PHOTO OF YOURSELF HERE** | **Name:** Iñaki Arango  **University/ Year / Major:** Harvard University ‘25 (Double Major in Computer Science and Statistics)  **Email:** inakiarango@college.harvard.edu  **Introduction:** |
| --- | --- |

**Japan-Side Mentor**

| **INSERT A PHOTO OF YOURSELF HERE** | **Name:**  **University/ Year / Major:**  **Email:**  **Introduction:** |
| --- | --- |

**3. Pre-Assignment**



Please watch these videos and be prepared for a casual discussion about them on the first day! (Bring any questions you have!)

A brief intro to Probability and Statistics:

​​<https://www.youtube.com/watch?v=zouPoc49xbk>

What is “Statistics”?:

<https://www.youtube.com/watch?v=sxQaBpKfDRk>

**4. Seminar Outline**



**DAY 1: Introductions, and Intro to Statistics, Data Types, and Probability Theory**

1. Course introduction and objectives
2. Importance of statistics in the real world
   1. Medical studies
   2. Sports statistics
   3. Weather forecasting
   4. Market research
   5. Quality control in manufacturing
   6. Opinion polls and surveys (Government, corporate, etc.)
3. Explain the building blocks of statistics (Give examples)
   1. What is a model and what is data, how do you go from one to the other
   2. Combination (nCk, nPk)
      1. I have 20 students in a classroom and I we are playing football, so I want to split them into 2 teams. How many ways are there of doing that?
      2. I have 10 performances at a dance show. How many sho orders are possible?
      3. Why is the answer to the two questions before different? We’ll look at that later.
   3. Probability
      1. I roll 2 6-sided dice. How likely am I to get 12? Is it the same probability of getting 6?
4. Detailed intro to combination
   1. N Choose K
   2. N Permute K
5. Detailed intro to probability
   1. Definition of simple probability
   2. The complement trick
6. Introduction to probability theory: independent events, conditional probability, and Bayes' theorem
7. Activity: Probability in games of chance (dice, coins, cards)
   1. There are 20 people at a chess club on a certain day. They each find opponents and start playing. How many possibilities are there for how they are matched up, assuming in each game it *does* matter who has the white pieces (in a chess game one player has the white pieces and the other player has the black pieces)?
   2. To fulfill the requirements for a certain degree, a student can choose to take any 7 out of a list of 20 courses, with the constraint that at least 1 of the 7 courses must be a statistics course. Suppose that 5 of the 20 courses are statistics courses.
      1. How many choices are there for which 7 courses to take?
      2. Explain intuitively why the answer to (a) is *not* C(5,1) \* C(19, 6.)

**DAY 2: Data Visualization, Descriptive Statistics, and Applications in Sports**

1. Icebreaker
2. Introduction to data visualization
3. Principles of effective data visualization
4. Types of charts and graphs: bar charts, pie charts, line graphs, histograms, and scatter plots
5. Measures of central tendency: mean, median, and mode
6. Measures of dispersion: range, variance, and standard deviation
7. Activity: Analyzing sports statistics and creating visualizations
   1. Have students use a website to find raw data about the teams/players and come up with ideas for potential questions they have about the data, and what interesting ways they think of presenting the information.

**DAY 3: Probability, Sampling, and Medical Statistics**

1. Icebreaker activity
2. Give time to finish the last activity if necessary.
3. What are the different ways in which we collect data?
   1. Introduction to sampling techniques: random, systematic, stratified, and cluster sampling
   2. Very brief and simple explanation about sampling distributions and the Central Limit Theorem -> how the data collected for very few explains the entire population.
4. Medical statistics: evaluating treatments, identifying risk factors, and predicting patient outcomes
5. Activity: Conduct a small-scale survey with their classmates to try and answer the questions from the previous class and start to make a presentation about what they found.

**Day 4: Putting Ideas into Action**

1. Icebreaker
2. Review of key concepts and techniques
3. Introduction to inferential statistics
4. Confidence intervals and hypothesis testing
5. Real-world applications of statistics: social media, big data, weather forecasting, and financial modeling
6. Group activity: Finish the presentation from the last class and present it to the other groups.
7. Reflection and discussion: Empowering students to apply statistical thinking in their daily lives
   1. Question what you are told (not disrespectfully)
   2. Dig deeper (do your own research)
   3. Care about others. If you find out something new that could help others, share it.

**5. Post-Assignment**



Video/Reading: (May give additional optional reading throughout course)

If you are interested in learning broader statistics topics, I recommend this playlist by the channel CrashCourse:

<https://www.youtube.com/playlist?list=PL8dPuuaLjXtNM_Y-bUAhblSAdWRnmBUcr>

If you are interested in learning more about a specific topic, you can come and talk to me, and I can recommend a resource specific to that topic and tailored to your background!

**6. Vocabulary List**



Statistics, Data, Qualitative, Quantitative, Nominal, Ordinal, Interval, Ratio, Probability, Visualization, Mean, Median, Mode, Variance, Regression