# Data Math and AP Statistics Outline for 2020 September
## Instructor: Sam Wu

**Time Frame**

(1:20 – 3:50 pm on Saturdays, totally 58 hours)

**AP Exam: May, 2021**

Note that the exam registration CANNOT be made at Olympiads School. Please register at your own school. If your school doesn't offer AP Exam registration, then you will need to find another school to register. If you have any issues with registration and require Olympiads School's help, please contact the school no later than the end of Jan. 2021

**Study Aids**

Please check the course website regularly for lecture notes and additional resources. Students are expected to bring the following to each class:

- Lecture notes/Homework
- Pen/Pencil/Eraser/Binder
- Graphing calculator

**Textbooks**

- "Mathematics of Data Management", by Barbara J. Canton and others, published by McGraw-Hill Ryerson, 2002, ISBN 0-07-552912-2/0-07-091714-0
- "The Practice of Statistics for the AP Exam", *by Moore,* published by W. H. Freeman and Company, 2014, Fifth edition, ISBN: 1464153876, 9781464153877

*\* Textbook can be rented or purchased from Olympiads School.*

**Resources**

AP Statistics Course and Exam Description

*http://media.collegeboard.com/digitalServices/pdf/ap/ap-statistics-course-description.pdf*

## Mathematics of Data Management (MDM4U)

**Chapter 1 Permutations and Organized Counting** **(6 hours)**

1.1 Organized Counting; 1.2 Factorials and Permutations;

1.3 Permutations with Some Identical Elements; 1.4 Pascal's Triangle; 1.5 Applying Pascal's Method

**Chapter 2 Combinations and the Binomial Theorem** **(6 hours)**

2.1 Organized Counting with Venn Diagrams; 2.2 Combinations;

2.3 Problem Solving with Combinations; 2.4 The Binomial Theorem

**Chapter 3 Introduction to Probability** **(6 hours)**

3.1 Basic Probability Concepts; 3.2 Odds; 3.3 Probabilities Using Counting Techniques

3.4 Dependent and Independent Events; 3.5 Mutually Exclusive Events;

3.6 Applying Matrices to Probability Problems

**Chapter 4 Probability Distributions** **(6 hours)**

4.1 Probability Distributions; 4.2 Binomial Distributions; 4.3 Geometry Distribution;

4.4 Hypergeometric Distributions

**Chapter 5 The Normal Distribution** **(8 hours)**

5.1 Continuous Probability; 5.2 Properties of the Normal Distribution;

5.3 Normal Sampling and Modelling; 5.4 Normal Approximation to the Binomial Distribution;

5.5 Repeated Sampling and Hypothesis; 5.6 Confidence Intervals

**AP Statistics Topic Outline**

Following is an outline of the major topics covered by the AP Statistics Exam. The ordering here is intended to define the scope of the course but not necessarily the sequence. The percentages in parentheses for each content area indicate the coverage for that content area in the exam.

**I.      Exploring Data: Describing patterns and departures from patterns (20%–30%)**
*Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns. Emphasis should be placed on interpreting information from graphical and numerical displays and summaries.*                          **(10 hours)**

A.      Constructing and interpreting graphical displays of distributions of univariate data (dotplot, stemplot, histogram, cumulative frequency plot)
1.      Center and spread
2.      Clusters and gaps
3.      Outliers and other unusual features
4.      Shape

B.      Summarizing distributions of univariate data
1.      Measuring center: median, mean
2.      Measuring spread: range, interquartile range, standard deviation
3.      Measuring position: quartiles, percentiles, standardized scores (z-scores)
4.      Using boxplots
5.      The effect of changing units on summary measures

C.      Comparing distributions of univariate data (dotplots, back-to-back stemplots, parallel boxplots)
1.      Comparing center and spread: within group, between group variation
2.      Comparing clusters and gaps
3.      Comparing outliers and other unusual features
4.      Comparing shapes

D.      Exploring bivariate data
1.      Analyzing patterns in scatterplots
2.      Correlation and linearity
3.      Least-squares regression line
4.      Residual plots, outliers and influential points
5.      Transformations to achieve linearity: logarithmic and power transformations

E.      Exploring categorical data
1.      Frequency tables and bar charts
2.      Marginal and joint frequencies for two-way tables
3.      Conditional relative frequencies and association
4.      Comparing distributions using bar charts

**II.      Sampling and Experimentation: Planning and conducting a study (10%–15%)**
*Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.*                                            **(10 hours)**

A.      Overview of methods of data collection
1.      Census
2.      Sample survey
3.      Experiment
4.      Observational study

B.      Planning and conducting surveys
1.      Characteristics of a well-designed and well-conducted survey
2.      Populations, samples and random selection
3.      Sources of bias in sampling and surveys
4.      Sampling methods, including simple random sampling, stratified random sampling and cluster sampling

C.      Planning and conducting experiments
1.      Characteristics of a well-designed and well-conducted experiment
2.      Treatments, control groups, experimental units, random assignments and replication
3.      Sources of bias and confounding, including placebo effect and blinding
4.      Completely randomized design
5.      Randomized block design, including matched pairs design

D.      Generalizability of results and types of conclusions that can be drawn from observational studies, experiments and surveys

**III.      Anticipating Patterns: Exploring random phenomena using probability and**
simulation (20%–30%)
*Probability is the tool used for anticipating what the distribution of data should look like under a given model.*                              `(already discussed in` **Data** `Math)`

A.      Probability
1.      Interpreting probability, including long-run relative frequency interpretation
2.      "Law of Large Numbers" concept
3.      Addition rule, multiplication rule, conditional probability and independence
4.      Discrete random variables and their probability distributions, including binomial and geometric
5.      Simulation of random behavior and probability distributions
6.      Mean (expected value) and standard deviation of a random variable, and linear transformation of a random variable

B.     Combining independent random variables
1.     Notion of independence versus dependence
2.     Mean and standard deviation for sums and differences of independent random variables

C.     The normal distribution
1.     Properties of the normal distribution
2.     Using tables of the normal distribution
3.     The normal distribution as a model for measurements

D.     Sampling distributions
1.     Sampling distribution of a sample proportion
2.     Sampling distribution of a sample mean
3.     Central Limit Theorem
4.     Sampling distribution of a difference between two independent sample proportions
5.     Sampling distribution of a difference between two independent sample means
6.     Simulation of sampling distributions
7.     t-distribution
8.     Chi-square distribution

**IV.     Statistical Inference: Estimating population parameters and testing hypotheses (30%–40%)**
*Statistical inference guides the selection of appropriate models.*     **(6 hours)**

A.     Estimation (point estimators and confidence intervals)
1.     Estimating population parameters and margins of error
2.     Properties of point estimators, including unbiasedness and variability
3.     Logic of confidence intervals, meaning of confidence level and confidence intervals, and properties of confidence intervals
4.     Large sample confidence interval for a proportion
5.     Large sample confidence interval for a difference between two proportions