AP Statistics
Mrs. Spano
2024-2025

Name _____

Summer Assignment

Welcome future statisticians! In August you will embark on a wonderful journey of exploring data analysis. Statistics is a math course unlike every other math course you have taken. You will never sit in class and wonder "When will I ever use this?" or "Why am I learning this?". Turn on the news, read the newspaper, look around you – statistics are everywhere! Real people use statistics in their real lives daily to make decisions. My average grocery bill is lower at *this* store, I have a better chance at winning *this* game, I'll probably make more money with *this* major… we say things like this all the time.

In AP Statistics you will do more writing than in any other math course. A statistic is a number with a context, and the context matters. The same percentage can mean very different things in different situations. To get you ready for the wonderful world of statistics, you will not complete a packet of Algebra review problems, rather, you need to be ready to see the context and write about it. The purpose of this assignment is to enhance your reading comprehension and writing skills in the mathematics classroom, as both reading and writing are important parts of the course.

All of your work is attached with all instructions. This work is due the first day of school and will be collected and graded as a formative assessment, so please take the assignment seriously.

If you have any questions or clarifications, email me (kspano@oxfordasd.org).

Late work will result in the following penalties:

Late	**Penalty**
1 day late	Loss of ¼ of total points
2 days late	Loss of ½ of total points
More than 2 days late	NO CREDIT

The summer assignment is composed of two parts.

1. **AP Statistics Exam.** Read the following information provided obtained from College Board regarding the AP Statistics Exam and answer the questions that follow. You will turn in the answered questions.

2. **Autobiographical Essay.** Compose a one-page (no longer) typed document. (NOTE: A partial page will **not** suffice.)

Part I: AP Statistics Course Information

Read the following and answer the questions at the end:

Highlights from the AP Statistics Course Description

(from http://apcentral.collegeboard.com/apc/public/repository/ap-statistics-course-description.pdf)

Introduction

The Advanced Placement Program offers a course description and exam in statistics to secondary school students who wish to complete studies equivalent to a one semester, introductory, non-calculus-based, college course in statistics.

Statistics and mathematics educators who serve as members of the AP Statistics Development Committee have prepared the Course Description and exam to reflect the content of a typical introductory college course in statistics. The exam is representative of such a course and therefore is considered appropriate for the measurement of skills and knowledge in the field of introductory statistics.

In colleges and universities, the number of students who take a statistics course is almost as large as the number of students who take a calculus course. A July 2002 article in the *Chronicle of Higher Education* reports that the enrollment in statistics courses from 1990 to 2000 increased by 45 percent — one testament to the growth of statistics in those institutions. An introductory statistics course, similar to the AP Statistics course, is typically required for majors such as social sciences, health sciences and business. Every semester about 236,000 college and university students enroll in an introductory statistics course offered by a mathematics or statistics department. In addition, a large number of students enroll in an introductory statistics course offered by other departments. Science, engineering and mathematics majors usually take an upper-level calculus-based course in statistics, for which the AP Statistics course is effective preparation.

The Course

The purpose of the AP course in statistics is to introduce students to the major concepts and tools for collecting, analyzing and drawing conclusions from data. Students are exposed to four broad conceptual themes:

1. Exploring Data: Describing patterns and departures from patterns
2. Sampling and Experimentation: Planning and conducting a study
3. Anticipating Patterns: Exploring random phenomena using probability and simulation
4. Statistical Inference: Estimating population parameters and testing hypotheses

AP Statistics Course Content Overview

The topics for AP Statistics are divided into four major themes: exploratory analysis (20–30 percent of the exam), planning and conducting a study (10–15 percent of the exam), probability (20–30 percent of the exam), and statistical inference (30–40 percent of the exam).

I. *Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns.* In examining distributions of data, students should be able to detect important characteristics, such as shape, location, variability and unusual values. From careful observations of patterns in data, students can generate conjectures about relationships among variables. The notion of how one variable may be associated with another permeates almost all of statistics, from simple comparisons of proportions through linear regression. The difference between association and causation must accompany this conceptual development throughout.

II. *Data must be collected according to a well-developed plan if valid information is to be obtained.* If data are to be collected to provide an answer to a question of interest, a careful plan must be developed. Both the type of analysis that is appropriate and the nature of conclusions that can be drawn from that analysis depend in a critical way on how the data was collected. Collecting data in a reasonable way, through either sampling or experimentation, is an essential step in the data analysis process.

III. *Probability is the tool used for anticipating what the distribution of data should look like under a given model.* Random phenomena are not haphazard: they display an order that emerges only in the long run and is described by a distribution. The mathematical description of variation is central to statistics. The probability required for statistical inference is not primarily axiomatic or combinatorial but is oriented toward using probability distributions to describe data.

IV. *Statistical inference guides the selection of appropriate models.* Models and data interact in statistical work: models are used to draw conclusions from data, while the data are allowed to criticize and even falsify the model through inferential and diagnostic methods. Inference from data can be thought of as the process of selecting a reasonable model, including a statement in probability language, of how confident one can be about the selection.

Topic Outline

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.

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.

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.

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.

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
    6. Confidence interval for a mean
    7. Confidence interval for a difference between two means (unpaired and paired)
    8. Confidence interval for the slope of a least-squares regression line
B. Tests of significance
    1. Logic of significance testing, null and alternative hypotheses; p-values; one- and two-sided tests; concepts of Type I and Type II errors; concept of power
    2. Large sample test for a proportion
    3. Large sample test for a difference between two proportions
    4. Test for a mean
    5. Test for a difference between two means (unpaired and paired)
    6. Chi-square test for goodness of fit, homogeneity of proportions, and independence (one- and two-way tables)
    7. Test for the slope of a least-squares regression line

The Use of Technology

The AP Statistics course adheres to the philosophy and methods of modern data analysis. Although the distinction between graphing calculators and computers is becoming blurred as technology advances, at present the fundamental tool of data analysis is the computer. The computer does more than eliminate the drudgery of hand computation and graphing — it is an essential tool for structured inquiry.

Data analysis is a journey of discovery. It is an iterative process that involves a dialogue between the data and a mathematical model. As more is learned about the data, the model is refined and new questions are formed. The computer aids in this journey in some essential ways. First, it produces graphs that are specifically designed for data analysis. These graphical displays make it easier to observe patterns in data, to identify important subgroups of the data and to locate any unusual data points. Second, the computer allows the student to fit complex mathematical models to the data and to assess how well the model fits the data by examining the residuals. Finally, the computer is helpful in identifying an observation that has an undue influence on the analysis and in isolating its effects.

In addition to its use in data analysis, the computer facilitates the simulation approach to probability that is emphasized in the AP Statistics course. Probabilities of random events, probability distributions of random variables and sampling distributions of statistics can be studied conceptually, using simulation. This frees the student and teacher from a narrow approach that depends on a few simple probabilistic models.

Because the computer is central to what statisticians do, it is considered essential for teaching the AP Statistics course. However, it is not yet possible for students to have access to a computer during the AP Statistics Exam. Without a computer and under the conditions of a timed exam, students cannot be asked to perform the amount of computation that is needed for many statistical investigations. Consequently, standard computer output will be provided as necessary and students will be expected to interpret it.

Currently, the graphing calculator is the only computational aid that is available to students for use as a tool for data analysis on the AP Exam.

Formulas and Tables

Students enrolled in the AP Statistics course should concentrate their time and effort on developing a thorough understanding of the fundamental concepts of statistics. They do not need to memorize formulas. [A] list of formulas and tables will be furnished to students taking the AP Statistics Exam.

The Exam

The AP Statistics Exam is 3 hours long and seeks to determine how well a student has mastered the concepts and techniques of the subject matter of the course. This paper-and-pencil exam consists of (1) a 90-minute multiple-choice section testing proficiency in a wide variety of topics, and (2) a 90-minute free-response section requiring the student to

answer open-ended questions and to complete an investigative task involving more extended reasoning. In the determination of the score for the exam, the two sections will be given equal weight.

Each student will be expected to bring a graphing calculator with statistical capabilities to the exam. The expected computational and graphic features for these calculators are described in an earlier section. Minicomputers, pocket organizers, electronic writing pads and calculators with qwerty (i.e., typewriter) keyboards will not be allowed. Calculator memories will not be cleared. However, calculator memories may be used only for storing programs, not for storing notes. A student may bring up to two calculators to the exam.

Multiple-Choice Questions

On the AP exam, there will be 40 multiple choice questions with five answer choices each.

Multiple-choice scores are based on the number of questions answered correctly. Points are not deducted for incorrect answers, and no points are awarded for unanswered questions. Because no points are deducted for incorrect answers, students are encouraged to answer all multiple-choice questions. On difficult questions, students should eliminate as many incorrect answer choices as they can, and then make an educated guess among the remaining choices.

Free-Response Questions

In the free-response section of the AP Statistics Exam, students are asked to answer five questions and complete an investigative task. Each question is designed to be answered in approximately 12 minutes. The longer investigative task is designed to be answered in approximately 30 minutes.

Statistics is a discipline in which clear and complete communication is an essential skill. The free-response questions on the AP Statistics Exam require students to use their analytical, organizational and communication skills to formulate cogent answers and provide students with an opportunity to:

- Relate two or more different content areas (i.e., exploratory data analysis, experimental design and sampling, probability, and statistical inference) as they formulate a complete response or solution to a statistics or probability problem.
- Demonstrate their mastery of statistics in a response format that permits the students to determine *how* they will organize and present each response.

The purpose of the investigative task is not only to evaluate the student's understanding in several content areas but also to assess his or her ability to integrate statistical ideas and apply them in a new context or in a nonroutine way.

Scoring of Free-Response Questions

The evaluation of student responses on the free-response section of the AP Statistics Exam reflects the dual importance of statistical knowledge and good communication. The free-response questions and the investigative task are scored "holistically"; that is, each question's response is evaluated as "a complete package." With holistic scoring, after reading through the details of a student's response, the scorer makes a judgment about the *overall quality* of the response. This is different from "analytic" scoring, where the individual components to be evaluated in a student's response are specified in advance, and each component is given a value counting toward the overall score.

The AP Statistics scoring guideline (rubric) for each free-response question has five categories, numerically scored on a 0 to 4 scale. Each of these categories represents a level of quality in the student response. These levels of quality are defined on two dimensions: statistical knowledge and communication. The specific rubrics for each question are tied to a general template, which represents the descriptions of the quality levels as envisioned by the Development Committee. This general template is given in the following table, "A Guide to Scoring Free-Response Statistics Questions."

A GUIDE TO SCORING FREE-RESPONSE STATISTICS QUESTIONS: THE CATEGORY DESCRIPTORS

Score Descriptors	Statistical Knowledge	Communication
	Identification of the important components of the problem Demonstration of the statistical concepts and techniques that result in a correct solution of the problem	Explanation of what was done and why, along with a statement of conclusions drawn in context
4 Complete	• shows complete understanding of the problem's statistical components • synthesizes a correct relationship among these components, perhaps with novelty and creativity • uses appropriate and correctly executed statistical techniques • may have minor arithmetic errors but answers are still reasonable	• provides a clear, organized, and complete explanation, using correct terminology, of what was done and why • states appropriate assumptions and caveats • uses diagrams or plots when appropriate to aid in describing the solution • states an appropriate and complete conclusion in context
3 Substantial	• shows substantial understanding of the problem's statistical components • synthesizes a relationship among these components, perhaps with minor gaps • uses appropriate statistical techniques • may have arithmetic errors but answers are still reasonable	• provides a clear but not perfectly organized explanation, using correct terminology, of what was done and why, but explanation may be slightly incomplete • may miss necessary assumptions or caveats • uses diagrams or plots when appropriate to aid in describing the solution • states a conclusion that follows from the analysis but may be somewhat incomplete
2 Developing	• shows some understanding of the problem's statistical components • shows little in the way of a relationship among these components • uses some appropriate statistical techniques but misses or misuses others • may have arithmetic errors that result in unreasonable answers	• provides some explanation of what was done, but explanation may be vague and difficult to interpret and terminology may be somewhat inappropriate • uses diagrams in an incomplete or ineffective way, or diagrams may be missing • states a conclusion that is incomplete
1 Minimal	• shows limited understanding of the problem's statistical components by failing to identify important components • shows little ability to organize a solution and may use irrelevant information • misuses or fails to use appropriate statistical techniques • has arithmetic errors that result in unreasonable answers	• provides minimal or unclear explanation of what was done or why it was done, and explanation may not match the presented solution • fails to use diagrams or plots, or uses them incorrectly • states an incorrect conclusion or fails to state a conclusion
0	• shows little to no understanding of statistical components	• provides no explanation of a legitimate strategy

- **The AP Statistics Exam will occur at noon on TBD.**

QUESTIONS

1) Name the four main content areas in AP Statistics and next to each, place the percentage of the exam that corresponds to that area

a)

b)

c)

d)

2) The AP Statistics exam is ___ hours long, in total.

3) 50% of your score on the exam comes from the _____ section and the remaining 50% comes from the _____ section.

4) There are ____ multiple choice questions and you have ___ minutes to complete them.

5) There are ___ standard free response questions followed by a larger free response question, called the _____.

6) True or False: Formulas are provided for you on the AP exam. *TRUE* *FALSE*

7) Free responses questions are scored out of __ points.

8) True or False: It is possible to receive full credit on a free response question if a minor mathematical error is present in your response. *TRUE* *FALSE*

9) True or False: There is a ¼ point penalty for each incorrect multiple-choice answer given. *TRUE* *FALSE*

10) True or False: AP Statistics is a calculus-based course. *TRUE* *FALSE*

11) True or False: Students are expected to know how to read and interpret computer output on the AP Statistics exam. *TRUE* *FALSE*

Part II: Autobiographical Essay

Content. The first paragraph should be a brief autobiographical statement about yourself that will help me get to know you. What are your likes and dislikes? Who are your role models? What makes you unique? In the second paragraph, tell me something you remember from a previous math course. It can be a particularly good experience or a particularly bad experience. Or describe a math teacher you've had who stood out in some way. In the third paragraph, tell me about your future plans. If you are planning on college and thinking about a particular major, tell me about that and why you are interested in that field. Alternatively, if you aren't planning on college, tell me about possible career interests in general terms. In the fourth and last paragraph, tell me something about this statistics course. Possible areas to address: Why are you taking AP Statistics? Is statistics an important tool in the field of study you are considering? Have you had any previous connection with statistics? Can you give an example from the "real world" where statistics is or has been important to you personally?

Requirements. Here are some specifics regarding your final paper:

1. Change the font to Times New Roman (if this is not available on your computer, use a simple newspaper-type font like the one you are reading). Make your font size 12 points.
2. Make the first line "Brief Autobiographical Sketch," and make the second line (the title) your name. Center both lines, and change your name to font size 18.
3. Insert the following paragraph headings, in order: Biographical, Retrospective, Future Plans, and AP Statistics, like the paragraph headings I have used on these directions, and then put these words in **bold**.
4. Leave one line of space between the title and the first paragraph, and leave one line between paragraphs.
5. Put something (or some things) in *italics*, and <u>underline</u> something else.
6. Number your page at the bottom of the page, centered.
7. Before you print the final copy, make sure you proofread your paper on the screen, and then check it for spelling and syntax.

Grading. I will evaluate you on: following instructions, spelling, grammar and syntax, general interest, and whether your report was submitted on time (see next page for grading rubric, which you are to attach to your essay).

Autobiographical Essay: Rubric Name: _____

4	Excellent
3	Good
2	Satisfactory
1	Unsatisfactory

General

Followed instructions for the investigation:

1	2	3	4	Format of title is as specified
1	2	3	4	Single-space within paragraphs; double-space between paragraphs.
1	2	3	4	Length is correct (one complete page-no more, no less).
1	2	3	4	Report is grammatically correct and free of all spelling errors.
1	2	3	4	Report avoids slang and clichés.
1	2	3	4	Report is clear and unambiguous.
1	2	3	4	Paragraph 1 Biographical: All content questions answered
1	2	3	4	Paragraph 2 Retrospective: All content questions answered
1	2	3	4	Paragraph 3 Future Plans: All content questions answered
1	2	3	4	Paragraph 4 AP Statistics: All content questions answered

Specifics

1	2	3	4	ITALICS/UNDERLINE/NUMBER/NAME

TOTAL: _____ /44