

Part I: Exploring Data

Exploring Data accounts for 20%-30% of the material covered on the AP Exam. “*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.*”

Check your understanding of the following topics:

I. EXPLORING DATA: DESCRIBE PATTERNS AND DEPARTURES FROM PATTERN

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

Part II: Producing Data

Producing Data accounts for 10%-15% of the material covered on the AP Exam. “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.”

Check your understanding of the following topics:

II. SAMPLING AND EXPERIMENTATION: PLANNING AND CONDUCTING A STUDY

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**

Part III: Anticipating Patterns

Probability and the study of randomness accounts for 20%-30% of the material covered on the AP Exam. "Probability is the tool used for anticipating what the distribution of data should look like under a given model."

Check your understanding of the following topics

III. ANTICIPATING PATTERNS: EXPLORING RANDOM PHENOMENA USING PROBABILITY AND SIMULATION

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

Part IV: Statistical Inference

Statistical Inference accounts for 30%-40% of the material covered on the AP Exam.

"Statistical Inference guides the selection of appropriate models."

Check your understanding of the following topics:

IV. STATISTICAL INFERENCE: ESTIMATING POPULATION PARAMETERS AND TESTING HYPOTHESES

A. Estimation (point estimators and confidence intervals)

- 1. Estimating population parameters and margins of error
- 2. Properties of point estimators, including unbiased-ness 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