WGU C207

Data-Driven Decision Making

OA TASK 1 — COMPLETE STUDY GUIDE

Topic Breakdowns • Worked Practice Problems • Formula Sheet

How WGU Tests This • Pre-Assessment Strategy

FOR REFERENCE USE ONLY — Educational resource to help you understand C207 concepts.
Need personalized help? WhatsApp: +1 564-544-6924 | gradevia.com

6	12+	40+	5
Core Topics	Worked Problems	Formulas	Exam Strategies

TABLE OF CONTENTS

SECTION 1 — Pre-Assessment Strategy: Read This First

WGU's OA model gives you a massive advantage — the Pre-Assessment (PA) tells you exactly which competencies you're weak on before the real exam. Smart students use it as a roadmap.

5-Step PA-to-OA Strategy

Step	Action	Why It Matters
1	Take the full Pre-Assessment cold (no studying first)	Reveals your true baseline, not your studied knowledge
2	Screenshot or note every question you got wrong or guessed	These are your high-ROI study targets
3	Map each wrong answer to a competency (A–F below)	C207 tests 6 distinct competencies; wrong answers cluster
4	Study ONLY the relevant sections of this guide for weak areas	Don't waste time on topics you already know
5	Retake PA — if you pass all competencies, schedule OA same week	Memory fades fast; strike while retention is high

Competency Map: The C207 OA tests six domains. Track your PA wrong answers against these:
A — Descriptive Statistics | B — Probability | C — Regression Analysis
D — Hypothesis Testing | E — Data Visualization | F — Decision Analysis
Tip: Most students are weakest on D (Hypothesis Testing) and C (Regression). Prioritize those first.

OA Format: The C207 OA is a proctored Objective Assessment via ProctorU. You face scenario-based multiple-choice questions presenting a business scenario and asking you to identify the correct analytical approach, interpret results, or choose the best decision given data. You are NOT asked to compute by hand — conceptual understanding and interpretation are what matter.

SECTION 2 — Topic Breakdown: Descriptive Statistics

Descriptive statistics summarize and describe the basic features of a dataset. Expect 10–15% of OA questions to touch this topic.

Core Concepts You Must Know

Mean — arithmetic average; sensitive to outliers

Median — middle value; best for skewed data

Mode — most frequent value; used for categorical data

Range — Max minus Min; simplest spread measure

Variance — average squared deviation from mean

Standard Deviation (SD) — sqrt of variance; same units as data

Skewness — positive = tail right; negative = tail left

Percentiles / Quartiles — divide data into equal parts

IQR — Q3 minus Q1; used to detect outliers

Outlier rule — below Q1 - 1.5*IQR or above Q3 + 1.5*IQR

Population vs Sample — denominator N vs n-1

Coefficient of Variation (CV) — SD / Mean; compares variability

How WGU Tests This — Descriptive Statistics

Question Pattern	What They're Testing	Watch Out For
'The sales data is right-skewed. Which measure of center is MOST appropriate?'	Understanding when median beats mean	Choosing mean because it's familiar — wrong for skewed data
'A dataset has outliers. Which measure of spread is MOST resistant?'	IQR vs standard deviation robustness	SD is NOT resistant to outliers; IQR is
'Sample SD uses n-1 in the denominator because...'	Degrees of freedom / Bessel's correction	Confusing population (N) and sample (n-1) formulas

SECTION 3 — Topic Breakdown: Probability

Probability underpins every statistical test in C207. Master these rules and you unlock both this section and hypothesis testing.

Core Concepts

Classical probability — favorable / total outcomes

Relative frequency — based on observed data

Addition Rule (OR) — P(A or B) = P(A) + P(B) - P(A and B)

Multiplication Rule (AND) — P(A and B) = P(A) x P(B|A)

Complement — P(A') = 1 - P(A)

Independent events — P(A and B) = P(A) x P(B)

Mutually exclusive — P(A and B) = 0

Conditional probability — P(A|B) = P(A and B) / P(B)

Bayes' Theorem — P(A|B) = [P(B|A) x P(A)] / P(B)

Expected Value — E(X) = sum of [x * P(x)]

Normal Distribution Key Facts

- 68% of values fall within 1 SD of the mean
- 95% of values fall within 2 SD of the mean
- 99.7% of values fall within 3 SD of the mean (Empirical Rule)
- Z-score formula: Z = (X - mean) / SD — converts any value to standard units
- Z-score of 1.96 = 97.5th percentile (used in 95% confidence intervals)

How WGU Tests This — Probability

Scenario type: 'A quality control manager knows 3% of products are defective. If two products are selected at random, what is the probability BOTH are defective?' Tests multiplication rule for independent events: 0.03 x 0.03 = 0.0009.

Common trap: Questions seeming to require addition rule that actually need the complement. 'What is the probability that AT LEAST ONE is defective?' — Use complement: 1 - P(neither) = 1 - (0.97 x 0.97) = 0.0591.

SECTION 4 — Topic Breakdown: Regression Analysis

Regression is one of the most heavily tested C207 topics. You must interpret regression output and understand what each coefficient means in a business context.

Simple Linear Regression (SLR)

- **Model:** Y = b0 + b1*X + error (Y = dependent, X = independent variable)
- **b0 (intercept)** — predicted Y when X = 0; often not meaningful on its own
- **b1 (slope)** — for every 1-unit increase in X, Y changes by b1 units
- **R (correlation)** — strength and direction of linear relationship; -1 to +1
- **R-squared** — proportion of variation in Y explained by X (0 to 1)
- **Residuals** — actual Y minus predicted Y; should be random (no pattern)
- **Assumptions:** linearity, independence, homoscedasticity, normality of residuals

Reading Regression Output (MedCore Pharma Example)

Coefficient	Value	Interpretation
Intercept (b0)	12.4	Baseline revenue when advertising spend = $0
Advertising Spend (b1)	3.2	Each $1K increase in ad spend predicts $3.2K more revenue
R-squared	0.78	78% of revenue variation explained by advertising spend
p-value (slope)	0.003	Statistically significant (p less than 0.05); relationship is real
Standard Error	1.1	Average prediction error of $1.1K

How WGU Tests This — Regression

Most common question type: Given regression equation and output, interpret the slope or predict Y for a new X value.

Example: Using the MedCore model (Revenue = 12.4 + 3.2 x AdSpend), what is predicted revenue when advertising spend is $10K? Answer: 12.4 + (3.2 x 10) = $44.4K

R-squared trap: R2 = 0.78 does NOT mean the model is 78% accurate. It means 78% of variance in the outcome is explained by the predictor. WGU will test this distinction.

SECTION 5 — Topic Breakdown: Hypothesis Testing

Hypothesis testing is the most commonly failed C207 competency. You are trying to DISPROVE the null hypothesis, not prove your research hypothesis.

The Hypothesis Testing Framework

Step	What You Do	Example
1. State hypotheses	H0 (null) = no effect; Ha (alternative) = effect exists	H0: mean defect rate = 5%; Ha: defect rate less than 5%
2. Set alpha	Significance level = acceptable false positive rate; typically 0.05	alpha = 0.05 means 5% chance of wrongly rejecting H0
3. Select test	Choose based on data type and number of groups	One sample vs known value — use one-sample t-test
4. Compute p-value	Probability of results this extreme if H0 is true	p = 0.02 means 2% chance of this result if H0 were true
5. Decision	If p less than alpha: reject H0 (statistically significant result)	p = 0.02 less than alpha = 0.05 — reject H0

Which Test to Use — Quick Selector

Situation	Correct Test
Compare one sample mean to a known value	One-sample t-test
Compare means of two independent groups	Independent samples t-test
Compare means before and after (same group)	Paired t-test
Compare means of 3 or more groups	One-way ANOVA
Test relationship between two categorical variables	Chi-square test of independence
Test if observed counts match expected counts	Chi-square goodness of fit
Test linear relationship between two quantitative variables	Simple Linear Regression / correlation

How WGU Tests This — Hypothesis Testing

- WGU will NEVER ask you to compute a t-statistic by hand — they test interpretation
- 'A p-value of 0.03 with alpha = 0.05 means...' — Reject H0; the result IS statistically significant
- Type I Error (false positive) = rejecting H0 when it is actually true; probability = alpha
- Type II Error (false negative) = failing to reject H0 when it is false; probability = beta
- Power of a test = 1 - beta = probability of correctly detecting a real effect
- Increasing sample size decreases both Type I and Type II errors simultaneously

SECTION 6 — Topic Breakdown: Data Visualization

C207 tests your ability to choose the RIGHT chart for a given data scenario and to interpret charts correctly. This is entirely conceptual — no drawing required.

Chart Type	Best Used When	What to Look For
Bar / Column Chart	Comparing categories (discrete data)	Height of bars; gaps between bars are normal
Histogram	Distribution of continuous data; shape and skew	Bars touch (no gaps); look for skew and outliers
Line Chart	Trends over time (time series data)	Slope direction; seasonal patterns; trend lines
Scatter Plot	Relationship between two quantitative variables	Direction (positive / negative) and strength of pattern
Box Plot	Distribution summary; comparing groups; spotting outliers	Median line, IQR box, whiskers, outlier dots
Pie Chart	Parts of a whole (proportions); fewer than 6 categories	Best when one category dominates; avoid for many slices
Heat Map	Patterns in a matrix or geographic data density	Color intensity indicates magnitude
Pareto Chart	Prioritizing — shows bars (frequency) plus cumulative percentage	80/20 rule: 80% of problems from 20% of causes

How WGU Tests This — Data Visualization

- 'A manager wants to show the distribution of customer wait times and identify skewness' — Histogram (not bar chart; wait time is continuous)
- 'Which chart is BEST for identifying the relationship between employee training hours and productivity?' — Scatter plot
- 'The data shows monthly sales over 3 years. Which chart best reveals the seasonal trend?' — Line chart
- Histogram vs Bar Chart is a classic WGU trap: histograms show continuous distributions; bar charts compare discrete categories

SECTION 7 — Topic Breakdown: Decision Analysis

Decision analysis provides structured frameworks for making choices under uncertainty. C207 focuses on expected value, decision trees, and sensitivity analysis.

Decision-Making Under Uncertainty — Criteria

Criterion	Rule	Use When...
Maximax (Optimistic)	Choose the option with the MAXIMUM possible payoff	Decision-maker is optimistic / risk-seeking
Maximin (Pessimistic)	Choose the option with the BEST worst-case payoff	Decision-maker is pessimistic / risk-averse
Minimax Regret	Minimize the maximum regret (opportunity loss)	Decision-maker focuses on avoiding regret
Expected Monetary Value (EMV)	EMV = sum of (payoff x probability) for each outcome	Probabilities are known; repeated decisions
EVPI	EVPI = Best EMV with perfect info minus Best EMV without	Deciding whether to buy a study or forecast

Decision Tree Basics

- **Square nodes** = decision points (you choose the branch)
- **Circle nodes** = chance / probability nodes (nature chooses the outcome)
- **Terminal nodes** = payoffs at the end of each branch
- **Rollback method:** work right to left — multiply payoffs by probabilities, sum at chance nodes, take max at decision nodes
- **Sensitivity analysis:** test how the best decision changes as probabilities or payoffs vary

How WGU Tests This — Decision Analysis

Typical question: MedCore Pharmaceuticals must decide whether to launch a new drug. If favorable (P = 0.6), profit = $500K. If unfavorable (P = 0.4), loss = $200K. What is the EMV of launching?

EMV = (0.6 x $500K) + (0.4 x -$200K) = $300K - $80K = **$220K**. Since EMV greater than 0, launch is the recommended decision.

SECTION 8 — How WGU Tests This: OA Exam Strategies

The C207 OA uses scenario-based multiple-choice questions. Understanding question patterns is as valuable as knowing the content.

Pattern	Example Stem	Strategy
'MOST appropriate' questions	'Which statistical test is MOST appropriate for comparing the mean scores of three training groups?'	Eliminate clearly wrong answers first. Use the test selector in Section 5.
'BEST interpretation' questions	'An R-squared of 0.65 BEST indicates that...'	Know exact definitions cold. These test whether you confuse correlation with causation.
'EXCEPT' or 'NOT' questions	'All of the following are assumptions of linear regression EXCEPT...'	Slow down. Underline NOT or EXCEPT. Verify each option individually.
Reading output tables	'Based on the regression output below, what does the slope coefficient mean?'	Read the table before the question. Know b0 vs b1 cold.
Business decision questions	'Given this EMV analysis, what should MedCore do?'	Work through the math step by step. Do not estimate.

5 OA Day Strategies

1. Read every question twice. The word MOST, LEAST, NOT, or EXCEPT changes the answer entirely.
2. Eliminate first. Most questions have two obviously wrong answers. Get to a 50/50 choice.
3. Trust your first instinct on conceptual questions. Second-guessing on vocabulary kills scores.
4. For calculation questions, write out each step. Do not do arithmetic in your head.
5. Flag and return. If a question takes more than 90 seconds, flag it and move on.

SECTION 9 — Master Formula Sheet and Reference Card

Print or screenshot this page. These are the formulas and definitions you need to have memorized for the C207 OA.

Descriptive Statistics Formulas

Formula	Expression	Notes
Population Mean	`mu = (sum of X) / N`	N = total population size
Sample Mean	`x-bar = (sum of X) / n`	n = sample size
Population Variance	`sigma^2 = sum(X-mu)^2 / N`	Average squared deviation
Sample Variance	`s^2 = sum(X-x-bar)^2 / (n-1)`	n-1 = Bessel's correction
Standard Deviation	`s = sqrt(s^2)`	Same units as original data
Z-Score	`Z = (X - mu) / sigma`	How many SDs from the mean
IQR	`IQR = Q3 - Q1`	Middle 50% of data
Outlier Boundaries	`Q1 - 1.5*IQR and Q3 + 1.5*IQR`	Tukey fence method
Coeff. of Variation	`CV = (s / x-bar) * 100%`	Relative variability measure

Probability Formulas

Formula	Expression	Notes
Addition Rule	`P(A or B) = P(A) + P(B) - P(A and B)`	General; works always
Mutually Exclusive	`P(A or B) = P(A) + P(B)`	P(A and B) = 0
Multiplication Rule	`P(A and B) = P(A) * P(B|A)`	General
Independent Events	`P(A and B) = P(A) * P(B)`	P(B|A) = P(B)
Complement	`P(A') = 1 - P(A)`	Faster for 'at least one'
Conditional Prob.	`P(A|B) = P(A and B) / P(B)`	Read: 'A given B'
Expected Value	`E(X) = sum [ x * P(x) ]`	Weighted average of outcomes
Z-Score	`Z = (X - mu) / sigma`	Use Z-table for probability

Regression and Hypothesis Testing Reference

Concept	Formula / Rule	Remember
Simple Linear Regression	`Y = b0 + b1*X`	b0 = intercept; b1 = slope
R-squared	`R2 = explained var / total var`	0 = no fit; 1 = perfect fit
Predicted value	`Y-hat = b0 + b1*X`	Plug in X; solve for Y

Concept	Formula / Rule	Remember
Residual	`e = Y_actual - Y_predicted`	Should be random; no pattern
Decision rule	`p less than alpha: REJECT H0`	p = 0.03, alpha = 0.05 -- reject
Type I Error rate	`alpha (typically 0.05)`	False positive probability
Power	`Power = 1 - beta`	Detecting a real effect
EMV	`EMV = sum (payoff_i * P_i)`	Pick option with highest EMV
EVPI	`EVPI = best EMV(perfect) - best EMV`	Max worth paying for info

SECTION 10 — Practice Problem Sets with Worked Solutions

All practice problems use MedCore Pharmaceuticals (fictitious) as the business context — the same scenario-based format WGU uses.

Practice Problem 1 — Descriptive Statistics

MedCore's regional sales team recorded the following monthly revenue figures ($K) for Q1: 42, 38, 51, 47, 39, 88, 44, 50, 43, 41. Calculate the mean, median, and identify whether the data is skewed.

Step-by-step solution:

- Sort the data: 38, 39, 41, 42, 43, 44, 47, 50, 51, 88
- Mean = (38+39+41+42+43+44+47+50+51+88) / 10 = 483 / 10 = 48.3
- Median = average of 5th and 6th values = (43+44) / 2 = 43.5
- Mean (48.3) greater than Median (43.5) — distribution is positively (right) skewed
- The outlier (88) pulls the mean upward; median is the better measure of center here

Answer: Mean = $48.3K, Median = $43.5K, Distribution is right-skewed; median is preferred.

Practice Problem 2 — Probability (Multiplication Rule)

A MedCore quality inspector knows 4% of medication capsules fail the seal test. If two capsules are randomly selected from a large batch, what is the probability BOTH capsules fail? Assume independence.

Step-by-step solution:

- P(one capsule fails) = 0.04
- Events are independent (large batch; sampling with replacement assumed)
- Multiplication rule for independent events: P(A and B) = P(A) * P(B)
- P(both fail) = 0.04 * 0.04 = 0.0016

Answer: P(both fail) = 0.0016, or 0.16% — a very low probability.

Practice Problem 3 — Regression Interpretation

MedCore's analytics team ran a regression of monthly sales revenue (Y, in $K) on the number of sales calls made (X). Output shows: Intercept = 15.2, Slope = 2.8, R-squared = 0.71, p-value for slope = 0.001. Interpret the slope and R-squared. Predict revenue for 30 sales calls.

Step-by-step solution:

- Slope (b1 = 2.8): Each additional sales call predicts $2.8K more revenue
- Intercept (b0 = 15.2): Predicted revenue when 0 sales calls are made is $15.2K
- R-squared = 0.71: 71% of variation in monthly revenue is explained by sales calls
- p-value = 0.001 less than 0.05: Slope is statistically significant; relationship is real
- Prediction: Y = 15.2 + (2.8 * 30) = 15.2 + 84 = $99.2K

Answer: Slope = $2.8K per call; R2 = 71% of variance explained; Predicted revenue = $99.2K for 30 calls.

Practice Problem 4 — Hypothesis Testing (p-value decision)

MedCore's operations team claims the average time to process an insurance claim is 14 days. A quality analyst tests 40 claims and finds a mean of 12.3 days. A one-sample t-test yields a p-value of 0.032. At alpha = 0.05, what is the conclusion?

Step-by-step solution:

- H0: mean processing time = 14 days (no change)
- Ha: mean processing time is NOT 14 days (two-tailed test)
- Alpha = 0.05 (significance level)
- p-value = 0.032
- Decision rule: If p less than alpha, reject H0
- 0.032 less than 0.05 — REJECT H0
- Conclusion: Statistically significant evidence that mean processing time differs from 14 days

Answer: Reject H0 (p = 0.032 < alpha = 0.05). Mean processing time is significantly different from 14 days.

Practice Problem 5 — Decision Analysis (EMV)

MedCore must decide whether to expand to a new region. If demand is high (P = 0.55), profit = $800K. If demand is low (P = 0.45), loss = $300K. Alternatively, they can license the product for a guaranteed $150K. Which option maximizes EMV?

Step-by-step solution:

- Option 1 (Expand): EMV = (0.55 x $800K) + (0.45 x -$300K)
- = $440K + (-$135K) = $305K
- Option 2 (License): EMV = $150K (guaranteed)
- Compare: $305K (expand) vs $150K (license)
- Decision: Expand has the higher EMV

Answer: Expand (EMV = $305K) vs License ($150K). EXPAND maximizes expected monetary value.

Practice Problem 6 — Chart Selection

MedCore's HR director wants to compare the distribution of employee satisfaction scores (continuous scale 1–100) across three departments: Sales, Operations, and R&D.; Which chart type is MOST appropriate?

Step-by-step solution:

- The data is continuous (1–100 scale), not categorical counts
- The goal is to COMPARE distributions across three groups
- Box plots display median, IQR, range, and outliers for each group in one view
- Histograms show one distribution at a time — less efficient for multi-group comparison
- Bar charts are for categorical data — wrong here

Answer: Side-by-side Box Plots — best for comparing distributions across multiple groups.

Need personalized help with WGU C207?

WhatsApp Us Now +1 564-544-6924

FOR REFERENCE USE ONLY — Educational supplement only. gradevia.com | WhatsApp: +1 564-544-6924