AP Statistics – Unit 2, Lesson 2

VIDEO FOLLOW-ALONG WORKSHEET

Two Categorical Variables: Two-Way Tables, Relative Frequencies & Association

Name: _____ **Period:** _____ **Date:** _____

How to Use This Worksheet

- Pause video at each **timestamp** to fill in blanks and work examples
- When you see "Of **ALL**..." → use **GRAND TOTAL**
- When you see "Of those in [row]..." → use **ROW TOTAL**
- When you see "Of [column] group..." → use **COLUMN TOTAL**

DENOMINATOR MAP (Keep this visible!)

- **JOINT** relative frequency (of ALL): cell ÷ **GRAND TOTAL**
  → "What % of ALL cases are in this cell?"
- **MARGINAL** relative frequency (row/column share of ALL):
  row total ÷ **GRAND TOTAL** OR column total ÷ **GRAND TOTAL**
- **CONDITIONAL** relative frequency (within a group):
  cell ÷ that **ROW's total** OR cell ÷ that **COLUMN's total**

Learning Targets

☐ I can compute **joint**, **marginal**, and **conditional** relative frequencies

☐ I can use conditional RFs and graphs to decide if variables are **associated**

Rounding: *Report % to 1 decimal; each conditional column should sum to 100% (±0.1%).*
Category Order: *No HS → HS diploma → Some college/AA → Bachelor's+*

[00:11-00:28] Learning Objectives

1. How do we calculate _____ for two categorical variables?
2. How do we use these to determine if there is an _____?

[00:42] Original Two-Way Table

Context: Age Groups (columns) vs. Educational Attainment (rows)

Table 1: Original Data from Video - Copy the values from the video

Education \ Age	25-34	35-54	55+	Row Total
No HS	_____	_____	_____	_____
HS diploma	_____	_____	_____	_____
Some college/AA	_____	_____	_____	_____
Bachelor's+	_____	_____	_____	_____
Column Total	_____	_____	_____	**Grand: _____**

[01:15-01:31] STEP 1: Find Your Totals (CRITICAL!)

ALWAYS compute totals BEFORE calculating any percentages!

- Column totals: _____ _____ _____
- Row totals: _____ _____ _____ _____
- **Grand total:** _____

[01:39-02:27] Joint Relative Frequency

Question: "What percent of **ALL people** are 25-34 with Bachelor's+?"

$$\textbf{Joint RF} = \frac{\text{single cell}}{\text{grand total}} = \frac{\_\_\_\_\_}{\_\_\_\_\_} \times 100 = \_\_\_\_\_\%$$

Interpretation Frame: "_____% of ALL people are [age group] AND [education category]."

JOINT Relative Frequency = A cell frequency divided by the grand total
Answers: "What % of ALL cases have BOTH characteristics?"

[02:50 & 03:32] Marginal Relative Frequency

Example 1 (Row): "What percent of **ALL** have only a HS diploma?"

$$\text{Marginal RF} = \frac{\text{HS row total}}{\text{grand total}} = \frac{\_\_\_\_\_}{\_\_\_\_\_} \times 100 = \_\_\_\_\_\%$$

Example 2 (Column): "What percent of **ALL** are 35-54 years old?"

$$\text{Marginal RF} = \frac{\text{35-54 column total}}{\text{grand total}} = \frac{\_\_\_\_\_}{\_\_\_\_\_} \times 100 = \_\_\_\_\_\%$$

Interpretation Frames:

- Row marginal: "_____% of ALL people have [education category]."
- Column marginal: "_____% of ALL people are [age group]."

MARGINAL Relative Frequency = Row/column total divided by grand total
Answers: "What % of ALL cases are in this row/column?"

[04:33 & 05:13] Conditional Relative Frequency

KEY PHRASES TO WATCH FOR:
"**Of those with...**" or "**Among people with...**" = CONDITIONAL!
Look at ONLY that row or column!

Common Pitfalls

- Using GRAND TOTAL when the prompt says "of those in this row/column..." (should use that row/column total)
- Reversing the condition (conditioning on a row when the question conditions on a column)
- Judging association by comparing COUNTS instead of CONDITIONAL PERCENTS

Example 1: "**Of those with** only a HS diploma, what percent are 35-54?"

$$\text{Conditional RF} = \frac{\text{cell in HS row}}{\text{HS row total}} = \frac{\quad\quad}{\quad\quad} \times 100 = \_\_\_\_\_\%$$

Example 2: "What percent **of 25-34 year olds** have no HS diploma?"

$$\text{Conditional RF} = \frac{\text{cell in 25-34 column}}{\text{25-34 column total}} = \frac{\quad\quad}{\quad\quad} \times 100 = \_\_\_\_\_\%$$

Interpretation Frames:

- Row-conditional: "Among [education category], _____% are [age group]."
- Column-conditional: "Among [age group], _____% have [education category]."

CONDITIONAL Relative Frequency = Cell within a specific row/column ÷ that row/column's total
Answers: "Within THIS group, what % have that characteristic?"

[06:20-06:40] Calculate ALL Conditional RFs

Education distribution WITHIN each age group (columns must = 100%!):

Table 2: Conditional Relative Frequencies (Education within Age Groups)

Education Level	25-34 (%)	35-54 (%)	55+ (%)
No HS	_____	_____	_____
HS diploma	_____	_____	_____
Some college/AA	_____	_____	_____
Bachelor's+	_____	_____	_____
TOTAL	**100.0**	**100.0**	**100.0**

[07:04-07:11] Segmented Bar Graph

Instructions: After calculating your conditional RFs, sketch the segmented bars below. Each bar should reach 100%, with segments proportional to your calculated percentages.

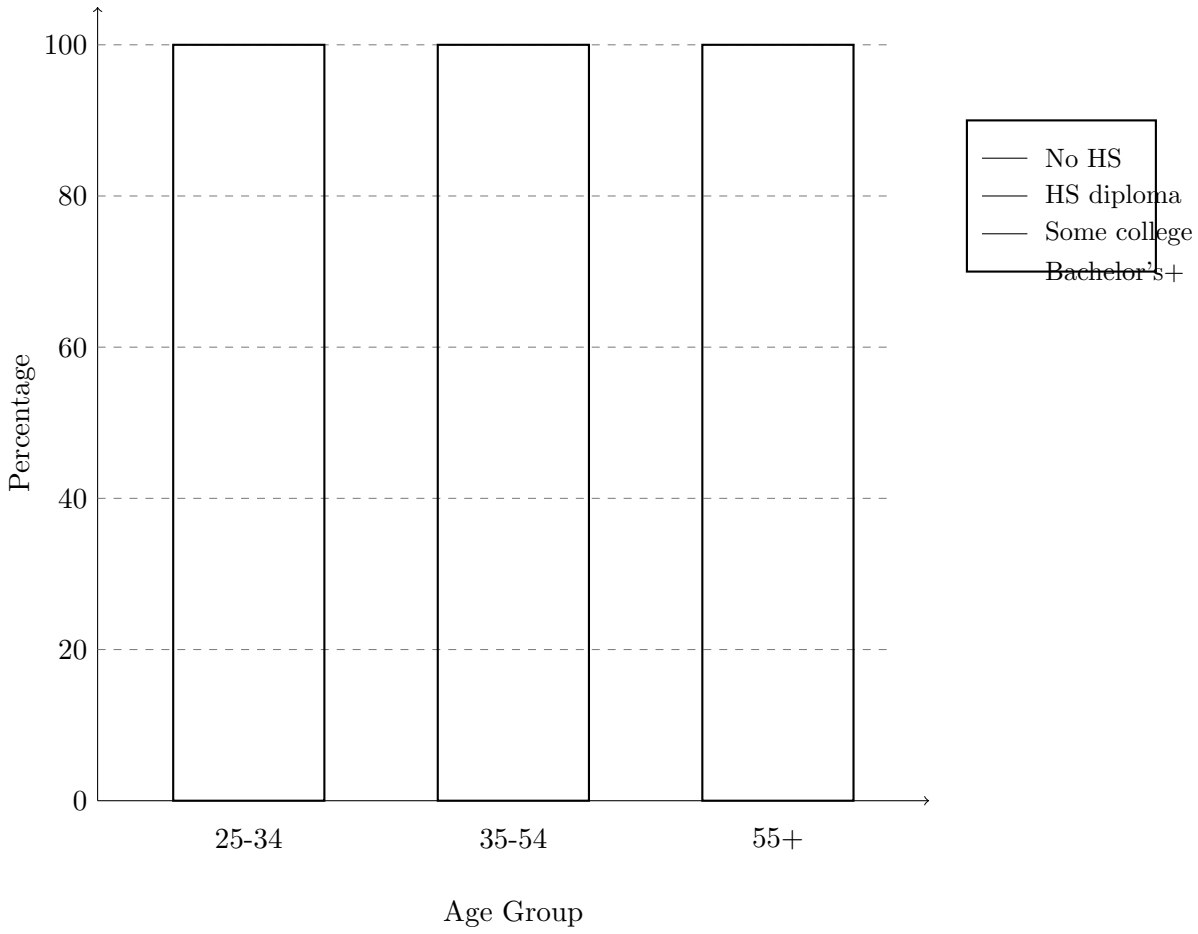


Figure 1: Segmented Bar Graph: Education Distribution by Age Group

Visual Check: Do the bars have the SAME shape or DIFFERENT shapes? _____

[07:29-08:03] THE BIG QUESTION: Association?

NOT Associated

Bars look the SAME (Similar proportions across groups)

Associated

Bars look DIFFERENT (Different proportions across groups)

Association Sentence Frame

"Because the distribution of conditional relative frequencies (education within age) appears [ ] **the SAME** [ ] **DIFFERENT** across age groups, age and education are [ ] **NOT associated** [ ] **ASSOCIATED**."

FINAL CONCLUSION

Write your complete conclusion using the frame above:

_ _

_ _

_ _

Note: *Association $\neq$ causation (today's work is descriptive).*

Visual Analysis: Side-by-Side Comparison

Instructions: Create a side-by-side bar graph using your conditional RF values.

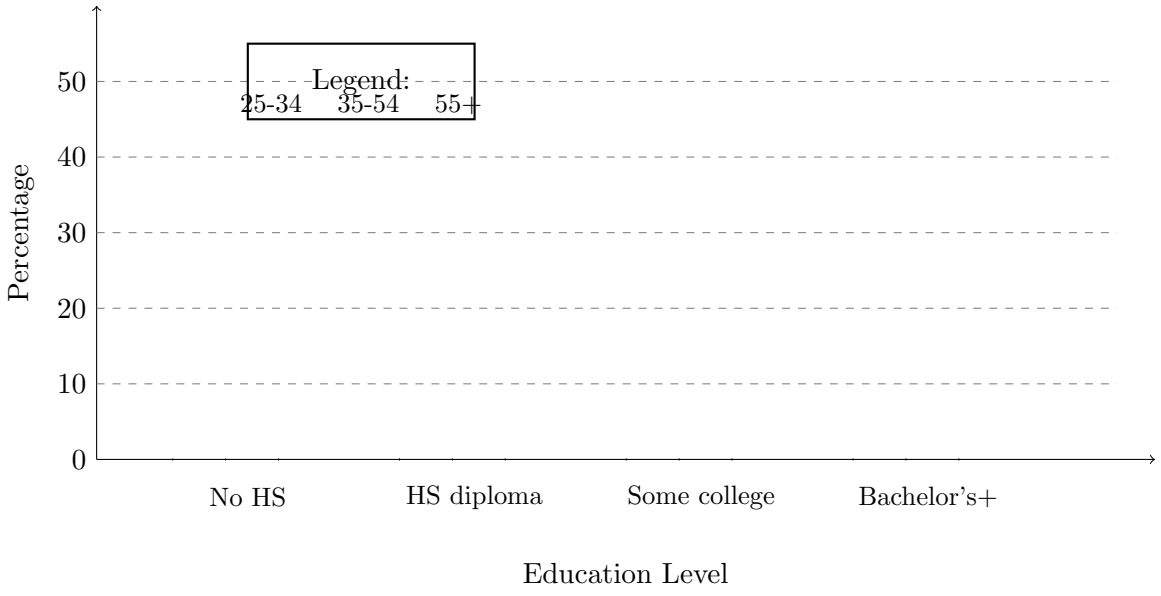


Figure 2: Side-by-Side Bar Graph: Draw bars based on your calculated values

When to Use Which Graph

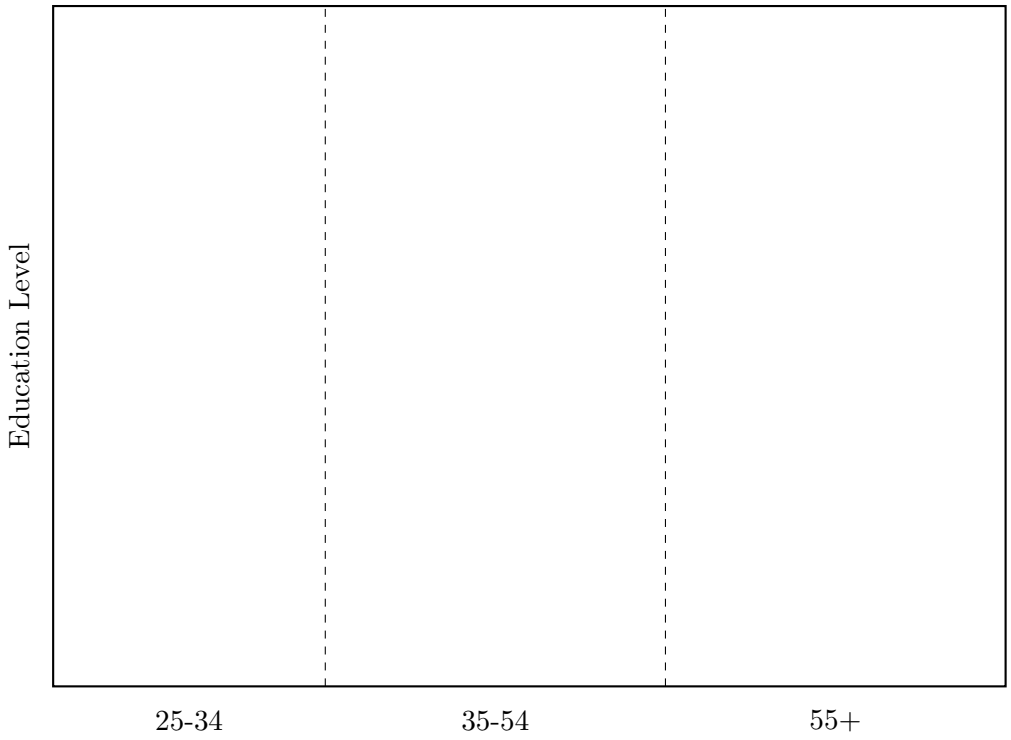
- **Side-by-side:** compare categories across groups (counts or percents)
- **Segmented:** compare distributions (each bar = 100%)
- **Mosaic:** adds group size context (tile area = joint RF)

Mosaic Plot Representation

Mosaic Plot Cheat Sheet

- Bar WIDTH $\propto$ group size (column total)
- Segment HEIGHT = conditional percent (within the bar)
- TILE AREA = joint relative frequency

Instructions: Sketch a mosaic plot below. Make bar widths proportional to column totals.



Adjust widths based on column totals, divide heights based on conditional %s

Figure 3: Mosaic Plot: Sketch your mosaic plot here

Quick Check: Denominator Detective

Question	Denominator	Answer
Of **ALL** people, what % are 35-54?	____________	_______%
Of those with **ONLY HS**, what % are 35-54?	____________	_______%
Among **25-34 year-olds**, what % have no HS?	____________	_______%

Exit Ticket (2 minutes)

1. Give one example of a **JOINT** relative frequency from today:

   _

   _

2. Give one example of a **CONDITIONAL** relative frequency (state which row or column you conditioned on):

   _

   _

3. Are age and education associated? Why?

   _

   _

   _

Teacher Notes:

- Mosaic plots preserve group sizes (width $\propto$ group size)
- Tile area = joint RF; segment height = conditional RF
- Always use conditional RFs to determine association!
- Check that each conditional column sums to 100% ($\pm$ 0.1%)

Quick Reference Summary

Type of RF	Numerator	Denominator	Answers the question...
Joint	Single cell	Grand total	What % of ALL cases are in this cell?
Marginal	Row/Column total	Grand total	What % of ALL cases are in this row/column?
Conditional	Cell in row/column	That row/column total	Within this group, what % have that characteristic?