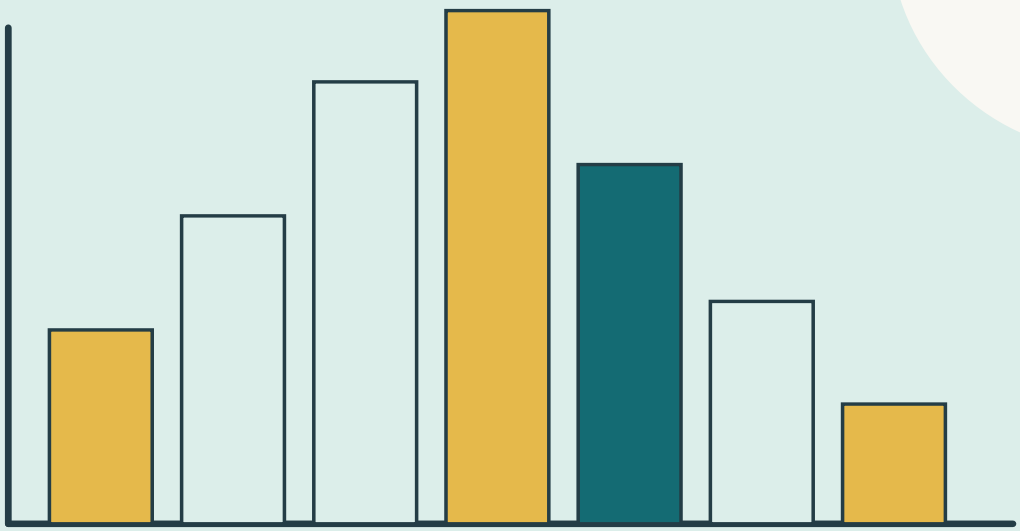
HISTOGRAM INTERPRETATION

Building and Reading Histograms



BINS / FREQUENCY

AI & DATA

Building and Reading Histograms

Learning outcomes

- Construct two histograms of the same continuous measurement using different bin widths.
- Assign boundary values exactly once and reconcile frequency totals.
- Explain how bin width changes appearance without changing observations.
- Translate a measurement definition into an auditable histogram with explicit interval rules.
- Reconcile counts, exclusions, repeated values and boundary observations before plotting.
- Compare width and origin choices while separating stable facts from design-sensitive appearance.
- Distinguish counts, relative frequency and density, including the area rule for unequal widths.
- Describe center, spread, tails and possible irregularities without inferring unsupported populations or causes.
- Produce an accessible interpretation card and reproducible frequency calculation from synthetic raw records.

Original teaching material. All organizations, people, records and operational rules in exercises are synthetic. Use only the supplied non-sensitive inputs. This is a practice course, not an accredited or regulated certification.

Study the foundations, work through the applied units, complete the practice tasks, and then attempt the assessment. Use only the supplied fictional practice data. Keep your original inputs and your revised work separate.

Core concepts

A histogram groups measurements rather than categories

A histogram summarizes a numerical distribution by grouping values into intervals and counting observations in each interval. The horizontal axis shows measurement values, while the vertical axis in this course shows counts. Unlike a category bar chart, neighboring histogram bars represent adjoining numerical ranges, so leave no artificial gaps between them. The supplied variable is request duration in minutes, a continuous measurement recorded to whole minutes. Keep the measurement unit and sample size in the title or subtitle. A chart of department names would be a category bar chart rather than the histogram practiced here.

Specify boundaries before counting

A bin needs both endpoints and a rule about inclusion. Use left-inclusive, right-exclusive intervals: a value equal to the lower boundary belongs in the bin, and a value equal to the upper boundary belongs in the next bin. Thus 10 belongs to [10, 15), not [5, 10). The brackets are shorthand for including the left endpoint and excluding the right. This convention prevents double counting. The supplied final boundaries exceed the largest observation, so no special final-bin exception is needed. Every observation must have exactly one destination, including repeated values and values exactly on a boundary.

Make a frequency table that can be audited

Create one row for every bin, including empty bins. Record lower boundary, upper boundary and count, then check that the counts sum to the number of raw observations. Reconcile this total before drawing the chart. A repeated measurement is a separate observation and contributes another count; do not deduplicate equal durations. For a spot check, list the observations assigned to a boundary-adjacent bin. A histogram drawn from a mistaken frequency table can look perfectly reasonable, so visual polish does not substitute for numeric reconciliation. Preserve the raw data beside the bin definitions for reproducibility.

Compare narrow and wide bins carefully

Narrow bins show finer detail but can make small samples look irregular. Wider bins combine neighboring counts and may conceal local differences. Neither choice changes the underlying observations or the sample size. In this exercise, make one chart with five-minute intervals and another with ten-minute intervals. Because widths are equal within each chart, bar heights may directly show frequency. Do not compare a taller bar in the wide-bin chart with a narrower bar as evidence of more observations overall: the wider bin simply covers more possible values. Unequal-width bins require more careful area or density treatment and are outside this exercise.

Write observations that survive reasonable design changes

Describe where observations concentrate, how far they extend and whether a tail seems longer, then check whether the statement remains sensible under both widths. A small sample can suggest a shape without proving a particular probability distribution or identifying a process mechanism. Avoid declaring two distinct populations based only on a bump that vanishes when bins change. Use actual frequency counts to support your observations. The capstone interpretation card should explicitly separate stable facts from width-sensitive appearances. Include a plain-text description and the frequency tables so someone who cannot see the chart can still understand the information.

Worked example

A separate dataset is 0, 4, 5, 9, 10 and 14 minutes. With bins [0, 5), [5, 10) and [10, 15), the counts are 2, 2 and 2. The boundary value 5 belongs in the second bin and 10 in the third. The frequencies sum to 6, matching the six records. If the first two bins are combined into [0, 10), their counts add to 4; no new observation has appeared. This reconciliation is the first check before rendering bars.

Foundation practice

Durations minutes

- 2
- 3
- 4
- 5
- 6
- 7
- 7
- 8
- 8
- 9
- 9
- 10
- 10
- 11
- 11
- 12
- 13
- 14
- 15
- 17
- 18
- 19
- 21
- 24

Sample size

24

Narrow bin edges

- 0
- 5
- 10
- 15
- 20
- 25

Wide bin edges

- 0
- 10
- 20
- 30

Boundary rule: Include lower boundary; exclude upper boundary. Every supplied value is below its chart final boundary.

Chart specification

X: Duration in minutes

Y: Number of requests

Y minimum

0

Bars: Contiguous, equal width within each chart

Required outputs: Two frequency tables, two histogram charts, one interpretation card

Foundation tasks

Guided preparation

Instructions

Read the complete input pack. Create the output columns specified in the capstone, keep the original IDs, and state the unit of work. Complete the first two records manually before using an optional AI assistant. Record any assumption that could change the result.

Independent capstone

Instructions

Assign all 24 observations to both bin systems. Produce the two frequency tables and two labeled histograms with a zero frequency baseline. Verify both total frequencies equal 24 and explain the placement of 5, 10, 15 and 20, even though 20 itself is not observed. Write three observations supported by counts and one caution about interpreting the apparent shape. Identify which bars merge when moving from five-minute to ten-minute intervals.

Peer verification and revision

Instructions

Exchange only your student work with a peer. Ask the peer to trace two ordinary results and every exception back to the supplied input. Record the exact item corrected, the original value, revised value and reason. Submit the revised artifact plus a short explanation of the most important remaining limitation.

01 Know what one observation represents

A histogram summarizes the distribution of one numerical variable. Each observation contributes to exactly one interval along a measurement axis. In this course, one observation is one completed synthetic service request, and the variable is elapsed processing duration recorded in whole minutes. A histogram can show how those durations are distributed, but it does not preserve the order in which requests arrived or identify which staff member handled them. Those are different questions requiring different records or displays.

The expansion case is the fictional Willow desk, with 36 completed requests. Its file assets/willow_durations.csv contains stable request IDs W01 through W36 and duration_minutes. Values range from zero to 44 minutes. A recorded zero is valid in this case: the recording convention rounds elapsed time to the nearest whole minute, so a very short positive duration can be recorded as zero. There are no missing values in the main Willow fixture. These conditions are part of the exercise definition, not facts to assume about all operational logs.

Before plotting, write a measurement sentence: 'Elapsed processing duration, in minutes rounded to whole minutes, for the 36 listed completed Willow requests.' This sentence identifies the unit, variable, precision and included population. It does not say that these are all arrivals, that uncompleted requests took zero minutes or that the sample represents every future request. An analysis of completed cases can omit requests still in progress; that distinction matters if the real question is how long all arrivals will eventually take.

A category bar chart answers a different question. Counts by department, request type or weekday use named categories rather than adjoining numerical intervals. A bar labeled 'support' is not the interval between two numerical endpoints. Similarly, plotting one bar per request ID would display individual values, not a frequency distribution. Identifying the intended statistic first helps prevent attractive but conceptually wrong charts. A histogram is appropriate here because the question concerns concentration and spread across ranges of a numerical measurement.

Field	Meaning	Local condition
request_id	One completed request	Unique W01–W36
duration_minutes	Elapsed processing duration	Whole-minute recorded measurement
Zero	Valid rounded short duration	Not missing

Field	Meaning	Local condition
Population	Listed completed requests	No claim about uncompleted or future requests

02 Inspect the raw values without destroying multiplicity

The sorted Willow values are 0, 1, 2, 3, 4, 5, 5, 6, 7, 8, 9, 9, 10, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 34, 39 and 44. Sorting this one-variable list helps inspect coverage, but preserve the request IDs in the source file. Equal measurements are not necessarily duplicated records. W06 and W07 each recorded five minutes and represent two separate requests, so both must contribute to a frequency count.

Differentiate a duplicate measurement from a duplicate record. A repeated request ID could indicate accidental copying, but a repeated duration is expected when measurements are rounded. Deduplicating the duration column would change the population from requests to distinct duration values. That would answer 'how many different durations occurred?' rather than 'how many requests had durations in this interval?' Neither a chart title nor a later percentage can repair that change of unit.

Inspect values for type and range before binning. A blank string is not zero. The text '12 min' is not automatically a numeric 12 unless an explicit, reliable conversion rule is applied. A negative elapsed duration may signal reversed timestamps, but do not silently take its absolute value. Preserve the raw record, flag the inconsistency and resolve it under a stated rule. In the main Willow case, all 36 supplied values are valid; an exercise later introduces deliberate exceptions so you can practice a reconciliation ledger.

Record the observation count, minimum and maximum before choosing bins. Willow has n = 36, minimum 0 and maximum 44. These values determine whether the proposed interval range covers all observations. A chart that stops at 40 would omit the 44-minute request unless an overflow policy is explicit. Clipping the horizontal axis can hide information even when the visible bars are drawn correctly. The count ledger, not the chart's visual neatness, is the first line of defense against that error.

E02-X1 Distinguish repeats from duplicate records

The following source rows are proposed: K01,5; K02,5; K02,5; K03,blank; K04,−2; K05,12. For this exercise only, a repeated identical ID/value row is a confirmed accidental duplicate; blank and negative durations remain unresolved and must be excluded from numeric analysis with reasons. Produce the included rows and a reconciliation ledger.

Submit: Included ID/value rows, excluded rows with reasons and counts.

03 Specify intervals with an executable boundary rule

A bin definition requires both endpoints and an inclusion convention. Throughout the manual exercises, [a,b) means a ≤ value < b: include the lower endpoint and exclude the upper endpoint. A five-minute scheme starting at zero uses [0,5), [5,10), [10,15) and so on. A value of five goes into the second interval, not the first and not both. A value of ten goes into the third. The notation makes adjacent intervals exhaustive over their covered range without overlapping.

Endpoints should be actual numerical boundaries, not ambiguous labels such as '0–5' and '5–10.' Those labels do not tell the reader which interval owns five. A table with lower, upper and count columns is often clearer than a long footnote. When the variable is rounded to whole minutes, a bin [0,5) contains the recorded integers 0, 1, 2, 3 and 4. It does not necessarily mean that every underlying unrounded duration was less than five, because rounding occurred before these recorded values were supplied.

The final interval deserves explicit attention. The Willow five-minute edges run from zero through 45, so the last bin is [40,45) and the maximum recorded value 44 is covered. No special final-bin exception is needed. If a future observation equaled 45, the strict convention would classify it as overflow unless another interval were added. Do not quietly change the convention only for that value while leaving the specification unchanged.

Software can implement a different last-edge rule. NumPy's histogram function, for example, documents half-open bins except for a last bin that includes its rightmost edge. That behavior is legitimate, but it differs from the uniform strict rule used in the manual exercise. Learn the tool's semantics rather than assuming the function name guarantees your chosen contract. The supplied student counting function implements every interval as left-inclusive and right-exclusive, reports overflow and tests exact edges explicitly.

Recorded value	Five-minute destination	Why
0	[0,5)	Lower boundary included
4	[0,5)	Below upper boundary
5	[5,10)	Previous upper boundary excluded
10	[10,15)	Starts the next interval
44	[40,45)	Below final upper boundary
45	Overflow for edges 0…45	Final upper boundary is excluded here

04 Construct and reconcile a frequency table

Begin with the bins, then assign each request. For Willow's five-minute bins, the counts are 5, 7, 6, 6, 5, 3, 2, 1 and 1. Their sum is 36, matching the number of valid raw observations. Keep empty bins in the table if the selected scheme produces any. A zero-count interval is information about the covered numerical range; removing it from the display can falsely bring separated clusters together.

Audit at least one interior boundary. The interval [5,10) contains 5, 5, 6, 7, 8, 9 and 9, for seven requests. The two tens belong in [10,15), which also contains 11, 12, 13 and 14, giving six. A mistaken rule that includes both endpoints would count the tens in two bins and could inflate the total. A mistaken rule that excludes both endpoints could lose boundary values entirely. Reconciliation catches the total error, while a membership check reveals its source.

A matching total is necessary but not sufficient. One request could be assigned to the wrong bin while another is omitted and another duplicated, leaving the same grand total. Use three checks together: every included record has one destination, no record has two destinations, and bin counts sum to the included count. Stable IDs make the first two checks practical. For a small teaching dataset, listing the members of a boundary-adjacent bin is an efficient spot check.

Distinguish a frequency table from a table of summary values. The row [20,25) count 5 says that five requests have recorded durations at least 20 and below 25. It does not say their average duration is 22.5, nor that five requests lasted exactly 20. The midpoint is a display or approximation device, not a recovered observation. Retaining the raw file allows exact statistics to be computed later without pretending that a grouped table preserves every value.

Interval (minutes)	Count	Cumulative count below upper edge
[0,5)	5	5
[5,10)	7	12
[10,15)	6	18
[15,20)	6	24
[20,25)	5	29
[25,30)	3	32
[30,35)	2	34
[35,40)	1	35
[40,45)	1	36

E04-X1 Audit a proposed Willow table

A proposed five-minute table has counts 5, 9, 4, 6, 5, 3, 2, 1, 1 for edges 0,5,…,45. Its total is 36. Compare it with the raw values and identify the specific boundary error consistent with these counts. Supply the corrected two affected rows.

Submit: An error explanation with exact memberships and corrected counts.

05 Draw a count histogram from the table

The five-minute Willow histogram uses duration in minutes on the horizontal axis and number of requests on the vertical axis. Each rectangle starts at its lower boundary, ends at its upper boundary and has height equal to its count. Adjacent intervals meet because they cover adjoining ranges of the same numerical scale. The vertical axis starts at zero, so a bar of height six is twice as tall as a bar of height three. The bin width is five minutes throughout this chart.

Use the numerical axis rather than equally spaced category names that happen to look like intervals. With equal widths the difference may be invisible, but the underlying representation matters when origins, gaps or widths change. A bar centered at 7.5 with width five represents [5,10). A bar centered at five with width five would instead span [2.5,7.5), a different range. Check the edges in the plot, not only the text labels beneath it.

The tallest Willow five-minute bar has count seven in [5,10). That is the modal interval for this binning, not a claim that every typical request takes 7.5 minutes. The raw data have several repeated integer values, and their individual modes need not equal a bin midpoint. The chart shows a concentration below 25 minutes and a sparse extension to 44. Exact statements about thresholds should be calculated from the raw values or aligned frequency boundaries, rather than estimated by eye.

Keep the chart visually restrained. Clear boundaries, readable axis labels, integer count ticks and a visible sample size matter more than gradients or three-dimensional effects. A caption should name the dataset and boundary rule. Supply the frequency table alongside the graphic so that someone who cannot see the bars can still recover the distribution. A histogram should make the evidence easier to inspect, not replace the evidence with a shape that can no longer be audited.

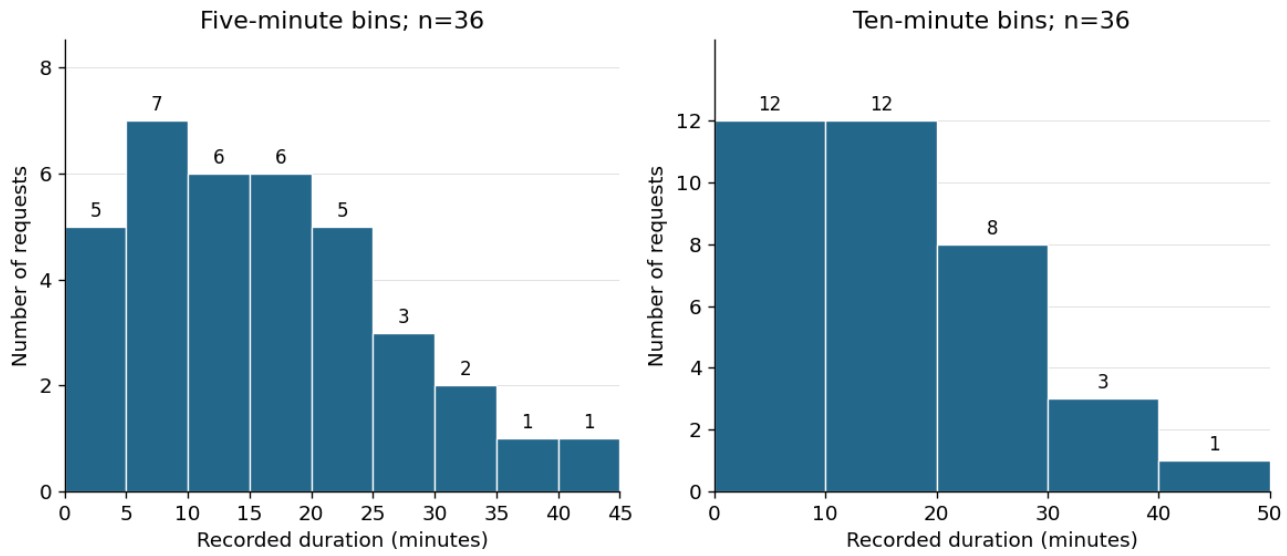


Figure 1. The same 36 synthetic Willow durations in five-minute and ten-minute bins. All intervals include their lower edge and exclude their upper edge; frequency tables provide exact values.

06 Change width without changing the data

Widening bins combines measurements over a larger range. Willow's ten-minute edges are 0, 10, 20, 30, 40 and 50. Counts are 12, 12, 8, 3 and 1. The first wide bin combines the first two five-minute counts, 5 + 7 = 12. The second combines 6 + 6 = 12. The third combines 5 + 3 = 8. The fourth combines 2 + 1 = 3, and the final bin contains the one observation at 44. The underlying 36 requests have not changed.

Aligned widths make this aggregation auditable. A ten-minute bin beginning at zero is exactly the union of two adjacent five-minute bins beginning at zero. If the origin changes, wide-bin counts cannot always be obtained by simply adding pairs of the old counts, because a new boundary may cut through an old bin. You would need the raw observations or a sufficiently detailed table. Grouped data lose information about where values sit inside each interval.

The wider chart has two equally tall first bars, while the five-minute chart has a single tallest interval [5,10). Neither view is the uniquely true shape. Both are faithful summaries under their declared binning choices. The finer view distinguishes local concentrations; the broader view smooths them. A statement that the 'most common duration is between five and ten' is specific to the five-minute intervals, while the fact that 24 of 36 requests are below 20 minutes survives both displayed schemes.

Do not compare raw bar heights across different widths as if each bar covered the same opportunity range. A count of twelve in a ten-minute bin versus seven in a five-minute bin does not indicate that the dataset grew. The broader interval has more room to accumulate observations. When comparing charts, keep the sample size, measurement unit, bin width and vertical measure visible. This four-part check prevents the most common misreading of a legitimate width comparison.

Ten-minute interval	Count	Derived from five-minute counts
[0,10)	12	5+7
[10,20)	12	6+6
[20,30)	8	5+3
[30,40)	3	2+1
[40,50)	1	1 plus no observations in [45,50)

07 Test the origin as well as the width

Two histograms can use the same width and still look different because their boundaries start at different places. Shift Willow's five-minute edges from 0,5,10,… to −2.5,2.5,7.5,…,47.5. The new counts are 3, 6, 7, 5, 6, 4, 2, 1, 1 and 1. The tallest interval is now [7.5,12.5), containing 8, 9, 9, 10, 10, 11 and 12. The data are unchanged, but different nearby observations are grouped together.

A negative first boundary does not imply negative observed durations. It is part of a display grid that covers the valid zero-minute observation. The first shifted bin [−2.5,2.5) contains 0, 1 and 2. A caption should make clear that durations are nonnegative and that the chosen bin extends below zero only to center the grid. If such a boundary would confuse the intended audience, use a zero-origin display as the primary chart and the shifted view only as a sensitivity demonstration.

Origin sensitivity is particularly strong for small samples or repeated rounded values near boundaries. A cluster of values just below and above ten can be split by a ten-minute boundary or united by a shifted bin. A dramatic claim about two separate populations should not depend on that arbitrary placement. Inspect neighboring choices and, when the sample is small, consider retaining a raw-value listing or a dot display as a companion rather than relying entirely on the histogram.

The purpose of sensitivity analysis is not to search for the prettiest chart or the strongest story. It is to identify which conclusions are robust to reasonable representation choices. Willow's observed range, sample size, count below 20 and sparse upper tail do not change with the origin. The exact location and height of the tallest bin do. Report those categories separately: stable numerical facts versus bin-dependent visual impressions. This habit makes a chart interpretation more defensible without requiring an arbitrary rule that one origin is always best.

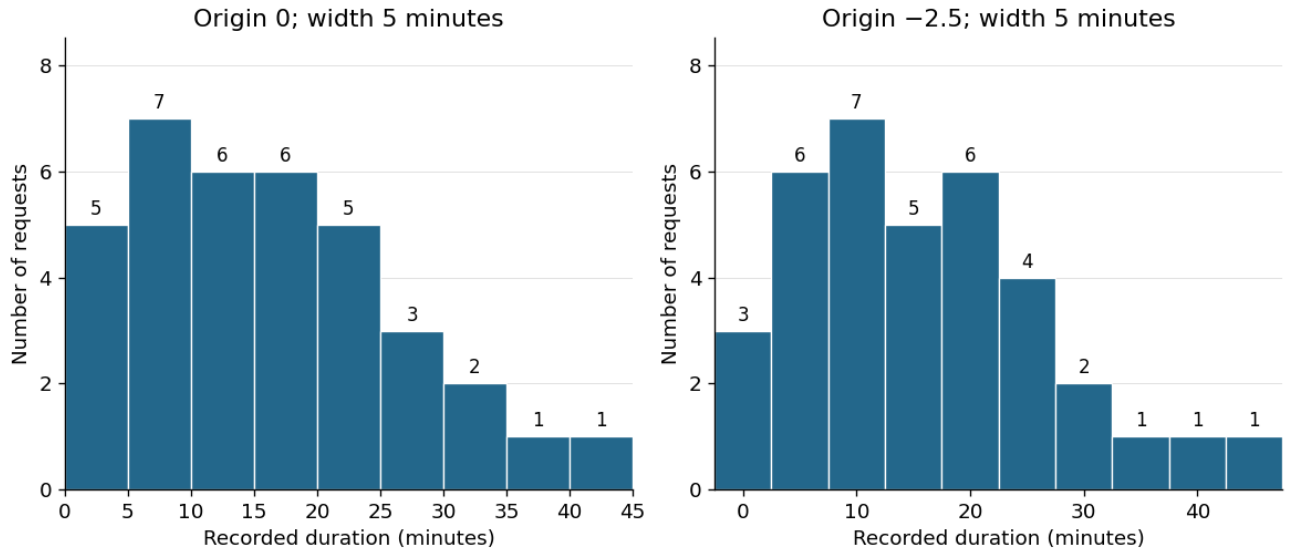


Figure 2. Five-minute width with two origins. Shifting boundaries changes the local peaks but not the 36 records, the 0–44 minute range or the exact count below 20 minutes.

Shifted interval	Count
[−2.5,2.5)	3
[2.5,7.5)	6
[7.5,12.5)	7
[12.5,17.5)	5
[17.5,22.5)	6
[22.5,27.5)	4
[27.5,32.5)	2
[32.5,37.5)	1
[37.5,42.5)	1
[42.5,47.5)	1

E07-X1 Separate stable facts from bin-sensitive impressions

Compare Willow’s five-minute zero-origin, ten-minute zero-origin and five-minute shifted-origin tables. Classify these statements: n=36; maximum=44; the tallest interval is [5,10); exactly 24 durations are below 20; there is one uniquely tallest bar. Explain each classification using a count or boundary fact.

Submit: Five labeled statements with supporting evidence.

08 Read concentration and spread in a disciplined order

Read a histogram in a consistent sequence: identify the variable and population, inspect the horizontal range, locate broad concentration, compare the tails, then examine gaps or isolated bins. This order reduces the temptation to treat the tallest bar as the whole story. For Willow, most recorded durations are in the lower part of the 0–44 minute range, while a small number extend into the thirties and forties. The five-minute frequency table supports that description without requiring a named probability model.

Center is not a single visual fact. Willow's median is 14.5 minutes, the average of the eighteenth and nineteenth sorted values, 14 and 15. Its mean is 571/36, approximately 15.86 minutes. These exact raw-data summaries are not read directly from the highest bar. The mean lies above the median in this case, consistent with the longer upper extension, but that relationship alone is not a universal diagnostic rule for every possible distribution. Use the actual shape and data rather than a slogan.

Spread also needs a specified measure. The observed range is $44 - 0 = 44$ minutes. The range is easy to calculate but depends only on the two extremes; it does not describe how tightly the other 34 observations cluster. Counting below meaningful thresholds can be more useful for an operational reader: 24 of 36 are below 20 minutes, and 32 of 36 are below 30. These facts distinguish the main concentration from the sparse tail without requiring a variance calculation.

Avoid calling 44 an error merely because it is far from the central values. It is a valid observation under the case conditions. An unusual value can motivate a source check or a question about process context, but deletion requires evidence and a stated rule. Removing it would change n, the maximum, the mean and the upper-tail count. If a sensitivity view excluding a value is ever justified, preserve the all-record view and clearly state what changed and why.

Raw-data summary	Willow result	Interpretation
Count	36	Included completed requests
Minimum / maximum	0 / 44 minutes	Observed endpoints, not bin edges
Median	14.5 minutes	Average of sorted positions 18 and 19
Mean	571/36 = 15.8611 minutes	Uses every recorded duration
Below 20 minutes	24/36 = 66.67%	Strict threshold excludes 20
Below 30 minutes	32/36 = 88.89%	Strict threshold excludes 30

09 Distinguish counts from relative frequency

A count histogram shows how many observations fall in each interval. A relative-frequency histogram divides each bin count by the number of included observations. With Willow's five-minute bins, the first two relative frequencies are 5/36 and 7/36, approximately 13.89 and 19.44 percent. The percentages across all nine bins sum to 100 percent before display rounding. The shape is the same as the count histogram because every bar is divided by the same denominator.

Relative frequency is useful when comparing groups with different sample sizes. A group with twice as many requests will often have taller count bars even if its distribution is similar. Converting to within-group percentages separates the distributional comparison from the total number observed. It does not make the groups representative or eliminate selection differences. Keep each group's sample size visible so that a 50 percent bar based on two requests is not mistaken for evidence as stable as a 50 percent bar based on hundreds.

Do not normalize each bar by its own interval width and still label the axis simply 'percent of requests.' That operation produces a density-like measure, discussed next. Also do not divide by the number of bins; five bins do not imply that each bar should be a fifth of the sample. The denominator for a relative frequency is the included observation count under the stated coverage policy. If out-of-range records are excluded, say whether percentages use all valid records or only the displayed subset.

Rounding creates a harmless presentation issue when percentages displayed to two decimals sum to 99.99 or 100.01. Calculate from unrounded values and state that totals may differ slightly due to rounding. Do not alter an individual bin count to force rounded percentages to total exactly 100. The underlying counts are authoritative. A reader who needs precise reconstruction should use the counts and n rather than infer exact fractions from rounded labels.

10 Use area, not raw height, when widths differ

The foundation exercises use equal-width bins, where count heights are straightforward. An advanced extension shows why unequal widths need different treatment. Group Willow into [0,10), [10,20), [20,30) and [30,50). Counts are 12, 12, 8 and 4. The last interval is twice as wide as each earlier interval. If raw counts are used as heights, the last rectangle's area represents twice as much per request as the others, breaking the usual area interpretation of a histogram.

Count density divides count by width. The four heights become 1.2, 1.2, 0.8 and 0.2 requests per minute of interval width. Multiplying height by width recovers each count: $0.2 \times 20 = 4$ for the last bin. Probability density divides count by n times width, giving 1/30, 1/30, 1/45 and 1/180 per minute. Each rectangle's area is the fraction of observations in that interval, and the four areas sum to one.

A density height is not the probability of one exact duration. For a continuous measurement, interval area is the relevant quantity. In this recorded, rounded teaching dataset, the density histogram remains a representation of grouped observations, not a claim that a particular

continuous population model is true. A bar height above one can be valid when measurement units or bin widths make densities numerically large; the requirement concerns total area, not a universal ceiling on height.

Label the vertical measure explicitly: 'Requests per minute of interval width' for count density, or 'Relative frequency density (per minute)' for normalized density. Do not compare a density chart to a count chart by reading equal heights as equal numbers of requests. In the supplied unequal-width figure, the final bar is low and wide because four observations are spread over a twenty-minute range. This extension clarifies the mathematics; it does not change the equal-width count specification of the original capstone.

Interval	Width	Count	Count density	Relative density
[0,10)	10	12	1.2	1/30
[10,20)	10	12	1.2	1/30
[20,30)	10	8	0.8	1/45
[30,50)	20	4	0.2	1/180

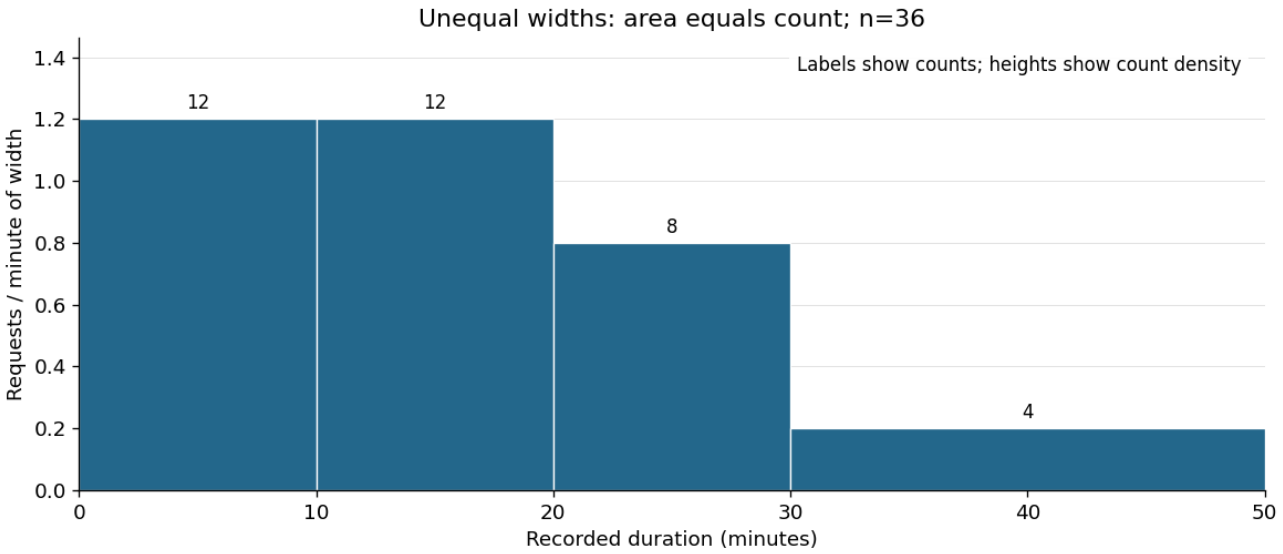


Figure 3. Unequal-width extension: count density makes each rectangle’s area equal its count. The final 20-minute interval has count 4 and height 0.2 requests per minute of width.

E10-X1 Repair an unequal-width histogram

A chart uses Willow bins [0,10), [10,20), [20,30), [30,50) with raw count heights 12,12,8,4 and claims area represents requests. Replace the heights with count densities, show all four recovered areas and state the normalized density for the final bin.

Submit: Four heights, four area calculations and a corrected axis label.

11 Answer threshold questions at the right level of detail

A histogram can answer exact threshold questions only when its boundaries and retained information support them. In Willow's zero-origin five-minute table, the count below 20 is the sum of the first four bins: 5 + 7 + 6 + 6 = 24. The count at least 20 is 36 − 24 = 12. Because 20 is a boundary and the convention is left-inclusive, this separation is exact. The request recorded as 20 belongs to the at-least-20 group.

'At most 20' is a different condition from 'below 20.' At most includes the value 20, so the raw Willow data give 25 requests at or below 20. The grouped [20,25) count of five does not reveal by itself how many equal 20. Without the raw observations, that bin could contain zero through five exact twenties. A grouped table can therefore support an exact below-20 count while not identifying an exact at-most-20 count. Small words in a question determine the required calculation.

For a threshold of 22, the five-minute table alone provides bounds. There are 24 requests below 20, and five in [20,25). Therefore the number below 22 is between 24 and 29, inclusive, based only on that table. The raw values resolve it to 26 because 20 and 21 are below 22 while 22, 23 and 24 are not. Do not distribute a bin's observations uniformly unless an explicit approximation assumption is justified and labeled.

Cumulative counts are useful companions. A cumulative frequency column at successive upper edges shows how many observations are below each boundary. It must never decrease and must end at the included sample size if all observations are covered. This column can reveal an addition error and supports threshold statements without requiring a second chart. Keep the strictness of each boundary in its label: 'below 20' is more precise than a vague 'up to 20' when equality matters.

E11-X1 Separate exact counts, bounds and raw-data answers

Using only Willow’s five-minute table, give the exact count below 20 and a valid bound for the count below 22. Then use the raw list to calculate below 22 and at most 20. Explain why the grouped table alone does not identify the latter two exact answers.

Submit: Four results with explicit less-than or less-than-or-equal conditions.

12 Compare groups using compatible bins and denominators

A comparison should hold the measurement and display definitions constant. The fictional Birch group has eight durations: 2, 4, 6, 8, 10, 12, 14 and 16. The fictional Maple group has sixteen durations: every integer from 0 through 15. Under common bins [0,10) and [10,20), Birch counts are four and four; Maple counts are ten and six. Maple has more observations in both bins, but that statement concerns counts, not the relative concentration of each group's distribution.

Divide by each group's own sample size. Birch's bin shares are 50 and 50 percent. Maple's are 62.5 and 37.5 percent. Maple has a larger share below ten, while Birch has a larger share from ten to below twenty. The count comparison and percentage comparison answer different questions, and both can be correct. A report that says Maple has 'more long requests' should specify whether it means six versus four requests or a larger within-group proportion; the latter would be false in this example.

Use shared bin edges when the goal is a direct visual comparison. If Birch uses five-minute bins and Maple ten-minute bins, unequal bar widths and different groupings complicate the comparison. Automatic bin selection can produce this mismatch without an obvious warning. Choose a common scheme suited to both datasets and disclose it. If one group has observations beyond the other's range, keep a common horizontal scale and show the full covered range rather than cropping each chart to make their shapes look alike.

Normalization does not solve representativeness. Birch and Maple may have different selection processes, time windows or proportions of unfinished work. In this exercise they are simply the listed completed synthetic requests. Do not infer that one service is faster in general from the histogram alone. A comparison of distributional shape becomes useful when paired with clear population definitions, sample sizes and relevant context. The chart should not silently expand the scope from listed observations to an entire organization.

Group	n	[0,10) count / share	[10,20) count / share
Birch	8	4 / 50%	4 / 50%
Maple	16	10 / 62.5%	6 / 37.5%

13 Understand measurement precision before interpreting spikes

Recorded whole-minute data often contain repeated values because many distinct underlying durations round to the same number. A spike at ten minutes can reflect a real concentration, a reporting habit, a system threshold or simple rounding. The histogram alone cannot decide among those explanations. In Willow, the recording convention is supplied, but no mechanism for particular repeated values is established. Describe the recorded distribution without inventing an operational cause for each bump.

Changing units requires changing both values and bin edges. A ten-minute duration is 600 seconds. A five-minute bin width becomes 300 seconds, and the same membership counts should be obtained when every value and every boundary is converted consistently. If values are multiplied by sixty but boundaries are left at 0,5,10,..., most observations will move outside the range. That is a unit mismatch, not evidence of a radically different distribution.

Do not add precision that the records do not contain. A value stored as 10 minutes does not support a statement that the actual duration was exactly 10.000 minutes. Under the case's nearest-minute convention, many underlying values could round to ten. The histogram summarizes recorded values; its interval labels should not imply more precise measurement

than the source supports. If a threshold decision depends on exact seconds, the rounded data may be insufficient even though a chart can still be drawn.

The distinction between measurement and binning is crucial. Rounding occurred before the course received the values. Binning is the analyst's later grouping choice. Narrower bins cannot recover precision lost through rounding; they may merely create many empty intervals between integer measurements. A one-second histogram of whole-minute records can look artificially spiky because every recorded value falls on a multiple of sixty seconds. Choose a display resolution that respects the measurement process as well as the sample size and analytical question.

14 Choose a useful binning plan rather than a magical rule

There is no single bin width that makes every dataset maximally informative. Very narrow intervals can produce many empty or singleton bars, making a small sample look erratic. Very wide intervals can conceal differences by combining most observations into one or two bars. The choice is a tradeoff between local detail and a stable summary. A useful plan starts with the question and measurement precision, then tests a few reasonable alternatives without selecting only the one that supports a preferred story.

For Willow, five-minute bins are interpretable because the variable is recorded in minutes and the range is 44 minutes. Ten-minute bins provide a broader comparison. The shifted five-minute view tests whether a local peak depends on alignment. These three views are deliberately limited and purposeful. Producing dozens of charts and reporting only the most dramatic one would make the analysis less transparent, even though every individual chart might be arithmetically valid.

Automated rules can provide starting points, especially for larger datasets, but they are not guarantees about the underlying population. Different rules respond differently to sample size, spread and extreme values. Record the actual resulting edges, not only a software option such as 'auto.' A future software version or a slightly changed dataset could select different boundaries. Explicit edges make the displayed result reproducible and let a reviewer check exact membership.

Use decision-relevant thresholds carefully. If the question concerns requests below twenty minutes, placing a boundary at twenty can make that count directly visible. This is a defensible design choice when disclosed. It becomes misleading if boundaries are chosen to hide unfavorable observations or if only a convenient subset of the range is displayed. A binning plan should state the covered range, width, origin, endpoint rule, vertical measure and treatment of values outside the range. Together these describe the actual transformation from records to chart.

15 Calculate bins with a small readable program

The supplied assets/histogram_counts.py program uses only Python's standard library. It reads Willow records, checks unique IDs and nonnegative whole-minute durations, and applies explicit left-inclusive, right-exclusive bins. The function returns counts plus underflow and overflow. Keeping those exceptions separate prevents an out-of-range value from disappearing silently. The program does not draw charts or modify the source file; it prints the frequency counts for two declared edge sets.

Read the comparisons in the code as a translation of the interval contract. The condition lower <= value < upper includes the lower edge and excludes the upper edge. A value below the first edge is underflow; a value at or above the final edge is overflow. The sorted-edge check rejects repeated or decreasing boundaries because those would create zero-width or invalid intervals. With valid strictly increasing edges and this conditional structure, every finite numeric value reaches one destination.

The code below is the counting function used by the asset. You can verify its logic manually with boundary values before running the complete script. The command is python assets/histogram_counts.py from the course folder. A successful run should print five-minute counts [5,7,6,6,5,3,2,1,1] and ten-minute counts [12,12,8,3,1], with zero underflow and overflow for each. These expected outputs are supplied for learning; state separately whether you executed the program in your own environment.

A small program can still be wrong if the input definition is wrong. Checking unique IDs will not detect two different IDs accidentally assigned to the same real request. Rejecting negative values does not establish that a zero is meaningful. The script encodes this exercise's contract, while a real data pipeline would need a documented relationship to the source system. Use automation to reproduce a justified rule, not to substitute for deciding what the rule should mean.

```
def strict_counts(values, edges):
    if len(edges) < 2 or any(a >= b for a, b in zip(edges, edges[1:])):
        raise ValueError("Edges must be strictly increasing")
    counts = [0] * (len(edges) - 1)
    underflow = overflow = 0
    for value in values:
        if value < edges[0]:
            underflow += 1
        elif value >= edges[-1]:
            overflow += 1
        else:
            for index, (lower, upper) in enumerate(zip(edges, edges[1:])):
                if lower <= value < upper:
                    counts[index] += 1
                    break
    return counts, underflow, overflow
```

16 Design boundary tests that can reveal a plausible error

A test should target the rule most likely to be implemented incorrectly. For edges 0,5,10,15, use values −1, 0, 4.999, 5, 9.999, 10, 14.999 and 15. Under the strict convention, the three bin counts are two, two and two, with one underflow and one overflow. The value fifteen is the critical final-edge check. A library function that includes the final right edge would instead place it in the final bin unless you explicitly adapt the behavior.

Also test a dataset with an empty interior interval. Values 1,2,11,12 under edges 0,5,10,15 produce counts 2,0,2. The zero middle bin must remain in the table and chart. If the implementation drops it and draws the remaining bars adjacent, it hides the gap from five to below ten. This is not a cosmetic issue: the numerical axis would no longer represent the actual spacing of the measurements.

Test invalid edge definitions as well. Edges 0,5,5,10 contain a repeated boundary and should be rejected. Edges 0,10,5 decrease and should also be rejected. A counting routine that accepts them may assign some values twice, leave holes or produce widths that do not make sense. Tests for obvious invalid inputs help distinguish a robust transformation from a function that happens to work on one friendly dataset.

Finally, reconcile every output: included bin counts plus underflow plus overflow must equal the number of numeric test values. When missing or invalid source values are handled before counting, maintain a separate source ledger so the total raw records can also be reconciled. State what was actually tested. A manually reasoned expected result is useful, but it is not evidence that a program executed successfully. A trustworthy analysis distinguishes expected outputs, performed checks and unresolved failures.

E16-X1 Run or reason through a boundary test pack

Use the strict_counts function with edges [0,5,10,15]. Test values [−1,0,4.999,5,9.999,10,14.999,15], then [1,2,11,12]. State the expected result for each and the expected behavior for edges [0,5,5,10]. If you execute the code, record observed outputs separately.

Submit: Two exact return triples, one error expectation and an honest execution statement.

17 Write an interpretation card with an equivalent text view

An interpretation card translates the chart into a few evidence-backed statements. Begin with the population and measurement definition. Then give one concentration statement, one spread or tail statement and one design-sensitivity statement. Include the frequency table or a link to the supplied data file for readers who need exact reconstruction. The card should convey what the chart means, not merely describe its colors or list every visual object.

A Willow model is: Among the 36 listed completed requests, recorded durations ranged from 0 to 44 minutes. Twenty-four requests, 66.67 percent, were below 20 minutes, and four were at least 30 minutes. The five-minute zero-origin view peaks at [5,10), with seven requests, while the ten-minute view has two equally tall first bins of twelve. This means the broad concentration below twenty is stable, but the exact modal interval depends on the grouping. The small selected set of completed requests does not establish a population distribution or explain the processing mechanism.

Separate observation from hypothesis. 'The upper tail contains few longer recorded durations' describes the data. 'A specialist team caused the long durations' invents a mechanism not supplied by the records. 'Two populations exist' is also too strong merely because a histogram has two bumps. If a possible explanation is useful, phrase it as a question for further evidence and specify what data would test it, such as request type or unresolved-case status measured consistently.

Accessibility is an analytical obligation, not an optional afterthought. Someone who cannot inspect the graphic should still learn the same main message and obtain the same counts. Use readable labels, a clear caption, a text description and a structured frequency table. Do not rely on a color difference alone to distinguish comparisons. A static image cannot provide every interactive accessibility feature, so the surrounding text and table carry essential information. Never claim a particular screen-reader or keyboard behavior unless it was actually tested in the final medium.

18 Build an independent histogram analysis

The fictional Juniper dataset supplies an independent application in assets/juniper_durations.csv. Its 24 recorded durations are 1,2,3,4,5,5,6,8,9,10,10,12,13,14,15,17,18,20,22,25,28,30,35 and 40 minutes. Each row is one completed request with a unique ID J01 through J24. All values are valid, recorded to whole minutes and included. There are no missing values, duplicate IDs or uncompleted cases in this exercise file. Do not merge these records with Willow or the foundation dataset.

Create a five-minute table with edges 0,5,10,…,45 and a ten-minute table with edges 0,10,20,30,40,50. Use the strict left-inclusive, right-exclusive rule for every interval. Produce one count histogram for each table with contiguous numerical intervals, a zero-based count axis and the same horizontal measurement units. Each chart should state n=24 and its bin width. Preserve every bin, including any zero count that occurs.

Then calculate exact raw-data counts below twenty, at least twenty and at most twenty minutes. These related conditions deliberately test whether equality is handled consistently. Compare the modal intervals in the two charts and identify one statement that survives the width change. Do not infer the original durations from interval midpoints, since the raw file is available. Your working table should retain lower and upper edges, counts and cumulative counts below the upper edge.

The final deliverable is an interpretation card plus the two frequency tables and charts. Include two supported observations, one bin-sensitive observation, one material limitation and a concise text alternative. You may draw by hand, use a spreadsheet or use an approved

plotting tool, but the membership rules and numerical outputs are the same. Document whether any automated calculation was actually executed. An attractive chart with incorrect boundary counts does not satisfy the task; a correct table without readable labels also leaves the communication incomplete.

E18-X1 Produce the Juniper distribution card

Use the complete Juniper list and the two specified edge sets. Deliver frequency and cumulative tables, two correctly labeled count charts, exact threshold counts and an interpretation card. Include the raw median as a cross-check of center, not as a claim about the tallest bin.

Submit: Two tables, two charts, threshold calculations and a bounded interpretation.

19 Audit the evidence before naming a distribution

A final histogram audit has three layers: records, transformation and interpretation. At the record layer, confirm the observation unit, valid count, measurement units, rounding convention and exclusions. At the transformation layer, confirm strictly increasing edges, complete interval coverage, the endpoint rule, one destination per observation and the selected vertical measure. At the interpretation layer, check that every sentence follows from the records or displayed grouping and does not imply an unsupported causal or population model.

Look for common numerical traps. A maximum printed axis value is not necessarily an observed maximum. A bar with height six is not six minutes when the vertical axis counts requests. A wider bin can have a higher count simply because it covers a wider range. A cumulative count is not an ordinary bin frequency. A density height is not an interval probability unless multiplied by width. Each trap can be prevented by naming the quantity and its unit before comparing magnitudes.

Named distributions require more evidence than visual resemblance. A roughly bell-shaped histogram does not prove normality, and a right tail does not prove an exponential process. Small samples, bin origins and rounded measurements can all affect appearance. The course's goal is descriptive competence: build a faithful grouped representation, test its sensitivity and communicate what is stable. If a later task needs a probability model, it should specify the modeling purpose and use appropriate checks rather than treating a histogram as a certificate.

The strongest submission is not the one with the most charts. It is the one whose charts can be reproduced from the supplied records and whose conclusions remain proportionate to what those records support. Keep the count table, bin specification and scope sentence together. If a chart is revised, recheck labels and interpretation as well as arithmetic, because a changed width or origin can invalidate an old sentence about the tallest interval. A trustworthy histogram makes its grouping decisions visible instead of asking the reader to mistake them for properties of the raw data.

20 Know what a grouped table cannot recover

A frequency table is a lossy summary. Many different raw datasets can produce exactly the same histogram. If five observations fall in [20,25), the table does not tell you whether they are all twenty, all twenty-four or spread through the interval. This lost detail matters when a reader asks for an exact mean, median or count at a threshold inside the interval. The appropriate response is to return to the raw data or state the limits of what the grouped table identifies.

A common approximation replaces every value in a bin with its midpoint. For Willow's five-minute table, the weighted midpoint total is 585 minutes: multiply counts 5,7,6,6,5,3,2,1,1 by midpoints 2.5,7.5,12.5,17.5,22.5,27.5,32.5,37.5,42.5 and sum. Dividing by 36 gives an approximate grouped mean of 16.25 minutes. The exact raw-data total is 571 and the exact mean is 15.8611 minutes. The difference, 14/36 or 0.3889 minutes, arises from replacing actual positions by midpoints.

The midpoint approximation is not wrong when clearly labeled and appropriate to the purpose. It becomes wrong when reported as an exact statistic recovered from the grouped table. Under this case's whole-minute convention, each observation in [a,a+5) can take recorded integer values a through a+4. The same frequency table therefore permits a mean as low as 13.75 or as high as 17.75 minutes. These are logical grouped-data bounds, not confidence limits and not forecasts. The actual mean lies within them.

This information-loss perspective also explains why you cannot reliably create a finer histogram from a coarser one. The ten-minute count 12 in [0,10) does not reveal how many belong in [0,5) versus [5,10). The Willow raw records happen to split it 5 and 7, but many other splits are possible. Never manufacture fine-grained detail merely because a charting tool accepts the desired number of bins. A reliable analyst distinguishes what the source preserves, what a declared approximation assumes and what remains unidentified.

Assessment

Choose one best answer per question. Questions 1–10 check the foundation concepts; questions 11–20 apply the extended study units. Complete and submit the practical tasks separately. A strong quiz score does not replace correct practical work.

1. With the foundation rule [lower,upper), where does a recorded duration of 10 minutes belong in the five-minute chart?

A. [5,10), because 10 is its printed endpoint

B. [10,15), because the lower edge is included

C. Half in each neighboring bin

D. In [5,10) only if it is the second repeated 10

2. The foundation data begin 2,3,4,5,6. How many of these observations belong in [0,5)?

A. 5

B. 4

C. 2

D. 3

3. The foundation values in [5,10) are 5,6,7,7,8,8,9,9. What is the bin count?

A. 8

B. 5

C. 10

D. 7

4. The foundation has 24 included requests, all covered by each declared range. What must the frequency total be?

A. 25, the final narrow boundary

B. 30, the final wide boundary

C. 24, with every request counted exactly once

D. 5, the number of narrow bins

5. Foundation narrow counts for [0,5) and [5,10) are 3 and 8. What is the aligned wide count for [0,10)?

A. 8

B. 11

C. 12

D. 14

6. When the foundation switches from five-minute to ten-minute bins without changing records, what changes?

A. The elapsed duration represented by each recorded value

B. The number of completed requests

C. The grouping and visible local detail

D. The measurement unit from minutes to tens of minutes

7. Why do adjacent equal-width histogram bars meet at their boundaries?

A. Every interval must have a positive count.

B. Touching bars guarantee that the counts are correct.

C. The measurement values have no repeated observations.

D. They represent adjoining ranges on one numerical scale.

8. Two distinct foundation requests both recorded 7 minutes. How should they enter [5,10)?

A. Count two observations in the same bin.

B. Place one in each adjacent bin to show repetition.

C. Count neither until the underlying seconds are known.

D. Count one because only one duration value is unique.

9. The foundation list has only 21 and 24 minutes at or above 20. How many requests meet duration≥20?

A. 2

B. 22

C. 4

D. 24

10. Which interpretation of the foundation two-width comparison is defensible?

A. The wider chart proves the population has two modes.

B. The narrower chart reveals the exact continuous population density.

C. The bin choice changes local appearance while the 24 observations stay fixed.

D. The wide chart contains more observations because its first bars are taller.

11. For strict bins [0,5), [5,10), [10,15), where is value 15 assigned by the supplied counting function?

A. The last bin, because all histogram tools include the final edge

B. It is silently discarded, so no returned total includes it

C. Underflow, because it is outside the specified bins

D. Overflow, because the final right edge is excluded

12. Willow’s [5,10) and [10,15) counts are 7 and 6. A table instead shows 9 and 4 but still totals 36. What does this demonstrate?

A. A correct total proves all interval counts are correct.

B. A matching grand total can hide a boundary misassignment.

C. Repeated values should always be removed.

D. The two intervals must be merged to avoid any valid comparison.

13. In Willow, which fact remains unchanged when the bin width and origin change?

A. The exact raw-data count below 20 is 24.

B. The tallest bin always begins at 5.

C. There is always one uniquely tallest bar.

D. The number of displayed bins remains nine.

14. For an unequal-width bin [30,50) containing 4 of 36 observations, what count-density height makes area equal count?

A. 4 requests per minute

B. 4/36 requests per minute

C. 0.2 requests per minute of interval width

D. 0.4 requests per minute of interval width

15. Using only Willow’s five-minute table, what can be stated about the number below 22?

A. Exactly 26, because every bin is internally uniform

B. Exactly 29, because [20,25) is partly below 22

C. Between 24 and 29; the raw records are needed for the exact count

D. Exactly 24, because partial bins must be discarded from all reasoning

16. Birch has 4 of 8 requests in [10,20), while Maple has 6 of 16 there. Which comparison is accurate?

A. Maple has the larger count, but Birch has the larger within-group share.

B. Maple has both the larger count and larger share.

C. Birch has the larger count because 50% exceeds 37.5%.

D. Their distributions cannot be compared unless n is equal.

17. A program returns [2,2] for values 1,2,11,12 with edges 0,5,10,15, omitting the empty middle interval. What is the best correction?

A. Keep it because the total remains 4.

B. Move 11 and 12 into [5,10).

C. Change the source values so there are no gaps.

D. Return [2,0,2] and retain the middle interval’s numerical space.

18. Willow’s midpoint-based grouped mean is 16.25 minutes, while its exact raw mean is 571/36≈15.86. How should 16.25 be labeled?

A. The exact mean, because bin midpoints are observed values

B. An approximation based on replacing observations with bin midpoints

C. A confidence limit for the true population mean

D. The median duration because it is near the center

19. The Juniper value 20 appears once. Which pair of counts is correct?

A. Below 20=18; at most 20=17

B. Below 20=17; at most 20=17

C. Below 20=18; at most 20=18

D. Below 20=17; at most 20=18

20. Which sentence best describes a small histogram with a sparse upper tail?

A. The listed observations include a few longer durations; the chart alone does not establish why.

B. The long durations must be recording errors and should be removed.

C. The tail proves a specific probability distribution for future requests.

D. The tallest interval identifies the exact duration of a typical future request.

Sources and further reading

NIST/SEMATECH: Histogram
https://www.itl.nist.gov/div898/handbook/eda/section3/histogra.htm

NumPy: numpy.histogram documentation
https://numpy.org/doc/stable/reference/generated/numpy.histogram.html

UK Government Analysis Function: Data visualisation charts
https://analysisfunction.civilservice.gov.uk/policy-store/data-visualisation-charts/