

§2.1 & 2.2 Frequency Distribution

Class (Bin) # of Items in Recycle	# Days Frequency Count	Relative Freq. = % = prop	\leq Cumm Freq.	Cumm Relative Frequency
0-4	5	7.35%	5	7.35%
5-9	20	29.4%	25	36.75%
10-14	25	36.76%	50	73.5%
15-19	10	14.7%	60	
20-24	5	7.35%	65	
25-29	3	4.4%	68	
	68	100%		

Relative frequency = proportion of Sample
= %

a) what proportion of Day were 0-4 items collected?

$$p = \frac{x}{n} = \frac{5}{68} = .0735 = 7.35\%$$

b) On what prop. of Day were fewer than 15 items collected?

$$\frac{50}{68} \text{ or } 7.35\% + 29.4\% + 36.76\%$$

$P(X \leq 14) = P(X < 15) = .735 = 73.5\%$
= Cumulative relative Frequency

Frequency Distribution Questions

- a) What proportion of classrooms had between 5 and 9 items? .294
- b) What percentage had less At Most than 14 items?

$$\text{Prop} = .368 + .294 + .074 = .736$$

73.6% = Cumulative percentage
for $X \leq 14$

$$\boxed{P(X \leq 14) = .736}$$

Add Columns to frequency Distribution

$$\text{Relative frequency} = \% = \text{proportion} = \frac{f}{n} = \frac{\# \text{ in class}}{\text{Total in Sample}}$$

$$n = \sum f = \text{Sample Size}$$

\sum = Capital Sigma
Mean take the Sum over
all categories

$$\sum \text{relative freq.} = 1$$

$$\sum p_i = 1$$

② Percentage = relative freq. $\cdot 100$

③ Cumulative frequency = Sum frequency to that class

Cumulative Relative frequency = Sum of proportions up to that class



Vocabulary

$$\text{Class width} = LCL_2 - LCL_1 = 5$$

To pick $\frac{\text{Max} - \text{Min}}{5} = \text{pick a Nice value close to } = \frac{29-0}{5} \approx 5$
this value

Lower class Limit (LCL) = Smallest data value in class

$$\text{are} = 0, 5, 10, 15, 20, 25,$$

$LCL_1, LCL_2 = 5$
 $LCL_1 = 0$

Upper Class limit (UCL) = 4, 9, 14, 19, 19+5, 29
↑
Add class width

Midpoints $\frac{UCL_1 + LCL_1}{2} = \frac{0+4}{2} = 2$

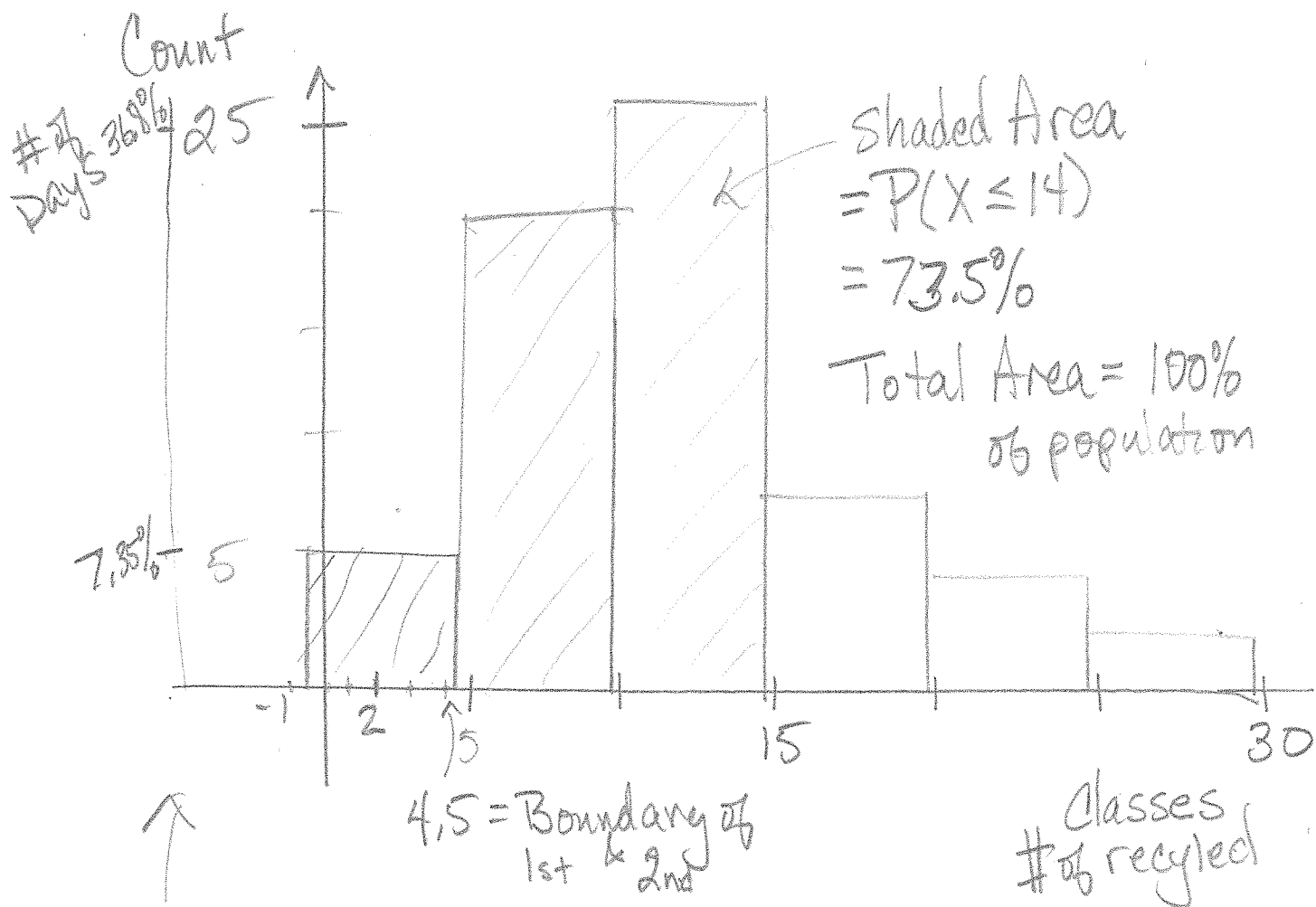
All Midpts = Mid₁ + Width, = 2, 7, 12, 17, 22, 27

Boundaries Split difference between LCL_2 and UCL_1

Upper boundary of first class = $\frac{UCL_1 + LCL_2}{2} = \frac{4+5}{2} = 4.5$

Lowest Class Boundary of 1st = $4.5 - 5 = -0.5$

Bin Start = $\frac{UCL_1 + LCL_2}{2} - \text{width}$



Shape is the Same when we use
 Counts or Relative frequency for
 y-axis

Boundaries = Edges of Rectangles

= width = 5

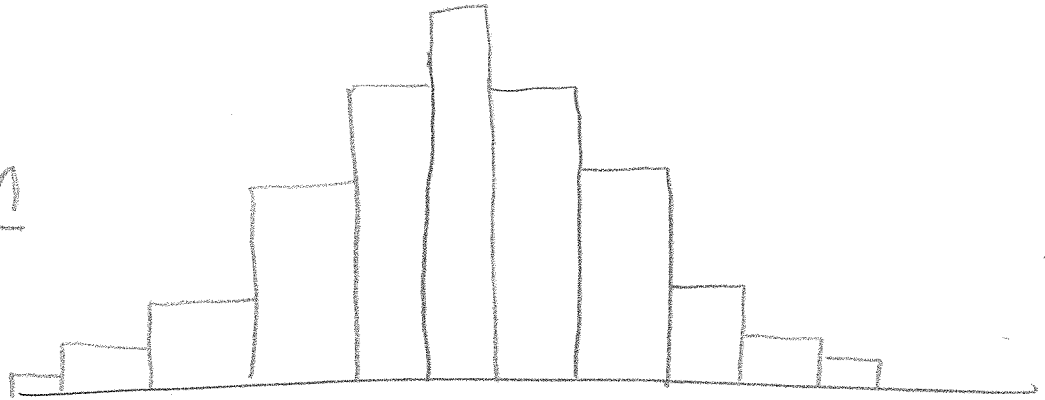
= height = count or relative frequency

- 0.5, 4.5, 9.5, 14.5, 19.5, 24.5, 29.5

Types of Distributions

Mean of Distribution = Balancing point of Histogram

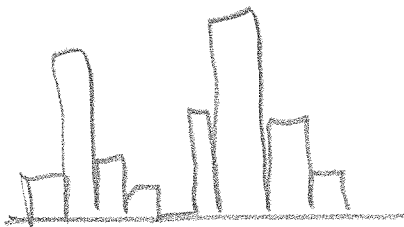
Normal Distribution



- ① Symmetric ② One high in Middle ③ Low on Sides

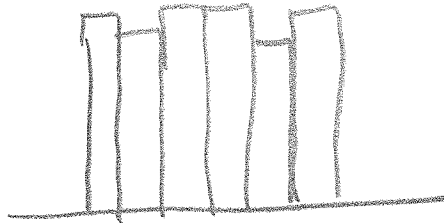
Close to Normal is good enough in Most of Stats

Bimodal



~~②~~

Uniform



~~③~~

Skewed Right
Tail on Right
↓



§2.1#16 or MML#7

You will have different #'s but same process

$$LCL_1 = 119.4 \quad \text{width} = .2$$

Class Voltage	freq
119.4 - 119.5	2
119.6 - 119.7	6
119.8 - 119.9	10
120.0 - 120.1	6
120.2 - 120.3	1
120.4 - 120.5	

$$\begin{aligned} \text{lowest class boundary} &= \frac{LCL_2 + UCL_1}{2} - \text{width} \\ &= \frac{119.6 + 119.5}{2} - .2 \\ &= 119.55 - .2 \\ &= \boxed{119.35} \end{aligned}$$

Use Statcrunch to make histogram

- ① Select Data List
- ② Stat Bins at = $\boxed{119.35}$ = lowest class boundary
Class width = $\boxed{.2}$

yes this Data Looks Normal

Use StatCrunch to make
Histogram and frequency Distribution

- ① Determine class width & where to start bins.

$$\text{Width} \approx \frac{\text{Max} - \text{Min}}{6} = \text{Choose a Nice Number close to this Value}$$

$$= \frac{121.6 - 120.7}{6} = .15 \text{ use } .2 \text{ so Nice}$$

$$\boxed{\text{Width} = .2}$$

- ② Start Bins $\frac{1}{2}$ unit below $\text{Min} = \text{Smallest Data Value}$

$$\text{Min} = 120.7 \text{ so}$$

$$\text{Start Bins} = 120.7 - .05 = 120.65$$

$\downarrow \frac{1}{2} \text{ unit}$
= lowest class Boundary

- ③ Hand write freq. Dist. Is it Normal or Uniform

Classes	Freq
120.7 - 120.8	6
120.9 - 121.0	5
121.1 - 121.2	4
121.3 - 121.4	5
121.5 - 121.6	5
121.7 - 121.8	5

→ Not Normal

→ It is uniform
Each outcome is
equally likely

← Print
Histogram

Add
width

2.1.16 # 7 in MML
~~2.1.16 # 4 in MML~~

Given that $LCL_1 = 124.0$ and width = 0.2 volts

① Start freq. Dist → Use To find $\text{Width} = 0.2$ = given

Class	freq
124.0 - 124.1	2
124.2 - 124.3	5
124.4 - 124.5	10
124.6 - 124.7	6
124.8 - 124.9	2

① $\text{Width} = 0.2$ = lowest class boundary

$$= \frac{LCL_2 + UCL_1}{2} - \text{width}$$

$$= \frac{124.2 + 124.1}{2} - 0.2$$

② Start bins At $= 123.95$ = Lowest class Boundary

② Use Statcrunch to get frequencies

☒ freq

☐ Graph

☐ Histogram

☒ open in statcrunch

☐ Graph → Histogram

① Select Data

② width = class width

Bins

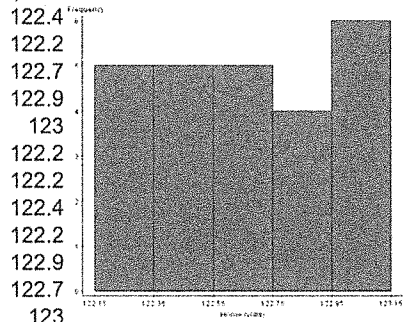
Start at = Lowest class boundary

③ Value Above Bar ☒

④ Copy → Paste special → Bit Map

Home
(volts)

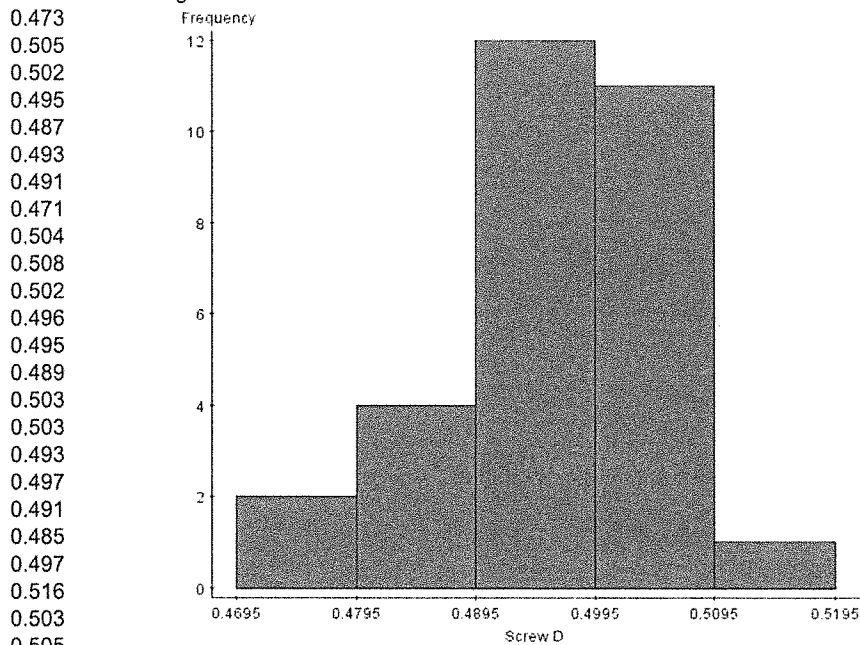
#6



122.4
122.2
122.7
122.9
123
122.2
122.2
122.4
122.2
122.9
122.7
123
122.7
122.7
122.9
123.1
122.4
122.5
122.3
123
122.9
123.1
123.1
122.5
122.7

#24

Screw Lengths



0.473
0.505
0.502
0.495
0.487
0.493
0.491
0.471
0.504
0.508
0.502
0.496
0.495
0.489
0.503
0.503
0.493
0.497
0.491
0.485
0.497
0.516
0.503
0.505
0.493
0.485
0.502
0.495
0.508
0.495

Sort Data: Click on [Data] and find [Sort Columns]. In the window that pops up select the column that you wish to sort with your cursor, and click on the button [Sort Column]. This creates a new column with the data sorted.

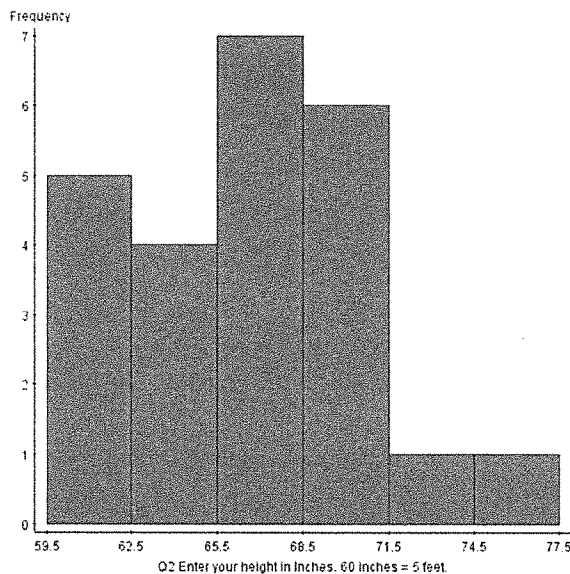
Frequency Distribution: If the bins or classes have more than one outcome you need to first BIN the DATA. To do this, click on [DATA] then find Bin Column. In the window that pops up, select the column that you wish to Bin. Next use [Stat] [Tables] [Frequency] applied to the *list of bins* that you just made to create a frequency distribution. You need to select where to start the bins and the width of each of the bins. Choose the start to be a nice number less than or equal to the minimum value. Choose the width so there are between 5 and 8 classes or bins. Width is a nice number approx equal to $(\max - \min)/6$. In statCrunch the bins include the lower class limit and do not include the upper class limit.

Frequency table results for Bin(Sort(Q2 Enter your height in Inches. 60 inches = 5 feet.)):

Bin(Sort(Q2 Enter your height in Inches. 60 inches = 5 feet.))	Frequency	Relative Frequency
60 to 63	5	0.20833333
63 to 66	4	0.16666667
66 to 69	7	0.29166666
69 to 72	6	0.25
72 to 75	1	0.041666668
75 to 78	1	0.041666668

Histogram: We want the histogram to match our frequency distribution. Click on [Graphics] then select the column with the *original data*. Set the bins to start one half units before the start of the first class, this is the initial class boundary. In the example, I set

START BINS AT to 59.5 and
Bin Width is 3.



Save: To save your work click on [Options] select copy, open a Word document, paste you table or graph. You can click on the corner to shrink the graph and change the font size to shrink your chart.

Using Statcrunch to Make a Histogram and a Frequency distribution

- ① Click the icon
- ② Open in Statcrunch

Graph

Histogram

- ① Select Data

Skip stuff

- ② Stat bins at

Lowest
class boundary

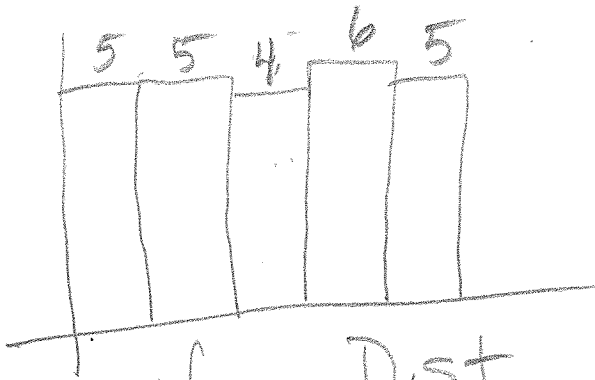
- ④ ☒

~~shows counts~~

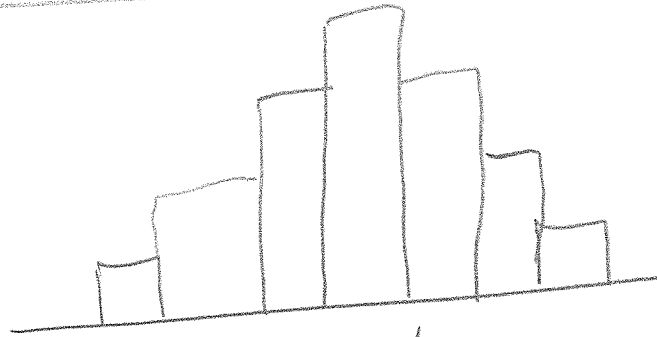
Value Above Bar

- ③ width = .2

Kinds of Distributions



Uniform Dist



Normal
Requires

- ① Start & end low
- ② One Maximum in Center
- ③ Symmetric

Height	
58 - 60	1 58
61 - 63	63 63 62 61 62 62, 62, 62
64 - 66	64 65, 61, 66, 66, 64, 64
67 - 69	67, 67, 67, 69, 68
70 - 72	71, 70, 72, 71, 72
73 - 75	73 73 75 74 74

In Calculator STAT EDIT LI Clear Enter

STAT PLOT 1: Enter

ION v

▷▷ ||||

XList: L1

Freq: 1

ZOOM 9: Zoom Stat

24
of 34 p

- of 34 patients

and a class distribution.

nts .

- and histogram

- n for each of

Frequency Distributions and Histograms

With calculator

① Enter Data into calculator

Stat

1 \rightarrow Edit

Highlight L1

Clear

29 30 36 41 45 50 57 61 28 50 36 58
60 38 36 47 40 32 58 46 61 40 55 32

② Make Histogram

Stat

Plot

Plot 1

Enter

Plot 1

On

off

Type

L1

Xlist

freq: 1

Window

Quiz fist

X-min = lowest class boundary = 24.5

X-max = largest class boundary = 64.5

X-scale = class width = 8

y-min = -2 \rightarrow may need to increase lots of data

y-max = 10

y-scale = 1

③ Make frequency Dist \rightarrow LCL = 25 width = 8

Graph **Trace**

class Bin	freq
25 - 32	5
33 - 40	6
41 - 48	4
49 - 56	3
57 - 64	6

Solve the problem.

- 1) A gardener has 75 clients, 45% of whom are businesses. Find the number of business clients.

Form a conclusion about statistical significance. Do not make any formal calculations. Either use the results provided or make subjective judgments about the results.

- 2) Last year, the average math SAT score for students at one school was 475. The headmaster introduced new teaching methods hoping to improve scores. This year, the mean math SAT score for a sample of students was 481. Is there statistically significant evidence that the new teaching method is effective? If the teaching method had no effect, there would be roughly a 3 in 10 chance of seeing such an increase. Does the result have statistical significance? Why or why not? Does the result have practical significance?

a) Use critical thinking to find the flaw in the sampling method.

- 3) You plan to make a survey of 200 people. The plan is to talk to every 10th person coming out of the school library. Is there a problem with your plan?

- 4) A researcher published this survey result: "74% of people would be willing to spend 10 percent more for energy from a non-polluting source". The survey question was announced on a national radio show and 1,200 listeners responded by calling in. What is wrong with this survey?

- 5) "38% of adults in the United States regularly visit a doctor". This conclusion was reached by a college student after she had questioned 520 randomly selected members of her college. What is wrong with her survey?