Density Curves

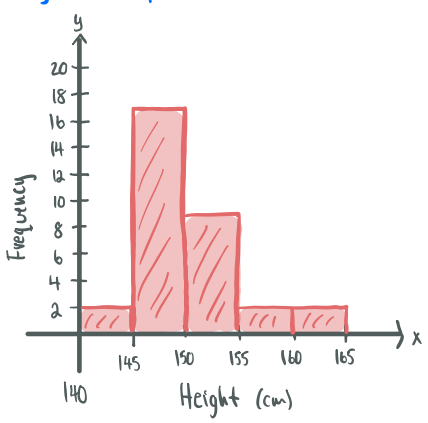
aka "Distribution" Curves

ex The heights, in cm, of 32 gymnasts were recorded:

148	152	147	149	150	147	151
142	156	148	148	149	150	152
155	154	151	154	148	150	149
145	147	148	161	152	162	149
146	151	150	157			

(a) Construct a Grouped Frequency Table, using groups of 5cm. Draw Histogram, then draw a curve connecting the top corners of each bar

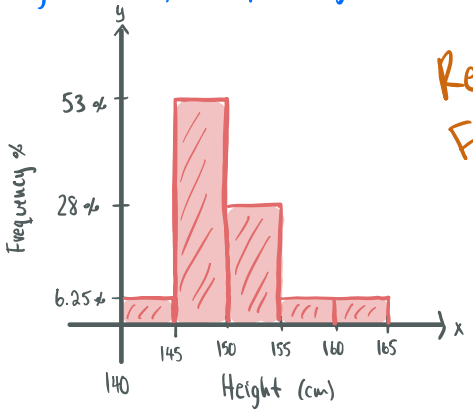
Class	F
$140 < h \leq 145$	2
$145 < h \leq 150$	17
$150 < h \leq 155$	9
$155 < h \leq 160$	2
$160 < h \leq 165$	2



Frequency Histogram

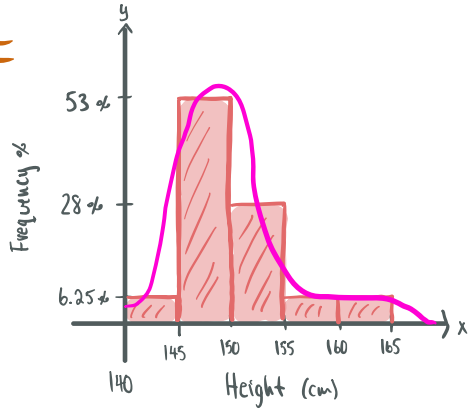
(b) Instead of using actual Frequency numbers, use percentages

Class	F %
$140 < h \leq 145$	6.25
$145 < h \leq 150$	53.12
$150 < h \leq 155$	28.12
$155 < h \leq 160$	6.25
$160 < h \leq 165$	6.25



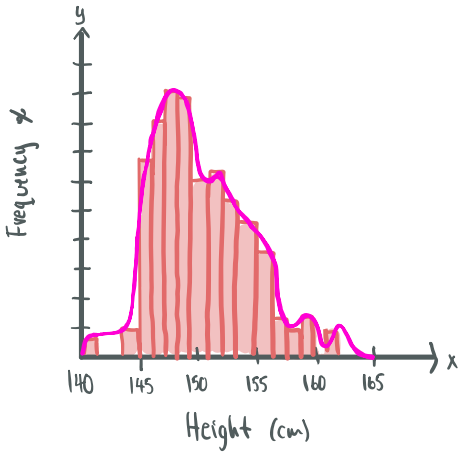
Relative Frequency Histogram

(C) Draw a curve connecting the tops of each bar
This is the **Density Curve**

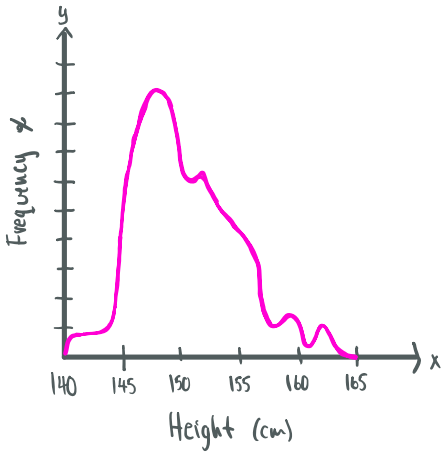


(D) Construct a Grouped Frequency Table, using groups of 2cm.
Draw Histogram, then draw a curve connecting the top corners of each bar

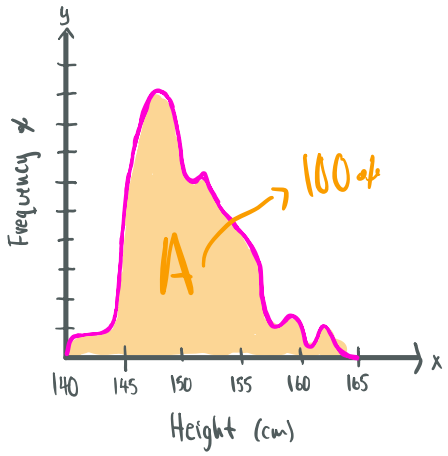
Class	F
$140 < h \leq 142$	1
$142 < h \leq 144$	0
$144 < h \leq 146$	2
$146 < h \leq 148$	8
$148 < h \leq 150$	8
$150 < h \leq 152$	6
$152 < h \leq 154$	2
$154 < h \leq 156$	2
$156 < h \leq 158$	1
$158 < h \leq 160$	0
$160 < h \leq 162$	2
$162 < h \leq 164$	0



⚠ You can imagine, the skinnier our Rectangles ⇒ the smoother our curve will be



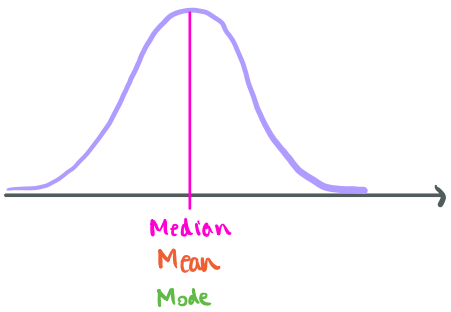
Let's take a deep dive into analysing Density Curves



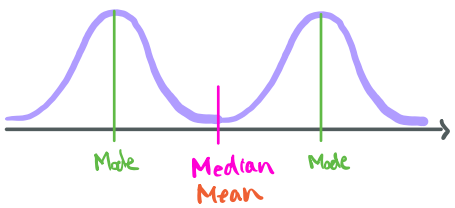
- Area under the curve is 100% of the data. Thus it will always equal 1

Consider the Distribution Curves Below
Approximate the Mean and Median for each curve,
Draw a vertical line to indicate these values

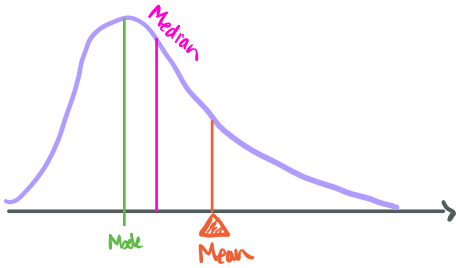
Normal Distribution



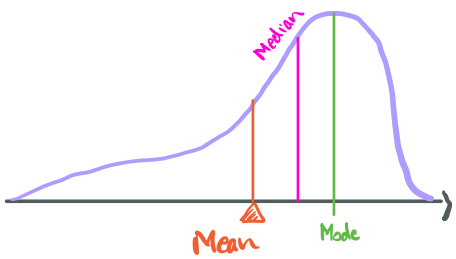
Bimodal Distribution



Right Skewed Distribution
Mean > Median



Left Skewed Distribution
Mean < Median



✡ Think of your Mean as the balancing point of the curve (Fulcrum)

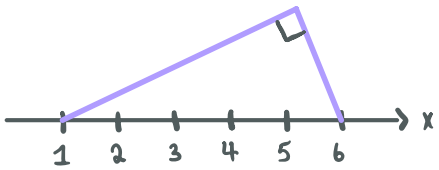
Skewed - An asymmetrical data set

✡ If a data's Mean ≠ Median it is skewed

Right Skew - Mean > Median tail points Right

Left Skew - Mean < Median tail points left

ex] Consider the Density Curve



(a) Approximate the Mean, Median, and Mode
↓ 3.5 ↓ 4.5 ↓ 5

(b) Is are data Skewed?

Yes, Mean < Median => left Skewed

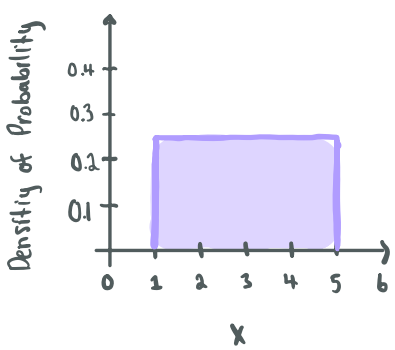
(c) Find the height of this curve

$$A_T = \frac{1}{2}bh = \frac{1}{2}(5)(h) = 1$$

$$\Rightarrow \boxed{h = \frac{2}{5}}$$

ex] Consider the Density Curve

Uniform Distribution



Find: (a) Min = 1

(b) Lower Quartile (Q_1) = 2

(c) Median = 3

(d) Mean = 3

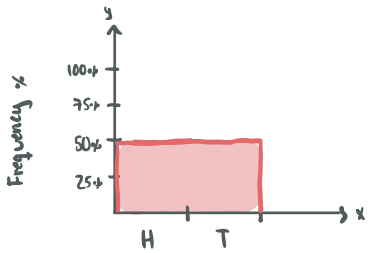
(e) Upper Quartile (Q_3) = 4

(f) Max = 6

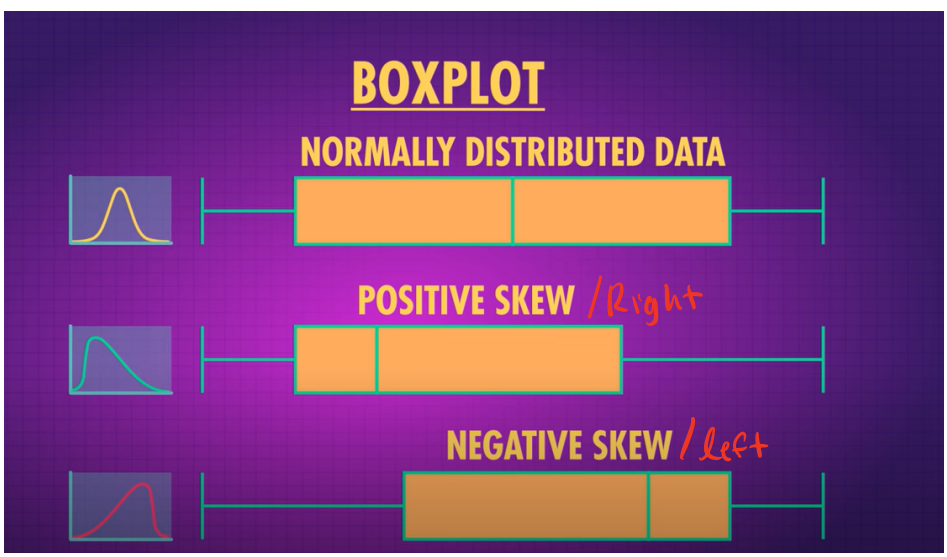
Uniform Distribution - each value in the Distribution has the same frequency

ex: dice roll, coin flip

ex] Draw the Distribution Curve for a coin flip

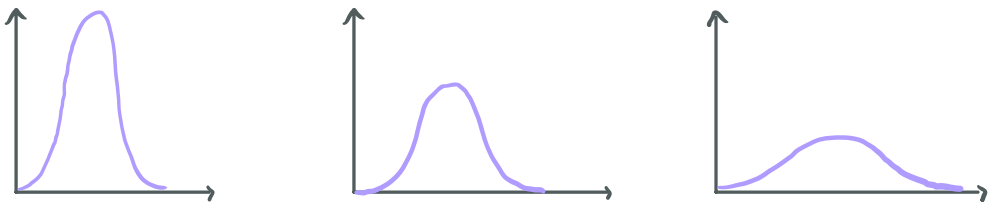


Box Plots



https://www.youtube.com/watch?v=bPFNxD3Yg6U

ex] Compare the Standard Deviations of the following Distributions

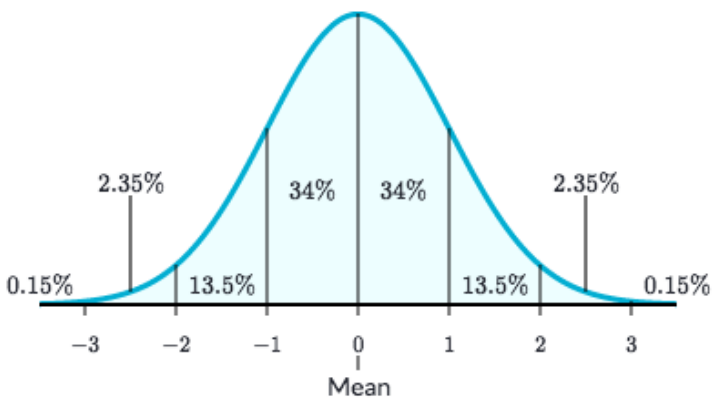


Smaller → Larger

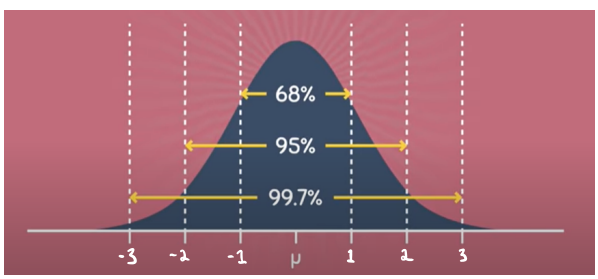
Standard Deviation

Empirical Rule / 68-95-99.7 Rule

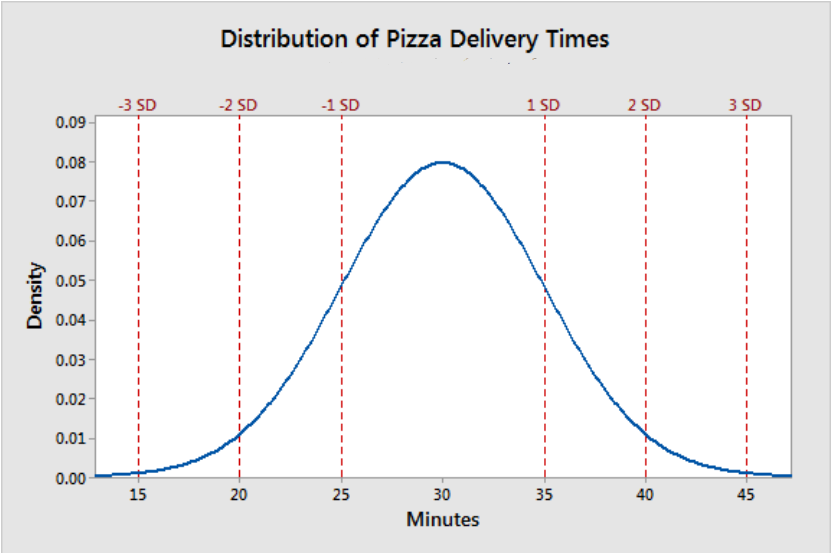
✱ Only with Normal Distributions (all types of Normal Distributions)



- 68% of our data lies within 1 standard Deviation
- 95% lies within 2 sd
- 99.7% lies within 3sd



ex] Consider the Distribution Curve below



(a) 68% of Pizza Delivery times are between 25 and 35 mins

(b) 95% of Pizza Delivery times are between 20 and 40 mins

(c) 99.7% of Pizza Delivery times are between 15 and 45 mins