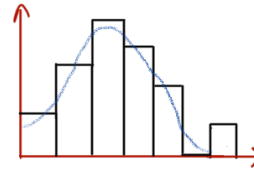


8 - graphical statistics

histograms

- shows the shape of the pmf / pdf
- split range of data into equal bins and count how many observations fall into each bin



stem and leaf plot

- they also show how the data are distributed within columns

ex: 1.2, 1.5, 1.8, 2.3, 2.4, 3.2, 3.6, 4.1, 5.3, 5.6

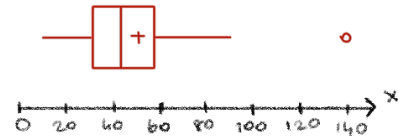
S	L
1	2 5 8
2	3 4
3	2 6
4	1
5	3 6

leaf unit: 0.1
stem unit: 1

boxplot

- the main descriptive statistics of a sample can be represented
- a box between the first and the third quartiles
- a line inside a box for a median
- extend whiskers to the smallest and the largest observations (biggie, favori)
- ↳ representing five-point summary
- five-point summary = $(\min X_i, \hat{\theta}_1, \hat{M}, \hat{\theta}_3, \max X_i)$
- a sample mean $\bar{X} \rightarrow$ a dot or a cross

$\bar{X} = 48.23$ $\min X_i = 9$ $\hat{\theta}_1 = 34$
 $\hat{M} = 42.5$ $\hat{\theta}_3 = 59$ $\max X_i = 139$
 outlier = 139 \rightarrow max: 89



scatter plots and time plots

- relationship between two variables

X	10	6	6	5
Y	0	2	0	4