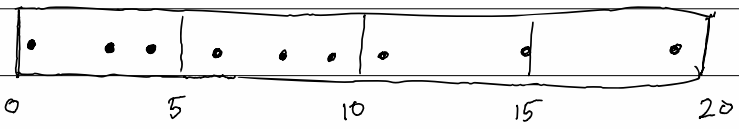
Chapter 4

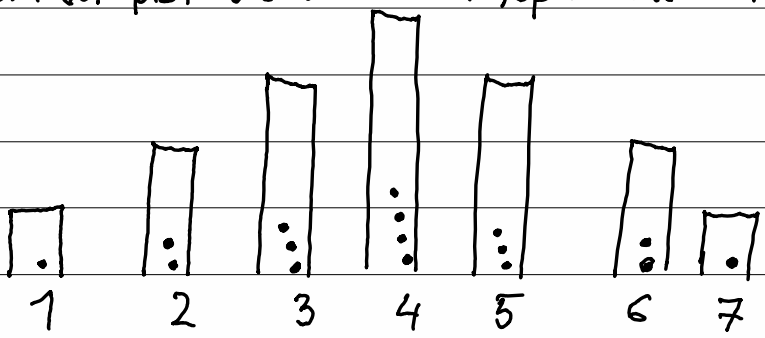
ex 4.1 make a dot plot

- provides good visual insight into the pattern and distribution of the data set.



ex 4.2 draw a bar chart

- in case there are more than 1 observations per value (dots from dot plot are stacked on top of each other to form bar



Histogram

- shapes of distribution - place where most data are located is called mode.
- mode is often located in or near the middle
- if mode has 1 peak ∩ = unimodial
- if mode has two peaks ∩∩

- we can also examine the symmetry of the distribution

a) symmetric

b) skewed

the ends of the data are called a tail

ex 4.3 2) a) frequency distribution table

- intervals	5.5 - 5.9 |||||||||	(9)
	6.0 - 6.4 |||||||||||	(10)
	6.5 - 7.0 ||	(2)

- remember to plot + describe in python notebook!

- import numpy + use arange on bins

intervals 0.1

Relational graphs

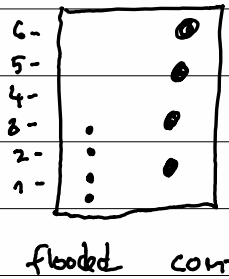
- used for 2+ more datasets
- good for examining relationship between variables

- types of graphs
  - stacked bar chart — f. ex categorical variables
  - stacked relative frequency

- both are good at whether the distrib. of observations are similar or different as a function of the other cat. variable

- for comparing the numerical var. to categorical var.
  side-by-side dot plot can be used

Ex)



- for larger data sets -> aligned histogram
- stacked histogram can be used to show counts

Scatter plot

- for comparing 2 numerical values
- data should be paired