





Nile - time series data  
variable measurements  
over periods of time

frequency

high frequency data

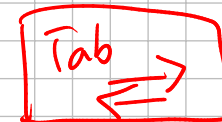
- finance  
e.g. stock prices

NYSE ATI ...

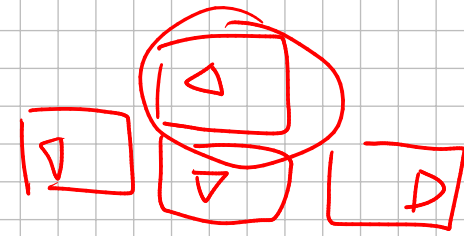
- trade transactions  
supply chain issues  
money transfers.

UNCTAD, WTO

auto complete



CSV text format  
comma separated values  
text editor



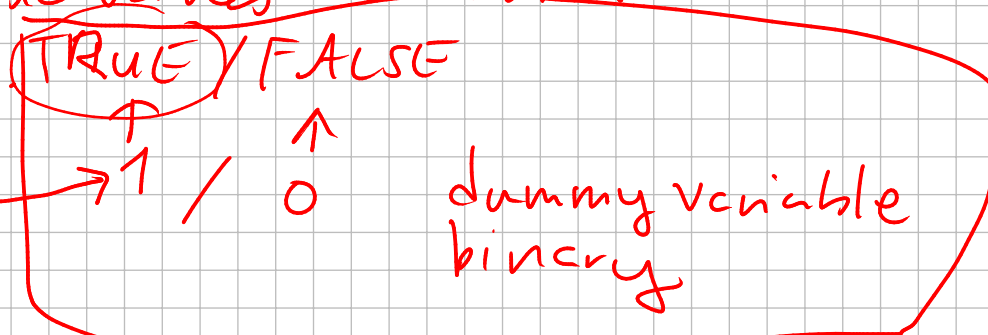
path  
Windows { C:\Windows | ...  
          { C:\Users\Andrew\Documents | ...

macOS /Users/Andrew | ...

name of file

ggg code

table : variables whose possible values are limited  
discrete variables e.g. TRUE / FALSE



recode

e.g. categorical variable

ask for a codebook!

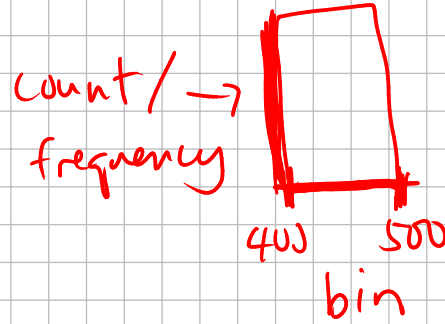
→ o/s? Windows, macOS, Linux, ...  
" 1 " 2 " 3 ...

principal-agent problem

histogram

visualization  
how to obtain?

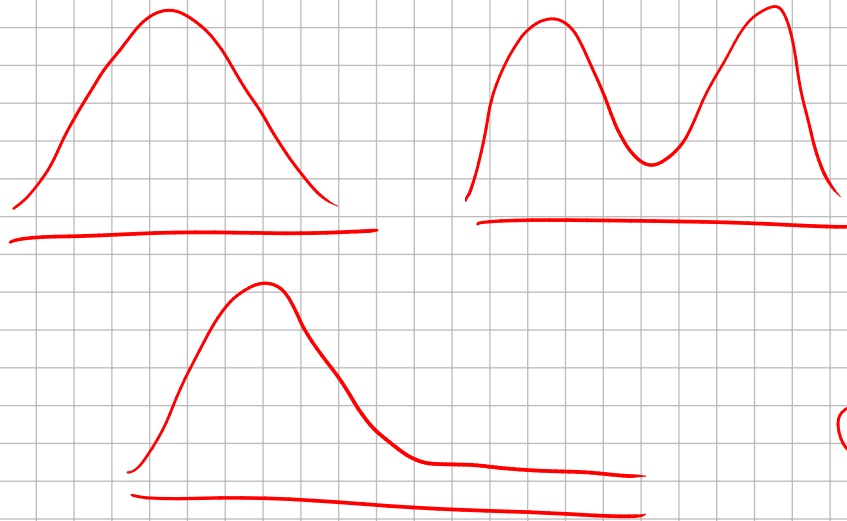
In R, NA missing



bins?  
frequency for every bin

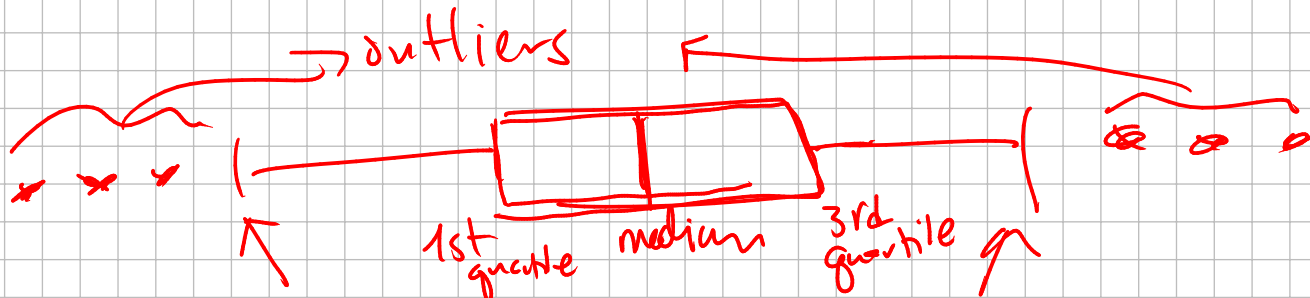
99

Old Faithful



$$\text{Area} = \underbrace{\text{length}} \times \underbrace{\text{width}}$$

count = length of interval width



whiskers → min/max or boundaries outside of which data will be classified as outliers.