Topics

- TECHNICAL
  - data types / structures
    - int, float, str, bool, datetime
    - list, tuple, set, dict
  - control flow
    - if / elif / else
    - for
    - while
    - comprehensions
  - functions
    - pass by reference / value
    - scope
    - variable number of inputs: *args, **kwargs
    - positional vs keyword
    - recursion
    - lambdas
    - decorator
  - classes
    - attributes / methods
    - self keyword
    - overloading
    - inheritance / composition
    - UML
  - packages / modules
    - import, from, as keywords
  - I/O
    - file formats: csv / tsv, json
    - open / close
    - context managers with / as keywords
  - numpy
    - shape
    - accessing entries
      - masking
      - iterating
      - indexing / slicing
    - aggregation along rows / columns
    - broadcasting
    - vectorized operations
  - matplotlib
    - One Quantitative Variable
      - Histogram
        - Rug plot
        - Density Estimate
      - Boxplot
    - One Qualitative Variable
      - Bar chart
    - One Qualitative / One Quantitative Variable

- ▪ Strip Plot
  - • Multiple Quantitative Variables
    - ▪ Scatter plots
      - • Color, Size, Transparency, Shape
    - ▪ Juxtaposed Boxplots
    - ▪ Superimposed Histograms
- ○ pandas
  - • Series and DataFrames
  - • loc / iloc
  - • missing / duplicate data
  - • multiple indices and stack / unstack
  - • dates and times
  - • merge / join
    - ▪ inner / outer
    - ▪ left / right
    - ▪ cross
  - • group
    - ▪ agg, filter, transform
    - ▪ pivot
    - ▪ melt
- ○ SQL
  - • schema / primary key
  - • select, from, where, order by, limit
  - • group by, having
  - • join, union
- ○ pattern matching
  - • glob-ing
    - ▪ *, ?
    - ▪ [characters],[!characters]
  - • regex
    - ▪ re: findall, sub, split, search
    - ▪ \s, \S, \d, \D, \w, \W
    - ▪ ., [], [^]
    - ▪ *, +, ?, {n}, {n,m}
    - ▪ |, \ , ^, $
- ○ bash
  - • Files and Directories
    - ▪ file
    - ▪ pwd
    - ▪ ls
    - ▪ cd
    - ▪ du
  - • Operations on Files
    - ▪ touch
    - ▪ mv
    - ▪ cp
    - ▪ rm
    - ▪ chmod

- Operations on Directories
  - rmdir
  - mkdir
  - grep
- View Files
  - less
  - head
  - cat
  - wc
- Keyboard Shortcuts
  - CTRL C
  - CTRL D
  - . and ..
- Scripts
  - redirect >>, >,
  - pipe |
  - for, while
  - if
  - variables: ${…}, ${1} and shift, ${@}
- git
  - clone / init / fork
  - working directory/ staging / local / remote + stash
  - add / commit
  - committed, staged, modified, untracked
  - diff and log
  - branch and checkout
  - merge and stash
  - push / pull
- exceptions
  - raise
  - assert
  - try / except / else / finally

- CONCEPTUAL
  - files and directories
    - path
    - environment variables
    - package manager
  - split / apply / combine
  - tabular vs nested data
  - database
    - scalable
    - declarative vs imperative
    - atomicity, consistency, isolation, durability
  - storage and running time
    - chunking
    - map / reduce

- lazy evaluation
  - references
    - mutable vs immutable
    - copy vs view
  - VCS
    - local, centralized, distributed
    - merging
      - merge vs rebase
      - merge conflict
      - three way merge vs fast forward merge
    - revert vs reset
    - push vs pull request
  - classes
    - inheritance
    - encapsulation
  - functions:
    - abstraction
    - decomposition
    - recursion: divide and conquer
  - error catching
    - debugging
    - logging
    - validating
    - testing