# Project: Presidential Election Database (PED)

2.1 Extract-Transform-Load

* Scraping → <http://docs.python-guide.org/en/latest/scenarios/scrape/>
* Parsing
* Cleansing
  + Eliminating duplicates,
  + Conflicts
  + ….
* Sites:
  + <https://www.loc.gov/rr/program/bib/elections/statistics.html>
  + <http://www.globalelectionsdatabase.com/index.php/index>
  + <http://www.270towin.com/historical-presidential-elections/>
  + Population: <https://fusiontables.google.com/DataSource?dsrcid=225439#rows:id=1>
  + <http://www.infoplease.com/ipa/A0781450.html>
* Guides:
  + Python <http://www.oracle.com/technetwork/articles/dsl/python-091105.html>
  + Ruby <http://www.oracle.com/technetwork/articles/marx-ruby-092465.html>

2.2 Queries

* 2.2.1 Query of a given year:
  + Candidates
    - <http://www.270towin.com/historical-presidential-elections/>
  + Their party affiliation
    - <http://www.infoplease.com/ipa/A0781450.html> (party/electoral/popular)
    - <http://www.270towin.com/historical-presidential-elections/>
  + Percentage of votes received
    - <https://fusiontables.google.com/DataSource?dsrcid=225439#rows:id=1>
    - <http://www.270towin.com/historical-presidential-elections/>
  + Winner - ‘Yes/No’
  + Polls prior to election
  + ….
  + Pseudocode:
    - QGY: Query a given year
      * Elections (year)
      * participants
* 2.2.2 Query of a given President/Candidate:
  + Name
  + Hometown
    - <https://en.wikipedia.org/wiki/List_of_Presidents_of_the_United_States_by_home_state>
  + Years he ran
    - Can be gotten from the previous query
  + The results
    - Won or lost previous query
  + Age → at least 35 years old
    - <https://en.wikipedia.org/wiki/List_of_Presidents_of_the_United_States_by_age>
  + gender
  + networth
* 2.2.3 Query for re-elected on non-contiguous times:
  + Re-elected after losing one or more elections in between
    - Can be gotten from previous
* 2.2.4 Swing Candidates:
  + Ran for one party one time
  + And another party another time
    - For such return the same info as in 2.2.2
* 2.2.5 Party Historical Query:
  + Results of a given party throughout the elections’ history.
  + Include performance stats (wins/losses) and electoral votes each time
    - <https://en.wikipedia.org/wiki/List_of_political_parties_in_the_United_States>
* 2.2.6 Miscellaneous Queries:
  + 4 Interesting queries of your choice
    - Candidate → slogans
    - Number of candidates from same hometown → google scraping
      * <https://en.wikipedia.org/wiki/List_of_United_States_presidential_candidates>
    - Donations
      * <http://www.fec.gov/disclosurep/pnational.do>
    - Difference of votes following the election
    - Net Worth of presidents
      * <http://www.huffingtonpost.com/2011/02/21/the-net-worth-of-the-amer_n_825939.html>
      * <http://www.usatoday.com/story/news/politics/elections/2015/08/26/24-7-wall-st-net-worth-presidential-candidates/32409491/>
    - Cost for campaign
      * <http://metrocosm.com/the-history-of-campaign-spending/> → spending from 1960-2012
      * <http://www.motherjones.com/mojo/2012/02/historic-price-cost-presidential-elections> → pres campaign cost
  + Subject to the approval of TA

3.1 Web-Server Setup (Apache HTTP Server 2.4.23)

* [~~https://httpd.apache.org/download.cgi~~](https://httpd.apache.org/download.cgi)
* [~~http://httpd.apache.org/docs/2.4/install.html~~](http://httpd.apache.org/docs/2.4/install.html)
* [~~https://jason.pureconcepts.net/2015/10/install-apache-php-mysql-mac-os-x-el-capitan/~~](https://jason.pureconcepts.net/2015/10/install-apache-php-mysql-mac-os-x-el-capitan/)
* [**https://www.apachefriends.org/download.html**](https://www.apachefriends.org/download.html)
* 3.1.1 Interaction PHP → Python (or HTML → Javascript)

MYSQL Reference Manual

* <https://dev.mysql.com/doc/refman/5.7/en/assigning-passwords.html>
* <http://mysql-python.sourceforge.net/MySQLdb.html> python-MySQLdb
* <https://www.tutorialspoint.com/mysql/mysql-select-query.htm> MySQL cmd guide