Summary of week ending 1/22/16:

* Based on analysis of calculations necessary to groom 2.5 MM+ records
  + Introduced the Python programming language to assist processing records.
  + Developed script(s) to address:
    - Contention/claims based records processing
    - Noise (errata) in resulting records
      * Processed 400 K of records in a few minutes
  + Important observation on how to perceive the relativity of prior when considering the target claim, therefore:
    - Currently developing and testing script(s) to:
      * Process the decision based information
      * Targeting a combination query to produce the best possible view which simulates the feature data vector for ear.

During the week, activities included:

* Working claim extraction workflows vs outcomes
* Figuring out and testing mannerisms & ways to get rid of noise
  + Most focused in the date/ dates
    - Determined high noise in deferred, not service connected
  + Working out flows, based on results of the fields and the ordering of fields
    - * Rating Corporate Claim Table – pulling from
* Engineering notebooks, present:
  + Several wrong assumptions – or- incorrectly written statements
    - Team has found a number of things that Notebooks stated
      * In comparing data to notebooks
        + to not be true.
* Identified high percentage of:
* One claim to many profiles
  + Two dates that matter most:
    - Most recent “begin” date – prior to; date of claim
    - Promulgation date
      * Base on Promulgation (The date it was actually authorized) date – as the driver
  + Date and id give you your unique id
  + Prom & agg date – by two things that represent decision points – prom date is better determinant
  + Claims can have more than 1 rating profile
  + 3 Variables for Ear
  + Comfortable at/ with 400 K records for Ear as basis