Inputs:

1. Unplatform json dumps from user sessions – Assessment, Log and Asset folders + sqlite db with school configuration and session ids
2. Item folder json files which have the information about the questions and answer choices – map this to ids

Outputs:

A: csv or xls ouput tables:

1. by session – by lesson CLIx Time
   1. CLIx Time : question, answer, feedback provided
   2. Word Play : sentences created by session
   3. Reflections : question, answer
2. by lesson summary statistics of questions served, correct answers, incorrect answers rolling up at the school level
3. by subject and by lesson summary statistics

B: summary statistics of artefacts created by session (how many OST creations, how many audio files were recorded and so on); rolled up at school level

How to do this:

1. read this document to get an understanding of how to traverse the folder structure for the json dumps: <https://docs.google.com/document/d/1OZ59bAxDM9Crn2glWpNjIA7QC9QLUOAwsNZ7wGHGc-4/edit#heading=h.2gx42pnx7ri9>
2. refer to the graphical data model in the last page of this document. Focus only on Assessment Taken, Assessment Section, Asset, Asset Content, Items folders for now
3. Scripting for Assessment section:
   1. Read the jsons and convert them into csvs
   2. Merge all csvs in a folder into 1 file (see merged csv file for example in attached)
   3. discard all fields in the merged csv, other than the ones highlighted in yellow in the attached xls (Item\_Merged\_CSV\_after\_talking\_to\_Cole.xls – worksheet “required\_fields”)
   4. need to get the question and response chosen ids from the Assessment section jsons – and map them back to the session ids. Session ids are in Assessment Taken jsons – link back recursively
   5. map the question ids and response choice ids with the actual question and response selected by the student, using the jsons in the “items” folder
   6. replace the question and response ids with the actual question and answer / feedback strings /correct or incorrect answer
   7. for each session id, need to have the complete list of questions and answers / correct / incorrect / feedback
   8. merge for all session ids in one school cumulatively
4. Scripting for Assets:
   1. Identify the Assets and AssetContents back to session id which is available in Assessment Taken jsons.
   2. Count the number of assets created by type for each session id and populate a csv table for the same
   3. Roll up this csv at the school level (for all sessions cumulatively)