**ETLDashboard Documentation:**

* Repository: [R:\Enterprise Risk Dashboard\Repo\ETLDashboard](file:///R:\Enterprise%20Risk%20Dashboard\Repo\ETLDashboard)
* Scripts used in application:
  + **GenerateNewETL.py: Using dataset present in file (csv, Excel, any file type)**

1. **Generate column report detailing appropriate T-SQL types, attributes, relationships between columns.**
2. **Generate .sql table definition file based on estimated column types.**
3. **Append new ETL to DynamicETL.Service appsettings-template.json file with default attributes.**
   * **TestETLPipeline.py: Once new ETL has been configured in DynamicETL,**

* **Every script has .json file with same name as script (ex: GenerateNewETL.json).**
* **Requirements**;
  + Python 3.6 installed with packages listed in “requirements.txt”.
* **Command line usage:**
  + “python <PathToScript>”
  + Requires python.exe be included in PATH environment variable.

**Script Arguments:**

* **GenerateNewETL.py:**
  + Required Arguments (in GenerateColumnAttributesReport.json):
    - **“data”**: Dictionary containing following keys:
      * **“path”**: Path to folder containing one or more files of different dates but same dataset.
      * **“sheets” (Optional):** List of strings containing sheets you wish to analyze. If used then “path” must point to an Excel file, and each sheet will be viewed as its own ETL.
    - **“reportpath”:** Pathto generated report detailing column attributes. Must point to Excel file.
    - **“filedatereg”:** Dictionary containing following keys:
* **TestETLPipeline.py:**
  + Required Arguments (in TestETLPipeline.json):
    - **“etlname”:** Name of ETL. Must be configured in DynamicETL.Service in QA/UAT/STG and in local copy.
    - **“filedate”**: Date of file used to test pipeline (listed in **postargs.json** or in optional **“comparefile”).**
    - **“reportpath”:** Path to output report comparing input data and output data. Must be an .xlsx file.
    - **“postargspath”:** Path to **postargs.json** file containing REST arguments POSTed to DynamicETL.WebAPI. Required whether testing locally or QA/UAT/STG.
    - **“testmode”:** One of LOCAL/QA/UAT/STG (case insensitive) denoting which version of the ETL you are testing.
  + Optional Arguments:
    - **“pkey”:** List ofstrings or string denoting which columns you want to use to compare two rows, i.e. the primary key of the dataset.
    - **“ignorecols”:** List of strings or string denoting which columns you want to ignore when comparing two rows. FileDate and RunDate columns are always ignored.
    - **“comparefile”:** Path to file containing data you want to compare to data loaded into ETL table, in case you do not want to use the file listed in postargs.json.