

Troubleshooting Failed File Uploads

After loading a file into the Data Lake you may see a status other than “SUCCESS”. Depending on the type of file you are trying to load you may see a failed status:



From a high level there are several reasons why this may occur. Most issues can be traced back to the file and how it is configured. For context on what is happening in the background when a file is loaded, visit [HERE](#) for a detailed visualization of the ingestion process.

High Level Process:

File Loaded	→	Loaded to S3	→	Schema Read	→	Process Magic	→	Table Created
-------------	---	--------------	---	-------------	---	---------------	---	---------------

Supported File Types

Structured

Structured files are files where the schema is detectable using standard relational query or analysis tools. The accepted file formats include CVS, XLSX, XLS and JSON. Files of these types will be processed using schema detection and will produce a queryable asset once the process is complete.

NOTE

Other structured files may be uploaded to the data lake for raw storage purposes (S3), but no schema detection will occur. This will result in not being able to query your data.

Unstructured

Unstructured files are considered to be lacking a detectable schema using standard relational query or analysis tools.

Things to Check:

Make sure the document you are trying to load does not have a header like in the example below. When the file is processed, the engine will fail trying to determine the schema due to the header.

