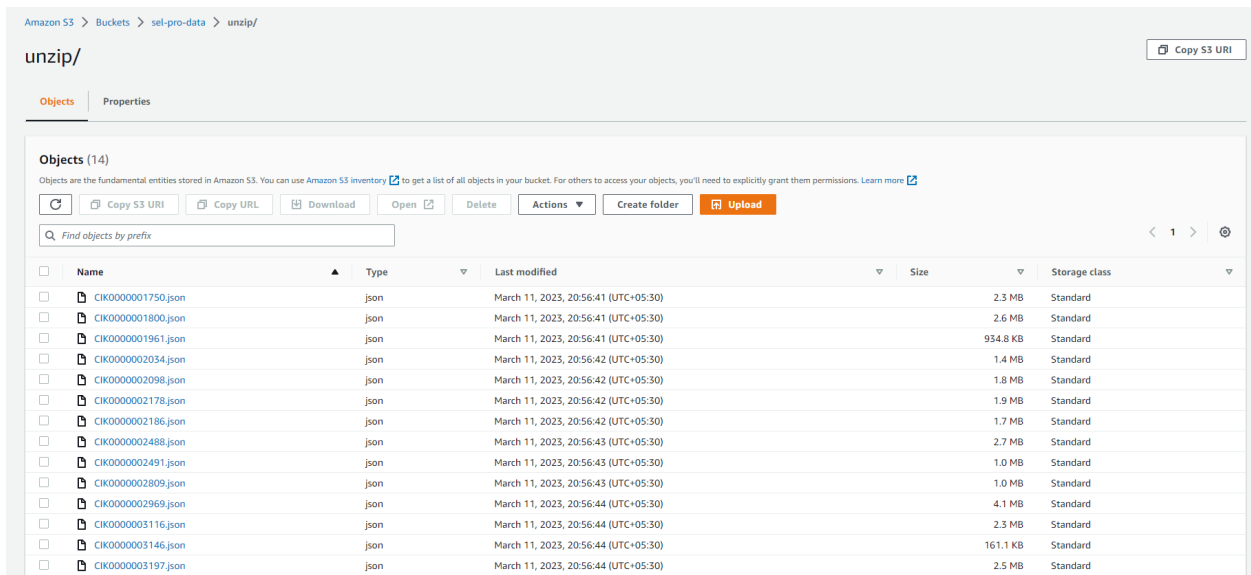
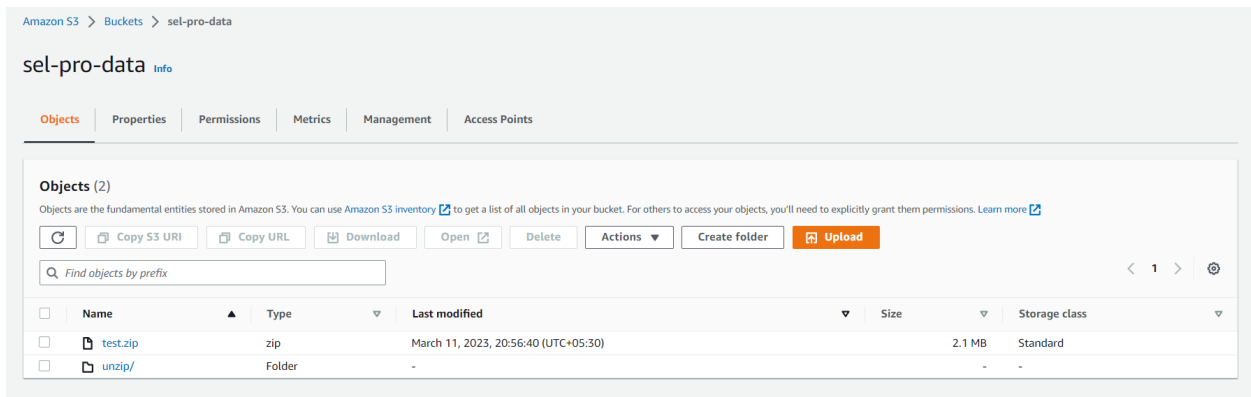


Project Title: Data Migration and Transformation Tool

Technologies: Python, Requests, Zipfile, boto3, Amazon (S3, DynamoDB), MongoDB

S3 To MongoDB

1. Uploading and unzipping in S3



2. Moved to MongoDB

project_de.big_json

Documents Aggregations Schema Explain Plan Indexes Validation

Filter Type a query: { field: 'value' } Reset Find More Options

ADD DATA EXPORT COLLECTION

1-14 of 14

#	big_json	_id ObjectId	cik Int32	entityName String	facts Object	
1	ObjectId("640c3dbf956ffde8d45...")	1750	"AAR CORP"	{ } 2 fields		
2	ObjectId("640c3dc2556ffde8d45...")	1800	"ABBOTT LABORATORIES"	{ } 2 fields		
3	ObjectId("640c3dc4556ffde8d45...")	1961	"JWORLDS INC."	{ } 3 fields		
4	ObjectId("640c3dc5556ffde8d45...")	2034	"ACETO CORP"	{ } 2 fields		
5	ObjectId("640c3dc7556ffde8d45...")	2098	"ACHE UNITED CORP"	{ } 2 fields		
6	ObjectId("640c3dc8556ffde8d45...")	2178	"ADAMS RESOURCES & ENERGY, INL..."	{ } 3 fields		
7	ObjectId("640c3dca556ffde8d45...")	2186	"BK TECHNOLOGIES CORPORATION"	{ } 2 fields		
8	ObjectId("640c3dc3556ffde8d45...")	2488	"ADVANCED MICRO DEVICES, INC"	{ } 3 fields		
9	ObjectId("640c3dc556ffde8d45...")	2491	"BALLY TECHNOLOGIES, INC."	{ } 2 fields		
10	ObjectId("640c3dce556ffde8d45...")	2889	"AGNICO EAGLE MINES LIMITED"	{ } 3 fields		
11	ObjectId("640c3dcf556ffde8d45...")	2969	"AIR PRODUCTS AND CHEMICALS, ..."	{ } 3 fields		
12	ObjectId("640c3dd2556ffde8d45...")	3116	"AKORN INC"	{ } 2 fields		
13	ObjectId("640c3dd3556ffde8d45...")	3146	"SPIRE ALABAMA INC"	{ } 2 fields		
14	ObjectId("640c3dd4556ffde8d45...")	3197	"CECO ENVIRONMENTAL CORP."	{ } 2 fields		

S3 To DynamoDB

1. Uploading and unzipping in S3

Amazon S3 Buckets sel-pro-data

sel-pro-data Info

Objects Properties Permissions Metrics Management Access Points

Objects (2)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload

Find objects by prefix

	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	test.zip	zip	March 11, 2023, 21:31:46 (UTC+05:30)	27.0 KB	Standard
<input type="checkbox"/>	unzip/	Folder	-	-	-

Amazon S3 > Buckets > sel-pro-data > unzip/

unzip/

Copy S3 URI

Objects | Properties

Objects (4)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload

Find objects by prefix

	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	CIK0000003982.json	json	March 11, 2023, 21:31:47 (UTC+05:30)	8.1 KB	Standard
<input type="checkbox"/>	CIK0000005187.json	json	March 11, 2023, 21:31:47 (UTC+05:30)	56.2 KB	Standard
<input type="checkbox"/>	CIK0000015393.json	json	March 11, 2023, 21:31:48 (UTC+05:30)	51.4 KB	Standard
<input type="checkbox"/>	CIK0000016104.json	json	March 11, 2023, 21:31:48 (UTC+05:30)	43.4 KB	Standard

2. Moved to DynamoDB

DynamoDB > Items > jsonData

Tables (1)

Any table tag

Find tables by table name

jsonData

jsonData

Autopreview View table details

▼ Scan or query items

Scan Query

Select a table or index

Table - jsonData

Select attribute projection

All attributes

Filters

Run Reset

Completed. Read capacity units consumed: 18

Items returned (4)

Copy Actions Create item

	clk	entityName	facts
<input type="checkbox"/>	15393	BUREAU OF NATIONAL AFFAIRS INC	{ "us-gaap": { "TM": { "IncreaseDecreaseOtherOperatingCapitalNet": { "description": { "S": "Amount of increase (decrease) in operating assets after deduction of operating liabilities class..."
<input type="checkbox"/>	5187	WYETH	{ "us-gaap": { "TM": { "AdjustmentsToAdditionalPaidInCapitalTaxEffectFromShareBasedCompensation": { "description": { "S": "Amount of increase in additional paid in capital (APIC) result..."
<input type="checkbox"/>	3982	ALLIS CHALMERS ENERGY INC.	{ "us-gaap": { "TM": { "PreferredStockSharesAuthorized": { "description": { "S": "The maximum number of nonredeemable preferred shares (or preferred stock redeemable solely at the op..."
<input type="checkbox"/>	16104	CAGLES INC	{ "us-gaap": { "TM": { "PropertyPlantAndEquipmentGross": { "description": { "S": "Amount before accumulated depreciation, depletion and amortization of physical assets used in the nor..."