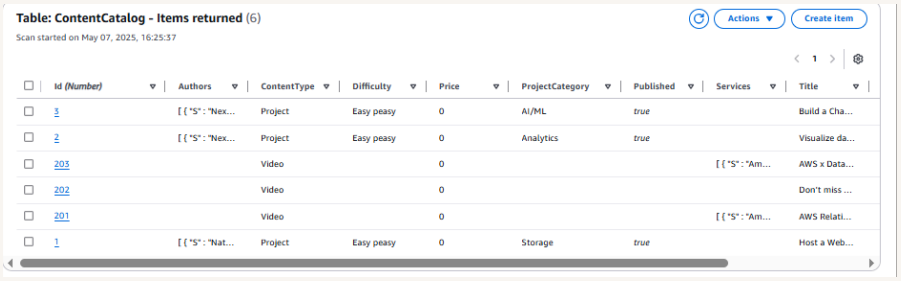
nextwork.org

Load Data into a DynamoDB Table

AM ampahben3@gmail.com



AM
ampahben3@gmail.com
NextWork Student
nextwork.org

Introducing Today's Project!

What is Amazon DynamoDB?

Amazon DynamoDB is a NoSQL database that provides fast performance, automatic scaling, and high availability. It’s great for real-time applications. Your project involved loading data into DynamoDB for efficient storage and retrieval.

How I used Amazon DynamoDB in this project

Amazon DynamoDB was used to load data into tables like ContentCatalog, Forum, Post, and Comment using batch-write-item operations. This ensured efficient storage and fast retrieval while supporting scalability.

One thing I didn't expect in this project was...

The data-loading process seemed like it wasn’t working, and I was stuck for a long time. Later, I realized it had actually worked all along. That was unexpected

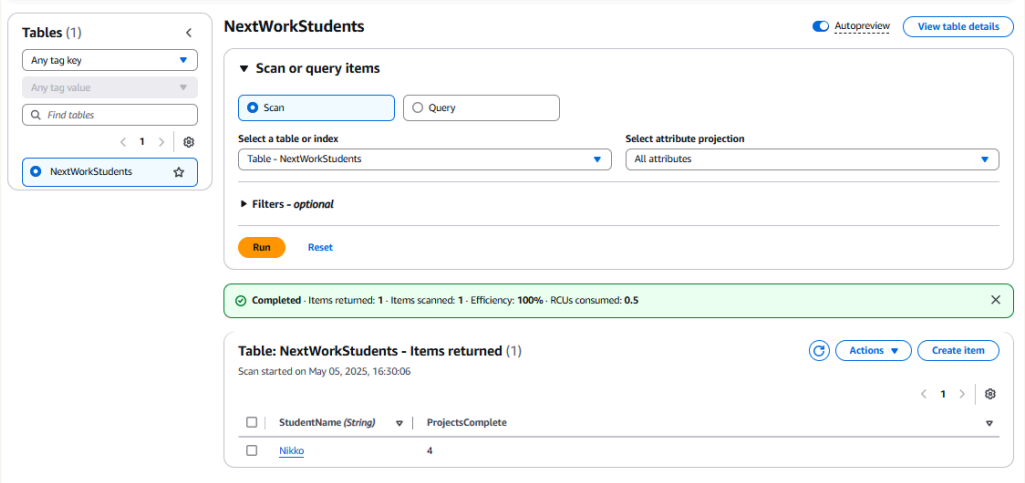
This project took me...

This project took more than an hour, mostly spent figuring out why there was no meaningful output while loading the database. I did a lot of Googling and prompting before realizing the issue.

Create a DynamoDB table

DynamoDB tables organize data using items, each with unique attributes. Unlike traditional databases, items aren't fixed in rows and columns. Instead, every item can have different attributes, making DynamoDB highly flexible for dynamic data storage.

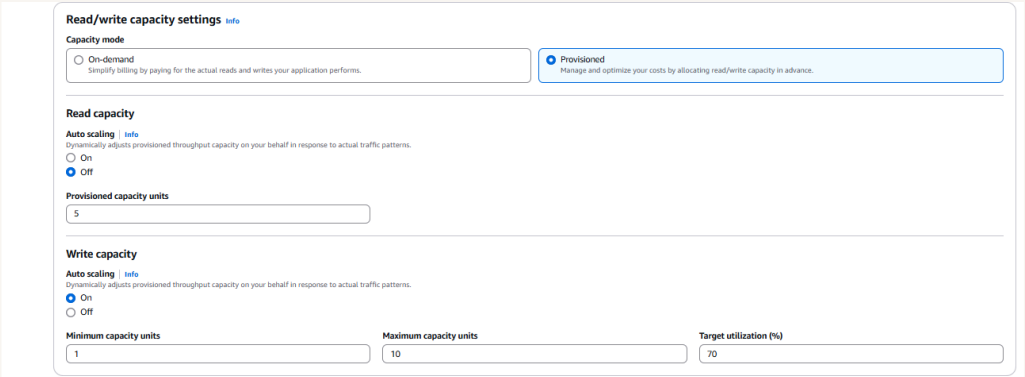
An attribute is an item that provides additional information about something, in this case, a key-value pair in Amazon DynamoDB. Attributes define an item's properties, helping store and organize data efficiently for easy querying and management.



Read and Write Capacity

Read capacity units (RCUs) and write capacity units (WCUs) are Amazon DynamoDB metrics defining database throughput. RCUs measure reads per second, while WCUs determine writes. These units help optimize performance and manage costs efficiently.

Amazon DynamoDB's Free Tier covers 25 RCUs, 25 WCUs, and 25GB of storage per month. I turned off auto scaling to prevent unexpected costs, as it automatically adjusts throughput. Managing capacity manually keeps expenses low.



Using CLI and CloudShell

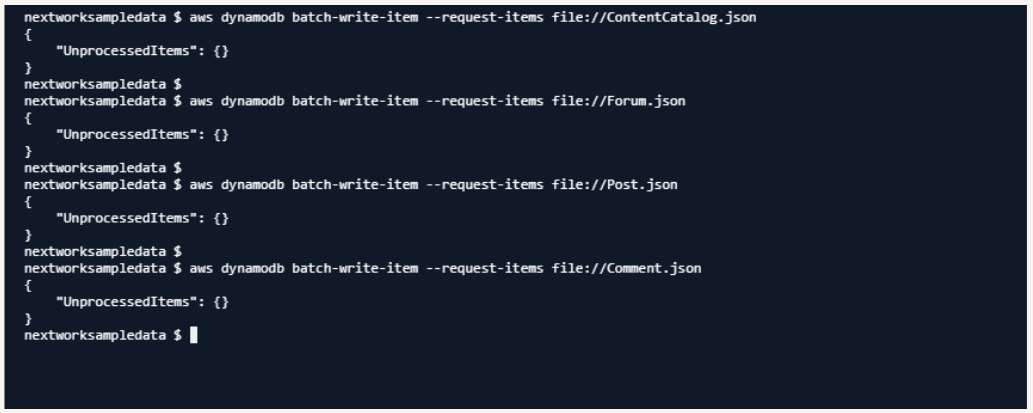
AWS CloudShell is an environment that uses commands to help perform tasks like accessing resources in AWS. It provides a browser-based terminal for running AWS CLI commands, making it easy to manage services without installing tools locally.

AWS CLI is a command-line tool for managing AWS services. It allows users to run commands to configure, control, and automate resources without using a graphical interface, making cloud management more efficient and flexible.

I ran a CLI command in AWS CloudShell that created four tables: ContentCatalog (Id attribute), Forum (partition key Name), Post (partition key ForumName, sort key Subject), and Comment (partition key Id, sort key CommentDateTime).

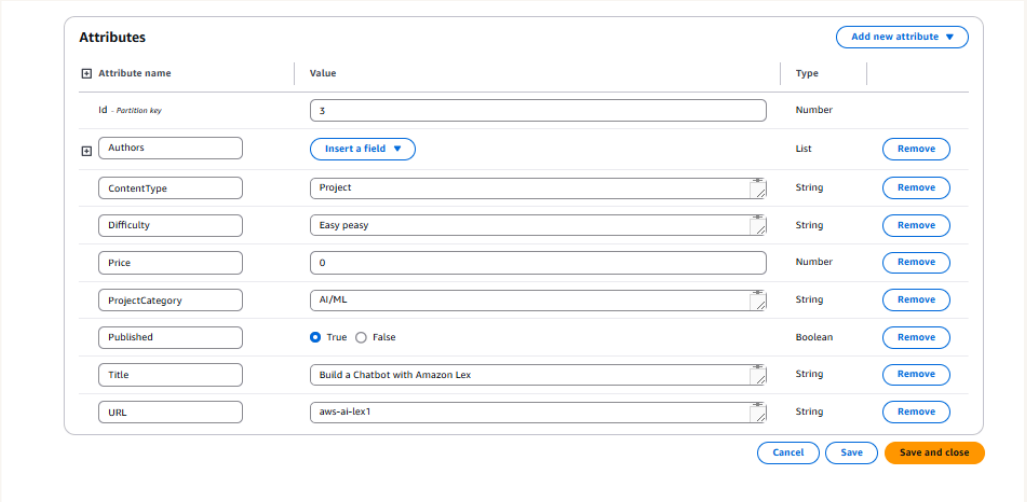
Loading Data with CLI

I ran a CLI(Command Line Interface) command in AWS CloudShell that executed multiple batch-write-item operations to insert data into ContentCatalog, Forum, Post, and Comment tables in DynamoDB. These commands were run twice to ensure data insertion.



AM
ampahben3@gmail.com
NextWork Student

Observing Item Attributes



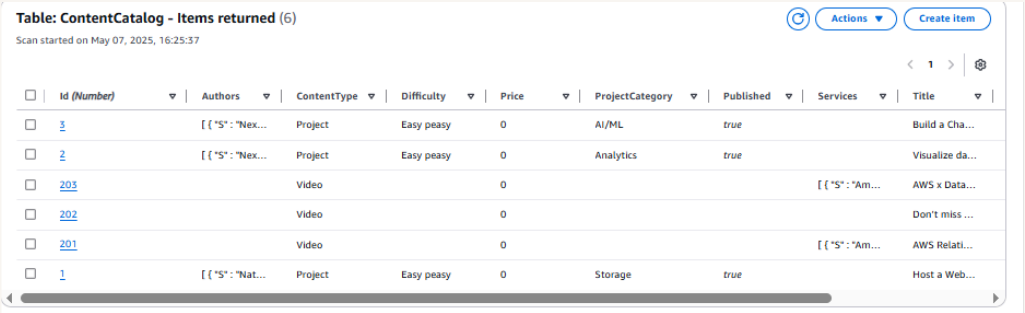
I checked a ContentCatalog item, which had the following attributes: Id, Authors, ContentType, Difficulty, Price, ProjectCategory, Published, Title, and URL. It describes a project titled Build a Chatbot with Amazon Lex, categorized under AI/ML.

I checked another ContentCatalog item, which had a different set of attributes: Id, Authors, ContentType, Duration, Price, Published, Title, and URL. It describes a video resource with specific metadata, including its duration and publication status.

Benefits of DynamoDB

DynamoDB allows items to have unique attributes, unlike relational databases where all rows must follow the same schema. It also uses partition keys for fast retrieval, while relational databases scan entire tables, which can slow performance.

Another benefit over relational databases is speed because DynamoDB uses partition keys to quickly locate data, while relational databases must scan entire tables, slowing performance.



The place to learn & showcase your skills

Check out nextwork.org for more projects