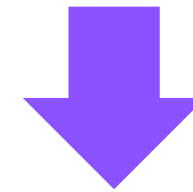


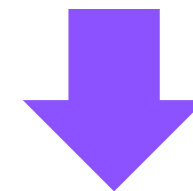
# Presentation on Amazon Kinesis Data Stream

By  
Tife Olatunji  
Sharon Mbaegbu

Put record API



**Amazon Kinesis**  
Data Streams



Get record API

# Data Upload Process



## Sign up for AWS

Explore Free Tier products with a new AWS account.

To learn more, visit [aws.amazon.com/free](https://aws.amazon.com/free).



Root user email address  
Used for account recovery and some administrative functions

AWS account name  
Choose a name for your account. You can change this name in your account settings after you sign up.

Create an AWS account  
Install AWS CLI



Configure AWS credentials  
in the terminal



Create Kinesis stream  
using AWS CLI



Upload data to the kinesis  
stream using Python script

```
aws configure  
  
# > accesskey id  
# > secret accesskey  
# > region name (default one you picked)
```

```
~/Documents/Fall_term_23/Big_data/project  
aws kinesis create-stream --stream-name amazon-books --shard-count 1  
  
~/Documents/Fall_term_23/Big_data/project  
aws kinesis describe-stream --stream-name amazon-books
```

```
response = kinesis_client.put_record(  
    StreamName=stream_name,  
    Data=record_data,  
    PartitionKey=partition_key  
)
```

# Processing the Stream

```
response = kinesis_client.get_records(  
    ShardIterator=shard_iterator,  
    Limit=1000  
)
```

Use Python script to get record from the stream



```
for record in response['Records']:  
    # Decode the record's data with padding  
    data = json.loads(record['Data'].decode('utf-8'))  
  
    # Check if the price is a valid number (not NaN)  
    price = data.get('price')  
    if isinstance(price, (int, float, np.number)):  
        total_books += 1  
        total_price += price  
        valid_prices.append(price)
```

Analyze the records

amazon-books [Info](#) Delete

**Data stream summary**

Status Active	Capacity mode Provisioned	ARN arn:aws:kinesis:ca-central-1:186963875312:stream/amazon-books	Creation time December 11, 2023 at 12:42 EST
	Data retention period 1 day		

Applications | Monitoring | Configuration | **Data viewer** | Enhanced fan-out (0) | Data stream sharing - new | EventBridge Pipes

Shard: shardId-000000000001 Starting position: [Info](#)

Trim horizon Get records

Next records

Find records

Partition key	Data	Approximate arrival timestamp	Sequence number
3Doodler Crea...	<a href="#">{"title": "3Doodler Create Flexy 3D Printin..."}</a>	December 11, 2023 at 12:43:46 EST	49647264146580104351367613691173...
Woodstock- Co...	<a href="#">{"title": "Woodstock- Collage 500 pc Puzzl..."}</a>	December 11, 2023 at 12:43:46 EST	49647264146580104351367613691176...
Terra by Battat...	<a href="#">{"title": "Terra by Battat \u2013 2013 4 Dinosau..."}</a>	December 11, 2023 at 12:43:46 EST	49647264146580104351367613691177...
ARTSCAPE Etc...	<a href="#">{"title": "ARTSCAPE Etched Glass 24\" x 36..."}</a>	December 11, 2023 at 12:43:46 EST	49647264146580104351367613691179...
Pokemon TCG: ...	<a href="#">{"title": "Pokemon TCG: Sun and Moon Cri..."}</a>	December 11, 2023 at 12:43:46 EST	49647264146580104351367613691180...

# Summary of the Result

```
Total books: 100
Average price: $17.45
Found book: Python for Data Science For Dummies
{'Title': ' Python for Data Science For Dummies ', 'Author(s)': 'John Paul Mueller, Luca Massaron'
, 'Price': 28.26, 'Description': 'The fast and easy way to learn Python programming and statistics'
}
Found book: Python for Data Science For Dummies
{'Title': ' Python for Data Science For Dummies ', 'Author(s)': 'John Paul Mueller, Luca Massaron'
, 'Price': 41.99, 'Description': 'Let Python do the heavy lifting for you as you analyze large data
sets'}
Total books: 100
Average price: $48.77
Total books: 12
Average price: $56.73
```

```
python process.py
al number of books: 8786
al price of books: 357758.995999999085
rage price of books: 40.71921192806739
```