

# ***converting csv file into a JSONL file***

## **What is JSONL?**

- JSONL (JSON Lines) = A text file where each line is a separate JSON object.
- Unlike a normal JSON file (which is one big object or array), JSONL is easier for streaming and handling large datasets.

## **Example Conversion**

**Suppose you have a data.csv:**

id,name,age

1,Alice,23

2,Bob,30

3,Charlie,25

**The converted data.jsonl will look like:**

```
{"id": "1", "name": "Alice", "age": "23"}
```

```
{"id": "2", "name": "Bob", "age": "30"}
```

```
{"id": "3", "name": "Charlie", "age": "25"}
```

## **◆ Methods to Convert**

### **1. Using Python (with pandas)**

```
import pandas as pd
```

```
# Load CSV
```

```
df = pd.read_csv("data.csv")
```

```
# Save as JSONL
```

```
df.to_json("data.jsonl", orient="records", lines=True)
```

✓ Each row becomes one line of JSON.

### **2. Using Python (without pandas)**

```
import csv
```

```
import json
```

```
with open("data.csv", "r") as csv_file, open("data.jsonl", "w") as jsonl_file:
```

```
    reader = csv.DictReader(csv_file)
```

```
    for row in reader:
```

```
        jsonl_file.write(json.dumps(row) + "\n")
```

👉 CSV → JSONL is useful for **LLMs, APIs, and dataset training** because they read line by line instead of loading a huge JSON at once.