

Loading Data

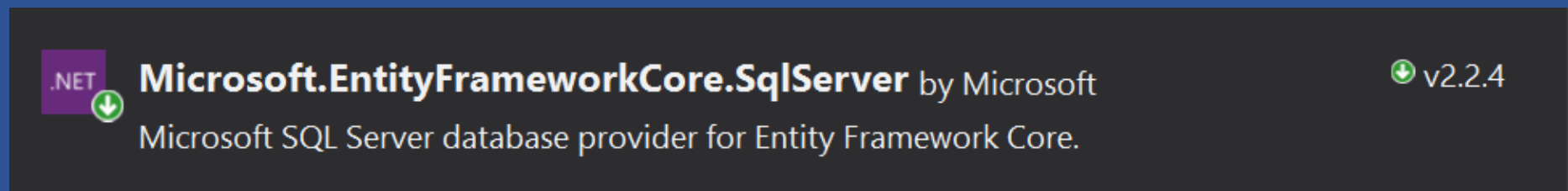
What's in this session?

- Create data model via Classes
- Load data from single file
- Load data from multiple files
- Load data from multiple folder
- Load data from Json

Create New Project

Create new .NET CORE console app project name = "Movie"

Add NuGet Packages



Create data model via Classes

```
public class HousingData
{
    // The LoadColumn attribute specifies your
    // properties' column indices
    [LoadColumn(0)]
    public float Size { get; set; }

    // Multiple columns at a time in the form of a vector
    [LoadColumn(1, 3)]
    [VectorType(3)]
    public float[] HistoricalPrices { get; set; }

    [LoadColumn(4)]
    [ColumnName("Label")]
    public float CurrentPrice { get; set; }
}
```

Load data from single file

```
//Create MLContext
MLContext mlContext = new MLContext();

//Load Data
IDataView data = mlContext.Data.LoadFromTextFile<HousingData>(
    @"E:\ml\house.csv",
    separatorChar: ',',
    hasHeader: true);

// show column Size
IEnumerable<float> sizeColumn = data.GetColumn<float>
    ("Size").ToList();
foreach (var v in sizeColumn)
    Console.Write(v + ",");


// show column Label
IEnumerable<float> labelColumn = data.GetColumn<float>
    ("Label").ToList();
foreach (var v in labelColumn)
    Console.Write(v + ",");
```

Load data from multiple files

```
//Create MLContext
MLContext mlContext = new MLContext();


//Load Data Files using * wildcard
IDataView data = mlContext.Data.LoadFromTextFile<Diamond>
    ("E:/ml/folder0/*",
     separatorChar: ',',
     hasHeader: true);

// show column Size
IEnumerable<float> sizeColumn = data.GetColumn<float>
    ("Size").ToList();
foreach (var v in sizeColumn)
    Console.WriteLine(v);
```



Load data from multiple folder

```
var path1 = @"E:/ml/Folder1/diamond1.csv";  
var path2 = @"E:/ml/Folder2/diamond2.csv";  
  
// Load Data  
IDataView data = textLoader.Load(path1, path2);  
  
// show column Size  
IEnumerable<float> sizeColumn = data.GetColumn<float>  
    ("Size").ToList();  
foreach (var v in sizeColumn)  
    Console.WriteLine(v);
```



Load data from Json

```
// Open the file to read from.
string json = File.ReadAllText(path);

// convert json to dataset
DataSet dataSet = JsonConvert.DeserializeObject
    <DataSet>(json);

// add dataset to a table
DataTable dataTable = dataSet.Tables["Table1"];

// convert dataTable to list
List<Diamond> myData = new List<Diamond>();
foreach (DataRow row in dataTable.Rows) ...

//Create MLContext
MLContext mlContext = new MLContext();

//Load Data
IDataView data = mlContext.Data.LoadFromEnumerable<Diamond>(myData);
```


What's next?

