

Modules

Retrieval

Document loaders

**CSV** 

## **CSV**

A comma-separated values (CSV) file is a delimited text file that uses a comma to separate values. Each line of the file is a data record. Each record consists of one or more fields, separated by commas.

Load CSV data with a single row per document.

```
from langchain_community.document_loaders.csv_l
import CSVLoader

loader =
CSVLoader(file_path='./example_data/mlb_teams_2
data = loader.load()
```

```
print(data)
```

```
[Document(page_content='Team:
Nationals\n"Payroll (millions)":
81.34\n"Wins": 98', lookup_str='', metadata=
{'source':
```

```
'./example_data/mlb_teams_2012.csv', 'row':
0}, lookup_index=0),
Document(page_content='Team: Reds\n"Payroll
(millions)": 82.20\n"Wins": 97',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
1}, lookup_index=0),
Document(page_content='Team:
Yankees\n"Payroll (millions)":
197.96\n"Wins": 95', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
2}, lookup_index=0),
Document(page_content='Team: Giants\n"Payroll
(millions)": 117.62\n"Wins": 94',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
3}, lookup_index=0),
Document(page_content='Team: Braves\n"Payroll
(millions)": 83.31\n"Wins": 94',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
4}, lookup_index=0),
Document(page_content='Team:
Athletics\n"Payroll (millions)":
55.37\n"Wins": 94', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
5}, lookup_index=0),
Document(page_content='Team:
Rangers\n"Payroll (millions)":
```

```
120.51\n"Wins": 93', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
6}, lookup_index=0),
Document(page_content='Team:
Orioles\n"Payroll (millions)": 81.43\n"Wins":
93', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
7}, lookup_index=0),
Document(page_content='Team: Rays\n"Payroll
(millions)": 64.17\n"Wins": 90',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
8}, lookup_index=0),
Document(page_content='Team: Angels\n"Payroll
(millions)": 154.49\n"Wins": 89',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
9}, lookup_index=0),
Document(page_content='Team: Tigers\n"Payroll
(millions)": 132.30\n"Wins": 88',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
10}, lookup_index=0),
Document(page_content='Team:
Cardinals\n"Payroll (millions)":
110.30\n"Wins": 88', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
11}, lookup_index=0),
Document(page_content='Team:
```

```
Dodgers\n"Payroll (millions)": 95.14\n"Wins":
86', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
12}, lookup_index=0),
Document(page_content='Team: White
Sox\n"Payroll (millions)": 96.92\n"Wins":
85', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
13}, lookup_index=0),
Document(page_content='Team:
Brewers\n"Payroll (millions)": 97.65\n"Wins":
83', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
14}, lookup_index=0),
Document(page_content='Team:
Phillies\n"Payroll (millions)":
174.54\n"Wins": 81', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
15}, lookup_index=0),
Document(page_content='Team:
Diamondbacks\n"Payroll (millions)":
74.28\n"Wins": 81', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
16}, lookup_index=0),
Document(page_content='Team:
Pirates\n"Payroll (millions)": 63.43\n"Wins":
79', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
17}, lookup_index=0),
```

```
Document(page_content='Team: Padres\n"Payroll
(millions)": 55.24\n"Wins": 76',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
18}, lookup_index=0),
Document(page_content='Team:
Mariners\n"Payroll (millions)":
81.97\n"Wins": 75', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
19}, lookup_index=0),
Document(page_content='Team: Mets\n"Payroll
(millions)": 93.35\n"Wins": 74',
lookup_str='', metadata={'source':
'./example data/mlb teams 2012.csv', 'row':
20}, lookup_index=0),
Document(page_content='Team: Blue
Jays\n"Payroll (millions)": 75.48\n"Wins":
73', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
21}, lookup_index=0),
Document(page_content='Team: Royals\n"Payroll
(millions)": 60.91\n"Wins": 72',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
22}, lookup_index=0),
Document(page_content='Team:
Marlins\n"Payroll (millions)":
118.07\n"Wins": 69', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
```

```
23}, lookup_index=0),
Document(page_content='Team: Red
Sox\n"Payroll (millions)": 173.18\n"Wins":
69', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
24}, lookup_index=0),
Document(page_content='Team:
Indians\n"Payroll (millions)": 78.43\n"Wins":
68', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
25}, lookup_index=0),
Document(page_content='Team: Twins\n"Payroll
(millions)": 94.08\n"Wins": 66',
lookup_str='', metadata={'source':
'./example data/mlb teams 2012.csv', 'row':
26}, lookup_index=0),
Document(page_content='Team:
Rockies\n"Payroll (millions)": 78.06\n"Wins":
64', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
27}, lookup_index=0),
Document(page_content='Team: Cubs\n"Payroll
(millions)": 88.19\n"Wins": 61',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
28}, lookup_index=0),
Document(page_content='Team: Astros\n"Payroll
(millions)": 60.65\n"Wins": 55',
lookup_str='', metadata={'source':
```

```
'./example_data/mlb_teams_2012.csv', 'row':
29}, lookup_index=0)]
```

## Customizing the CSV parsing and loading

See the csv module documentation for more information of what csv args are supported.

```
loader =
CSVLoader(file_path='./example_data/mlb_teams_2csv_args={
    'delimiter': ',',
    'quotechar': '"',
    'fieldnames': ['MLB Team', 'Payroll in mill.'Wins']
})
data = loader.load()
```

```
print(data)
```

```
[Document(page_content='MLB Team:
Team\nPayroll in millions: "Payroll
(millions)"\nWins: "Wins"', lookup_str='',
```

```
metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
0}, lookup_index=0),
Document(page_content='MLB Team:
Nationals\nPayroll in millions: 81.34\nWins:
98', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
1}, lookup_index=0),
Document(page_content='MLB Team:
Reds\nPayroll in millions: 82.20\nWins: 97',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
2}, lookup_index=0),
Document(page_content='MLB Team:
Yankees\nPayroll in millions: 197.96\nWins:
95', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
3}, lookup_index=0),
Document(page_content='MLB Team:
Giants\nPayroll in millions: 117.62\nWins:
94', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
4}, lookup_index=0),
Document(page_content='MLB Team:
Braves\nPayroll in millions: 83.31\nWins:
94', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
5}, lookup_index=0),
Document(page_content='MLB Team:
Athletics\nPayroll in millions: 55.37\nWins:
94', lookup_str='', metadata={'source':
```

```
'./example_data/mlb_teams_2012.csv', 'row':
6}, lookup_index=0),
Document(page_content='MLB Team:
Rangers\nPayroll in millions: 120.51\nWins:
93', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
7}, lookup_index=0),
Document(page_content='MLB Team:
Orioles\nPayroll in millions: 81.43\nWins:
93', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
8}, lookup_index=0),
Document(page_content='MLB Team:
Rays\nPayroll in millions: 64.17\nWins: 90',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
9}, lookup_index=0),
Document(page_content='MLB Team:
Angels\nPayroll in millions: 154.49\nWins:
89', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
10}, lookup_index=0),
Document(page_content='MLB Team:
Tigers\nPayroll in millions: 132.30\nWins:
88', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
11}, lookup_index=0),
Document(page_content='MLB Team:
Cardinals\nPayroll in millions: 110.30\nWins:
88', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
```

```
12}, lookup_index=0),
Document(page_content='MLB Team:
Dodgers\nPayroll in millions: 95.14\nWins:
86', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
13}, lookup_index=0),
Document(page_content='MLB Team: White
Sox\nPayroll in millions: 96.92\nWins: 85',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
14}, lookup_index=0),
Document(page_content='MLB Team:
Brewers\nPayroll in millions: 97.65\nWins:
83', lookup_str='', metadata={'source':
'./example data/mlb teams 2012.csv', 'row':
15}, lookup_index=0),
Document(page_content='MLB Team:
Phillies\nPayroll in millions: 174.54\nWins:
81', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
16}, lookup_index=0),
Document(page_content='MLB Team:
Diamondbacks\nPayroll in millions:
74.28\nWins: 81', lookup_str='', metadata=
{'source':
'./example_data/mlb_teams_2012.csv', 'row':
17}, lookup_index=0),
Document(page_content='MLB Team:
Pirates\nPayroll in millions: 63.43\nWins:
79', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
```

```
18}, lookup_index=0),
Document(page_content='MLB Team:
Padres\nPayroll in millions: 55.24\nWins:
76', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
19}, lookup_index=0),
Document(page_content='MLB Team:
Mariners\nPayroll in millions: 81.97\nWins:
75', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
20}, lookup_index=0),
Document(page_content='MLB Team:
Mets\nPayroll in millions: 93.35\nWins: 74',
lookup_str='', metadata={'source':
'./example data/mlb teams 2012.csv', 'row':
21}, lookup_index=0),
Document(page_content='MLB Team: Blue
Jays\nPayroll in millions: 75.48\nWins: 73',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
22}, lookup_index=0),
Document(page_content='MLB Team:
Royals\nPayroll in millions: 60.91\nWins:
72', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
23}, lookup_index=0),
Document(page_content='MLB Team:
Marlins\nPayroll in millions: 118.07\nWins:
69', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
24}, lookup_index=0),
```

```
Document(page_content='MLB Team: Red
Sox\nPayroll in millions: 173.18\nWins: 69',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
25}, lookup_index=0),
Document(page_content='MLB Team:
Indians\nPayroll in millions: 78.43\nWins:
68', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
26}, lookup_index=0),
Document(page_content='MLB Team:
Twins\nPayroll in millions: 94.08\nWins: 66',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
27}, lookup index=0),
Document(page_content='MLB Team:
Rockies\nPayroll in millions: 78.06\nWins:
64', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
28}, lookup_index=0),
Document(page_content='MLB Team:
Cubs\nPayroll in millions: 88.19\nWins: 61',
lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
29}, lookup_index=0),
Document(page_content='MLB Team:
Astros\nPayroll in millions: 60.65\nWins:
55', lookup_str='', metadata={'source':
'./example_data/mlb_teams_2012.csv', 'row':
30}, lookup_index=0)]
```

## Specify a column to identify the document source

Use the source\_column argument to specify a source for the document created from each row. Otherwise file\_path will be used as the source for all documents created from the CSV file.

This is useful when using documents loaded from CSV files for chains that answer questions using sources.

```
loader =
CSVLoader(file_path='./example_data/mlb_teams_2
source_column="Team")
data = loader.load()
```

```
print(data)
```

```
[Document(page_content='Team:
Nationals\n"Payroll (millions)":
81.34\n"Wins": 98', lookup_str='', metadata=
{'source': 'Nationals', 'row': 0},
lookup_index=0), Document(page_content='Team:
Reds\n"Payroll (millions)": 82.20\n"Wins":
```

```
97', lookup_str='', metadata={'source':
'Reds', 'row': 1}, lookup_index=0),
Document(page_content='Team:
Yankees\n"Payroll (millions)":
197.96\n"Wins": 95', lookup_str='', metadata=
{'source': 'Yankees', 'row': 2},
lookup_index=0), Document(page_content='Team:
Giants\n"Payroll (millions)": 117.62\n"Wins":
94', lookup_str='', metadata={'source':
'Giants', 'row': 3}, lookup_index=0),
Document(page_content='Team: Braves\n"Payroll
(millions)": 83.31\n"Wins": 94',
lookup_str='', metadata={'source': 'Braves',
'row': 4}, lookup_index=0),
Document(page_content='Team:
Athletics\n"Payroll (millions)":
55.37\n"Wins": 94', lookup_str='', metadata=
{'source': 'Athletics', 'row': 5},
lookup_index=0), Document(page_content='Team:
Rangers\n"Payroll (millions)":
120.51\n"Wins": 93', lookup_str='', metadata=
{'source': 'Rangers', 'row': 6},
lookup_index=0), Document(page_content='Team:
Orioles\n"Payroll (millions)": 81.43\n"Wins":
93', lookup_str='', metadata={'source':
'Orioles', 'row': 7}, lookup_index=0),
Document(page_content='Team: Rays\n"Payroll
(millions)": 64.17\n"Wins": 90',
lookup_str='', metadata={'source': 'Rays',
'row': 8}, lookup_index=0),
Document(page_content='Team: Angels\n"Payroll
```

```
(millions)": 154.49\n"Wins": 89',
lookup_str='', metadata={'source': 'Angels',
'row': 9}, lookup_index=0),
Document(page_content='Team: Tigers\n"Payroll
(millions)": 132.30\n"Wins": 88',
lookup_str='', metadata={'source': 'Tigers',
'row': 10}, lookup_index=0),
Document(page_content='Team:
Cardinals\n"Payroll (millions)":
110.30\n"Wins": 88', lookup_str='', metadata=
{'source': 'Cardinals', 'row': 11},
lookup_index=0), Document(page_content='Team:
Dodgers\n"Payroll (millions)": 95.14\n"Wins":
86', lookup_str='', metadata={'source':
'Dodgers', 'row': 12}, lookup_index=0),
Document(page_content='Team: White
Sox\n"Payroll (millions)": 96.92\n"Wins":
85', lookup_str='', metadata={'source':
'White Sox', 'row': 13}, lookup_index=0),
Document(page_content='Team:
Brewers\n"Payroll (millions)": 97.65\n"Wins":
83', lookup_str='', metadata={'source':
'Brewers', 'row': 14}, lookup_index=0),
Document(page_content='Team:
Phillies\n"Payroll (millions)":
174.54\n"Wins": 81', lookup_str='', metadata=
{'source': 'Phillies', 'row': 15},
lookup_index=0), Document(page_content='Team:
Diamondbacks\n"Payroll (millions)":
74.28\n"Wins": 81', lookup_str='', metadata=
{'source': 'Diamondbacks', 'row': 16},
```

```
lookup_index=0), Document(page_content='Team:
Pirates\n"Payroll (millions)": 63.43\n"Wins":
79', lookup_str='', metadata={'source':
'Pirates', 'row': 17}, lookup_index=0),
Document(page_content='Team: Padres\n"Payroll
(millions)": 55.24\n"Wins": 76',
lookup_str='', metadata={'source': 'Padres',
'row': 18}, lookup_index=0),
Document(page_content='Team:
Mariners\n"Payroll (millions)":
81.97\n"Wins": 75', lookup_str='', metadata=
{'source': 'Mariners', 'row': 19},
lookup_index=0), Document(page_content='Team:
Mets\n"Payroll (millions)": 93.35\n"Wins":
74', lookup_str='', metadata={'source':
'Mets', 'row': 20}, lookup_index=0),
Document(page_content='Team: Blue
Jays\n"Payroll (millions)": 75.48\n"Wins":
73', lookup_str='', metadata={'source': 'Blue
Jays', 'row': 21}, lookup_index=0),
Document(page_content='Team: Royals\n"Payroll
(millions)": 60.91\n"Wins": 72',
lookup_str='', metadata={'source': 'Royals',
'row': 22}, lookup_index=0),
Document(page_content='Team:
Marlins\n"Payroll (millions)":
118.07\n"Wins": 69', lookup_str='', metadata=
{'source': 'Marlins', 'row': 23},
lookup_index=0), Document(page_content='Team:
Red Sox\n"Payroll (millions)":
173.18\n"Wins": 69', lookup_str='', metadata=
```

```
{'source': 'Red Sox', 'row': 24},
lookup_index=0), Document(page_content='Team:
Indians\n"Payroll (millions)": 78.43\n"Wins":
68', lookup_str='', metadata={'source':
'Indians', 'row': 25}, lookup_index=0),
Document(page_content='Team: Twins\n"Payroll
(millions)": 94.08\n"Wins": 66',
lookup_str='', metadata={'source': 'Twins',
'row': 26}, lookup_index=0),
Document(page_content='Team:
Rockies\n"Payroll (millions)": 78.06\n"Wins":
64', lookup_str='', metadata={'source':
'Rockies', 'row': 27}, lookup_index=0),
Document(page_content='Team: Cubs\n"Payroll
(millions)": 88.19\n"Wins": 61',
lookup_str='', metadata={'source': 'Cubs',
'row': 28}, lookup_index=0),
Document(page_content='Team: Astros\n"Payroll
(millions)": 60.65\n"Wins": 55',
lookup_str='', metadata={'source': 'Astros',
'row': 29}, lookup_index=0)]
```