

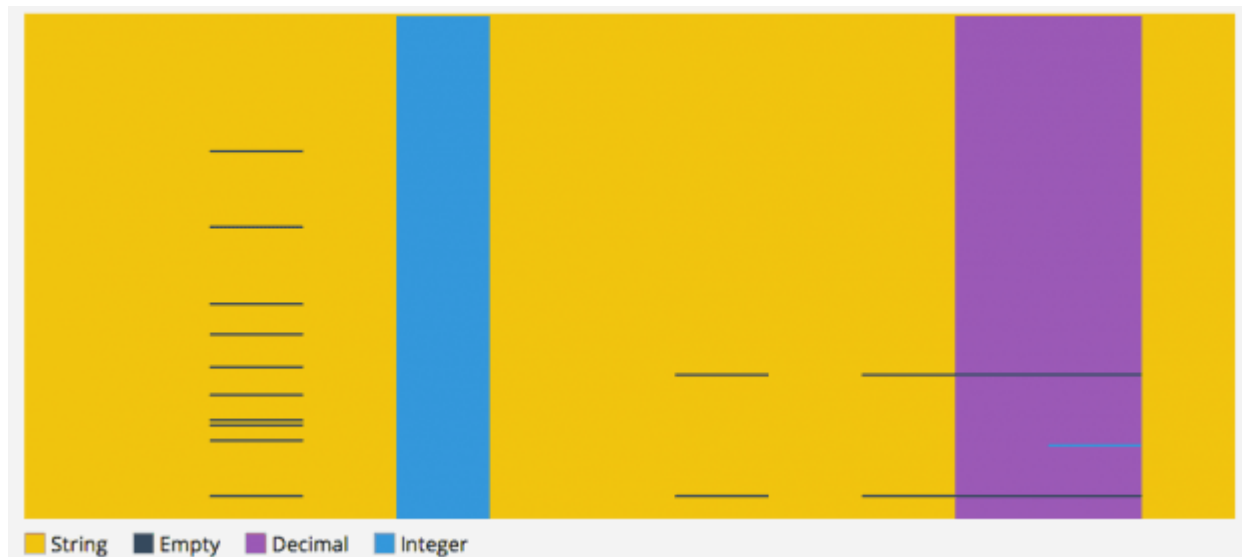


# CSV Fingerprint - инструмент для проверки файлов, содержащих данные

Алексей Ница

<https://te-st.org/2014/11/03/csv-fingerprint/>

Статья обновлена 12 декабря 2021



С помощью CSV Fingerprint можно быстро просмотреть CSV-файл с данными на предмет ошибок, некорректных данных или незаполненных ячеек в визуальном формате без большого количества отвлекающих деталей.

CSV является простым и распространенным форматом для табличных данных, который использует запятые для разделения строк и столбцов. Почти все электронные таблицы и программы баз данных позволяют пользователям импортировать и экспортировать в CSV. Но часто программы по-разному воспринимали нестандартные случаи, например, когда запятая есть в самих данных.

Легко сделать ошибку, когда необходимо сделать так, чтобы CSV-файл соответствовал конкретному формату. Инструмент CSV Fingerprint позволяет просмотреть файл с данными в визуальном формате без большого количества отвлекающих деталей. Цвета соответствуют типам данных. Для просмотра отдельных элементов можно нажать на ячейки. При этом появляется увеличительное стекло, которое позволит посмотреть данные в конкретной ячейке.

## Пример

Для этого анализа был использован файл CSV с данными, собранными в Калифорнийском департаменте водного хозяйства. Благодаря такому анализу можно увидеть, что второй столбец «lake» имеет несколько пустых ячеек, которые отображаются темно-серым цветом. Также можно увидеть, что все значения в столбце «capacities» являются числами и окрашены в синий цвет.

Провести анализ своего файла CSV. Для этого необходимо сохранить данные в формате CSV и открыть в любом текстовом редакторе. После чего нужно скопировать и вставить данные в поле инструмента.

Инструмент CSV Fingerprint имеет открытый исходный код, который можно найти на [github](https://github.com).