Чтение файлов с помощью Pandas

pd.read_excel

```
class pandas.read_excel(
io, sheet_name=0, header=0, names=None,
index_col=None, usecols=None, squeeze=False,
dtype=None, engine=None, converters=None,
true_values=None, false_values=None,
skiprows=None, nrows=None, na_values=None,
keep_default_na=True, na_filter=True,
verbose=False, parse_dates=False,
date_parser=None, thousands=None,
comment=None, skipfooter=0,
convert_float=None, mangle_dupe_cols=True,
storage options=None)
```

CSV— текстовый формат, предназначенный для хранения табличных данных.

CSV

Comma-Separated Values (значения, разделённые запятыми).

CSV



partner_data.csv

id,age,job,marital,education,balance,housing,loan,default 1,26.0,unemployed,married,tertiary,454,yes,yes,no 2,25.0, technician, married, secondary, 517, no, no, no 3,36.0, technician, single, secondary, 1044, yes, no, no 4,45.0,blue-collar,married,secondary,146,yes,yes,no 5,27.0, management, married, tertiary, 145, no, no, no 6,45.0,unemployed,divorced,primary,11,no,no,yes 7,35.0, management, married, tertiary, 298, yes, no, no 8,27.0, technician, single, secondary, 905, yes, yes, no 9,,management,married,tertiary,-485,yes,no,yes 10,32.0,unemployed,married,secondary,10,yes,no,no 11,46.0,blue-collar,married,secondary,78,yes,yes,no 12,53.0, management, married, tertiary, 0, no, no, no 13,32.0, management, single, tertiary, 223, no, no, no 14,34.0, student, tertiary, 387, yes, no, no 15,58.0, management, married, tertiary, 96, no, no, no 16,77.0, retired, married, tertiary, 0, no, no, no 17,31.0,blue-collar,single,secondary,-161,no,no,no 18,47.0,blue-collar,,secondary,3630,yes,no,no 19,,blue-collar,married,secondary,302,no,yes,no

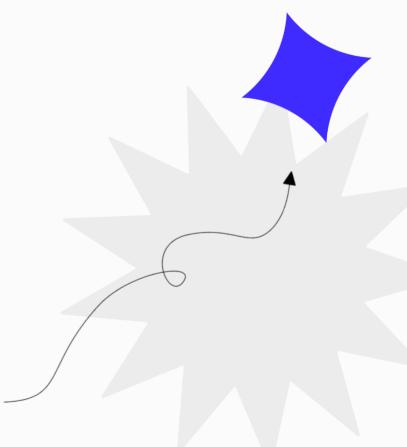
pd.read_csv

class pandas.read csv(filepath or buffer, sep=NoDefault.no default, delimiter=None, header='infer', names=NoDefault.no default, index col=None, usecols=None, squeeze=False, prefix=NoDefault.no default, mangle dupe cols=True, dtype=None, engine=None, converters=None, true values=None, false values=None, skipinitialspace=False, skiprows=None, skipfooter=0, nrows=None, na values=None, keep default na=True, na filter=True, verbose=False, skip blank lines=True, parse dates=False, infer datetime format=False, keep date col=False, date parser=None, dayfirst=False, cache dates=True, iterator=False, chunksize=None, compression='infer', thousands=None, decimal='.', lineterminator=None, quotechar='"', quoting=0, doublequote=True, escapechar=None, comment=None, encoding=None, encoding errors='strict', dialect=None, error bad lines=None, warn bad lines=None, on bad lines=None, delim whitespace=False, low memory=True, memory map=False, float precision=None, storage options=None)

Выбор элементов

- iloc по индексу
- 2 .loc по названию





Выбор элементов

df.iloc[row_indexer,

loc

Идентификатор (индекс или название) строки column_indexer]

Идентификатор (индекс или название) столбца

Полный срез

