

# Pandas basics

## Loading the key libraries

```
import sys
import pandas as pd
import numpy as np

print(f"pandas version {pd.__version__}")
print(f"numpy version {np.__version__}")
print(f"python version {sys.version}")
```

```
pandas version 2.0.0
numpy version 1.24.2
python version 3.11.3 | packaged by conda-forge | (main, Apr 6 2023, 08:57:19) [GCC 11.3.0]
```

## A comment on file formats

The most common simple format for tabular data is comma separated or tab separated (`csv` or `tsv`).

Newer formats such as `arrow` and `parquet` are more efficient in storage and faster to load.

Pandas 2.0 can handle these newer formats.

**Reading and writing a dataframe**

**Series and Dataframes**

**File formats**

**Columns and Rows**

**Selecting columns and rows**

**Filtering**

**Missing values**

**Grouping**

**Summarizing**

**Functions**