

1 Data Wrangling: Clean, Transform, Merge, Reshape

1.1 Combining and Merging Data Sets

Database-style DataFrame Merges

Merging on Index

Concatenating Along an Axis

Combining Data with Overlap

1.2 Reshaping and Pivoting

Reshaping with Hierarchical Indexing

Pivoting “long” to “wide” Format

1.3 Data Transformation

Removing Duplicates

Transforming Data Using a Function or Mapping

Replacing Values

Renaming Axis Indexes

Discretization and Binning

Detecting and Filtering Outliers

Permutation and Random Sampling

Computing Indicator/Dummy Variables

1.4 String Manipulation

String Object Methods

Regular expressions

Vectorized string functions in pandas

1.5 Example: USDA Food Database