

# Data Manipulation with pandas

## 1 Transforming Data

### DataFrame: homelessness

#### Introducing DataFrames

- Inspecting a DataFrame
- Parts of a DataFrame

#### Sorting and subsetting

- Sorting rows
- Subsetting columns
- Subsetting rows
- Subsetting rows by categorical variables

#### New columns

- Adding new columns
- Combo-attack!

## 2 Aggregating Data

### DataFrame: sales

#### Summary statistics (summarize numerical data)

- Mean and median
- Summarizing dates
- Efficient summaries
- Cumulative statistics

#### Counting (summarize categorical data)

- Dropping duplicates
- Counting categorical variables

#### Grouped summary statistics

- What percent of sales occurred at each store type?
- Calculations with `.groupby()`
- Multiple grouped summaries

#### Pivot tables

- Pivoting on one variable
- Fill in missing values and sum values with pivot tables

## 3 Slicing and indexing

### DataFrame: temperatures

#### Explicit indexes

- Setting & removing indexes
- Subsetting with `.loc[]`
- Setting multi-level indexes
- Sorting by index values

#### Slicing and subsetting with `.loc` and `.iloc`

- Slicing index values
- Slicing in both directions
- Slicing time series
- Subsetting by row/column number

#### Working with pivot tables

- Pivot temperature by city and year
- Subsetting pivot tables
- Calculating on a pivot table

## 4 Creating and Visualizing DataFrames

### DataFrame: avocados

#### Visualizing your data

- Which avocado size is most popular?
- Changes in sales over time
- Avocado supply and demand
- Price of conventional vs. organic avocados

#### Missing values

- Finding missing values
- Removing missing values
- Replacing missing values

#### Creating DataFrames

- List of dictionaries
- Dictionary of lists

#### Reading and writing CSVs

- CSV to DataFrame
- DataFrame to CSV