Pandas Data Glossary:

Reference Table:

Unit	Examples	Exercises
Unit 1 - Series	-	-
Unit 2 - DataFrames	-	-
Unit 3 - Selection	-	-
Unit 4 - Formatting	insurance.csv	bike_rental.csv
Unit 5 - Element-wise	-	-
Unit 6 - Missing Data	-	-
Unit 7 - Combining Data	-	-
Unit 8 - Aggregations	-	-
Unit 9 - Time Series	-	-
Unit 10 - Preprocessing	-	bike_rental.csv

Datasets:

1. Insurance.csv:

Description:

A dataset on personal and health-related details across individuals. This information can be used to predict the cost of medical insurance claims by the individuals.

Overview:

The dataset consists of 1346 records of individuals' personal details across 7 categories (columns):

- **Client ID**: A technical key, uniquely identifying customers
- **age:** Age of individual
- **bmi**: Body Mass Index
- **children:** Number of children covered by health insurance / Number of dependents
- **smoker:** Does the individual smoke (yes; no)
- **region:** The individual's residential area in the US (northeast; southeast; southwest; northwest)
- **CHARGES**: Individual's medical costs billed by health insurance

2. Bike_rental.csv

Description:

A dataset on bike rentals and weather-related information for different days of the year. This dataset can be used to predict the number of bike rentals on a certain day.

Overview:

The dataset consists of 732 records of detailing the number of bike rentals on a day as well as the environmental and seasonal settings for that day. The data consists of 12 categories (columns):

- **date**: Full date column in format 'mm/dd/yyyy'
- **day**: Day of the month (e.g. 1,2,3,...,31)
- **mnth**: Month of the year (e.g. 1 = Janurary; ...; 12 = December)
- **year**: Year
- **season**: (1 = Spring; 2 = Summer; 3 = Autumn; 4 = Winter)

- **holiday**: whether day is holiday or not (extracted from http://dchr.dc.gov/page/holiday-schedule)
- **weathersit**: Weather condition categories. (1 = Good weather conditions for cycling, e.g. clear to partly cloudy; 2 = Medium weather conditions for cycling, e.g. misty; 3 = Poor weather conditions for cycling e.g. snow, thunderstorms, fog
- **temp**: Temperature in Celsius
- **atemp**: "Feels like" temperature in Celsius
- **HUMIDITY:** Humidity measured in percentage
- **windspeed**: Windspeed
- **Rentals**: Count of total rental bikes that day