Introduction[¶](https://www.kaggle.com/code/residentmario/renaming-and-combining" \l "Introduction" \t "_self)

Çoğu zaman, veriler bize memnun olmadığımız sütun adları, indeks adları veya diğer adlandırma kurallarıyla gelir. Bu durumda, rahatsız edici girdilerin adlarını daha iyi bir şeye dönüştürmek için pandas fonksiyonlarını nasıl kullanacağınızı öğreneceksiniz.

Ayrıca, birden fazla DataFrame ve/veya Series'ten gelen verileri nasıl birleştireceğinizi de keşfedeceksiniz.

# Renaming

Burada tanıtacağımız ilk fonksiyon, indeks ve/veya sütun adlarını değiştirmenize olanak tanıyan **rename()**'dir. Örneğin, veri setimizdeki **points** sütun adını **score** olarak değiştirmek için şunu yaparız:

import pandas as pd

pd.set\_option('display.max\_rows', 5)

reviews = pd.read\_csv("../input/wine-reviews/winemag-data-130k-v2.csv", index\_col=0)

reviews.rename(columns={'points': 'score'})

| country | description | designation | score | price | province | region\_1 | region\_2 | taster\_name | taster\_twitter\_handle | title | variety | winery |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | Italy | Aromas include tropical fruit, broom, brimston... | Vulkà Bianco | 87 | NaN | Sicily & Sardinia | Etna | NaN | Kerin O’Keefe | @kerinokeefe | Nicosia 2013 Vulkà Bianco (Etna) | White Blend | Nicosia |
| 1 | Portugal | This is ripe and fruity, a wine that is smooth... | Avidagos | 87 | 15.0 | Douro | NaN | NaN | Roger Voss | @vossroger | Quinta dos Avidagos 2011 Avidagos Red (Douro) | Portuguese Red | Quinta dos Avidagos |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 129969 | France | A dry style of Pinot Gris, this is crisp with ... | NaN | 90 | 32.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Marcel Deiss 2012 Pinot Gris (Alsace) | Pinot Gris | Domaine Marcel Deiss |
| 129970 | France | Big, rich and off-dry, this is powered by inte... | Lieu-dit Harth Cuvée Caroline | 90 | 21.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Schoffit 2012 Lieu-dit Harth Cuvée Car... | Gewürztraminer | Domaine Schoffit |

**rename()** fonksiyonu, sırasıyla **index** veya **columns** anahtar kelime parametrelerini belirterek indeks veya sütun adlarını değiştirmenize olanak tanır. Çeşitli girdi formatlarını destekler, ancak genellikle bir Python sözlüğü en uygun olanıdır.

İndeksin bazı elemanlarını yeniden adlandırmak için kullanılan bir örnek aşağıdadır:

reviews.rename(index={0: 'firstEntry', 1: 'secondEntry'})

| country | description | designation | points | price | province | region\_1 | region\_2 | taster\_name | taster\_twitter\_handle | title | variety | winery |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| firstEntry | Italy | Aromas include tropical fruit, broom, brimston... | Vulkà Bianco | 87 | NaN | Sicily & Sardinia | Etna | NaN | Kerin O’Keefe | @kerinokeefe | Nicosia 2013 Vulkà Bianco (Etna) | White Blend | Nicosia |
| secondEntry | Portugal | This is ripe and fruity, a wine that is smooth... | Avidagos | 87 | 15.0 | Douro | NaN | NaN | Roger Voss | @vossroger | Quinta dos Avidagos 2011 Avidagos Red (Douro) | Portuguese Red | Quinta dos Avidagos |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 129969 | France | A dry style of Pinot Gris, this is crisp with ... | NaN | 90 | 32.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Marcel Deiss 2012 Pinot Gris (Alsace) | Pinot Gris | Domaine Marcel Deiss |
| 129970 | France | Big, rich and off-dry, this is powered by inte... | Lieu-dit Harth Cuvée Caroline | 90 | 21.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Schoffit 2012 Lieu-dit Harth Cuvée Car... | Gewürztraminer | Domaine Schoffit |

Sütunları muhtemelen sık sık yeniden adlandırırsınız, ancak indeks değerlerini çok nadiren yeniden adlandırırsınız. Bunun için **set\_index()** genellikle daha kullanışlıdır.

### rename\_axis() ile İndeks Adlarını Değiştirme

Hem satır indeksi hem de sütun indeksi kendi **name** niteliğine sahip olabilir. Bu adları değiştirmek için tamamlayıcı **rename\_axis()** metodu kullanılabilir. Örneğin:

reviews.rename\_axis("wines", axis='rows').rename\_axis("fields", axis='columns')

| fields | country | description | designation | points | price | province | region\_1 | region\_2 | taster\_name | taster\_twitter\_handle | title | variety | winery |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| wines |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | Italy | Aromas include tropical fruit, broom, brimston... | Vulkà Bianco | 87 | NaN | Sicily & Sardinia | Etna | NaN | Kerin O’Keefe | @kerinokeefe | Nicosia 2013 Vulkà Bianco (Etna) | White Blend | Nicosia |
| 1 | Portugal | This is ripe and fruity, a wine that is smooth... | Avidagos | 87 | 15.0 | Douro | NaN | NaN | Roger Voss | @vossroger | Quinta dos Avidagos 2011 Avidagos Red (Douro) | Portuguese Red | Quinta dos Avidagos |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 129969 | France | A dry style of Pinot Gris, this is crisp with ... | NaN | 90 | 32.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Marcel Deiss 2012 Pinot Gris (Alsace) | Pinot Gris | Domaine Marcel Deiss |
| 129970 | France | Big, rich and off-dry, this is powered by inte... | Lieu-dit Harth Cuvée Caroline | 90 | 21.0 | Alsace | Alsace | NaN | Roger Voss | @vossroger | Domaine Schoffit 2012 Lieu-dit Harth Cuvée Car... | Gewürztraminer | Domaine Schoffit |

# Combining[¶](https://www.kaggle.com/code/residentmario/renaming-and-combining" \l "Combining" \t "_self)

Bir veri kümesinde işlem yaparken, bazen farklı DataFrame ve/veya Series'leri önemsiz olmayan şekillerde birleştirmemiz gerekir. Pandas'ın bunu yapmak için üç temel metodu vardır. Artan karmaşıklık sırasına göre bunlar **concat()**, **join()** ve **merge()**'dir. **merge()**'in yapabildiği çoğu şey, **join()** ile daha basit bir şekilde yapılabildiğinden, bu metodu atlayıp ilk iki fonksiyona odaklanacağız.

### concat() Metodu

En basit birleştirme metodu **concat()**'tır. Bir elemanlar listesi verildiğinde, bu fonksiyon bu elemanları bir eksen boyunca birleştirir.

Bu, verilerin farklı DataFrame veya Series nesnelerinde olmasına rağmen aynı alanlara (sütunlara) sahip olduğumuzda kullanışlıdır. Örneğin: **YouTube Videoları** veri seti, verileri menşe ülkeye göre ayırır (örneğin bu örnekte Kanada ve İngiltere). Aynı anda birden fazla ülkeyi incelemek istersek, onları **concat()** kullanarak birleştirebiliriz:

canadian\_youtube = pd.read\_csv("../input/youtube-new/CAvideos.csv")

british\_youtube = pd.read\_csv("../input/youtube-new/GBvideos.csv")

pd.concat([canadian\_youtube, british\_youtube])

| video\_id | trending\_date | title | channel\_title | category\_id | publish\_time | tags | views | likes | dislikes | comment\_count | thumbnail\_link | comments\_disabled | ratings\_disabled | video\_error\_or\_removed | description |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | n1WpP7iowLc | 17.14.11 | Eminem - Walk On Water (Audio) ft. Beyoncé | EminemVEVO | 10 | 2017-11-10T17:00:03.000Z | Eminem|"Walk"|"On"|"Water"|"Aftermath/Shady/In... | 17158579 | 787425 | 43420 | 125882 | https://i.ytimg.com/vi/n1WpP7iowLc/default.jpg | False | False | False | Eminem's new track Walk on Water ft. Beyoncé i... |
| 1 | 0dBIkQ4Mz1M | 17.14.11 | PLUSH - Bad Unboxing Fan Mail | iDubbbzTV | 23 | 2017-11-13T17:00:00.000Z | plush|"bad unboxing"|"unboxing"|"fan mail"|"id... | 1014651 | 127794 | 1688 | 13030 | https://i.ytimg.com/vi/0dBIkQ4Mz1M/default.jpg | False | False | False | STill got a lot of packages. Probably will las... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 38914 | -DRsfNObKIQ | 18.14.06 | Eleni Foureira - Fuego - Cyprus - LIVE - First... | Eurovision Song Contest | 24 | 2018-05-08T20:32:32.000Z | Eurovision Song Contest|"2018"|"Lisbon"|"Cypru... | 14317515 | 151870 | 45875 | 26766 | https://i.ytimg.com/vi/-DRsfNObKIQ/default.jpg | False | False | False | Eleni Foureira represented Cyprus at the first... |
| 38915 | 4YFo4bdMO8Q | 18.14.06 | KYLE - Ikuyo feat. 2 Chainz & Sophia Black [A... | SuperDuperKyle | 10 | 2018-05-11T04:06:35.000Z | Kyle|"SuperDuperKyle"|"Ikuyo"|"2 Chainz"|"Soph... | 607552 | 18271 | 274 | 1423 | https://i.ytimg.com/vi/4YFo4bdMO8Q/default.jpg | False | False | False | Debut album 'Light of Mine' out now: http://ky... |

Karmaşıklık açısından ortada yer alan birleştirme metodu **join()**'dir.

### join() Metodu ile Birleştirme

**join()**, ortak bir indekse sahip farklı DataFrame nesnelerini birleştirmenize olanak tanır. Örneğin, hem Kanada'da hem de Birleşik Krallık'ta aynı gün trend olan videoları çekmek için aşağıdakileri yapabiliriz:

left = canadian\_youtube.set\_index(['title', 'trending\_date'])

right = british\_youtube.set\_index(['title', 'trending\_date'])

left.join(right, lsuffix='\_CAN', rsuffix='\_UK')

lsuffix ve rsuffix parametreleri burada gereklidir çünkü veriler hem İngiliz hem de Kanada veri kümelerinde aynı sütun adlarına sahiptir. Eğer bu doğru olmasaydı (örneğin, daha önce yeniden adlandırmış olsaydık), bunlara ihtiyacımız olmazdı.