

Transforming

Transform adalah ketika mengubah dataset yang ada menjadi entitas baru, dapat dilakukan dengan:

1. konversi dari satu data type ke data type yang lain,
2. transpose dataframe,
3. atau yang lainnya.

Hal yang biasa dilakukan pertama kali setelah data dibaca adalah mengecek tipe data di setiap kolomnya apakah sesuai dengan representasinya. Untuk itu dapat menggunakan atribut `.dtypes` pada dataframe yang telah kita baca tadi,

```
[nama_dataframe].dtypes
```

Untuk konversi tipe data, secara default system akan mendeteksi data yang tidak bisa di render as date type or numeric type sebagai object yang basically string. Tidak bisa di render oleh system ini karena berbagai hal, mungkin karena formatnya asing dan tidak dikenali oleh python secara umum (misal: date type data → '2019Jan01').

Data contoh tersebut tidak bisa di render karena bulannya Jan tidak bisa di translate menjadi in form of number (00-12) dan tidak ada '-' di antara tahun, bulan dan harinya. Jika seluruh data pada kolom di order_date sudah tertulis dalam bentuk 'YYYY-MM-DD' maka ketika dibaca, kolom order_date sudah langsung dinyatakan bertipe data datetime.

Untuk merubah kolom date_order yang sebelumnya bertipe object menjadi kolom bertipe datetime, cara pertama yang dapat dilakukan adalah menggunakan

```
pd.to_datetime(argumen)
```

dengan argumen adalah isi kolom dari dataframe yang akan dirubah tipe datanya, misal dalam format umum

```
nama_dataframe["nama_kolom"]
```

Sehingga lengkapnya dapat ditulis sebagai

```
nama_dataframe["nama_kolom"] = pd.to_datetime(nama_dataframe["nama_kolom"])
```