



顧客の声

我が社では、顧客台帳をオリジナルのExcelフォーマットを使って管理しています。おかげさまで商売のほうは順調で、小売店の売り上げも安定しています。なので、Excelのデータが今、かなりのボリュームになっていて、恐らくデータを分析できると、いろんな知見がわかると思うんですが、試しに少し分析をお願いできませんか？

前提条件

この小売店では、商品A～Zの26商品を取り扱っています。

集計期間における商品の単価は変動していません。

売上履歴は簡易のシステムからCSV出力していますが、システムへの入力レジを担当したスタッフが入力しています。

顧客台帳は店舗管理者が週次で集計し、Excelで管理しています。

■表2-1：データ一覧

| No. | ファイル名 | 概要 |
|-----|---------------------|----------------------------|
| 1 | uriage.csv | 売上履歴 期間は2019年1月～2019年7月 |
| 2 | kokyaku_daicho.xlsx | 手入力で店舗が管理している顧客台帳 |

ノック11： データを読み込んでみよう

まずは、データをそれぞれ読み込んで、先頭5行のみ表示し、データを眺めてみましょう。現場のデータはExcelやCSV等で管理されている事も多いです。第1章のECサイトのケースはリレーショナルデータベースの構造を想定していたため、データの整合性が担保されている前提で読み込みが行えました。今回はどうでしょうか。

基本的なデータの読み込み方は第1章とほぼ同じです。

まずは、uriage.csvとkokyaku_daicho.xlsxをJupyter-Notebookで読み込んでみましょう。

```
import pandas as pd
uriage_data = pd.read_csv("uriage.csv")
uriage_data.head()

kokyaku_data = pd.read_excel("kokyaku_daicho.xlsx")
kokyaku_data.head()
```

※Excelの読み込みは「pd.read_excel」となっている点に注意

■図2-1：データの読み込み結果(uriage.csv)

ノック11：データを読み込んでみよう

```
In [1]: import pandas as pd
uriage_data = pd.read_csv("uriage.csv")
uriage_data.head()

Out [1]:
```

| | purchase_date | item_name | item_price | customer_name |
|---|---------------------|-----------|------------|---------------|
| 0 | 2019-06-13 18:02:34 | 商品A | 100.0 | 深井葉々美 |
| 1 | 2019-07-13 13:05:29 | 商品S | NaN | 浅田翼二 |
| 2 | 2019-05-11 19:42:07 | 商品a | NaN | 南部翼二 |
| 3 | 2019-02-12 23:40:45 | 商品Z | 2600.0 | 麻生莉緒 |
| 4 | 2019-04-22 03:09:35 | 商品a | NaN | 平田鉄二 |

■図2-2：データの読み込み結果(kokyaku_daicho.xlsx)

```
In [2]: kokyaku_data = pd.read_excel("kokyaku_daicho.xlsx")
kokyaku_data.head()

Out [2]:
```

| | 顧客名 | かな | 地域 | メールアドレス | 登録日 |
|---|-------|----------|----|----------------------------|------------|
| 0 | 須賀ひとみ | すがひとみ | H市 | suga_hitomi@example.com | 2018/01/04 |
| 1 | 岡田 歌也 | おかた としや | E市 | okada_toshiya@example.com | 42782 |
| 2 | 芳賀 希 | はが のぞみ | A市 | haga_nozomi@example.com | 2018/01/07 |
| 3 | 荻野 愛 | おぎの あい | F市 | ogino_ai@example.com | 42872 |
| 4 | 栗田 幸一 | くりた けんいち | E市 | kurita_kenichi@example.com | 43127 |