

ノック 86：画像内の人がどこに顔を向けているのかを 検出してみよう	220
ノック 87：検出した情報を統合し、タイムラプスを作ってみよう	224
ノック 88：全体像をグラフにして可視化してみよう	226
ノック 89：人通りの変化をグラフで確認しよう	228
ノック 90：移動平均を計算することで ノイズの影響を除去しよう	229

第10章 アンケート分析を行うための自然言語処理 10本ノック 233

ノック 91：データを読み込んで把握しよう	234
ノック 92：不要な文字を除外してみよう	236
ノック 93：文字数をカウントしてヒストグラムを表示してみよう	239
ノック 94：形態素解析で文章を分割してみよう	241
ノック 95：形態素解析で文章から「動詞・名詞」を抽出してみよう	243
ノック 96：形態素解析で抽出した頻出する名詞を確認してみよう	244
ノック 97：関係のない単語を除去してみよう	246
ノック 98：顧客満足度と頻出単語の関係を見てみよう	248
ノック 99：アンケート毎の特徴を表現してみよう	252
ノック 100：類似アンケートを探してみよう	254

付録 257

Appendix ① データ結合とデータ正規化	258
Appendix ② 機械学習	262
Appendix ③ 最適化問題	265
おわりに	269
索引	270

第1部 基礎編：データ加工

データ分析の入門を終えた学習者がビジネス現場で実際のデータを分析しようとすると、いくつかの壁にぶつかります。その最初の壁は、「まず何をするのか」です。多くの入門書では、Pythonの環境設定や、ひとつひとつの関数の使い方については手厚く解説されている一方で、そもそもビジネスの現場でデータ分析の業務を始める際に、最初に何をするのかが書かれているものはほとんどありません。現場で最初に行うことは「必要なデータを集めること」なのですが、何が必要かは、目的や状況によって変わってしまうため、入門書で一般論を解説することは難しいのです。

そこで、第1部では、ウォーミングアップとして、比較的「基礎的な」場合を扱います。ECサイトや小売店といった、比較的データが揃いやすく、かつ、その状況に読者自身が置かれていることを想像しやすいようなテーマを設定することで、実際のビジネス現場でデータ分析業務を行う際、「まず何をするのか」のイメージをつかむことを目標とします。また、技術的には、データの読み込みや加工など、「データを扱う」という最も基礎的な作業を行なっていきます。これらのイメージをつかむことが、第2部以降の様々な状況に対応していく基礎力になります。

第1部で取り扱うPythonライブラリ

データ加工：Pandas

可視化：Matplotlib