

Python 実践データ分析 100 本ノック

【正誤表】

●21 ページ 4～5 行目

【誤】 実行すると、5000、1737、6737 と出力されるはずですが、5000 と 1737 を足して、6737 となりますので

【正】 実行すると、5000、1786、6786 と出力されるはずですが、5000 と 1786 を足して、6786 となりますので

●22 ページ 上段のコード 3 行目

【誤】 `print(len(transaction))`

【正】 (この行、不要につき削除)

●22 ページ 下から 3 行目

【誤】 実行すると、7144、6737、7144 と出力されます。

【正】 実行すると、7144、6786、7144 と出力されます。

●33 ページ 1 行目

【コード挿入】

本書は、「自分でやってみましょう！」の精神に基づいて、全ての実習コードを掲載してはいません。また、コードは図 1-13 に表示されていますが、本文に掲載されていないのはわかりにくいという声もありますので、本文に次のコードを挿入します(第 4 刷で追加)。

```
graph_data = pd.pivot_table(join_data, index='payment_month', columns='item_name', values='price', aggfunc='sum')
graph_data.head()
```

●95 ページ 8 行目

【誤】 `ccustomer_clustering.groupby("cluster").mean()`

【正】 `customer_clustering.groupby("cluster").mean()`

●96 ページ コードの 1 行目

【誤】 `cfrom sklearn.decomposition import PCA`

【正】 `from sklearn.decomposition import PCA`

●98 ページ コードの最終行

【誤】 `customer.join.head()`

【正】 (この行、不要につき削除)

●102 ページ 1～2 行目

【誤】 これは、4 ヶ月で退会してしまったためです。

【正】 これは、**まだ入会してからの期間が短く、データが存在しない**ためです。

●107 ページ 2～4 行目

【誤】 1 人目は、6 ヶ月前から 1 ヶ月毎に 8 回、7 回、8 回、6 回、4 回、4 回、3 回来ている顧客で、2 人目は、8 回、6 回、4 回、3 回、3 回、2 回、2 回来っている顧客の翌月の来店回数を予測します。

【正】 1 人目は、6 ヶ月前から 1 ヶ月毎に 7 回、8 回、6 回、4 回、4 回、3 回来っている顧客で、2 人目は、6 回、4 回、3 回、3 回、2 回、2 回来っている顧客で、**どちらも 8 ヶ月の在籍期間**の顧客の翌月の来店回数を予測します。

●129 ページ 中段コードの 9～12 行目

【誤】 `elif campaign_name == "デイトタイム":`

`campaign_name_list = [0, 1]`

`elif campaign_name == "ナイト":`

`campaign_name_list = [0, 0]`

【正】 `elif class_name == "デイトタイム":`

`class_name_list = [0, 1]`

`elif class_name == "ナイト":`

`class_name_list = [0, 0]`

●130 ページ 5-18: 退会の予測結果

【訂正】 129 ページの修正に伴い、右上の図と差し替えます。

●142 ページ 図 6-7: 輸送実績の総コスト集計結果

【訂正】 提供データ「tbl_transaction.csv」に誤りがありました。このデータを修正後の集計結果は下の図のようになります。

```
In [10]: # 支社のコスト合計を算出
print("関東支社の総コスト: " + str(kanto["Cost"].sum()) + "万円")
print("東北支社の総コスト: " + str(tohoku["Cost"].sum()) + "万円")

関東支社の総コスト: 2189.3万円
東北支社の総コスト: 2062.0万円
```

ノック50: 顧客の退会を予測しよう

```
In [21]: count_1 = 3
routing_flg = 1
period = 10
campaign_name = "入会費無料"
class_name = "オールタイム"
gender = "M"
```

```
In [22]: if campaign_name == "入会費半額":
campaign_name_list = [1, 0]
elif campaign_name == "入会費無料":
campaign_name_list = [0, 1]
elif campaign_name == "通常":
campaign_name_list = [0, 0]
if class_name == "オールタイム":
class_name_list = [1, 0]
elif class_name == "デイトタイム":
class_name_list = [0, 1]
elif class_name == "ナイト":
class_name_list = [0, 0]
if gender == "F":
gender_list = [1]
elif gender == "M":
gender_list = [0]
input_data = [count_1, routing_flg, period]
input_data.extend(campaign_name_list)
input_data.extend(class_name_list)
input_data.extend(gender_list)
```

```
In [23]: print(model.predict([input_data]))
print(model.predict_proba([input_data]))
```

```
[1.]
[[0.01219512 0.98780488]]
```

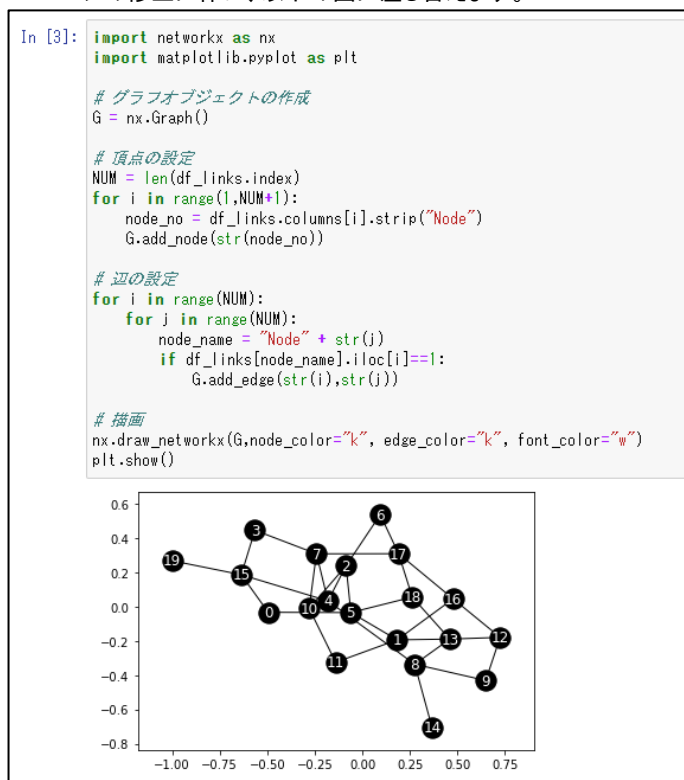
●185 ページ 下のコードの最終行

【誤】 `if df_links.iloc[i][j]==1:`

【正】 `node_name = "Node" + str(j)`
`if df_links[node_name].iloc[i]==1:`

●186 ページ 図 8-2: ネットワークの可視化

【訂正】 185 ページの修正に伴い、以下の図に差し替えます。



●187 ページ 三段目のコードの 5 行目

【誤】 `if df_links.iloc[i][j]==1:`

【正】 `node_name = "Node" + str(j)`
`if df_links[node_name].iloc[i]==1:`

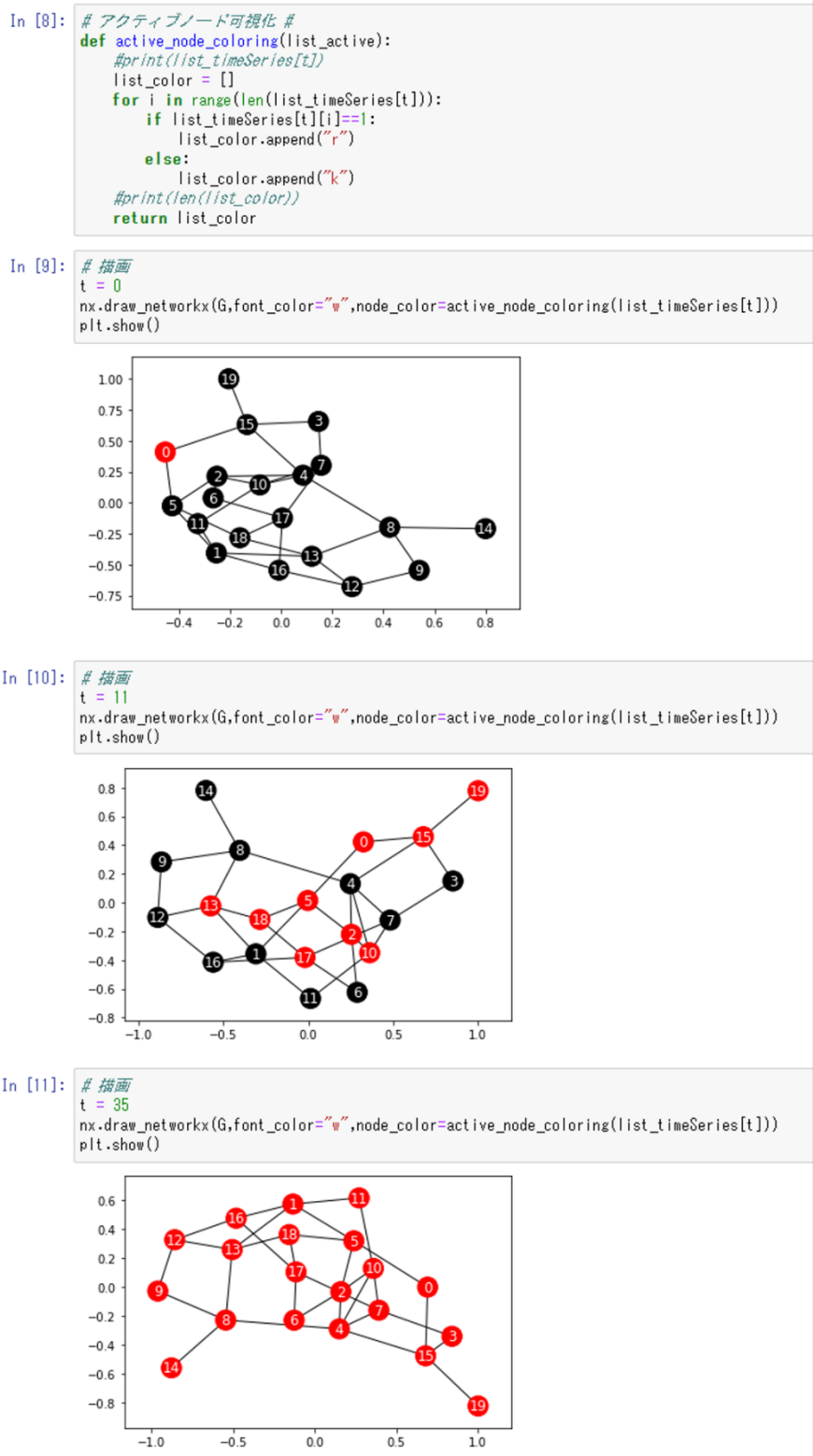
●188 ページの図 8-3: ロコミ伝播の計算

【訂正】 187 ページの修正に伴い、差し替えます。



●190 ページの図 8-4: 口コミ伝播の可視化

【訂正】 187 ページの修正に伴い、以下の図に差し替えます。



<本書紹介サイト>

<https://www.shuwasystem.co.jp/book/9784798058757.html>

<秀和システム>

<http://www.shuwasystem.co.jp/>