

# 第1章：データを俯瞰する

---

## データ列ベクトルとその表現

---

以下は「Xフードサービス」および「Yコーポレーション」それぞれの各チェーン店における、口コミグルメサイト「Foody」でのユーザ評価の平均値です。

下に続く問題について答えてみてください。

	店舗スコアX	店舗スコアY
店舗id	Xフードサービス	Yコーポレーション
1	3.5	2.8
2	2.9	3.3
3	2.9	3.1
4	2.6	3.2
5	2.3	2.7
6	3.1	2.9
7	3.1	3.1
8	3.4	3.3
9	3.3	2.9
10	3.1	2.8
11	3.0	2.8
12	2.8	3.1
13	3.0	3.3
14	3.5	3.4
15	3.1	3.3
16	3.0	3.3
17	3.2	2.8
18	3.3	3.3
19	3.2	2.8
20	2.8	3.0
21	3.1	
22	3.1	
23	2.8	
24	3.2	
25	3.2	

- Q1. 各社の店舗スコアの平均を計算してください。
- Q2. 各社の店舗スコアの分散、標準偏差を計算してください。
- Q3. 各社の店舗スコアの変動係数を計算し、データのばらつき具合の大きさを比較してください。

- Q4. 各社の店舗スコアの最大値を計算してください。
- Q5. 各社の店舗スコアの最小値を計算してください。
- Q6. 各社の店舗スコアの中央値を計算してください。
- Q7. 各社の店舗スコアの第1四分位数および第3四分位数を計算してください。
- Q8. Q4~Q7の内容をまとめて、両社の「ボックスプロット（箱ひげ図）」を作成し、比較してください。フリーハンドの手書きでも構いません。

## 度数分布表

自社の商品「Xプライム」の購入者に対してWeb上で行ったアンケート結果が、分析担当者であるあなたのもとに届きました。あなたは媒体資料を作成するために、「Xプライム」の購入者世帯の平均年収（合算）を算出しようとしたが、届いた購入者世帯の年収データは1.1での生データとは異なり、度数分布表・ヒストグラムによるものでした。

階級下限	階級上限	階級値	頻度
0	100	50	0
100	200	150	0
200	300	250	3
300	400	350	25
400	500	450	94
500	600	550	82
600	700	650	45
700	800	750	4
800	900	850	2
900	1000	950	0
		<b>合計</b>	<b>255</b>

- Q9. 得られた度数分布・ヒストグラムデータから、「Xプライム」購入者の世帯年収のモード（最頻値）を求めてください。
- Q10. 得られた度数分布・ヒストグラムデータから、「Xプライム」購入者の平均世帯年収の近似値を求めてください。