

# 統計・機械学習モデル宿題 1

廣江友哉 2125178\*

2024 年 10 月 30 日

## 目次

1	課題 1	1
2	課題 2	3
	参考文献	3

## 1 課題 1

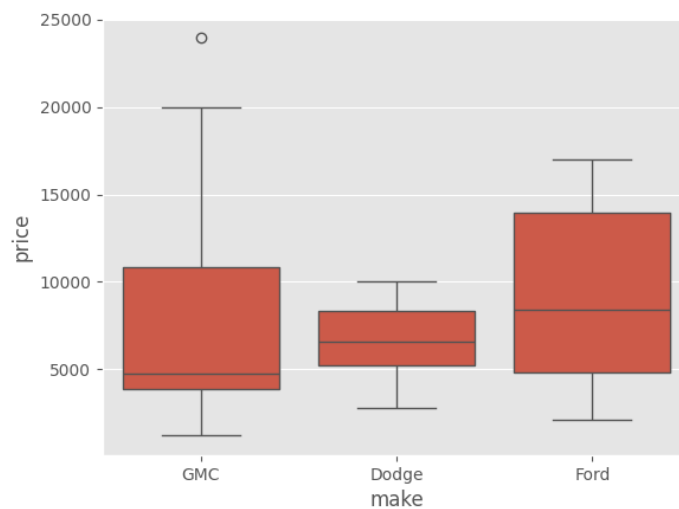


図 1 箱ひげ図

箱ひげ図からメーカーにより価格帯の広がりには違いがあることがわかる。例えば、最低価格と最高価格のトラックのメーカーは共に GMC 社である。また、製品価格の中央値が最も小さいのも GMC 社である。ここから GMC 社は製品価格の分散が他社と比較して大きいことがわかる。一方

---

\* Email: [hiro-e-tomoya-yp@ynu.jp](mailto:hiro-e-tomoya-yp@ynu.jp)

で、価格の広がり最も小さいのは Dodge 社である。四分位範囲も狭く、最高価格は他社と比較して安く、最低価格も他社と比較して高いことがわかる。Ford 社は価格の中央値が 3 社の中で最も高い。

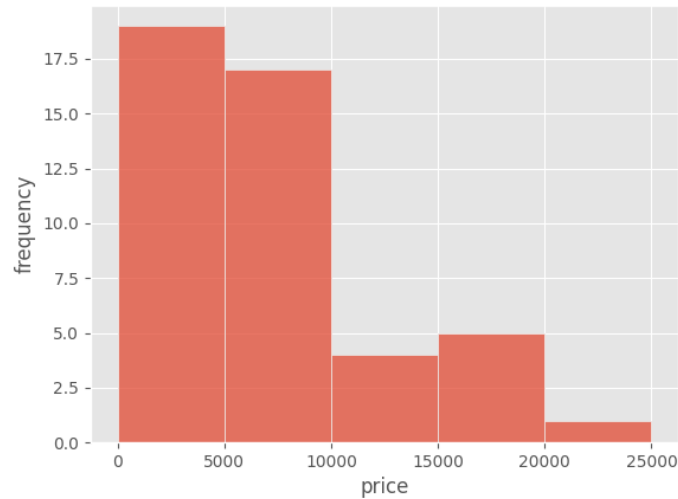


図 2 ヒストグラム

図 2 のヒストグラムは、価格を 5 つの階級に分けて、0 ドル以上 25000 ドル未満の範囲でプロットしている。0 ドル以上 5000 ドル未満の階級の度数が最も高く、19 台である。元のデータは 46 行からなる CSV データであるから、中央値が 5000 ドル以上 10000 ドル未満の階級に属していることがわかる。

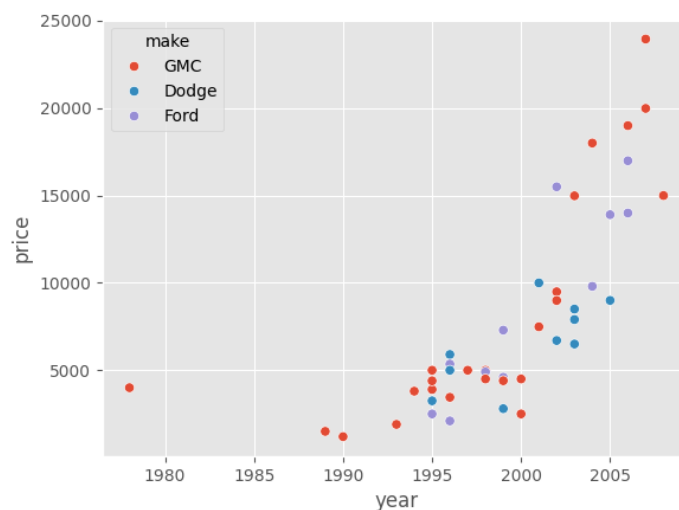


図 3 散布図

図3の散布図を確認すると、年々トラックの製品価格が高くなる傾向が3社ともに見て取れる。また、図1の箱ひげ図でも確認したように、GMC社の製品には3社の中で最低価格のものと最高価格のものが存在する。さらに、年代に注目すると、最も古いトラックはGMC社のもので、Ford社とDodge社が同時期からデータとして存在することがわかる。

## 2 課題2

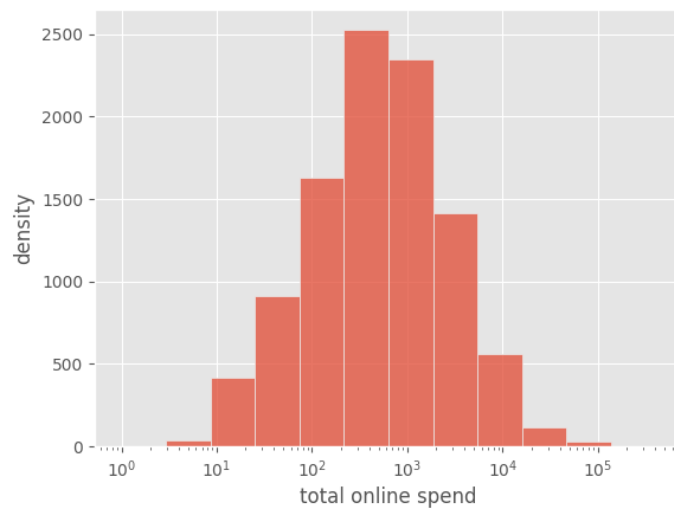


図4 ヒストグラム

図4のヒストグラムは、