

# どう貢献できるか

- 生物情報学でのデータ分析経験や業務で培った環境構築の経験を活かせる。
- データ分析についてのPDCAサイクル（フィードバックを得る、再解析）を行える。
- Anacondaを用いたデータ可視化・解析。

## 解析計画

- 方針：意思決定支援に向けて、標準的なBI手法によりデータを可視化する。
- データセット：自動車ローンの債務不履行データ（Kaggle）

### L&T Vehicle Loan Default Prediction

<https://www.kaggle.com/datasets/mamtadhaker/lt-vehicle-loan-default-prediction>

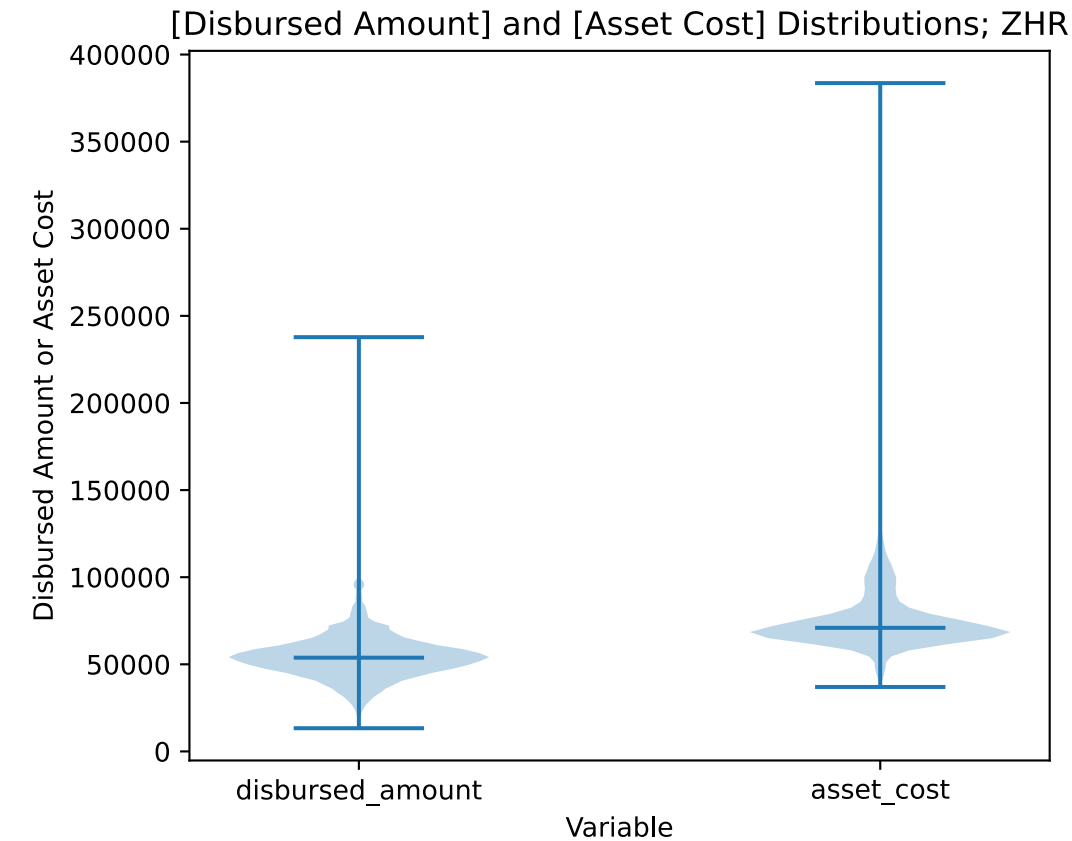
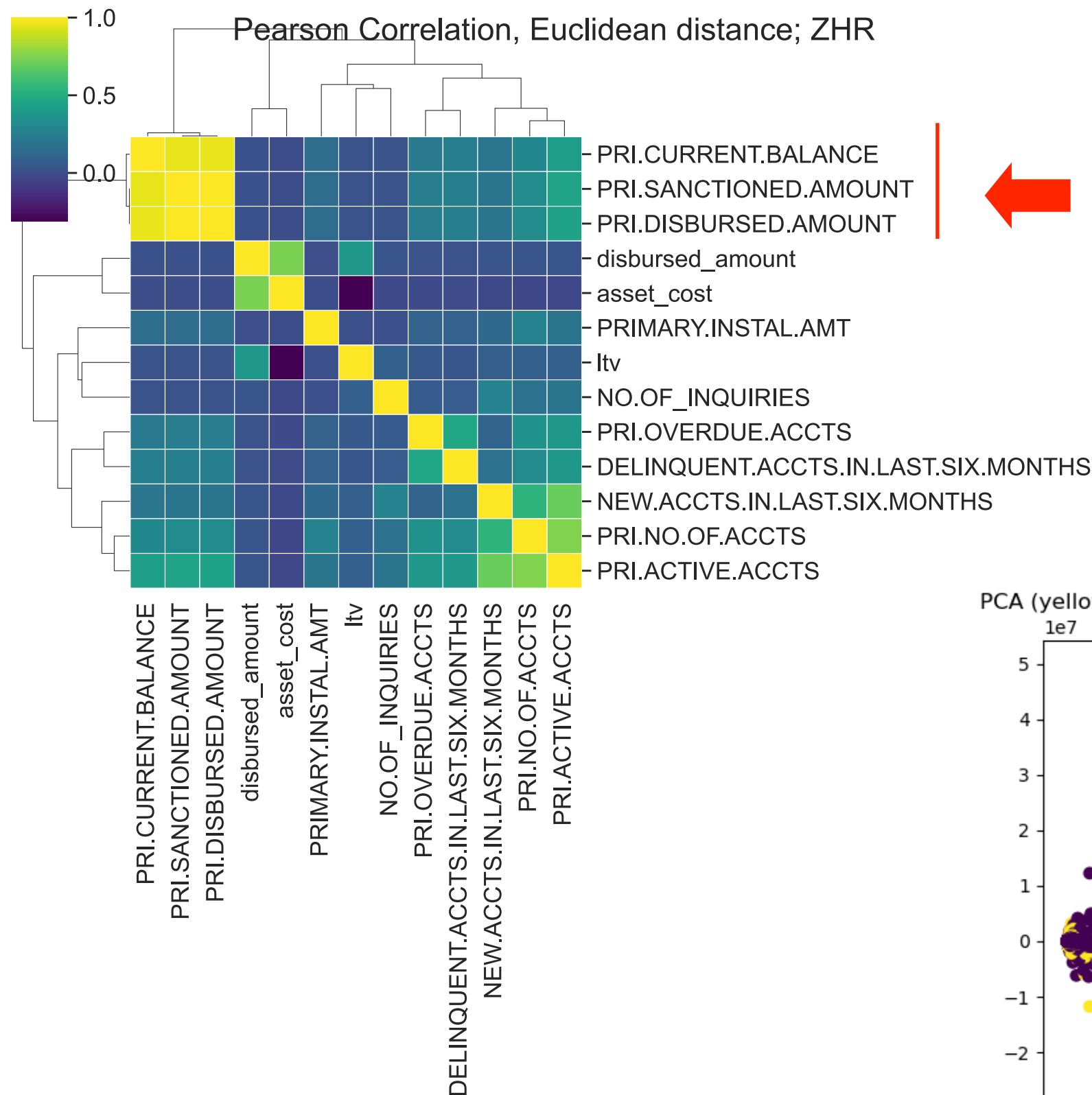
約23万件の一意的顧客ID（匿名化済み）

- 期間：正味1～2週間
- 成果物：図、ソースコード（1200行程度、言語：Python 3.9.12）

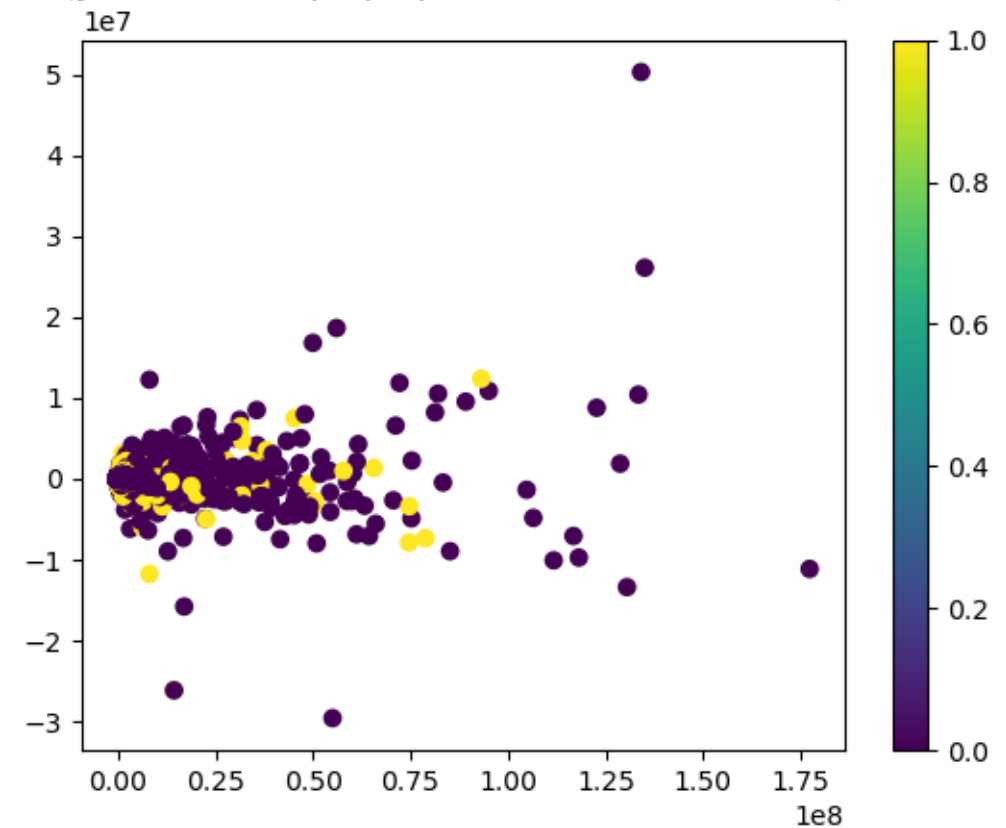
<https://github.com/sakana9305/Financials01>

- 外れ値の扱い：PCAプロットの外れ値7件程度を解析から除外。

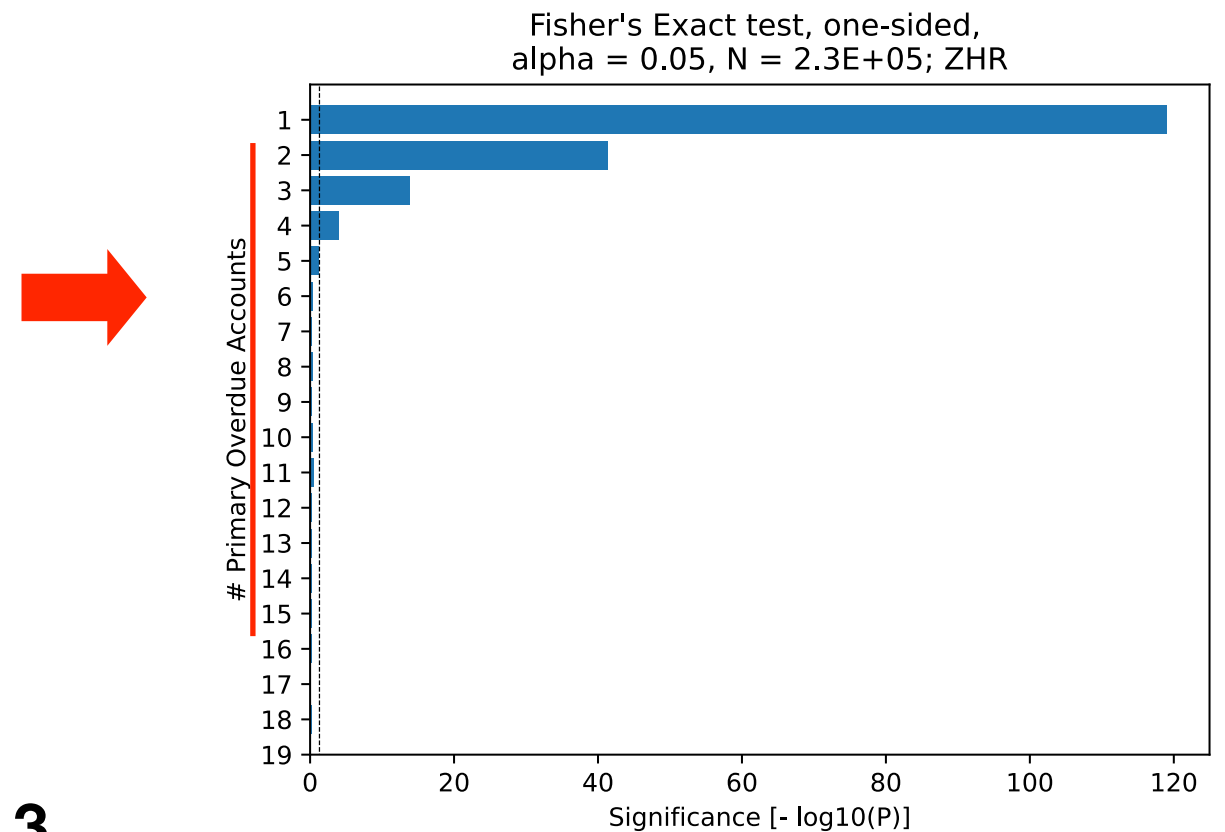
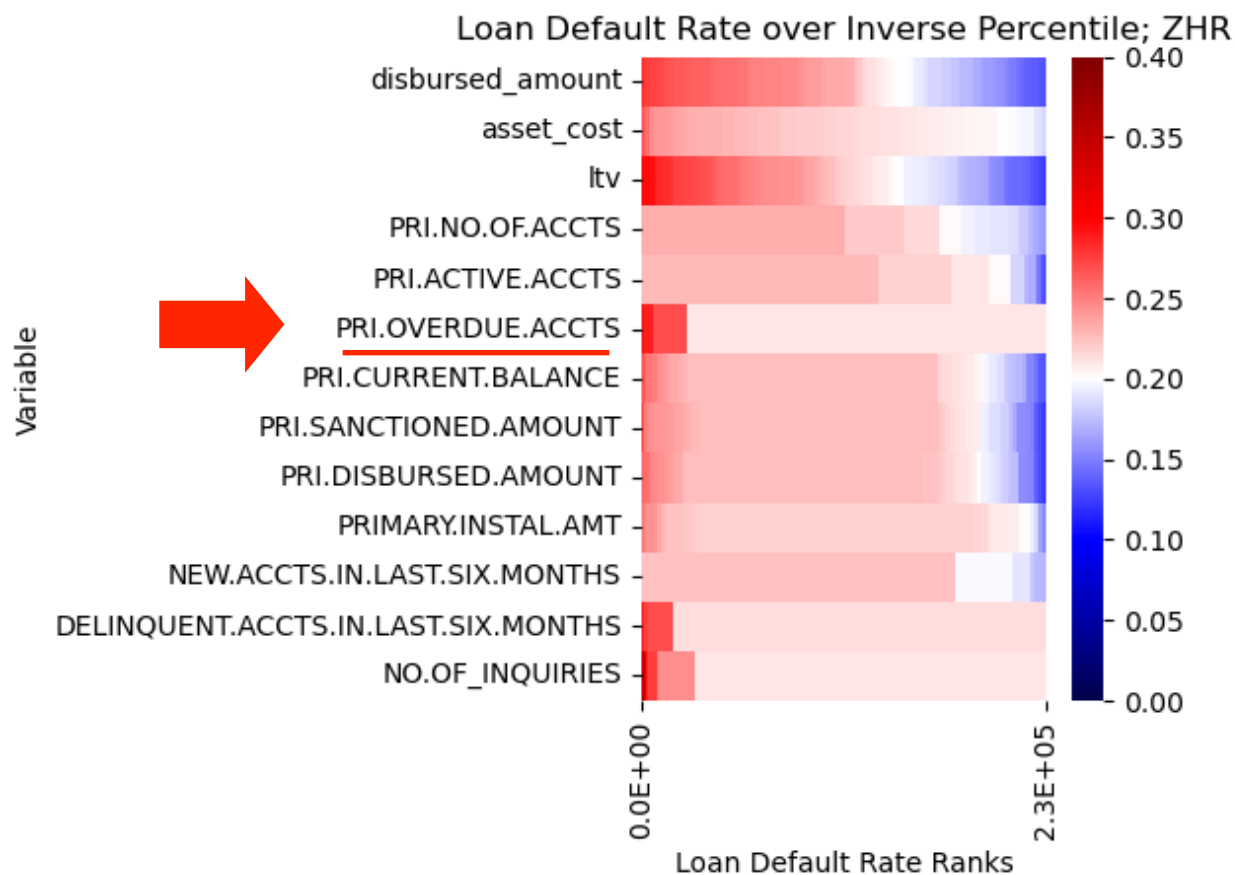
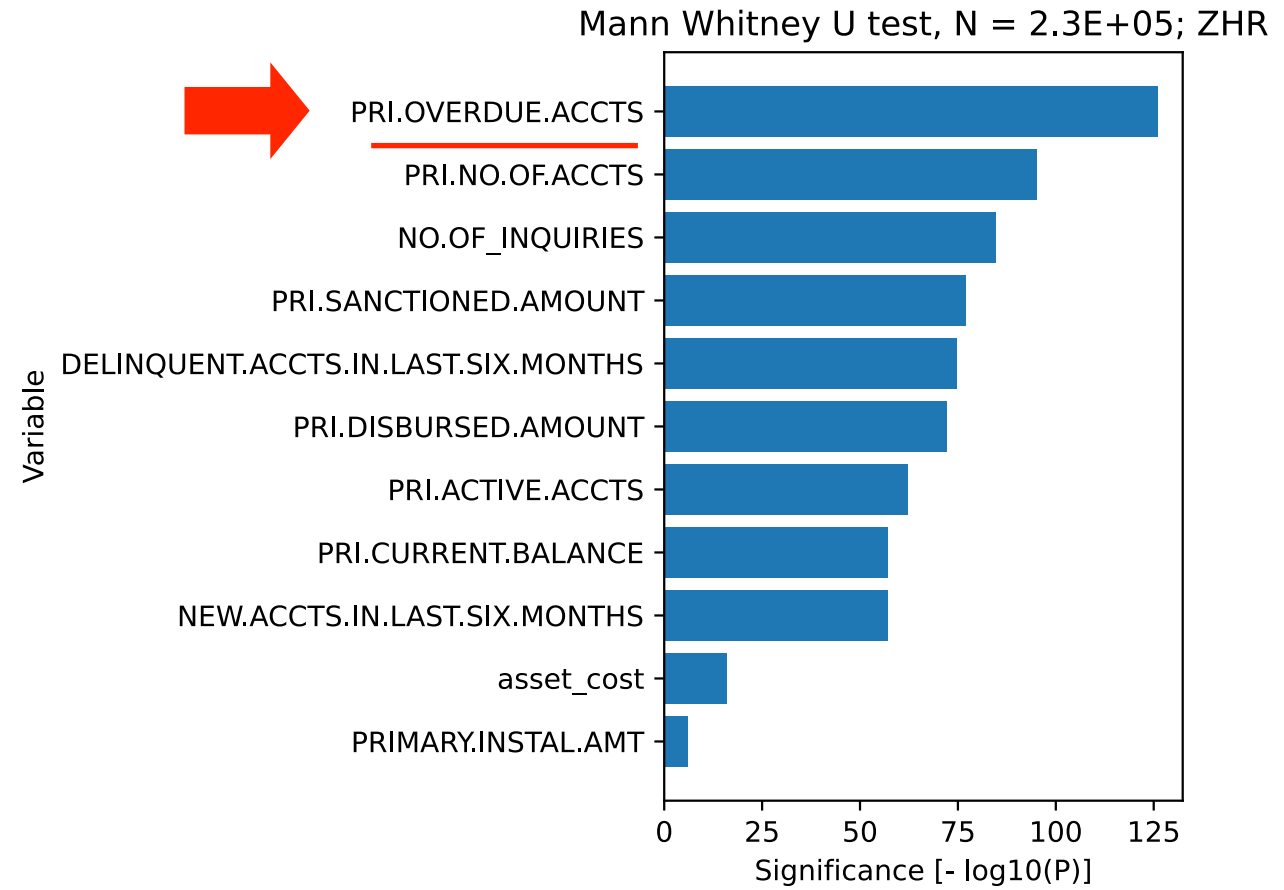
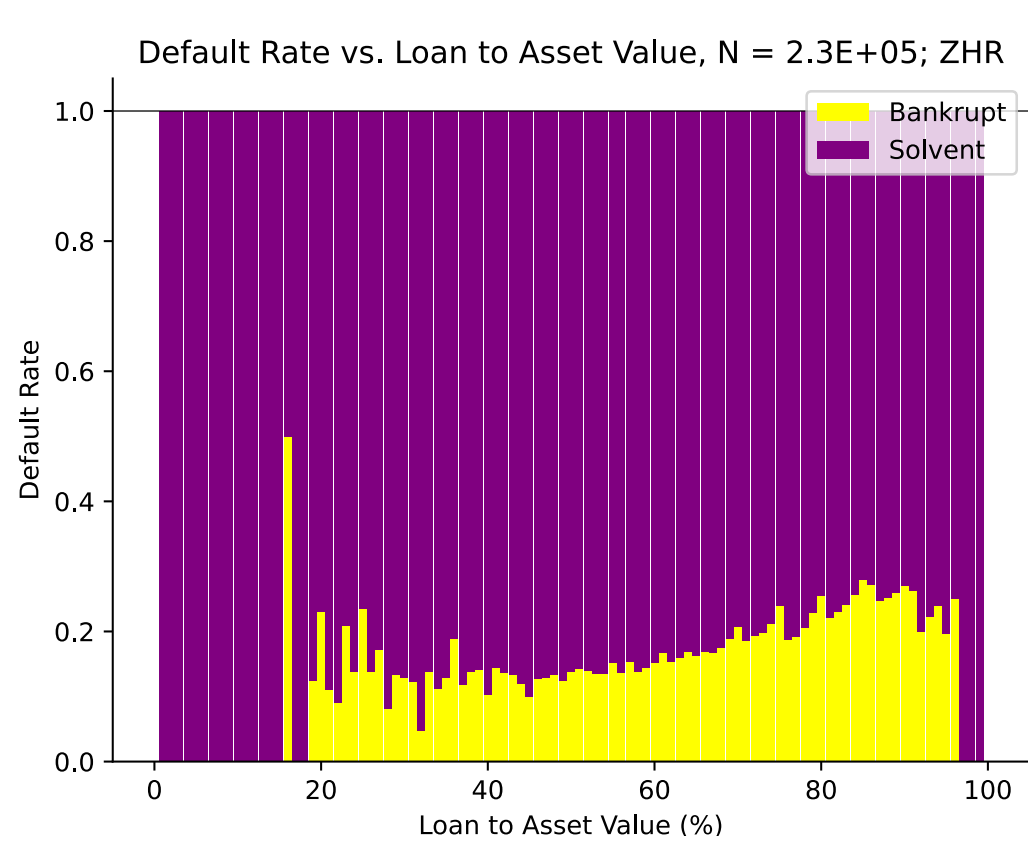
# ローン支払額と自動車価格の密度分布（右上）、 ピアソン相関行列（左）、主成分分析（右下）



PCA (yellow: bankrupt, purple: solvent, N = 2.3E+05); ZHR



# LTVごとの滞納率（左上）、滞納群と非滞納群の比較（右上）、 下側パーセント点ごとの滞納率（左下）、滞納口座数閾値ごとの分割表（右下）



**Title1.**

**Title2.**

**Title1.**

**Title2.**