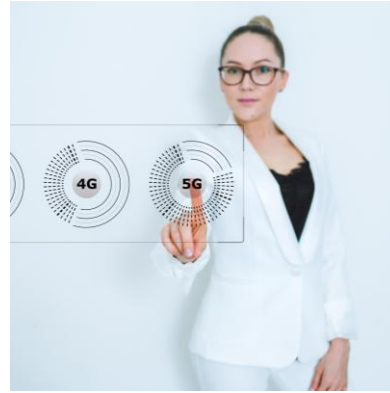
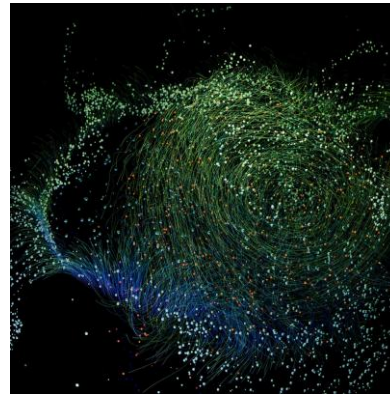


AIによる革新的与信スコアリング



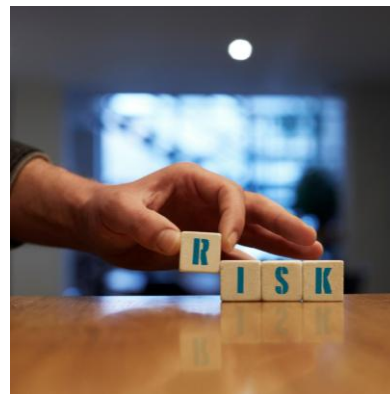
従来手法との差別化

AIによる与信スコアリングは、従来のルールベース手法に比べて高い新規性と進歩性を持っています。



LightGPMモデルの活用

LightGPMモデルを使うことで、大量データから複雑なパターンの抽出と高精度なスコアリングが可能です。



効率化とリスク低減

AIスコアリングにより、与信判断の効率が向上し、リスクを効果的に低減できます。

AI与信スコアリングの技術

高度なデータ解析

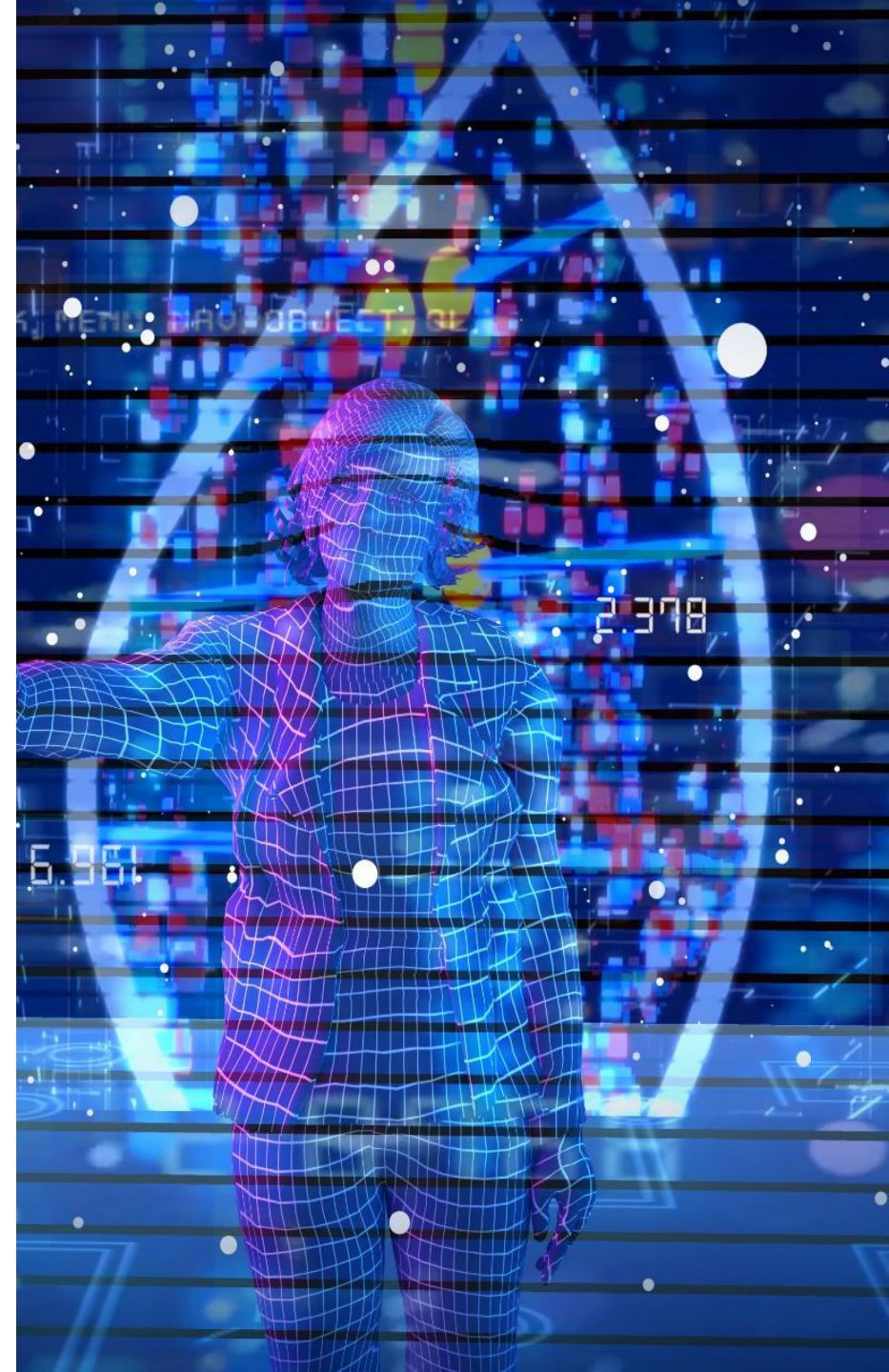
AIが膨大なデータを解析し、従来よりも正確な与信判断が可能になります。

リアルタイム処理と精度向上

多様な情報をリアルタイムで分析し、与信スコアの精度が大幅に向上しています。

パターン検出と透明性

AIは見落とされていたパターンも発見し、審査のスピードと透明性が向上します。



斬新な分析と予測力



高い分類性能

lightGBMモデルはAccuracyやAUCが高く、優れた分類能力を発揮しました。

バランスの良い指標

precision・recall・f1-scoreの各指標も全体的に高く、信頼できる予測力を示しています。

安定した結果

サポート値も十分に分布しており、モデルの結果は一貫して安定しています。

LightGBMの革新的 分析・予測・評価力

高速・高精度な分析・予測力

LightGBMは高速かつ高精度な分類を実現し、効率的なAI分析を可能にします

モデル性能評価指標

Accuracyは全体の正答率を正しく当てた割合で示します

AUCは延滞の可能性を識別した評価指標を表わします

分類レポート

Precisionは延滞の危険ありと予測した時に本当に延滞した割合を示します

Recallは本当に延滞した人を何人見つけられたかの割合を示します

f1-scoreはprecision と recall のバランスを取った数値を表します。

support、macro avg、weighted avgで分類性能を多面的に評価します。



AI与信スコアリング発明と先行技術の進歩性新・新規性 比較表

| 項目 | 従来の与信システム | 与信関連特許 | | | AI与信スコアリング |
|------------------------------------|---|-----------------------|------------------|--------------------------|---|
| | | 特開2023-114573 | 特開2024-083177 | 特開2025-052617 | |
| AIモデルの実装 | 公式にはほぼなし (IRにAI導入方針を公開してる所もあるが、結果については情報がない) | 記載なし | 記載なし | 記載なし | XGBoost、 LightGBM(今回のサンプルに実装) 、Random Forest、CatBoost、Logistic Regression、Neural Networks (MLP),etc |
| 利用データ種別 | 属性データ中心 | 賃貸借契約関連データ 過去の審査結果 | 属性データ中心 | 電子マネー履歴 | 属性データ+動的データ (IoT、公共料金、通信履歴、銀行取引履歴、アンケート、etc) |
| 動的データの利用 | 不明 | なし | ほぼなし | 行動履歴(詳細不明) | 口座入出金履歴 ：(キャッシュフロー) 月ごとの入金額、出金額の変動パターンを分析 カード利用状況 ・返済状況：リボ払い残高、遅延履歴、利用限度額に対する使用率、過剰利用を即座にリスクシグナルとして反映 公共料金の支払い履歴 ：電気・ガス・携帯代の滞納履歴、通信データ量の急減 |
| 定性情報の数値化 | 不明 | 記載なし | 記載なし | 記載なし | アンケート、カテゴリ変数のmap化し数値化 |
| ノイズ除去・欠損補完 | 不明 | 記載なし | 記載なし | 記載なし | 実装済み(欠損補完、数値変換でLightGBMに最適化) |
| 統合スコアリング | 不明なるも可能性は低い | 異なるモデル段階的使用? | 信用スコア算出+分割支払条件設定 | セグメント最適化が主でありスコア計算は参考程度? | 属性データ×動的データの統合 ：両方を同時に判断材料にすることで精度が向上 全項目を多変量解析的アプローチ：従来の経験則やルールベースのみならず総合力で判断可能 SHAPによる統合スコアの根拠提示：なぜこのスコアなのかを具体的な項目(借入残高や借入比率など)のSHAP値で説明可能 |
| 正解率評価指標 (Performance Metrics) | 不明なるも可能性は極めて低い | 記載なし | 記載なし | 記載なし | Accuracy ：全体の正答率 AUC ：延滞の可能性を識別出来た確率 |
| 分類評価指標 (Classification Metrics) | 不明なるも可能性は極めて低い | 記載なし | 記載なし | 記載なし | Precision ：延滞の危険ありと予測した時、本当に延滞した割合 Recall ：本当に延滞した人をどれくらい発見し得たかの確率 f1-score ：precision と recall のバランスを取った数値 |
| XAI(根拠説明機能) | 不明なるも可能性は低い | 一部アラート説明あり | 記載なし | 記載なし | SHAP ：個人別スコアの根拠となる説明変数と解説 |
| 個別スコアの出力 | 不明なるも可能性は低い | 承認/否決判定が中心 | スコアはあるが条件決定用 | セグメント判別が中心 | 個人ごとに根拠と共に延滞確率を出力する |
| 即時処理性 | 不明なるも可能性は低い | 処理即時性の明記は薄い | 記載薄い | 特に言及なし | 即時スコア算出を志向しAPI化が可能 Webシステム REST API Serverless API(AWS Lambda) gRPC,etc |

従来の主な与信データ（某信用情報機関のデータ）

| 項目 | 内容例 |
|--------|--------------------------|
| 氏名 | 山田 太郎 |
| 生年月日 | 29221 |
| 契約種別 | クレジット・ローン・割賦販売・携帯端末割賦 など |
| 契約会社名 | 〇〇クレジット |
| 契約日 | 44362 |
| 契約額 | 500,000円 |
| 残債額 | 200,000円 |
| 入金状況 | 正常、延滞（〇日）、異動 |
| 完済状況 | 完済、解約、異動解約 |
| 保証人有無 | 有・無 |
| 照会履歴件数 | 過去6ヶ月で何件か |
| 異動情報 | 延滞、債務整理、破産など |
| 更新日 | 45016 |

AI与信スコアリングシステムで実際に使用したサンプルデータ抜粋（実際は1000件）

| 申請者 | 性別 | 婚姻 | 年齢 | 年収 | 職種 | 職業 | 借入比率 | 借入申込額 | 他借入残高 | 異動有無 | 債券整理 | 業界 | 学歴 | 扶養家族 | 持家 | 外人 | 固定電話 | 勤続年数 | 保証人 | 担保 |
|-----|----|----|----|---------|----|----------|------|---------|---------|------|------|------|----|------|----|----|------|------|-----|----|
| 1 | 女 | 未 | 41 | 3760000 | 4 | ホールスタッフ | 0.36 | 1030000 | 1350000 | 0 | 0 | 飲食 | 3 | 0 | 0 | 0 | 0 | 13 | 0 | 0 |
| 2 | 女 | 未 | 58 | 3890000 | 3 | 介護スタッフ | 0 | 2240000 | 0 | 0 | 0 | サービス | 4 | 1 | 1 | 0 | 1 | 2 | 0 | 0 |
| 3 | 女 | 既 | 61 | 6750000 | 3 | プログラマ | 0 | 1220000 | 0 | 0 | 0 | IT | 3 | 0 | 1 | 0 | 0 | 42 | 0 | 0 |
| 4 | 女 | 既 | 46 | 3780000 | 2 | 接客スタッフ | 0.36 | 4010000 | 1360000 | 0 | 0 | サービス | 3 | 1 | 1 | 0 | 1 | 20 | 1 | 0 |
| 5 | 女 | 既 | 41 | 2710000 | 3 | 介護スタッフ | 0 | 340000 | 0 | 0 | 0 | サービス | 4 | 2 | 0 | 0 | 0 | 20 | 1 | 1 |
| 6 | 男 | 未 | 50 | 3620000 | 4 | ホールスタッフ | 0 | 1750000 | 0 | 0 | 0 | 飲食 | 2 | 0 | 0 | 0 | 1 | 16 | 0 | 1 |
| 7 | 女 | 未 | 44 | 4750000 | 1 | 市役所職員 | 0.45 | 3990000 | 2150000 | 0 | 0 | 行政 | 2 | 1 | 1 | 0 | 0 | 18 | 0 | 1 |
| 8 | 男 | 既 | 39 | 3180000 | 4 | ホールスタッフ | 0.26 | 2500000 | 830000 | 0 | 0 | 飲食 | 4 | 2 | 0 | 0 | 1 | 6 | 0 | 1 |
| 9 | 女 | 未 | 35 | 4920000 | 2 | 接客スタッフ | 0.21 | 1870000 | 1010000 | 0 | 0 | サービス | 3 | 2 | 0 | 0 | 0 | 11 | 0 | 0 |
| 10 | 男 | 未 | 57 | 6540000 | 3 | プログラマ | 0 | 390000 | 0 | 0 | 0 | IT | 4 | 1 | 0 | 0 | 0 | 16 | 1 | 0 |
| 11 | 男 | 未 | 40 | 4300000 | 3 | プログラマ | 0 | 4420000 | 0 | 0 | 0 | IT | 4 | 2 | 0 | 0 | 0 | 17 | 0 | 0 |
| 12 | 男 | 未 | 37 | 5060000 | 3 | プログラマ | 0.22 | 2800000 | 1120000 | 0 | 0 | IT | 3 | 0 | 1 | 0 | 0 | 15 | 1 | 1 |
| 13 | 男 | 既 | 25 | 4280000 | 2 | 接客スタッフ | 0.09 | 3770000 | 400000 | 0 | 0 | サービス | 2 | 2 | 1 | 0 | 1 | 6 | 1 | 1 |
| 14 | 女 | 未 | 60 | 5260000 | 3 | プログラマ | 0 | 3320000 | 0 | 0 | 0 | IT | 4 | 3 | 1 | 0 | 1 | 2 | 1 | 1 |
| 15 | 男 | 未 | 40 | 4240000 | 2 | 製造オペレーター | 0.12 | 3300000 | 510000 | 0 | 0 | 製造 | 4 | 1 | 1 | 0 | 0 | 19 | 0 | 1 |
| 16 | 女 | 未 | 35 | 6400000 | 1 | 市役所職員 | 0.32 | 560000 | 2070000 | 0 | 0 | 行政 | 2 | 2 | 0 | 0 | 0 | 17 | 0 | 1 |
| 17 | 男 | 未 | 59 | 3960000 | 2 | 販売スタッフ | 0.17 | 4550000 | 680000 | 0 | 0 | 小売り | 3 | 0 | 0 | 0 | 1 | 9 | 1 | 1 |
| 18 | 女 | 未 | 45 | 3730000 | 4 | ホールスタッフ | 0 | 1400000 | 0 | 0 | 0 | 飲食 | 2 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| 19 | 女 | 未 | 52 | 2500000 | 2 | 接客スタッフ | 0 | 1180000 | 0 | 0 | 0 | サービス | 1 | 1 | 0 | 0 | 0 | 24 | 1 | 0 |

処理フロー

1. データ取得

スコアリング（与信審査）の元となる信用情報を取得する

データ取得先の例

| | |
|--------------------|-----------------------------|
| 株式会社日本信用情報機構（JICC） | 信販会社・消費者金融・クレジットカード・各種銀行など |
| 株式会社シー・アイ・シー（CIC） | 信販会社・消費者金融・クレジットカード・携帯電話など |
| 一般社団法人全国銀行協会（JBA） | 銀行・労働金庫(ろうきん)・信用金庫・JA(農協)など |

2. データクレンジング

取得したデータの品質向上を図る

取得したデータの欠損値やノイズ除去、外れ値の処理、正規化、不整合の修正などを通してAIモデルの精度と効率の向上を目指す

3. エンコーディング

文字データを数値に変換する

AIは数値データしか扱えないため、エンコーディングすることでAIが学習できるよう文字データをカテゴリ変数として数値を割り当てる

4. スケーリング

データの値を特定の範囲に変換する

データの値を特定の範囲に変換したり、異なる単位のデータをの単位を揃える事でAIの学習を効率的に行う

5. 次元削除

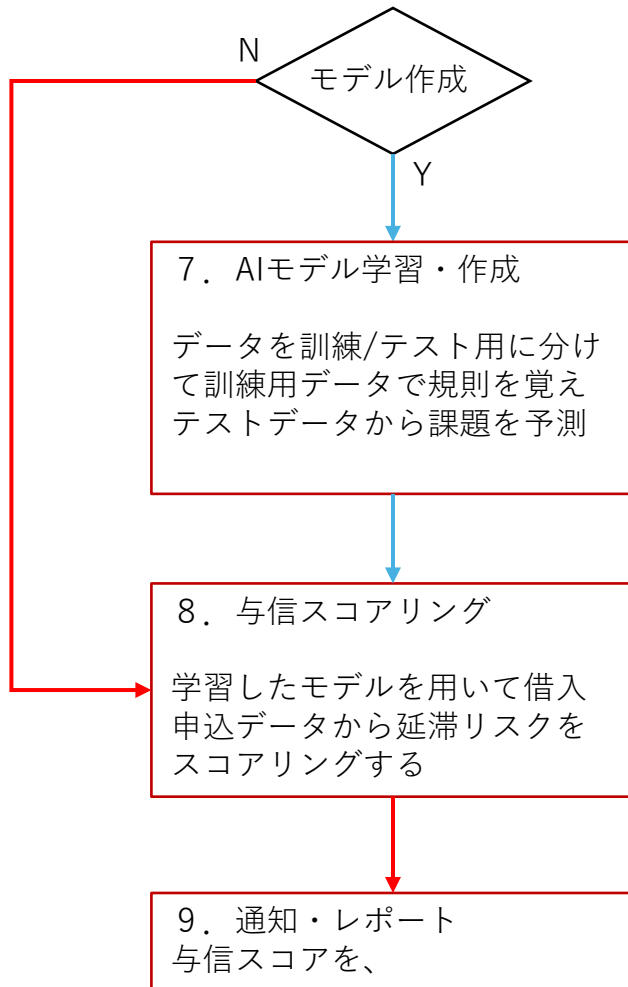
データを圧縮し、計算を効率化する

データ内の相関関係を見つけ、相関を最大化する新しい軸を定義することで元のデータセットの次元を減らしデータを圧縮して計算を効率化する

6. 特徴選択

重要な特徴量だけを取り出す

多数の特徴量（分析対象のデータの特徴を数値化したもの）の中から関連性の強いものを選択する



目的変数を説明変数で予測できるように、訓練用データから勾配ブーストと呼ばれる決定木の集まり（アンサンブル）で学習し、テストデータでその効果をAccuracyやAUCの値によって精度を検証することで、未知のデータに対する予測性能を高める。また、そのさい使用したカテゴリ変数のマップを保存する。

7で学習したAIモデルに対し、7で保存したカテゴリ変数マップでロジット値とシグモイド関数の積で与信判定スコアを求め、SHAP値によってその判定予測に至った根拠を属性値（年収、勤続年数など）で説明する。

予測結果をWebシステムや、REST API、Serverless APIなどで一覧表を作成し表示する

AIモデル学習プログラム実行結果（個人別スコアリング評価をSHAP値が証明）

■ 個人 No.298 の予測スコア: 0.2106
▼ 個人別 SHAP 上位3項目 ▼
年収: SHAP = 1.9489 (プラス要因)
勤続年数: SHAP = -0.6753 (マイナス要因)
申込額: SHAP = -0.2645 (マイナス要因)

この人は与信スコア0.2106で延滞リスクが隠れ潜んでいると予測されている。
年収のSHAP値が2に近い値であることから延滞リスクを上げており年収が低い事でリスクありと評価されている。
逆に勤続年数が-0.67と延滞リスクを下げているので、勤続年数はそこそこ長いのでリスクはそれ程でもないと評価された。

■ 個人 No.299 の予測スコア: 0.0075
▼ 個人別 SHAP 上位3項目 ▼
業種: SHAP = -0.9981 (マイナス要因)
借入比率: SHAP = -0.4741 (マイナス要因)
年収: SHAP = -0.3502 (マイナス要因)

この人は与信スコア0.0075と延滞リスクが極めて低いと予測されている。
業種のSHAP値が1に近い値であり、延滞リスクを上げている。
職業による貴賤はないものの、一般社会通念上あるいは、少なくともこのAIモデルではあまり高評価を得ていないようである。
同様に借入比率も少なく年収もそこそこあるのでリスクを下げる要因となっていると評価された。

■ 個人 No.300 の予測スコア: 0.0270
▼ 個人別 SHAP 上位3項目 ▼
勤続年数: SHAP = -0.8371 (マイナス要因)
年収: SHAP = -0.3177 (マイナス要因)
婚姻状況: SHAP = 0.2495 (プラス要因)

モデル全体の評価結果
Accuracy (正解率) : 0.9167
AUC: 0.8066

分類レポート:

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.9500 | 0.9603 | 0.9551 | 277 |
| 1 | 0.4500 | 0.3913 | 0.4186 | 23 |
| accuracy | | | 0.9167 | 300 |
| macro avg | 0.7000 | 0.6758 | 0.6869 | 300 |
| weighted avg | 0.9117 | 0.9167 | 0.9140 | 300 |

✓ モデルとスコア出力が完了しました。individual_scores.csv に保存しました。

Web (Flask APIプログラム) で実際にブラウザに表示したところ (20件抜粋)



| No | Sex | Age | Income | BorrowingRatio | Industry | CreditAppAmount | OtherDebts | DebtRestruct | DelinquencyInfo | Score |
|----|-----|-----|---------|----------------|----------|-----------------|------------|--------------|-----------------|---------------|
| 1 | 男 | 50 | 3690000 | 0.00 | サービス | 3210000 | 0 | 0 | 0 | 0.0796 |
| 2 | 女 | 60 | 3380000 | 0.00 | 製造 | 2440000 | 0 | 0 | 0 | 0.0126 |
| 3 | 女 | 48 | 3860000 | 0.00 | サービス | 2760000 | 0 | 0 | 0 | 0.0824 |
| 4 | 男 | 60 | 4080000 | 0.47 | 製造 | 3880000 | 1900000 | 0 | 0 | 0.5532 |
| 5 | 女 | 53 | 4200000 | 0.24 | 製造 | 2790000 | 990000 | 0 | 0 | 0.0319 |
| 6 | 男 | 40 | 3860000 | 0.00 | 小売り | 740000 | 0 | 0 | 0 | 0.0632 |
| 7 | 男 | 64 | 3130000 | 0.00 | 飲食 | 4930000 | 0 | 0 | 0 | 0.0736 |
| 8 | 男 | 25 | 5620000 | 0.28 | 行政 | 1840000 | 1600000 | 0 | 0 | 0.0616 |
| 9 | 女 | 36 | 4470000 | 0.00 | 行政 | 3260000 | 0 | 0 | 0 | 0.037 |
| 10 | 男 | 53 | 3920000 | 0.18 | 小売り | 2170000 | 720000 | 0 | 0 | 0.0112 |
| 11 | 男 | 41 | 6860000 | 0.31 | 行政 | 2770000 | 2110000 | 0 | 0 | 0.0663 |
| 12 | 女 | 52 | 5900000 | 0.11 | 教育 | 1070000 | 620000 | 0 | 0 | 0.0064 |
| 13 | 女 | 29 | 4920000 | 0.00 | サービス | 530000 | 0 | 0 | 0 | 0.024 |
| 14 | 女 | 43 | 4190000 | 0.00 | 小売り | 5000000 | 0 | 0 | 0 | 0.038 |
| 15 | 男 | 41 | 3660000 | 0.00 | 製造 | 1410000 | 0 | 0 | 0 | 0.0907 |
| 16 | 女 | 56 | 4980000 | 0.15 | 教育 | 3340000 | 730000 | 1 | 1 | 0.9704 |
| 17 | 男 | 64 | 3140000 | 0.19 | 飲食 | 3980000 | 590000 | 0 | 0 | 0.0412 |
| 18 | 女 | 34 | 5680000 | 0.13 | 教育 | 880000 | 720000 | 0 | 0 | 0.0211 |
| 19 | 女 | 40 | 3200000 | 0.00 | 飲食 | 530000 | 0 | 0 | 0 | 0.0068 |
| 20 | 男 | 56 | 4470000 | 0.00 | 小売り | 3980000 | 0 | 0 | 0 | 0.0145 |