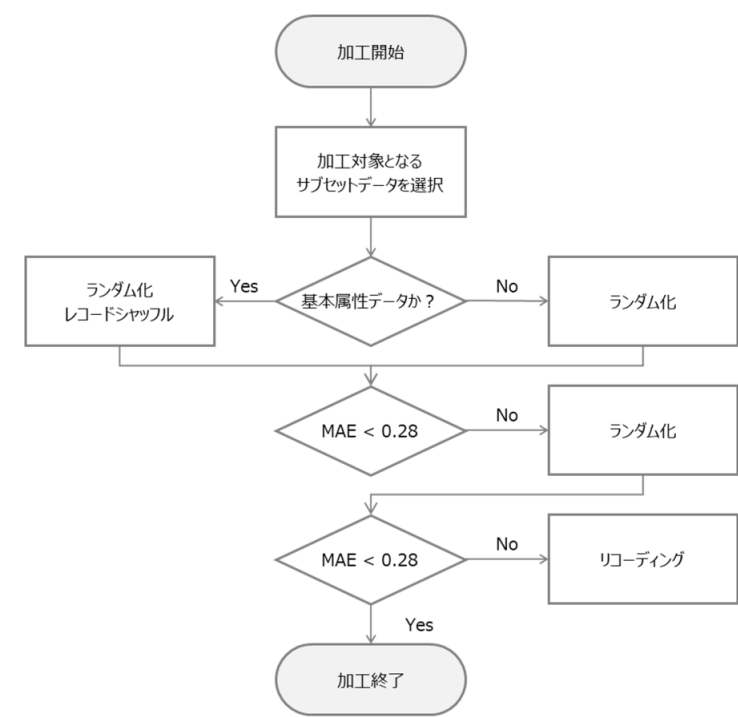
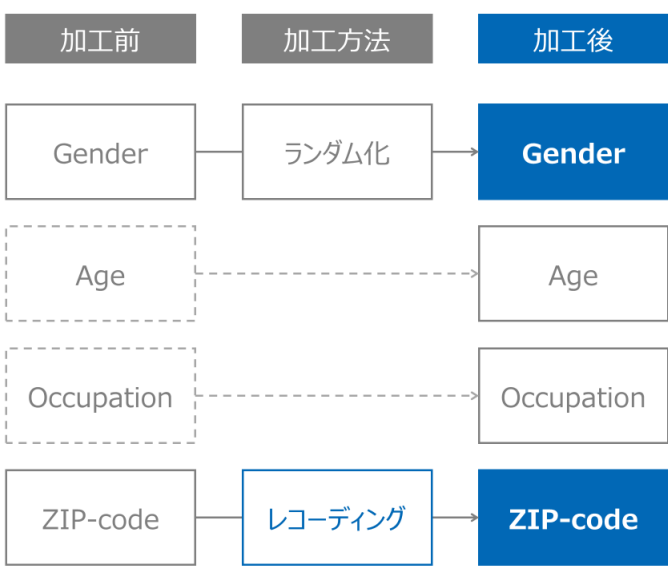


匿名化フェーズ

匿名化フェーズのフローチャート

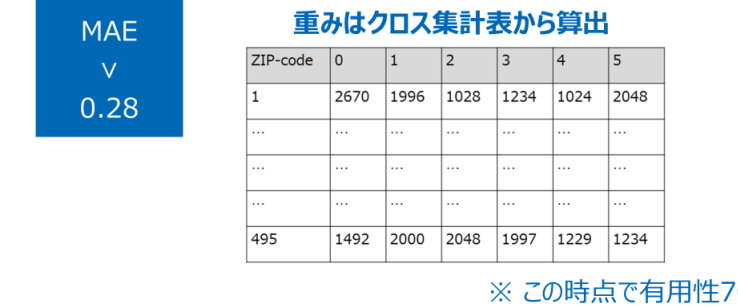
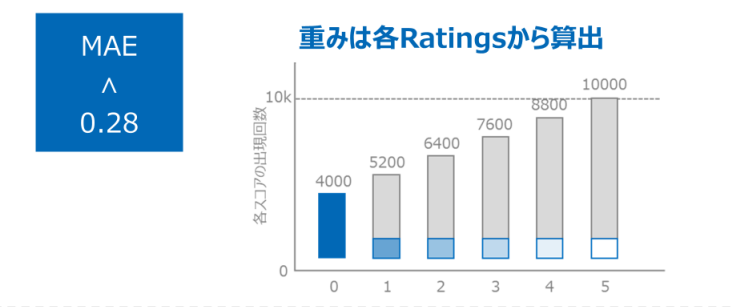
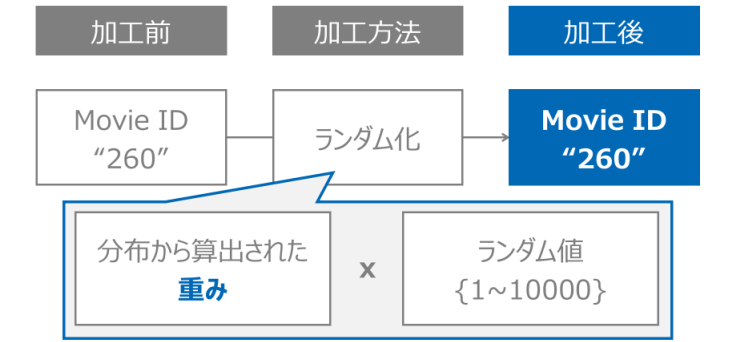


① ランダム化 / レコードシャッフル



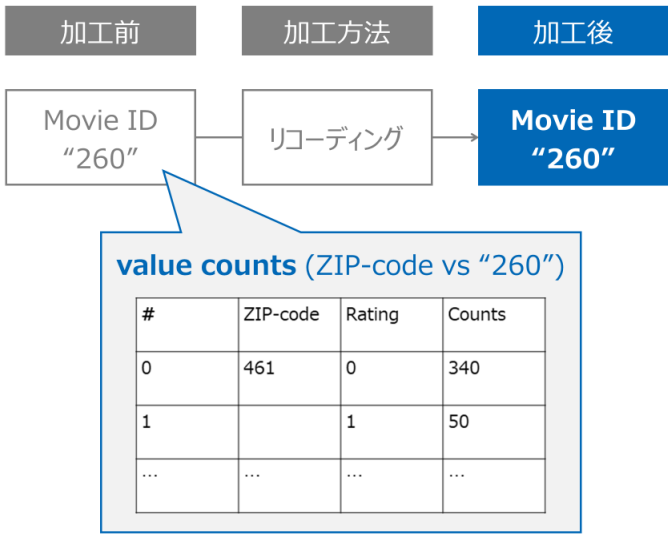
※ この時点で有用性65

② ランダム化



※ この時点で有用性70

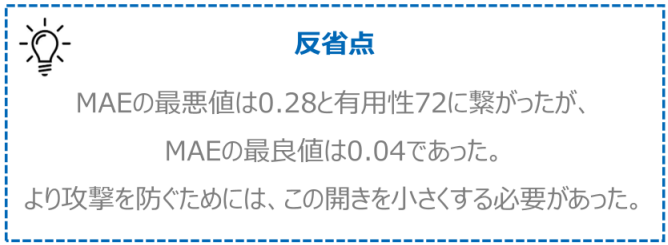
③リコーディング



オリジナルデータからZIP-codeごとのユニーク数をカウント

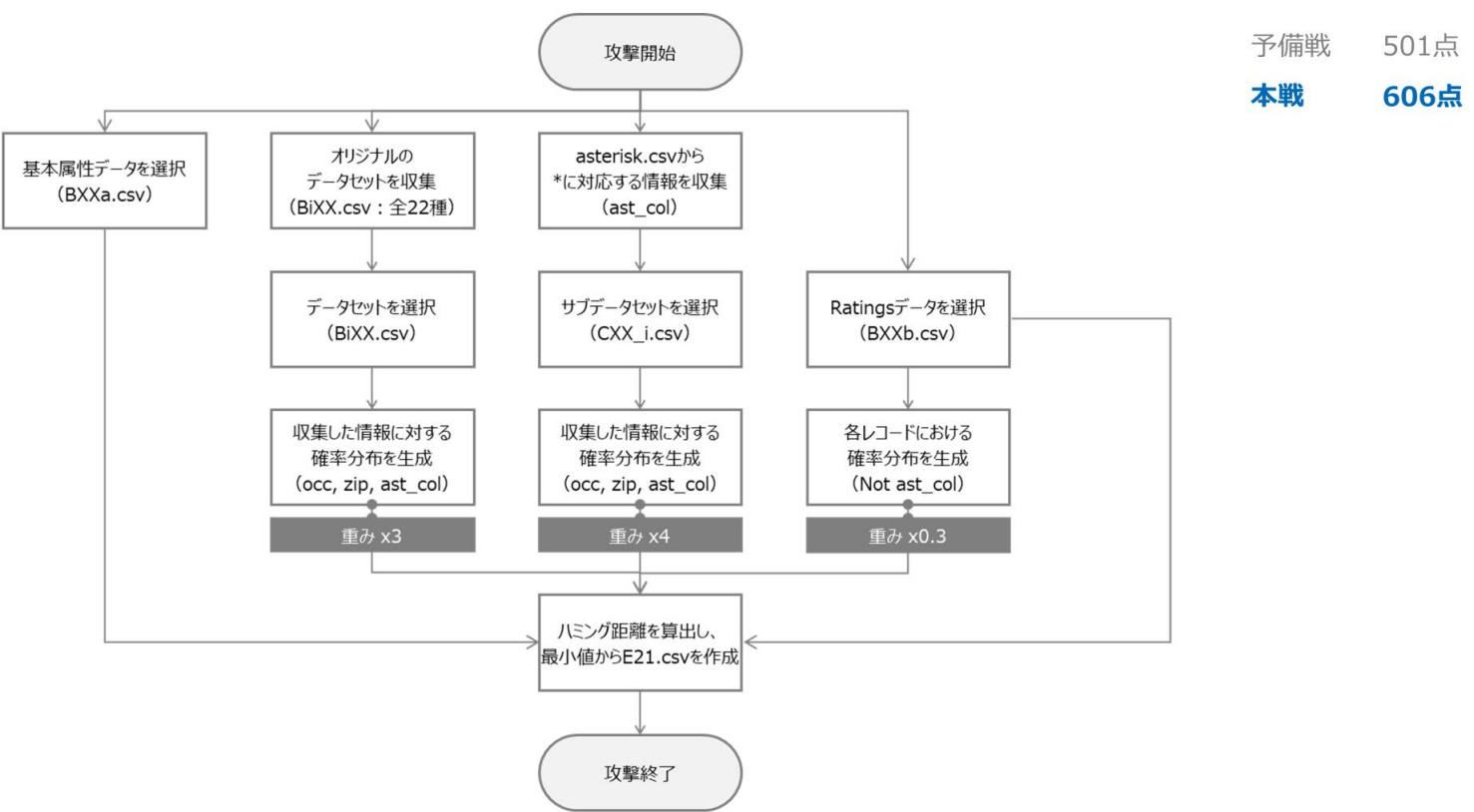
Ratingの一部を最頻値に置換

※ この時点で有用性72



攻撃フェーズ

匿名化フェーズのフローチャート

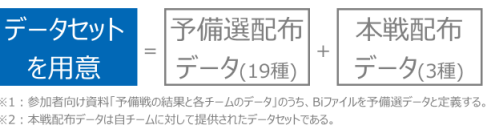


予備戦 501点

本戦 606点

確率分布を生成 #1 (オリジナルデータから生成)

匿名化サブデータ以外の信頼できるスコアを算出するのが狙い



3つのデータセットから総合的な確率分布を生成 (Occ / ZIPごとの)

確率分布を生成 #2 (匿名化サブデータから生成)

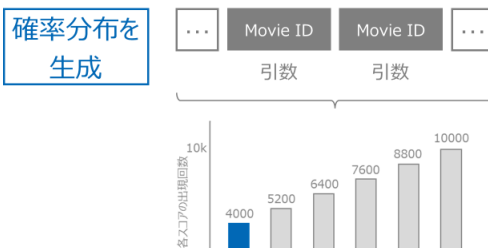
CodaBenchにて公開される有用性を反映するのが狙い



ハミング距離

確率分布を生成 #3 (Ratingsデータから生成)

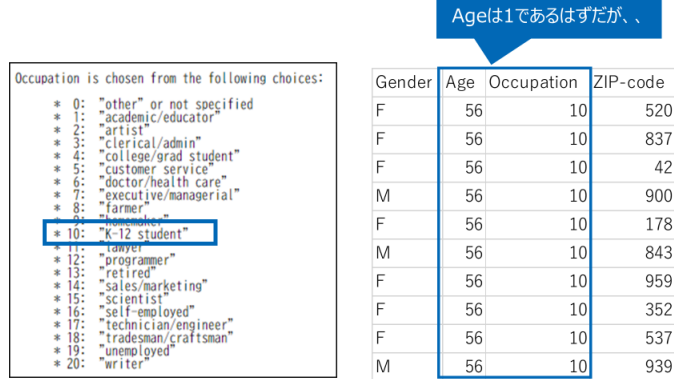
個人ごとのRatingの傾向を反映するのが狙い



BXXb.csv

データ分析の前提

分析対象カラムの選定



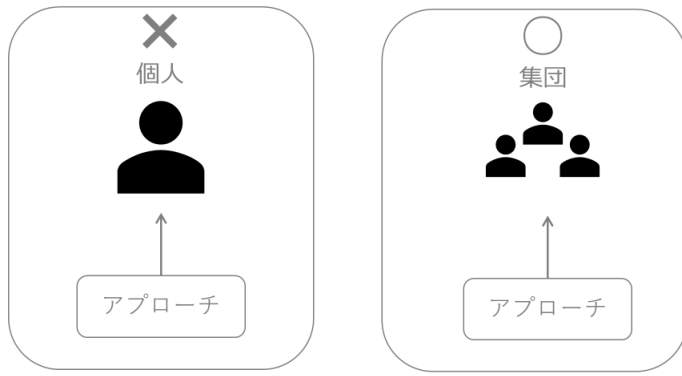
Ageはすでに丸められているので正しいと仮定するとOccupationは加工処理が実行されていると想定

ZIP-codeははややこしいため分析対象から除外Gender, Ratingsは有用性72.01より利用できると仮定

以上より、Gender, Age, Ratingsを使用

合成データの特徴

匿名加工処理により個人に限定してアプローチすることは不可能



今回与えられているデータセットのみでは有用性のある分析が困難であったため仮説を元にデータ分析や利活用を検討

データ分析 #1

データ分析①：Ratingsを元にしたEC商品のレコメンド広告

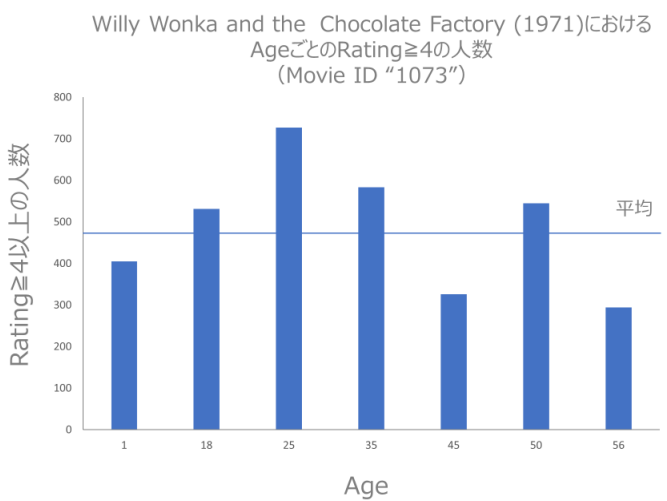
仮説

Willy Wonka and the Chocolate Factory
放映当時チョコレートの売り上げが上昇

原因

- チョコレート工場が舞台の映画
- 魅力的な映像による感覚的刺激
- ストーリーへ共感
- 映画を見た人はチョコレートが食べくなる

高評価を付与したユーザー数の多い年齢層へチョコレート商品のレコメンド



グラフを見ると25(25歳〜34歳)が一番多い。ニッポンの結婚適齢期が20代後半〜30代前半と思うと、結婚式場の引き出物で送るのが良いと結論付け。

データ分析 #2

データ分析②：映画の舞台の国への旅行ツアー作成

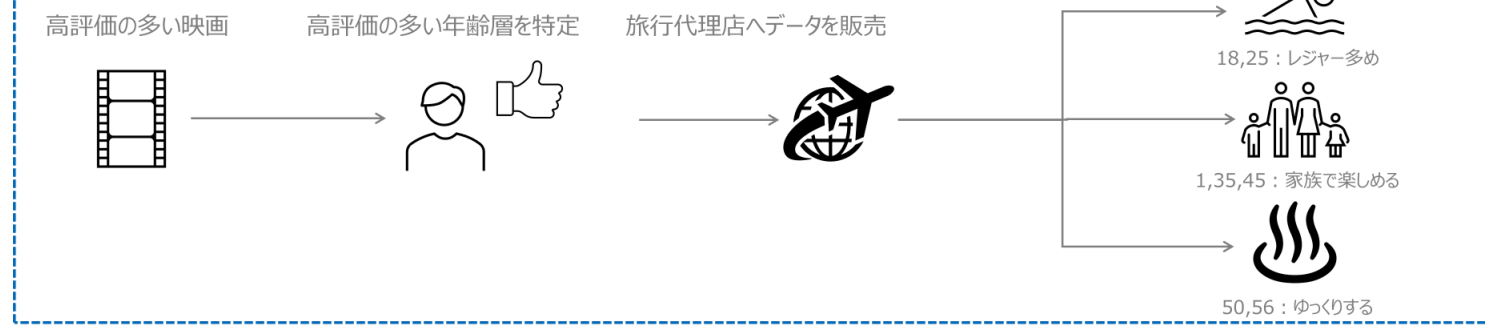
仮説

高評価を付けた映画の舞台に行ってみたい

現実にある街が舞台となっている映画 (一部抜粋)

映画名	舞台/関連場所	国名
Space Jam	ニューヨーク	アメリカ合衆国
Darby O'Gill and the Little People	ダブリン	アイルランド
FairyTale, A True Story	ヨークシャー	イギリス
E.T. the Extra-Terrestrial	ロサンゼルス	アメリカ合衆国

高評価を付与した特定の年齢層を対象とした旅行ツアーを作成



データ分析 #3

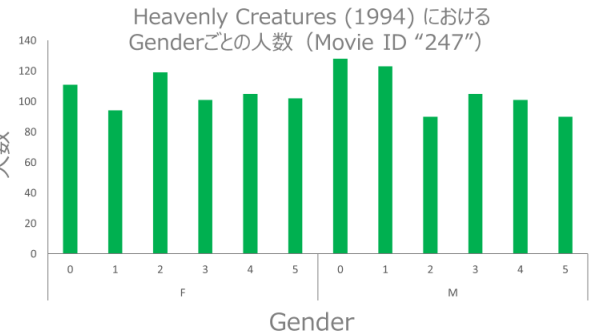
データ分析③：未成年の心の状態を推測する

仮説

未成年者が啓発的な内容の映画を観た際、その映画のレーティングを基に彼らの心理状態を推測することが可能である。

映画名	啓発的なテーマ
Heavenly Creatures	人間の心の闇、友情の影響
Star Wars, Episode IV - A New Hope	勇気、自己発見、チームワークの重要性
Willy Wonka and the Chocolate Factory	創造性、好奇心、ルールを守ることの重要性
E.T. the Extra-Terrestrial	異なる存在との共存、友情、思いやり
Labyrinth	自己成長、挑戦、解決力
Legend	善悪の対立、自然との調和、勇気と犠牲
Beetlejuice	人生と死の価値観、生活の見直し
Toys	戦争と平和、テクノロジーに対する批判的視点

データ分析により、極端なRatingを付けるなどの傾向を把握状況に応じて、親やカウンセラーに連携し、それに応じた適切な対応を実施



Movie ID "247" vs "260"のクロス集計表 (Age=1)

		Star Wars, Episode IV - A New Hope					
		0	1	2	3	4	5
Heavenly Creatures	0	23	29	27	19	32	25
	1	30	26	18	20	31	25
	2	23	13	27	20	28	18
	3	19	23	26	30	20	24
	4	25	23	25	31	23	28
	5	21	17	26	18	23	15

良作品に低評価をつける原因があります。

