



# 北海道大学

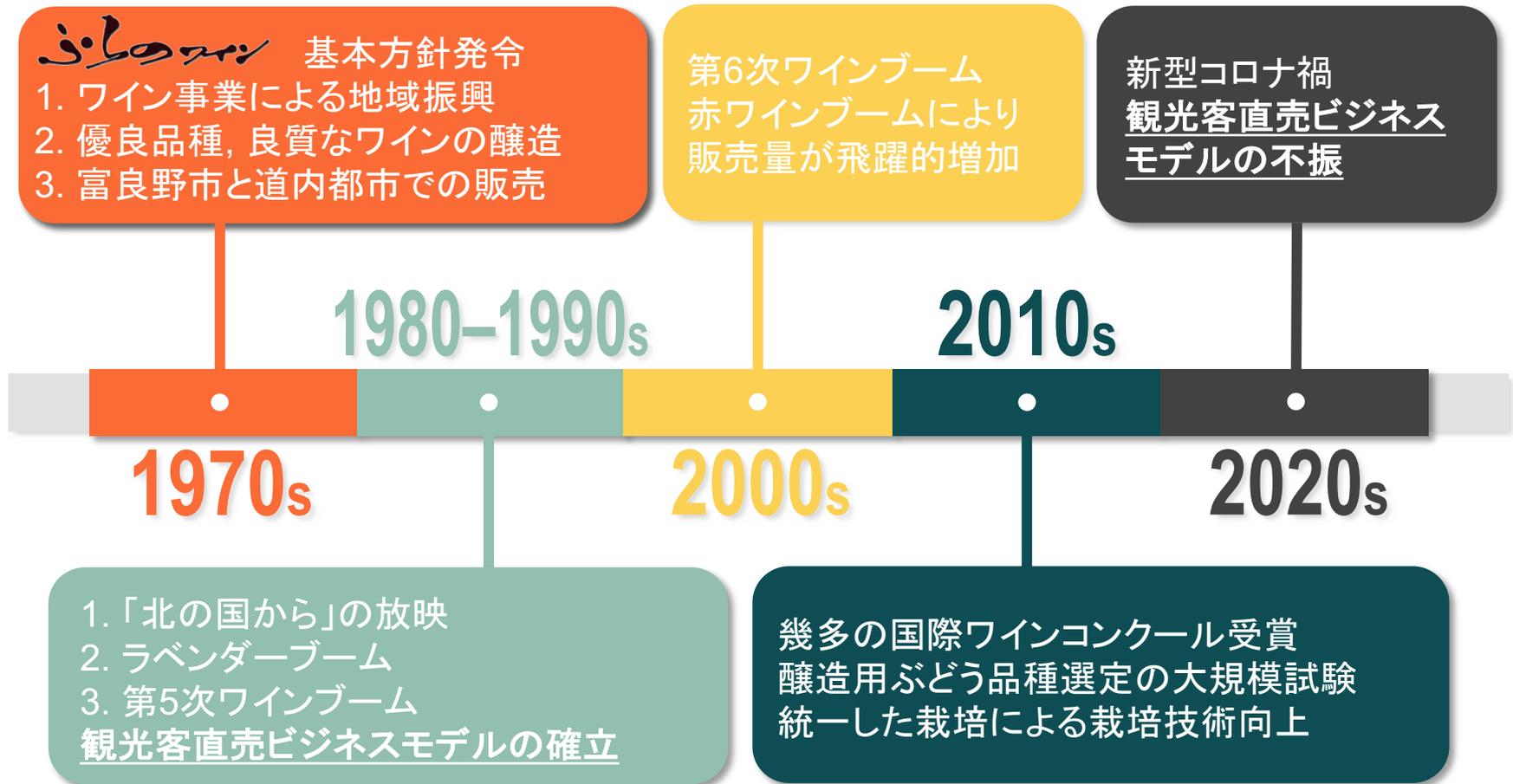
## おうちでお試しワイン

～オンライン試飲セットで売上げアップ～

2022年3月25日

北海道大学 スマート物質科学プログラム 1期生  
川向 ほの香、坂口 周弥、棚橋 慧太、椿 啓司

# コロナ禍での観光客直売ビジネスモデルの不振



# オンラインショップを重視した販売モデルに着目

p.8【令和3年度から令和7年度に向けて製造本数  
26万本を目標にするための戦略】より

- 1 直売店等売り上げ増を目標にキャンペーンや限定商品の販売  
リピーターの確保をめざしたイメージ戦略の強化
- 2 道内卸への販売強化
- 3 道外への販売強化(オンラインショップ等)**
- 4 道内・外物産展への積極的な出展
- 5 市民への販売推進  
(市内限定、オリジナル商品の販売、市民キャンペーン)
- 6 ふらのワインの強みを生かした販売
- 7 観光客誘致と観光客向けキャンペーン
- 8 果汁を使った販売戦略(ジュース・加工品)
- 9 収量実績にあった買い取り単価の検討
- 10 新規ぶどう耕作地(者)の調査及び確保の推進
- 11 ぶどう耕作者の助成
- 12 ぶどう栽培技術の指導
- 13 ぶどう品種改良等の研究
- 14 普及センター・JA ふらの・市農林課との連携と技術指導



 データので観光客直売ビジネスモデルの課題を解決しやすい

# オンラインショップの顧客データを利用したDX

“オンラインショップ”  
顧客データがあるためDXを適用可能 “ターゲットとなる人物像”

顧客データ

分析

はじめにオンラインショップの顧客像を分析  
➡ 顧客の購入意欲を促進するアイデアを考えた

# オンラインショップの顧客データを利用したDX

“オンラインショップ”  
顧客データがあるためDXを適用可能

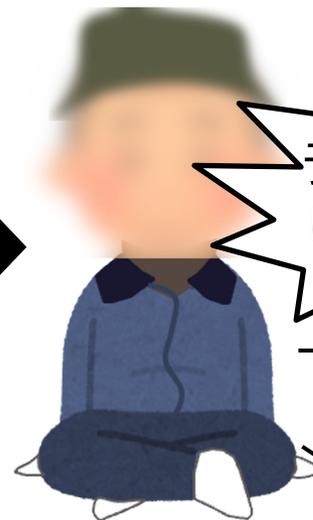


顧客データ



分析

“ふらのワインが飲みたく  
なってしまった顧客”



データを利用  
したアイデア

熊の晩酌  
買ってきてー



はじめにオンラインショップの顧客像を分析  
➔ 顧客の購入意欲を促進するアイデアを考えた



# データの分析

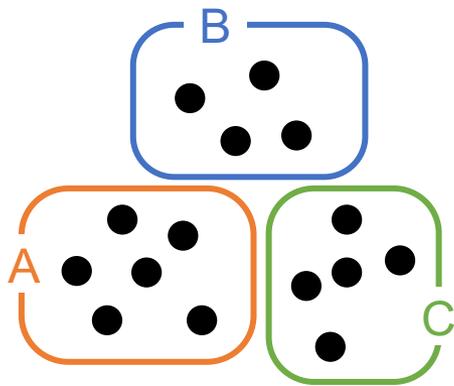
## ○分析データ

「オンラインショップ利用状況（2017年11月～2019年12月）」

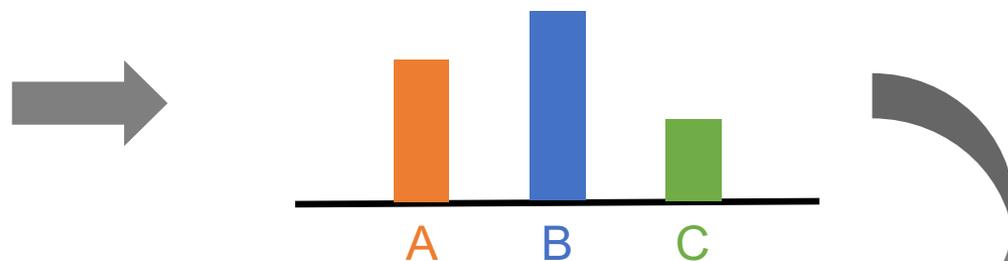
- ・分析に用いた項目：郵便番号、購入端末、支払方法、配送時間、商品名、商品単価、送料、手数料、合計金額（事前に、項目を数値化する「データ準備」を行った）
- ・ふらのワインの認知度を鑑み、富良野市からの注文を分析から除外

## ○分析の方針

① データをグループに分類



② 各グループの特徴量を比較



③ 各グループの傾向を把握し  
販売戦略につなげる

# ① データをグループに分類

## ○分類方法1 「リピートの有無での分類」

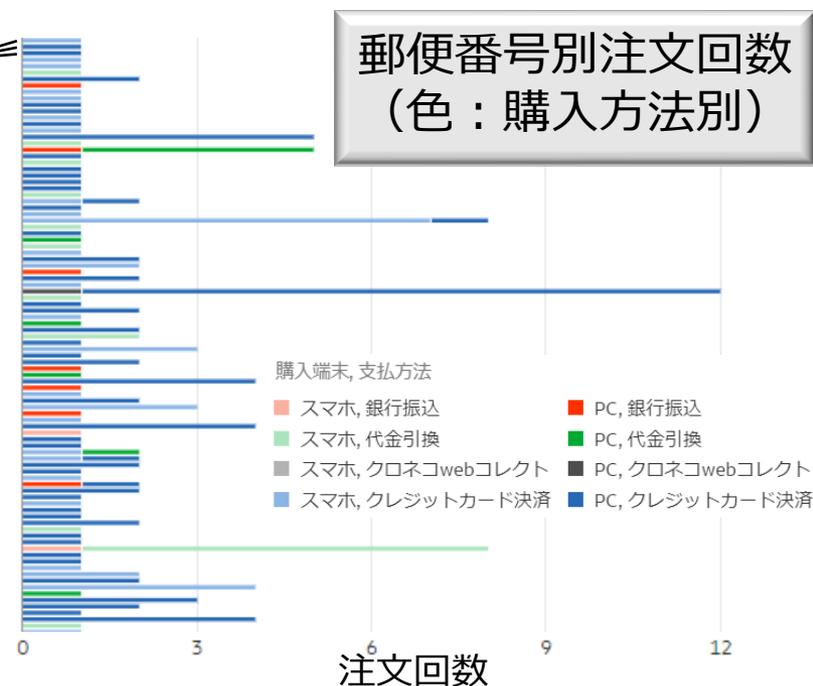
- ・「買わなかった人」に関するデータが存在しないため、買わなかった人と買った人との違いを分析し、新規購入者増加を検討することは不可能
- ・一方でデータを「1回きりの購入者」と「リピーター」に分け、リピーター増加を検討することは可能

## ○リピーターの判断法

- ・今回は氏名及び番地等の情報を使用できないため、正確なリピートの有無の判断は困難
- ・同一郵便番号から複数回注文があった場合、購入方法が同じであることが多い。

➡ 同一郵便番号を同一購入者とみなし  
リピートの有無でグループ分け

各郵便番号  
(一部抜粋)



# ① データをグループに分類

## ○分類方法2 「機械学習が示す指標による分類」

- ・機械学習の導入で多くの特徴量を同時に考慮したデータの分類が可能
- ・数理的アルゴリズムによって、人間がデータを眺めるだけでは判断できない、また人間の主観に左右されない分類が可能

## ○行った分類の手順

特徴量

X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
その他ワイ果汁	セット	箱	合計本数	合計量	ギフト	ラベル	送料	購入全本数	合計金額	
0	1	0	0	12	8640	0	0	0	41	20385
0	0	1	0	2	720	1	0	500	2	6098
0	0	0	0	12	8640	0	0	0	57	15804

### 1. データの散らばり方を見る

➡ 主成分分析

グラフ化不可能な、多くの特徴量を持つデータを  
情報量の損失を抑えつつ少ない次元で表現可能

### 2. 数学的アルゴリズムでグループ分けをする

➡ k-平均法

各特徴量を軸とする高次元空間に散らばるデータを、  
その距離を基準として任意の数のグループに分類

### 3. グループ分けに大きく寄与した特徴量を考察



# ① データをグループに分類

## ○分類方法2 「機械学習が示す指標による分類」

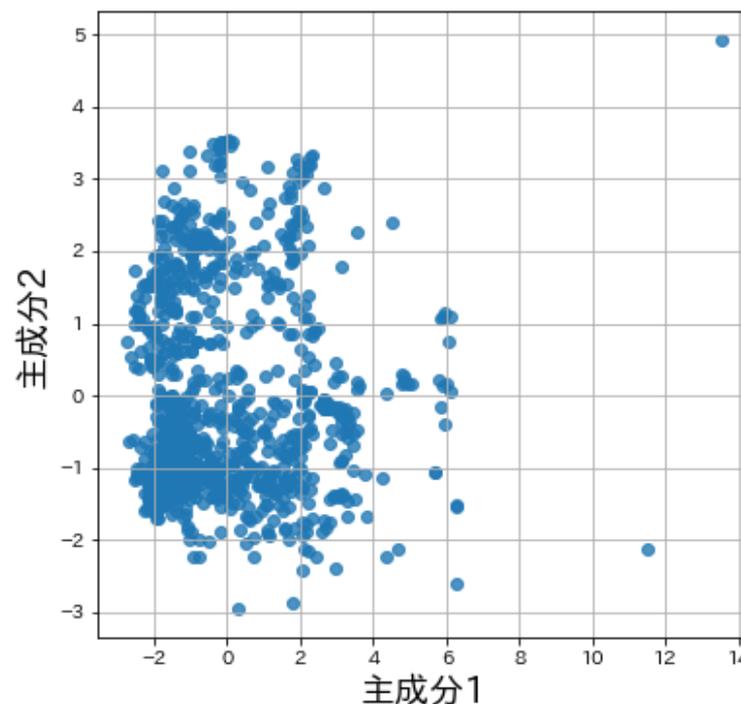
### 1. データの散らばり方を見る ➡ 主成分分析

分析用の準備データは29の  
特徴量を持つ  
→29軸の可視化は不可能

情報量が多くなる（ばらつきが  
大きくなる）順に29次元空間で  
軸を取り直す

順に2軸を取り出して可視化

散らばりを見るだけでは  
分類の基準を決められず



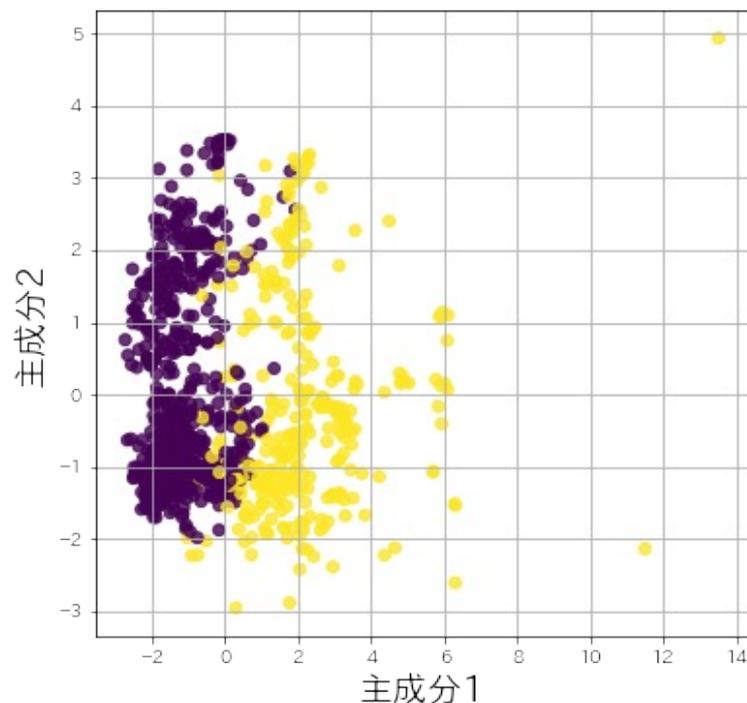
# ① データをグループに分類

○分類方法2 「機械学習が示す指標による分類」

2. 数学的アルゴリズムでグループ分けをする → k-平均法

各データをk-平均法を用いて  
2グループに分け、主成分分析で  
作成した散布図で分類結果を表示

主に軸「主成分1」の方向で  
2グループに分類された。



どの特徴量が分類に大きく寄与したのか？



# ① データをグループに分類

## ○分類方法2 「機械学習が示す指標による分類」

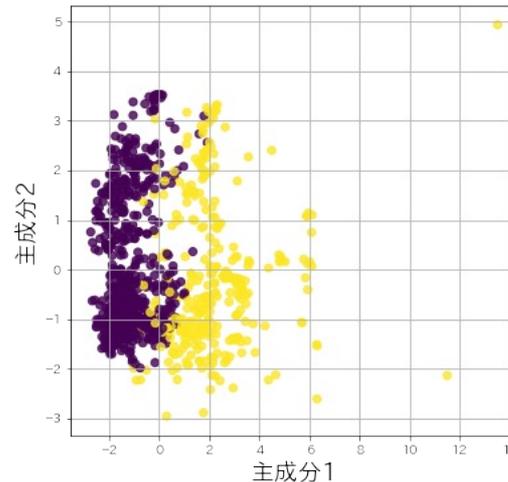
### 3. グループ分けに大きく寄与した特徴量を考察

軸「主成分1」への寄与が大きかった特徴量の中で「郵便番号毎の全購入本数」に注目

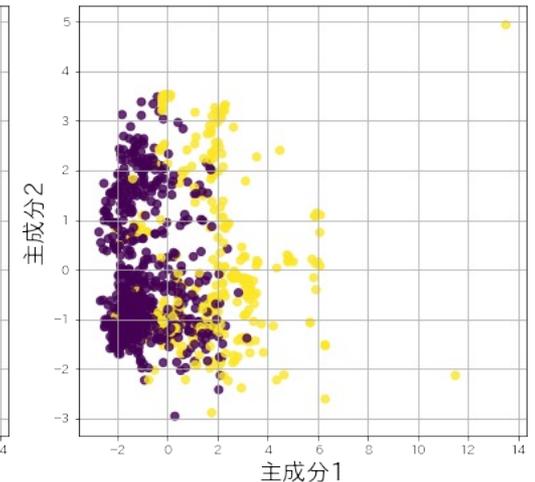
この特徴量に関し「〇本以上 / 未満」の2つに要素を分類、k-平均法のクラスターと比較

22本以上 / 未満で分類すると最も高い82.2%の一致

k-平均法での分類



購入本数計22本以上 / 未満での分類



2年間で合計22本以上購入した「大量購入者」と22本未満の「少量購入者」で購入の傾向が異なる可能性が高い！



# ① データをグループに分類

## ○分類方法



※ “2年間”での“オンラインショップ内”での「1回きり」であることに注意

※ リピーターは郵便番号判断のため、異なる購入者が含まれる可能性に注意

## ○グループ毎に比較するデータ

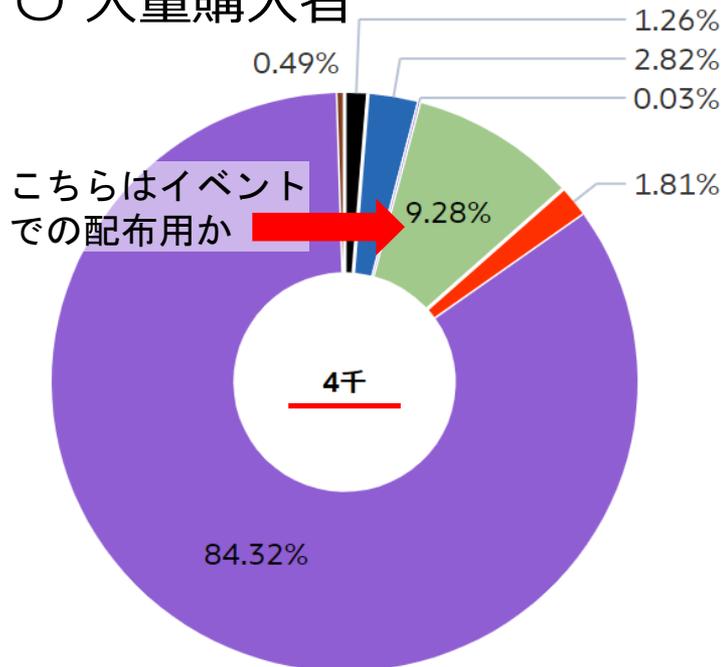
1. 「購入本数」と「瓶の容量別の購入割合」
2. 「商品別の購入本数」



# データの解析

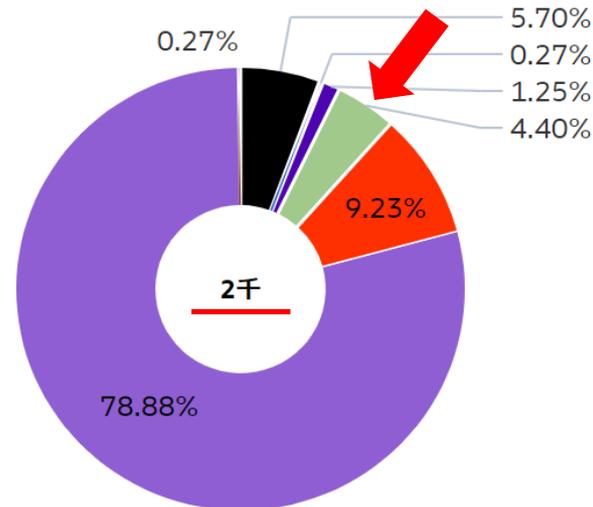
## 「購入本数」と「瓶の内容容量別の購入割合」

○ 大量購入者

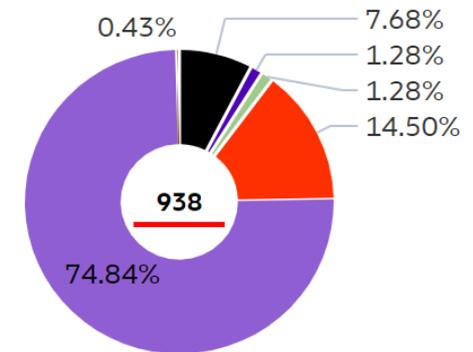


○ 少量購入者

・ 1回きりの購入者



・ リピーター



■ 0 ■ 180 ■ 200 ■ 360 ■ 500 ■ 720 ■ 750 (mL)

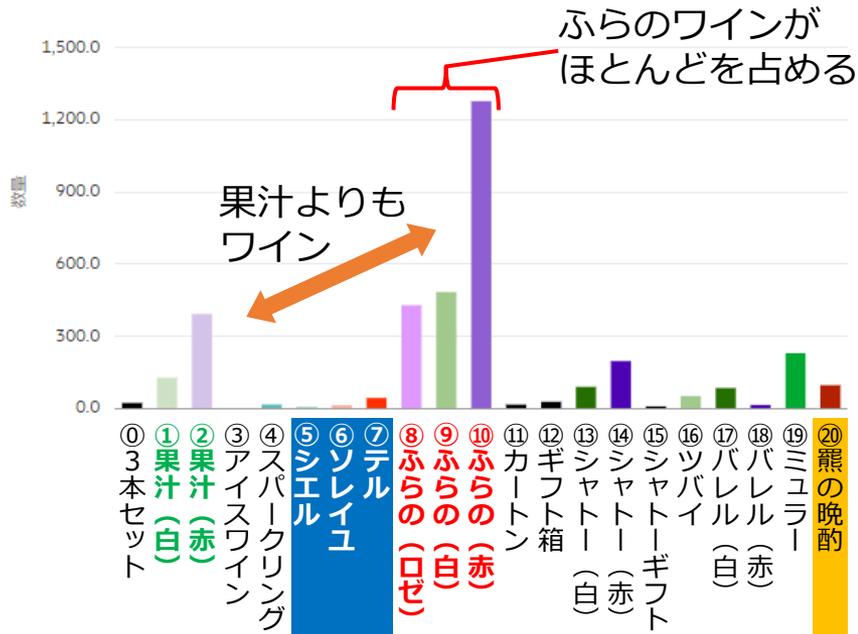
- ・ 1回きりの購入者とリピーターの大きな差は**360mL瓶の購入傾向**  
→ 「ワインを試したいが720mLでは多い」という客層がふらのワインのみ買って満足か？



## ② 各グループの特徴量を比較

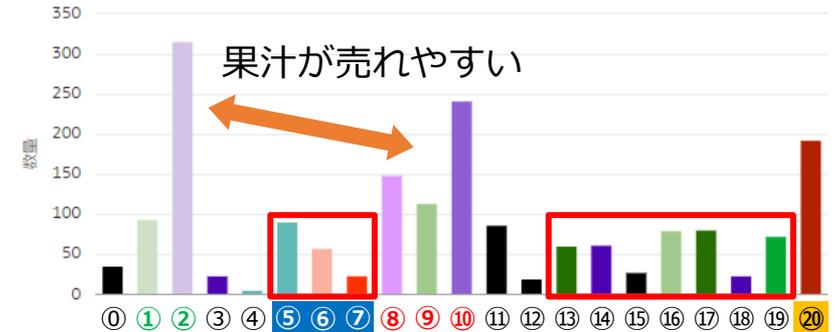
### 「商品別の購入本数」

#### ○ 大量購入者



#### ○ 少量購入者

##### ・ 1回きりの購入者



##### ・ リピーター



- ・ 1回きりの購入者は「果汁、ふらのワイン、熊の晩酌」の購入量の多さが特徴的であったが、**リピーターはその他のワイン（赤粋）の割合が増加**  
→いろいろな飲みたいと思った人がリピートした可能性がある。



# 提案

## データ

大量購入者はおそらく業者  
 コロナ禍の現在、購入量を増やすのは難  
 ⇒少量購入者（個人）に注目



## 現地で視察

試飲した銘柄が売れやすい傾向

試飲すると  
欲しくなる



→現地に来れなくても  
試飲できる方法はないか？

オンラインショップで試飲セットを販売



# 新規購入者増加のための試飲セットの具体例

**内容：**試飲セットを数パターン用意する。

ふらのワイン赤、白、ロゼ  
(基本セット)



テル、シエル、ソレイユ  
(甘め)



シャトーふらの赤、白  
(高級感)



**容器：**ミニボトル、パウチ

**送料：**アンケートの回答者は無料



ビンよりも軽量  
送料が安い？

イメージ図

- ・回答率アップ
- ・データ解析に利用
- ・リピーターを狙った情報提供

## 今後のデータ解析のための提案

- ・オンラインショップを利用してもらうためには、そもそもWebページを閲覧してもらわなければいけない。  
→SNSの積極的な活用や広告にも重点を置く。



- ・オンラインショップでの売上向上のため利用してもらうWebページを新しくすることは非常に有効な手段  
→閲覧履歴などを追跡できるようにするとより多様な種類のデータが集まる。

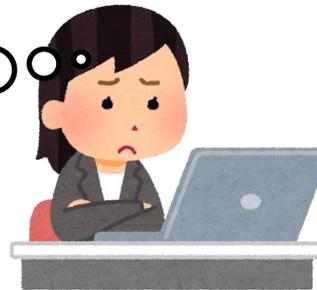


# 現在のWebページでの課題点

- ・オンラインショップはわざわざ別ページとして開き、煩雑に感じる。



- ・古い印象があると、個人情報の入力にためらってしまう可能性が出てくる。

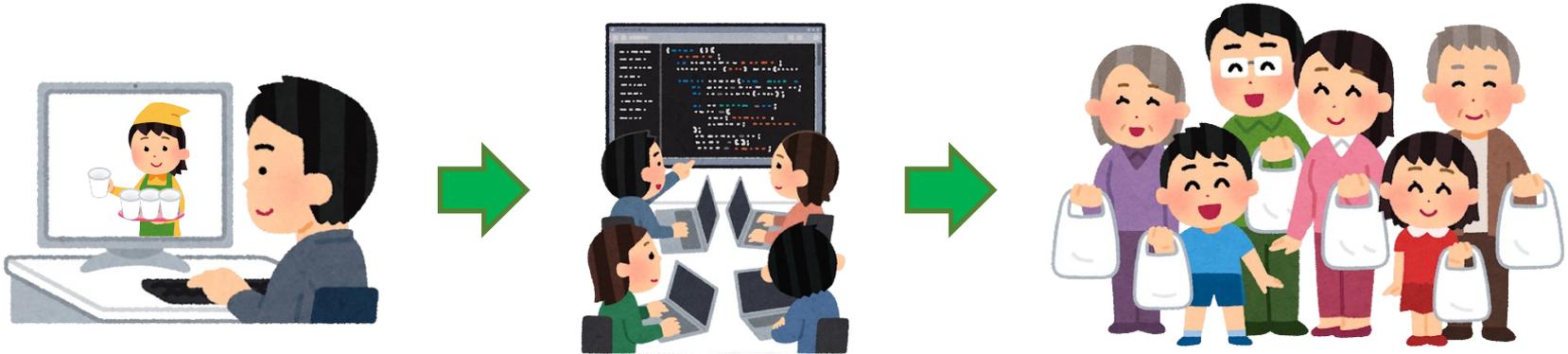


クレジット情報    住所  
家族構成    利用用途



# まとめ

- ・ **試飲セットを提供する**ことで、**新規購入者**を増やす。



- ・ 試飲セット購入の際に、**アンケート**を答えてもらう。  
今後のデータ解析に用いるデータを得る。  
→売上の傾向や次回購入の提案をすることで  
リピーターを増やす。



- ・ **Webページの改修**によりオンラインショップの利便性の向上。  
→Webページの閲覧時間などの新しい種類のデータを得ることで  
解析をより正確にする。



## データ準備

- ・「商品名」と「内容量」が一体化されていたため、それら二つに分解
- ・ギフト箱やカートンを購入した人に「ギフト」と、ラベンダーラベルを選んで購入した人に「ラベル」とラベル付けを行った
- ・変数が多くなることを防ぐため、熊の晩酌及びミュラートウルガフ以外のワインを「ふらのワイン」「果汁」「高めのワイン」「甘めのワイン」「その他ワイン」で分類した
- ・郵便番号から「緯度」と「経度」を算出し、さらに主成分分析で1次元化したものを居住地の特徴量とした
- ・関連を考慮し、購入者居住都道府県のワイン購入量を特徴量に加えた
- ・「注文日時」から「注文月」と「注文時間(hour)」を抜き出し、さらに平日休日の判別(1 or 0)を特徴量に加えた
- ・「購入端末」「配送時間」「支払方法」はそのままダミー変数化

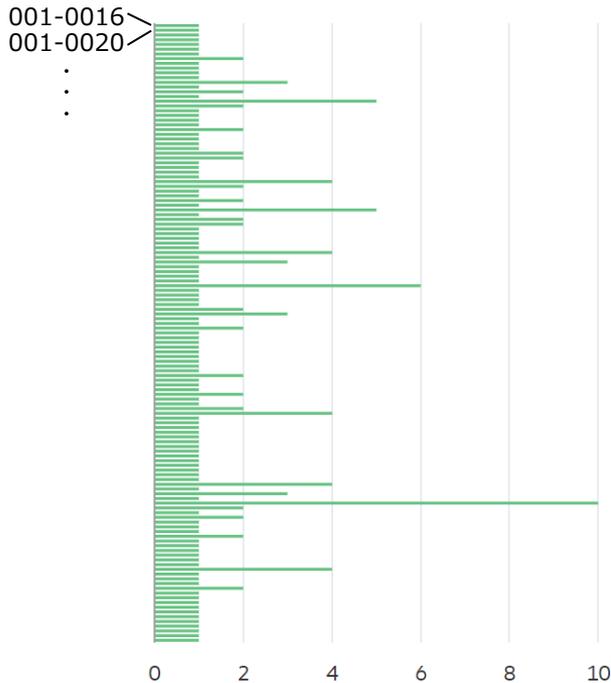


# データの解析

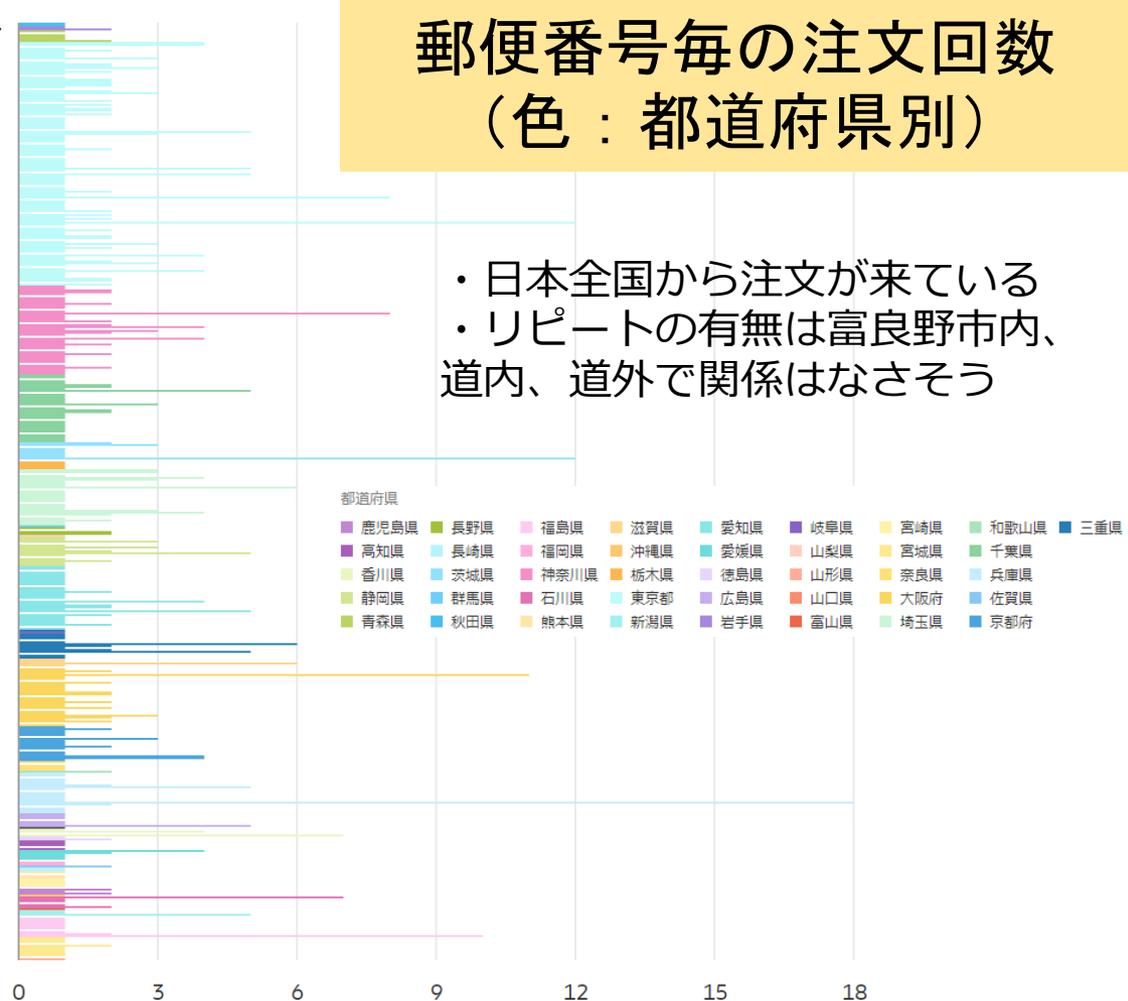
### 富良野市内



### 道内 (富良野市外)

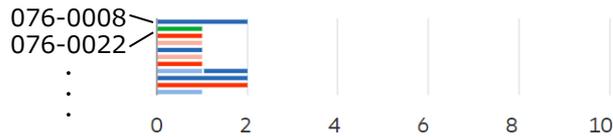


### 道外

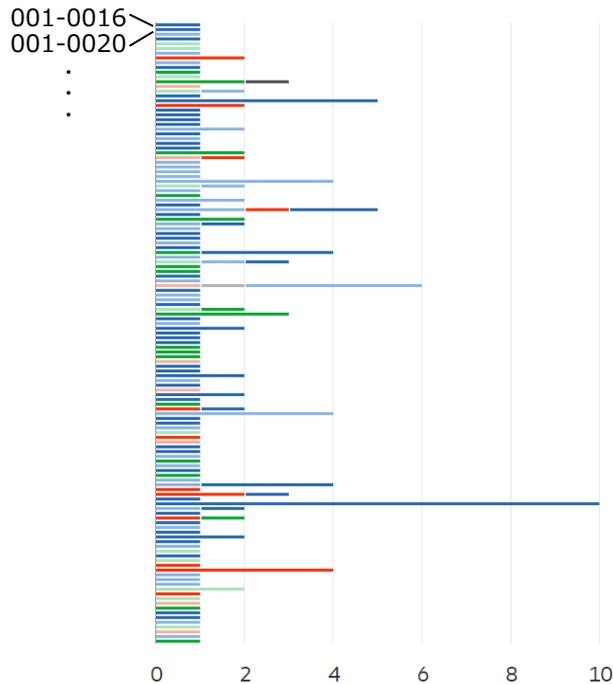


# データの解析

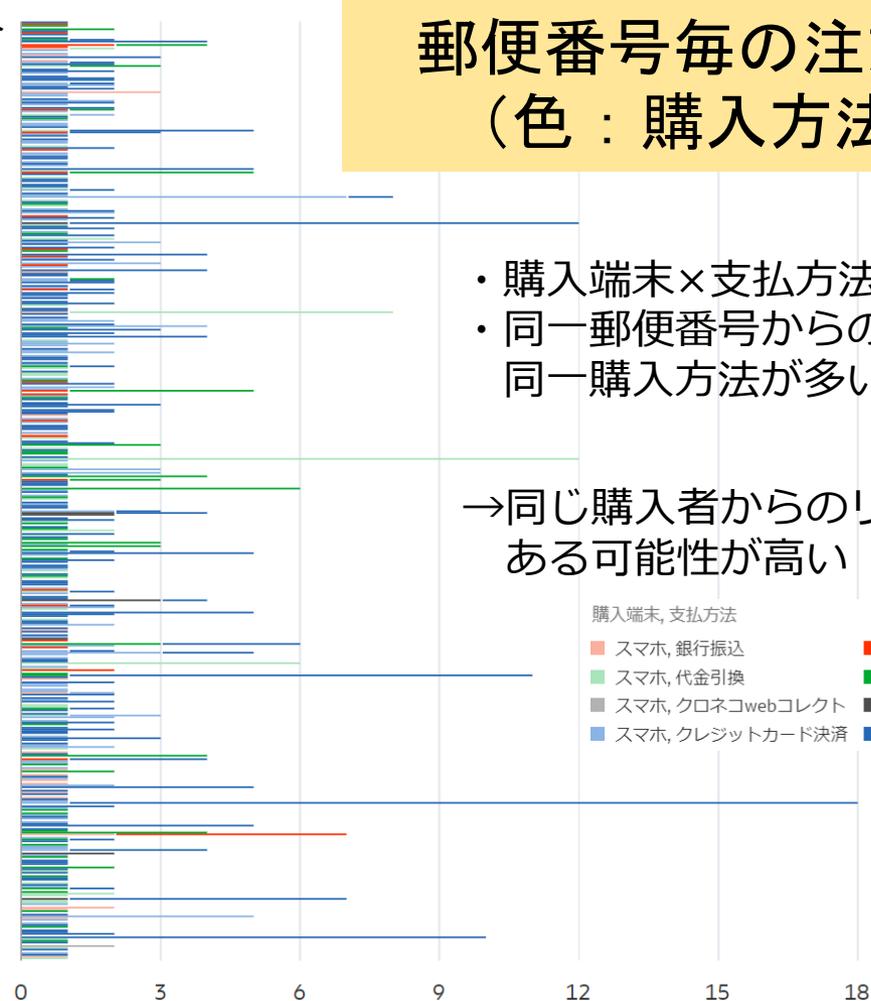
## 富良野市内



## 道内 (富良野市外)



## 道外



## 郵便番号毎の注文回数 (色：購入方法別)

- ・購入端末×支払方法で色分け
- ・同一郵便番号からの注文は同一購入方法が多い

→同じ購入者からのリピートである可能性が高い

購入端末, 支払方法

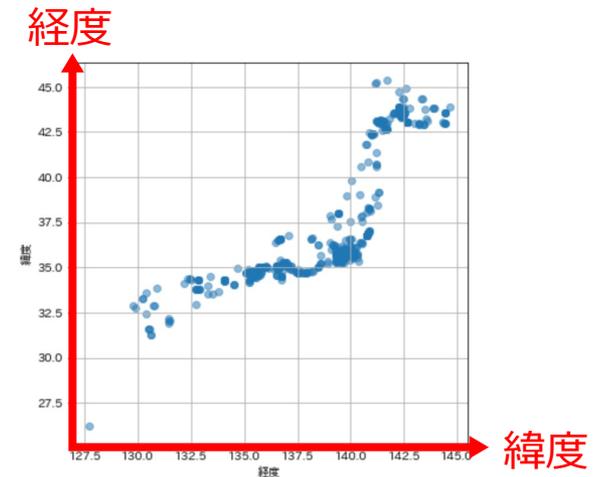
- |                    |                   |
|--------------------|-------------------|
| ■ スマホ, 銀行振込        | ■ PC, 銀行振込        |
| ■ スマホ, 代金引換        | ■ PC, 代金引換        |
| ■ スマホ, クロネコwebコレクト | ■ PC, クロネコwebコレクト |
| ■ スマホ, クレジットカード決済  | ■ PC, クレジットカード決済  |



# 主成分分析

散布図でデータの散らばり方を見たい

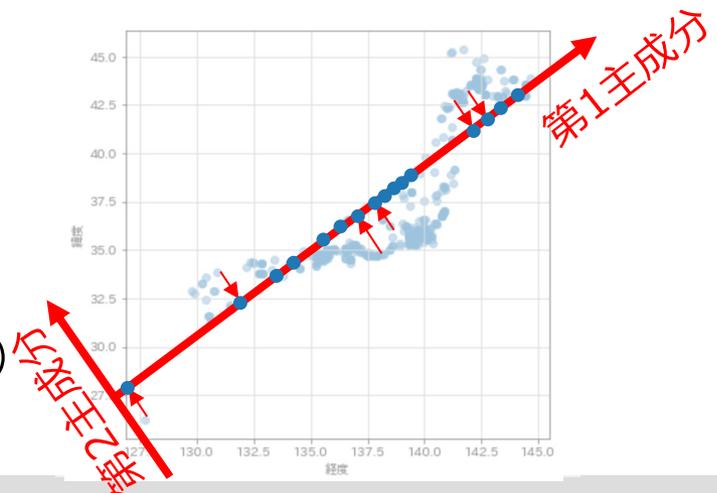
例えば購入者の居住地情報は「緯度」と  
「経度」（郵便番号から変換した）の2変数で示せる  
→2次元の散布図で可視化できる



しかし、分析用のデータは、約30の項目を  
持つため、これらの情報を全て含めた散布図は作れない

→「主成分分析」で軸を減らす

例えば先の2変数で表された居住地情報は  
右のように87.8%の情報を維持したまま  
1次元化（軸: 第1主成分）できる  
（軸: 第2主成分を合わせると情報量は100%）



# k-平均法

## k-平均法でクラスタリング

k-平均法は任意の数のクラスター（集団）に各要素（データ点）を分ける

### アルゴリズム

1. 人が指定するクラスター数だけデータ点をランダムに選択
2. ルールを満たすようにクラスターを拡大
3. クラスターに未所属のデータ点がなくなったら終了
3. 任意の回数1~3を繰り返し、最もルールを満たす状態をクラスタリングの結果とする

### ルール

- 1つのクラスター内で重心から各要素までの距離が小さくなる
- クラスター同士の重心の距離が遠くなる



# k-平均法

「k-平均法」でクラスタリング

k-平均法は

- ・ 1つのクラスター内で重心から各要素までの距離が小さくなる
- ・ クラスター同士の重心の距離が遠くなる

ように、各要素（データ点）を任意の数のクラスター（集団）に分ける

例えば以下の要素を3つのクラスターに分類



# 新規購入者増加のための試飲セットの具体例

