

# PLASMA 新機能説明会

2020年2月28日(金)

14:00～17:00

産総研 臨海副都心センター

産業技術総合研究所  
人工知能研究センター  
確率モデリング研究チーム



# 説明会の流れ

1. PLASMAについて
2. 5つのユースケース紹介
  1. ID付データからの顧客理解
  2. 医療データ活用事例
  3. 産業系IoTデータ活用
  4. 健康・生活支援
  5. 大規模集客・観光
3. 実データを用いた一連のPLASMA利用・操作例  
(最新機能の紹介を含む)

# PLASMA:

## Probabilistic Latent Semantic Structure Modeling API

PLASMAって何？



「確率的潜在意味構造モデリング」のための Java 言語による API セット  
です

<https://youtu.be/u2cXLdTAYNs>

のPLASMA紹介ムービーをまずは御覧ください

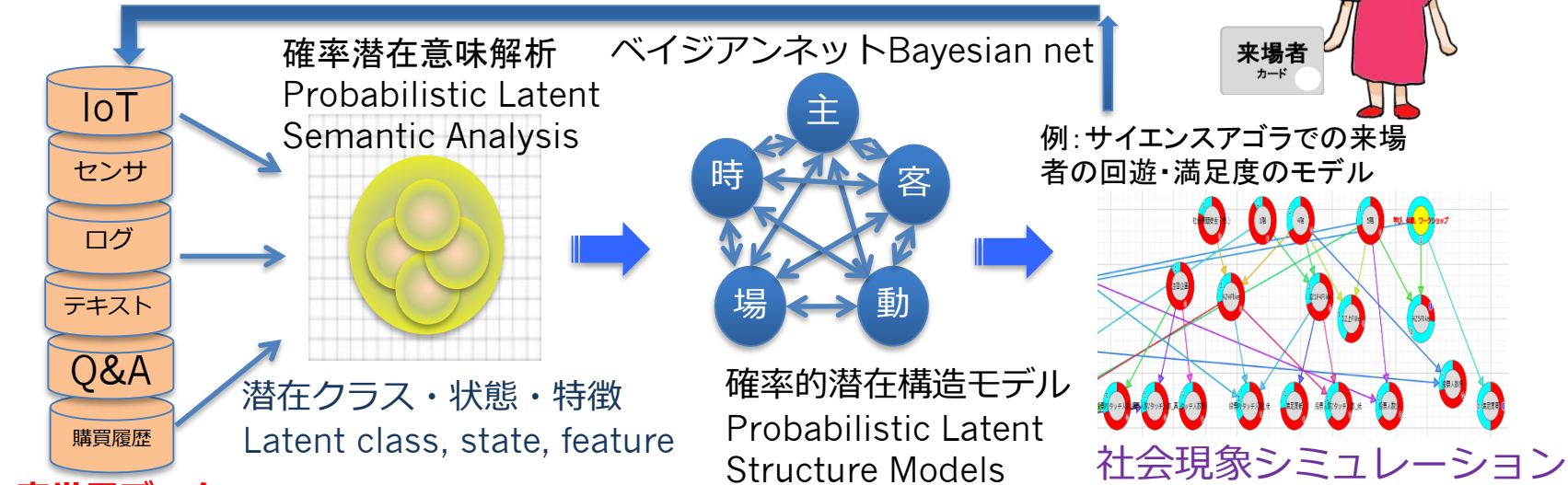


# PLASMAを用いた、価値創出のデジタルトランスフォーメーション

PLASMAで何が出来るの？

- **実社会のデータ**から**計算モデル**を自動構築
- **未来の現象生成(予測)と制御**
- **継続的な価値の向上**

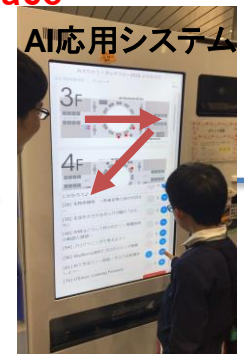
が出来ます



実世界データ  
Physical data

サイバーフィジカルモデリング

計算空間  
Cyber space



未来の現象生成・制御へ



# PLASMAのユースケース

PLASMAの具体的活用例は？



1. 小売店にて店長が: ID-POSデータを用いて真のロイヤルカスタマーの発見
2. 病院にて**医師や院長**が: **DPCデータ**の可視化とモデル化、各病院や地域の診療行為の特性の把握と改善へ
3. 状況や環境に応じた**工場**の管理
4. **健康**イベント: 継続してイベントに参加してもらい病気になるのを防ぐ
5. 回遊型展示大規模**イベント**にて: イベントが産み出す価値の継続的な向上→観光振興にも応用

例: 大規模イベント・サイエンスアゴラでのAIタッチラリー



あなたと同じタイプの人には、これらブースが人気です



# 1. ID-POSデータを用いて真のロイヤルカスタマーの発見

ロイヤルカスタマーを見つけたい



# 活用: 真のロイヤルカスタマーの発見

所属確率(%)

x: 顧客

z1

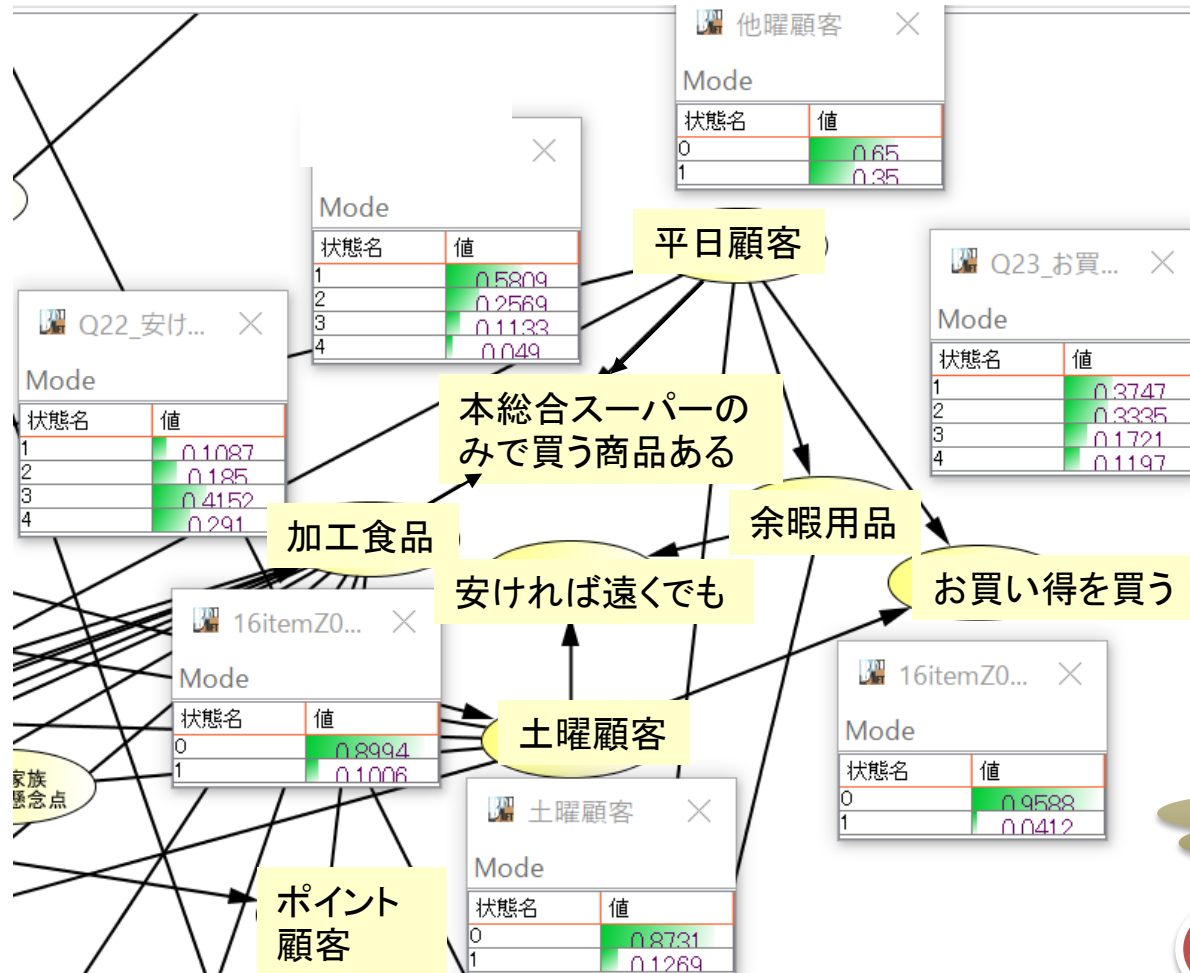
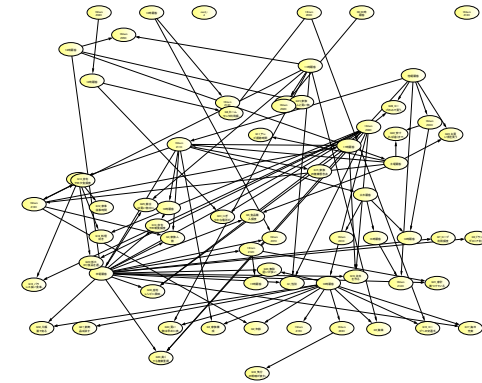
z2

z3

ID34***	2%	83%	15%
---------	----	-----	-----

顧客ID	timeZ1	timeZ2	timeZ3	itemZ1	itemZ2	itemZ3	itemZ4	...
ID34***	0	1	0	0	0	1	0	...

ベイジアンネットワーク全体図



本総合スーパーのみで買う商品あり

- 高くても健康重視
- 余暇用品を買う傾向
- 土・日・ポイントデー顧客
- お買い得品を買う

品質を重視しながらもかしくお  
買い物のお客さんか





# ID-POSデータの活用



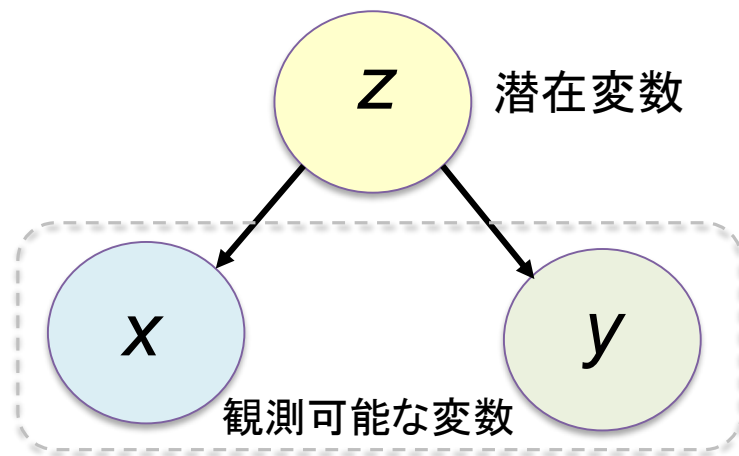
ID-POS: 「顧客ID」「商品ID」を「いつ買ったか」

||

人の「来店行動」・「購買行動」

ID-POSデータ

顧客ID	年月日	商品コード
10029	20090926	213
10046	20090129	446
10078	20090714	2779
10114	20090802	4003
10114	20090802	446
10133	20090221	446
10134	20090217	102
10134	20090217	446
10136	20090119	2779



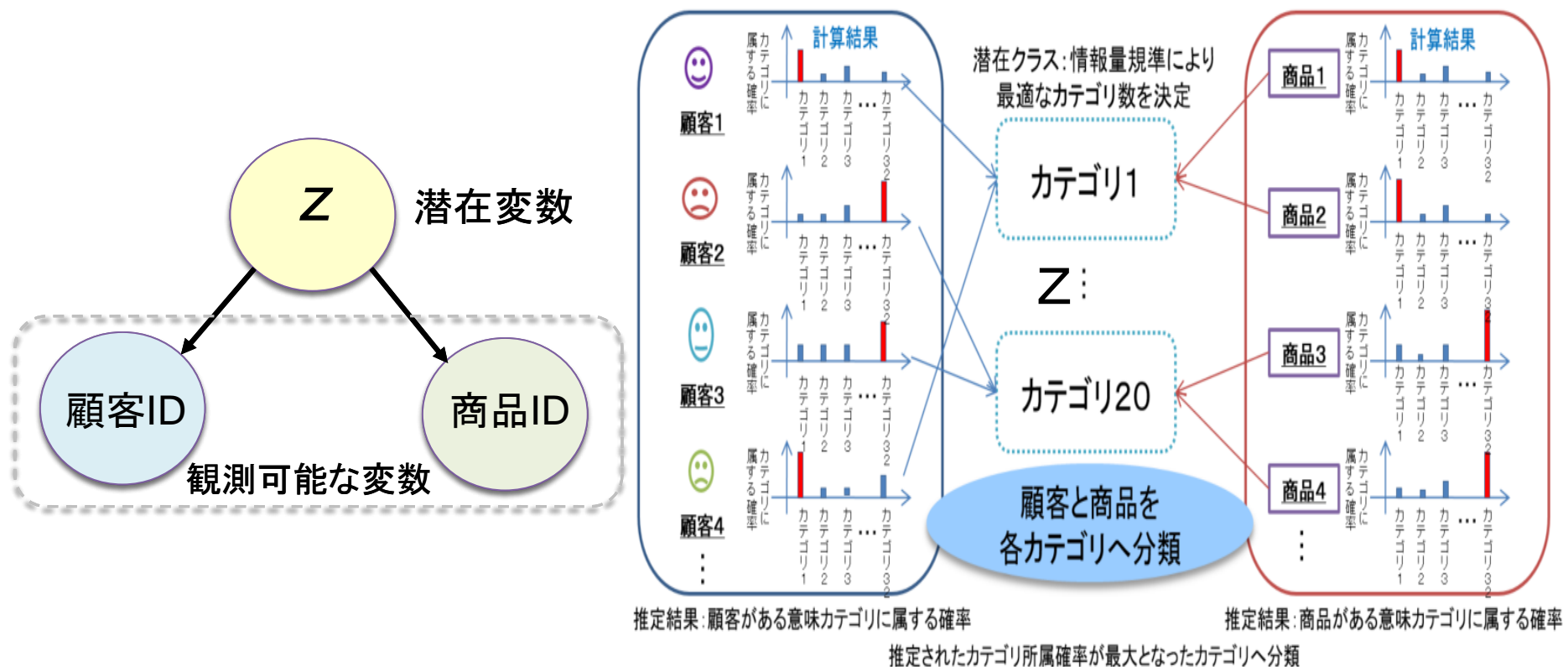
# 顧客と商品のpLSA



ID-POS: 「顧客ID」「商品ID」を「いつ買ったか」

||

人の「来店行動」・「購買行動」



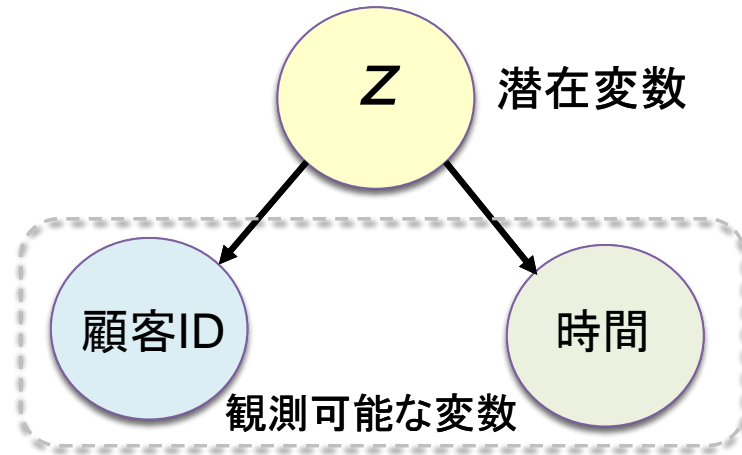
# 顧客と時間のpLSA



ID-POS: 「顧客ID」「商品ID」を「いつ買ったか」

||

人の「来店行動」・「購買行動」



# 来店行動の潜在的時間意味

y: 日	所属確率(%)	z		
		z1	z2	z3
2010/5/22 土	91	8	1	
2010/5/23 日	17	83	0	
2010/5/24 月	30	0	70	
2010/5/25 火	4	36	60	
2010/5/26 水	34	8	59	
2010/5/27 木	34	11	54	
2010/5/28 金	5	25	70	
2010/5/29 土	94	4	1	
2010/5/30 日	20	80	0	
2010/5/31 月	27	0	73	
2010/6/1 火	7	34	59	
2010/6/2 水	33	7	60	
2010/6/3 木	35	11	55	
2010/6/4 金	6	23	71	
2010/6/5 土	96	4	0	
2010/6/6 日	20	80	0	
2010/6/7 月	26	0	74	
2010/6/8 火	5	35	60	
2010/6/9 水	34	7	59	
2010/6/10 木	35	10	54	
2010/6/11 金	7	23	70	
2010/6/12 土	92	7	1	
2010/6/13 日	18	82	0	
2010/6/14 月	28	0	72	
2010/6/15 火	6	32	61	
2010/6/16 水	31	11	58	
2010/6/17 木	34	12	54	
2010/6/18 金	6	23	71	

結果的に  
曜日で分かれた

z1: "土曜" クラスタ  
z2: "日曜" クラスタ  
z3: "平日" クラスタ

と呼べそうだ

## 例外一覧

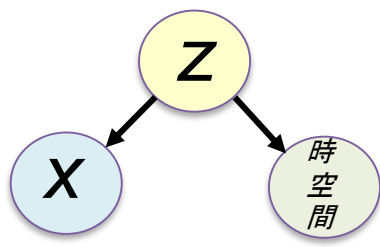
z1なのに土以外  
z2なのに日以外  
z3なのに平日以外は無し

例外は祝日と、  
年末年始・お盆のみ

☆は国民の祝日

☆	2009/1/1 木	70	30	0
	2009/1/2 金	44	46	10
☆	2009/1/12 月	55	17	27
☆	2009/2/11 水	53	28	19
☆	2009/3/20 金	35	40	24
☆	2009/4/29 水	53	30	18
☆	2009/5/4 月	55	18	28
☆	2009/5/5 火	30	54	17
☆	2009/5/6 水	48	31	21
☆	2009/7/20 月	57	17	26
	2009/8/13 木	31	44	25
	2009/8/14 金	28	35	37
☆	2009/9/21 月	58	12	30
☆	2009/9/22 火	33	48	19
☆	2009/9/23 水	50	27	23
☆	2009/10/12 月	56	17	28
☆	2009/11/3 火	31	54	15
☆	2009/11/23 月	61	14	25
☆	2009/12/23 水	47	38	15
	2009/12/29 火	27	41	32
	2009/12/30 水	32	44	24
	2009/12/31 木	34	47	19

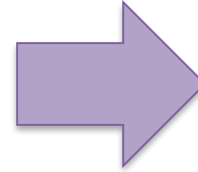
# 潜在的時空間意味解析



記号

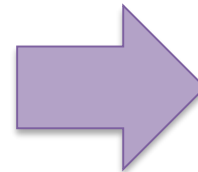
再利用可能な知識

時間  
7時半  
12時半  
19時



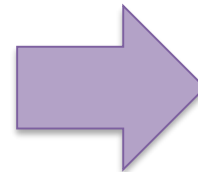
潜在的時間意味  
食事中

空間  
〒450-0002  
〒060-0001  
.  
.



潜在的空間意味  
大きな駅から徒歩圏内

時空間  
2017年5月23日16時@〒450-0002  
2015年7月12日10時@〒060-0001  
.  
.



潜在的時空間意味  
天気:晴れ  
気温:22~24°C  
快適



# 活用: 真のロイヤルカスタマーの発見

所属確率(%)

x: 顧客

z1

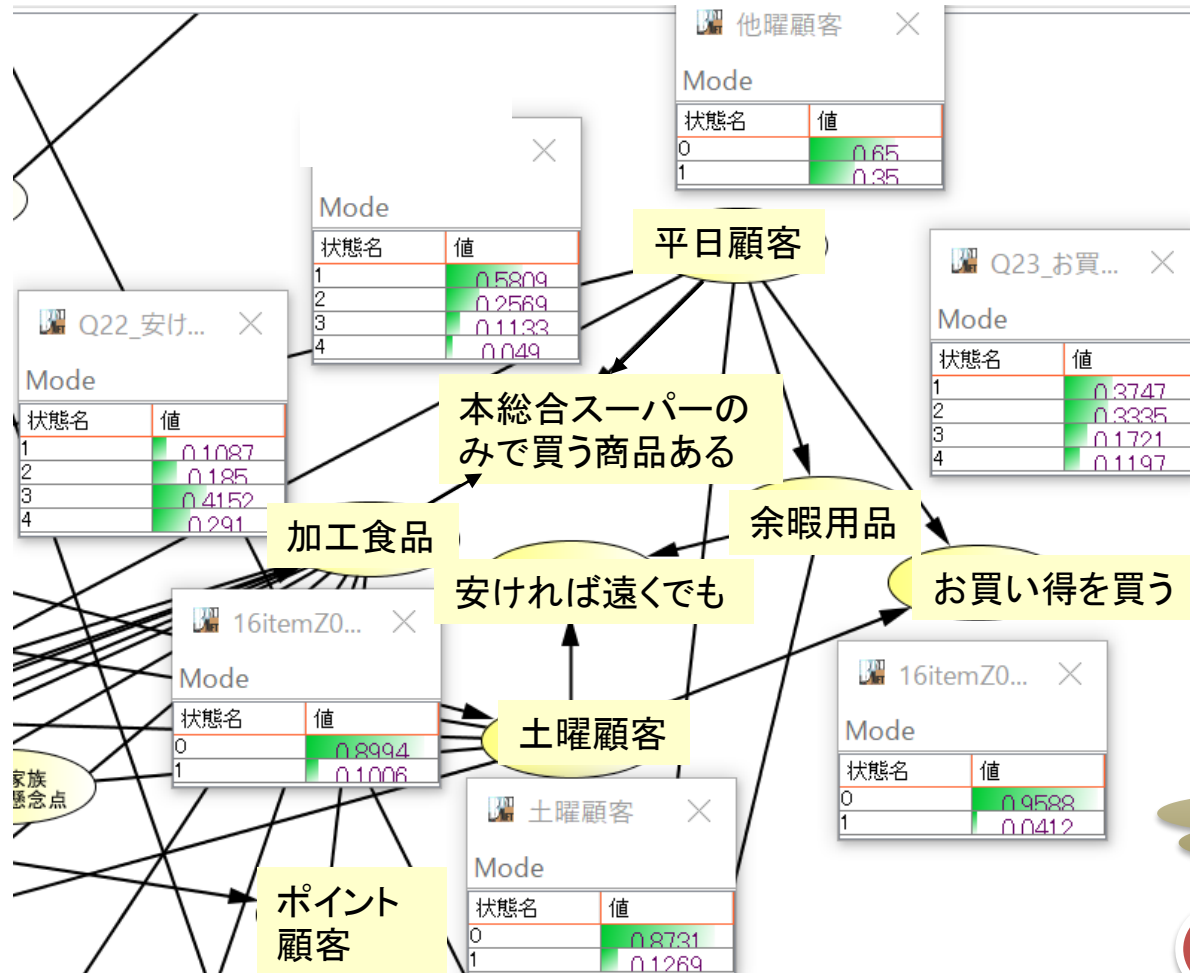
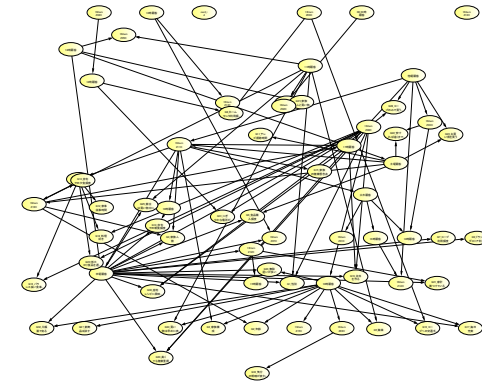
z2

z3

ID34***	2%	83%	15%
---------	----	-----	-----

顧客ID	timeZ1	timeZ2	timeZ3	itemZ1	itemZ2	itemZ3	itemZ4	...
ID34***	0	1	0	0	0	1	0	...

ベイジアンネットワーク全体図



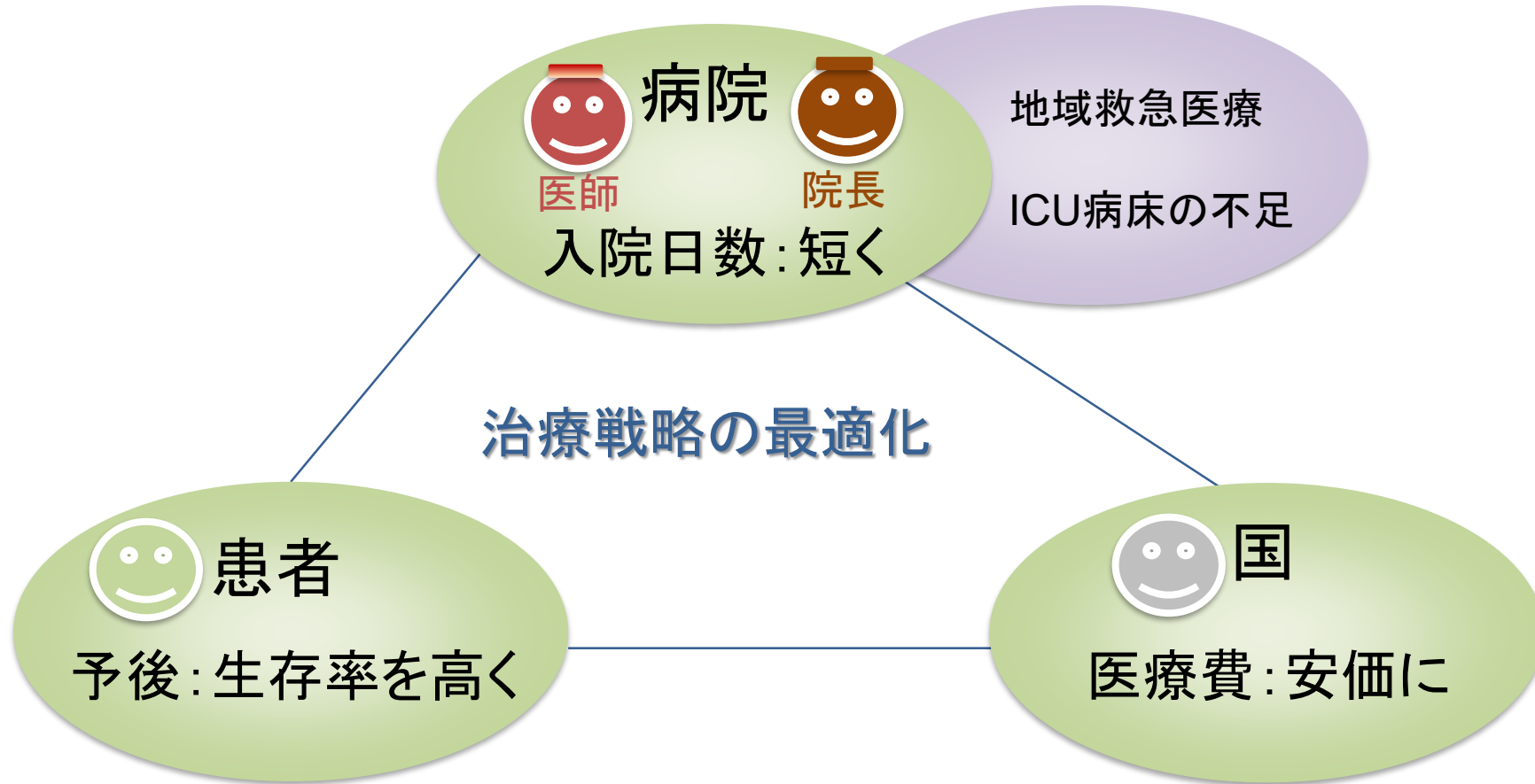
本総合スーパーのみで買う商品あり

- 高くても健康重視
- 余暇用品を買う傾向
- 土・日・ポイントデー顧客
- お買い得品を買う

品質を重視しながらもかしくお  
買い物のお客さんか

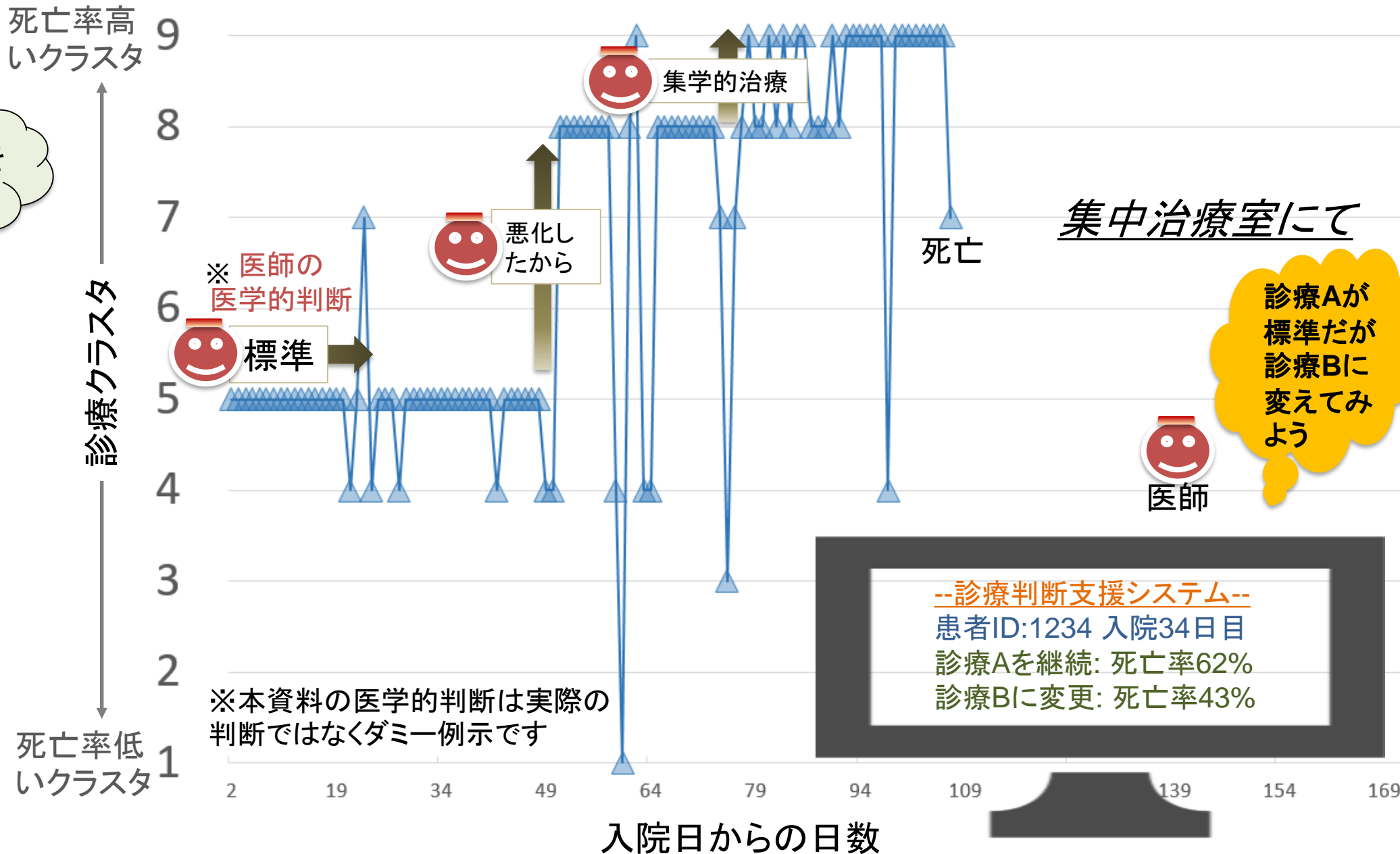


## 2. DPCデータとpLSAを用いた 患者と診療行為の同時クラスタリング



# 同一患者の時間経過に伴うクラスタ遷移

次の治療を  
どうするか





# クラスタ移行パターン

## 全ケース

z6

うちの病院の  
治療状況  
は？



前日と同じクラス  
タに留まる場合  
が多い

### 当日のクラスタ

t-1\t	1	2	3	4	5	6	7	8	9	total
1	596	30	76	237	26	6	15	10	15	1011
2	25	1271	87	812	77	86	49	30	36	2473
3	63	56	2984	686	94	27	40	19	105	4074
4	254	863	731	467	375	78	288	168	149	3373
5	24	28	91	313	1820	26	27	26	11	2366
6	16	103	19	182	11	78	21	181	14	625
7	12	41	70	244	45	18	1718	62	36	2246
8	8	52	19	214	38	32	83	1271	96	1813
9	16	49	105	150	18	12	47	61	761	1219
Enter	14	47	4	105	1	266		12	9	458
total	1028	2540	4186	3410	2505	629	2288	1840	1232	19658

前日のクラスタ

初日のz6が多い  
z6は入院＆手術が  
多い

死亡率が  
低いクラス  
タが多い

### 生存ケース

t-1\t	1	2	3	4	5	6	7	8	9	total
1	588	24	87	226	17	6	15	6	12	961
2	19	1148	75	726	85	69	47	24	30	2203
3	54	48	2515	622	86	24	36	13	52	3450
4	239	775	666	324	307	60	226	79	58	2734
5	15	24	81	247	1443	18	26	12	5	1871
6	15	82	16	143	6	70	19	129	11	491
7	12	34	64	199	42	16	1279	39	26	1711
8	8	41	17	127	21	16	57	578	43	908
9	14	41	54	61	11	8	30	21	282	522
Enter	13	38	1	75	1	206		7	5	346
total	977	2255	3556	2750	1999	493	1735	908	524	15197

### 死亡ケース

t-1\t	1	2	3	4	5	6	7	8	9	total
1	8	6	9	11	9			4	3	50
2	6	123	12	86	12	17	2	6	6	270
3	9	8	469	64	8	3	4	6	53	624
4	15	88	65	143	68	18	62	89	91	639
5	9	4	10	66	377	8	1	14	6	495
6	1	21	3	39	5	8	2	52	3	134
7		7	6	45	3	2	439	23	10	535
8		11	2	87	17	16	26	693	53	905
9	2	8	51	89	7	4	17	40	479	697
Enter	1	9	3	30		60		5	4	112
total	51	285	630	660	506	136	553	932	708	4461

死亡率が高い  
クラスタが多い

- 入院\_がん診療連携拠点病院加算
- 入院\_医療安全対策加算1
- 入院\_妊産婦緊急搬送入院加算
- 入院\_患者サポート体制充実加算
- 入院\_感染防止対策加算
- 入院\_感染防止対策加算1
- 入院\_感染防止対策地域連携加算
- 入院\_新生児特定集中治療室管理料1
- 処置\_連続携帯式腹膜灌流
- 処置\_酸素吸入(手術当日)
- 処置\_酸素吸入(手術時)
- 処置\_<カテ交換>
- 処置\_<ドレーン固定>
- 処置\_<人工呼吸中>
- 処置\_<人工呼吸>
- 処置\_<胆嚢穿刺>
- 処置\_<追加切開鼠径部~大腿>
- 処置\_<op当日>
- 手術\_▽休日加算(処置手術・牽引)
- 手術\_▽新生児加算(輸血)
- 手術\_▽時間外加算(処置手術・牽引)
- 手術\_▽深夜加算(処置手術・牽引)
- 手術\_▽3才未満幼児加算(手術)
- 手術\_アスパラカリウム注10mEq17. 12%10m
- 手術\_アセリオ静注液1000mg1, 000mg100m
- 手術\_アートニー-O注5単位
- 手術\_アートセレブ脳脊髄手術用洗浄灌流液500r
- 手術\_不規則抗体検査加算
- 手術\_中心静脈用カテーテル(標準・シングルルー
- 手術\_中心静脈用カテーテル(標準・マルチルーメ
- 手術\_人工呼吸を行っている患者
- 手術\_人工呼吸中
- 手術\_人工弁輪・三尖弁用¥
- 手術\_人工弁輪・僧帽弁用¥
- 手術\_人工弁輪(3)僧帽弁・三尖弁兼用¥
- 手術\_人工心肺回路(メイン回路・抗血栓性あり)¥
- 手術\_人工心肺回路(メイン回路・抗血栓性なし)¥
- 手術\_人工心肺回路(個別機能品・カーディオミー
- 手術\_人工心肺回路(個別機能品・ラインフィルター
- 手術\_人工心肺回路(個別機能品・安全弁)¥
- 手術\_人工心肺回路(個別機能品・心筋保護用貯
- 手術\_人工心肺回路(個別機能品・熱交換器)¥
- 手術\_人工心肺回路(個別機能品・血液学的パラメ
- 手術\_人工心肺回路(分離体外循環回路)¥
- 手術\_人工心肺回路(心筋保護回路)¥
- 手術\_人工心肺回路(血液濃縮回路)¥
- 手術\_人工心肺回路(補助循環回路・抗血栓性なし



# DPCデータ

- DPC(Diagnosis Procedure Combination=診断群包括分類)データの蓄積  
 診断(病名) 治療・処置 組み合わせ

## DPCデータ

### 患者データ

(1患者=1行データ)

患者ID	診療科名	性別	年齢	MDC6名(主要診断群)	死亡の有無	医療費(円)	在院日数合計
***65	心臓血管外科	0	62	敗血症	0	14613070	158
***19	血液科	1	71	肺の悪性腫瘍	1	2470700	89

⋮

### 明細データ

(1行為毎に1行データ)

患者ID	実施日	分類名	診療行為名称
***65	22	注射	サンドスタチン皮下注用50 $\mu$ g 1mL
***65	23	注射	アキネトン注射液5mg 0.5%1mL

⋮

DPCデータを用いたこれまでの報告: 治療法の効果の検証が大半、治療戦略への応用の報告は認められない

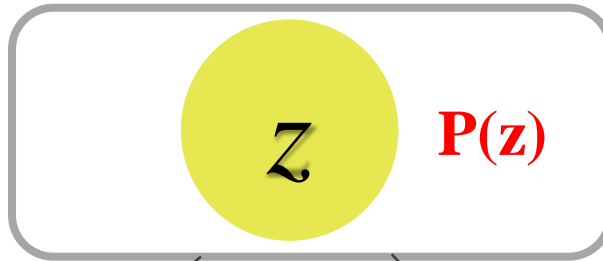
医療ビッグデータの利活用

# pLSA(Probabilistic Latent Semantic Analysis)

確率的潜在意味解析

2つの変数xとyの背後に共通の特性,潜在変数zがあると仮定

潜在変数:  $z$



$$P(x_i, y_j) = \sum_k P(x_i|z_k)P(y_j|z_k)P(z_k)$$

$$L = \sum_i \sum_j n(i, j) \log P(x_i, y_j)$$

$P(x|z)$

$P(y|z)$

$x$ :患者

患者<sub>1</sub>  
患者<sub>2</sub>  
⋮  
患者<sub>457</sub>  
患者<sub>458</sub>

458名の患者

$y$ :診療

注射\_セピドナリン  
投薬\_スチバーガ錠  
処置\_ベネトリン吸入液  
手術\_血管塞栓術  
⋮

4406の診療行為名称

同じような診療を受けている似ている患者

同じ患者に行われている似ている診療

# pLSAによるクラスタ分類

産業技術総合研究所の  
知財ソフトウェア「PLASMA」

計算時間: ノートPCで30分程度

x: 患者、y: 診療行為名称 の共起行列作成

クラスタ数(k=1~12)でそれぞれ初期値10種類(計120通り)のpLSAを実行

AIC最小のクラスタ数、初期値の結果を採用

所属確率値が最大の潜在変数 z に 各x と 各y を分類する

各潜在クラスタへの各患者の所属確率

		潜在クラスタ			
		クラスタ z1	クラスタ z2	...	クラスタ z9
患者ID	0000001	0.2	0.7	...	0.03
	0000002	0.8	0.1	...	0.05
	0000003	0.03	0.07	...	0.22
	...				

各潜在クラスタへの各診療行為の所属確率

		潜在クラスタ			
		クラスタ z1	クラスタ z2	...	クラスタ z9
診療行為	診療行為1	0.5	0.2	...	0.1
	診療行為2	0.05	0.01	...	0.07
	診療行為3	0.1	0.15	...	0.6
	...				

クラスタ1

クラスタ2

クラスタ9

# 患者と診療行為を同時に9クラスタに分類

z9

処置_局所陰圧閉鎖処置(被覆材貼付・100平方cm未満)
手術_血管塞栓術(頭部)
投薬_グリセリン「マルイシ」
検査_白血球中サイトメガロウイルスpp65抗原
注射_ピーエヌツインー3号輸液
画像_CTスキャン(CT撮影・16列未満マルチスライス型機器)2回目以降
その他_急性硬膜下血腫・頭蓋内に達す
入院_無菌治療室管理加算1

z1

処置_ベネトリン吸入液0.5%
手術_造血幹細胞移植(臍帯血移植)
投薬_インデラル錠10mg
検査_アルドラーゼ
入院_看護補助加算1
その他_精神科退院指導料

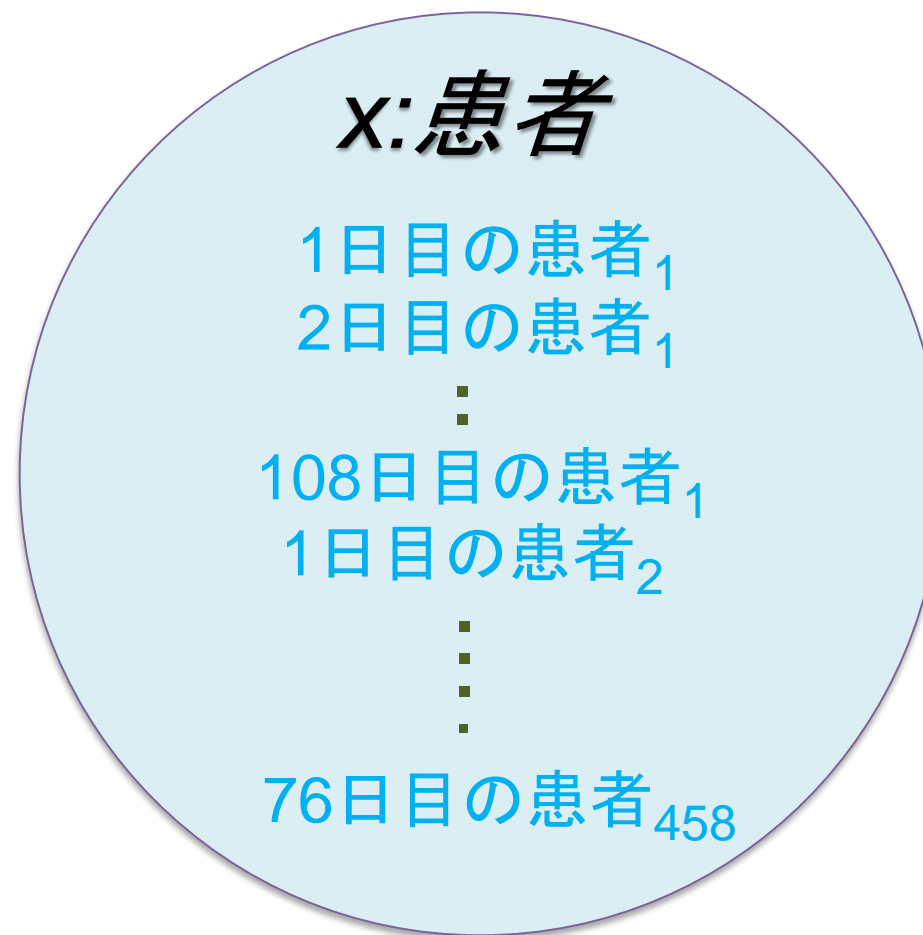
z2

処置_皮膚科軟膏処置(6000平方cm以上)
検査_唾液腺
注射_セピドナリン静注用1g
画像_腎シンチグラム
その他_骨盤骨折
入院_オショック

在院日数

死亡率

# 入院初日からの日数と患者IDを組み合わせる

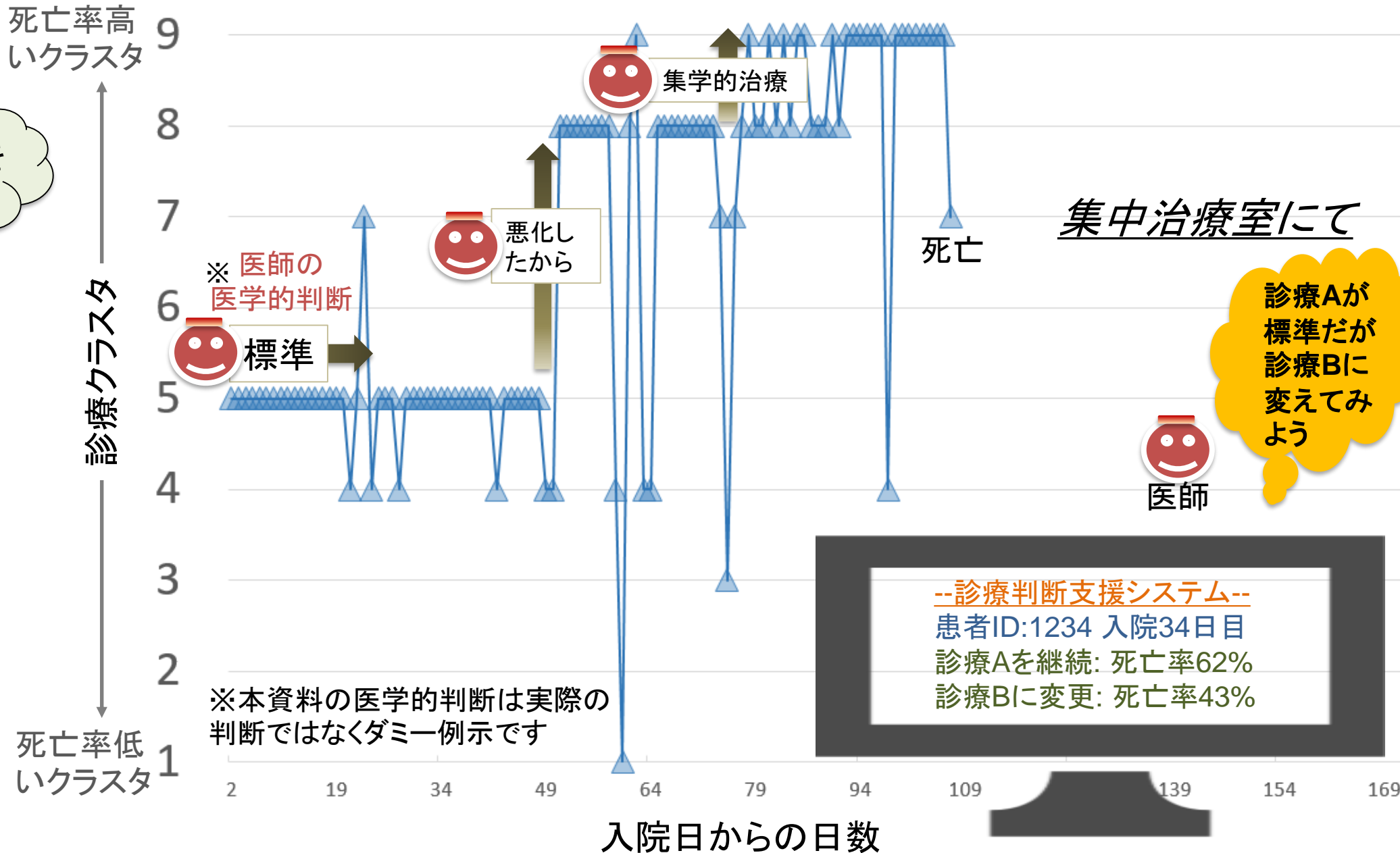


1日目の患者と2日目の患者を別患者のように扱う



# 同一患者の時間経過に伴うクラスタ遷移

次の治療を  
どうするか

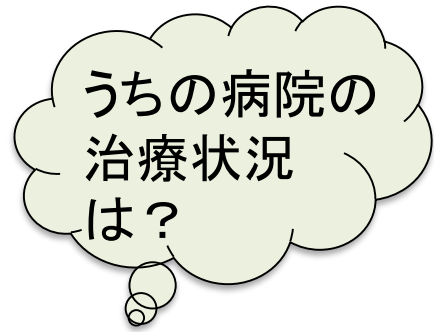




# クラスタ移行パターン

全ケース

z6



前日と同じクラス  
タに留まる場合  
が多い

当日のクラスタ

前日のクラスタ

t-1\t	1	2	3	4	5	6	7	8	9	total
1	596	30	76	237	26	6	15	10	15	1011
2	25	1271	87	812	77	86	49	30	36	2473
3	63	56	2984	686	94	27	40	19	105	4074
4	254	863	731	467	375	78	288	168	149	3373
5	24	28	91	313	1820	26	27	26	11	2366
6	16	103	19	182	11	78	21	181	14	625
7	12	41	70	244	45	18	1718	62	36	2246
8	8	52	19	214	38	32	83	1271	96	1813
9	16	49	105	150	18	12	47	61	761	1219
Enter	14	47	4	105	1	266		12	9	458
total	1028	2540	4186	3410	2505	629	2288	1840	1232	19658

初日のz6が多い  
z6は入院&手術が  
多い

死亡率が  
低いクラスタ  
が多い

生存ケース

t-1\t	1	2	3	4	5	6	7	8	9	total
1	588	24	87	226	17	6	15	6	12	961
2	19	1148	75	726	85	69	47	24	30	2203
3	54	48	2515	622	86	24	36	13	52	3450
4	239	775	666	324	307	60	226	79	58	2734
5	15	24	81	247	1443	18	26	12	5	1871
6	15	82	16	143	6	70	19	129	11	491
7	12	34	64	199	42	16	1279	39	26	1711
8	8	41	17	127	21	16	57	578	43	908
9	14	41	54	61	11	8	30	21	282	522
Enter	13	38	1	75	1	206		7	5	346
total	977	2255	3556	2750	1999	493	1735	908	524	15197

死亡ケース

t-1\t	1	2	3	4	5	6	7	8	9	total
1	8	6	9	11	9			4	3	50
2	6	123	12	86	12	17	2	6	6	270
3	9	8	469	64	8	3	4	6	53	624
4	15	88	65	143	68	18	62	89	91	639
5	9	4	10	66	377	8	1	14	6	495
6	1	21	3	39	5	8	2	52	3	134
7		7	6	45	3	2	439	23	10	535
8		11	2	87	17	16	26	693	53	905
9	2	8	51	89	7	4	17	40	479	697
Enter	1	9	3	30		60		5	4	112
total	51	285	630	660	506	136	553	932	708	4461

死亡率が高い  
クラスタが多い

- 入院\_がん診療連携拠点病院加算
- 入院\_医療安全対策加算1
- 入院\_妊産婦緊急搬送入院加算
- 入院\_患者サポート体制充実加算
- 入院\_感染防止対策加算
- 入院\_感染防止対策加算1
- 入院\_感染防止対策地域連携加算
- 入院\_新生児特定集中治療室管理料1
- 処置\_連続携帯式腹膜灌流
- 処置\_酸素吸入(手術当日)
- 処置\_酸素吸入(手術時)
- 処置\_<カテ交換>
- 処置\_<ドレーン固定>
- 処置\_<人工呼吸中>
- 処置\_<人工呼吸>
- 処置\_<胆嚢穿刺>
- 処置\_<追加切開鼠径部~大腿>
- 処置\_<op当日>
- 手術\_▽休日加算(処置手術・牽引)
- 手術\_▽新生児加算(輸血)
- 手術\_▽時間外加算(処置手術・牽引)
- 手術\_▽深夜加算(処置手術・牽引)
- 手術\_▽3才未満幼児加算(手術)
- 手術\_アスパラカリウム注10mEq17. 12%10m
- 手術\_アセリオ静注液1000mg1, 000mg100m
- 手術\_アートニー-O注5単位
- 手術\_アートセレブ脳脊髄手術用洗浄灌流液500r
- 手術\_不規則抗体検査加算
- 手術\_中心静脈用カテーテル(標準・シングルルー
- 手術\_中心静脈用カテーテル(標準・マルチルーメ
- 手術\_人工呼吸を行っている患者
- 手術\_人工呼吸中
- 手術\_人工弁輪・三尖弁用¥
- 手術\_人工弁輪・僧帽弁用¥
- 手術\_人工弁輪(3)僧帽弁・三尖弁兼用¥
- 手術\_人工心肺回路(メイン回路・抗血栓性あり)¥
- 手術\_人工心肺回路(メイン回路・抗血栓性なし)¥
- 手術\_人工心肺回路(個別機能品・カーディオミー
- 手術\_人工心肺回路(個別機能品・ラインフィルター
- 手術\_人工心肺回路(個別機能品・安全弁)¥
- 手術\_人工心肺回路(個別機能品・心筋保護用貯液
- 手術\_人工心肺回路(個別機能品・熱交換器)¥
- 手術\_人工心肺回路(個別機能品・血液学的パラメ
- 手術\_人工心肺回路(分離体外循環回路)¥
- 手術\_人工心肺回路(心筋保護回路)¥
- 手術\_人工心肺回路(血液濃縮回路)¥
- 手術\_人工心肺回路(補助循環回路・抗血栓性なし

### 3. 工場にて

## 歩留まり

(実際に得られた製品生産量/原料の投入量から期待される生産量)

## を高めたい

→異常が発生する条件を知りたい



製造業

# PLASMAでのPLSA(複雑な形式のデータ)

## 入力: センサーデータ

目的変数 y

X

事例id	故障発生	センサA	部材素材	部材温度	圧力
1	0	正常	A	22.9	1008.8
2	0	正常	A	12.0	1018.6
3	0	正常	A	34.0	1001.8
4	0	異常検知	A	25.3	992.9
5	0	正常	A	29.1	1011.0
6	0	正常	A	21.7	1012.5
7	0	正常	A	17.8	1008.8
8	0	正常	B	14.0	1006.9
9	0	正常	B	23.1	995.0
10	0	正常	B	17.2	1007.5
11	1	異常検知	B	26.5	997.0
12	0	正常	B	23.9	994.1
13	0	正常	B	27.9	1012.2



出力: 最適クラスタ数と  
事例と状態のクラスタリング結果

3クラスタ分けが最適でした

クラスタ		クラスタ	
No.	所属確率	No.	所属確率
事例id	Z	状態	Z
1	Z002	センサA_正常	Z002
2	Z003	センサA_異常検知	Z001
3	Z002	部材素材_A	Z002
4	Z001	部材素材_B	Z003
		部材温度_高め	Z002
		部材温度_低め	Z003
		圧力_高い	Z002
		圧力_普通	Z003
		圧力_低い	Z001

Z	故障発生率
Z001	77%
Z002	4%
Z003	11%

Z001クラスタは 故  
障発生率が高い

事例を状態(説明変数)でクラスタリングし、各クラスタ毎の故障発生率(目的変数)を算出、連続値は離散化

## 4. 健康イベント

健康寿命を延ばす為  
継続的に健康イベントに  
来ていただくには



健康管理医・保健師



# 健康イベント

◇ xperia touchを用いて



築99年の別荘で



千葉市美術館で

◇ デジタルサイネージを用いて



雰囲気の良い場所で  
手芸などと組み合わせか



健康管理医

第2問  
どちらの考え方が心の健康に良いと思いますか？


半分しかない  半分もある

Copyright © National Institute of Advanced Industrial Science and Technology (AIST). All rights reserved.



## 5.回遊型展示大規模イベントにて: イベントが産み出す価値の継続的な向上

### 来場者の考える価値

- 
- ・興味ある事をより深く知りたい
  - ・最新情報を得たい
  - ・話のネタにしたい
  - ・イベントでしか得られないものへの期待
  - ・新しい興味の発見


来場者とコミュニケーションしたい

来場者が来ない事態は避けたい

### ブース出展者の考える価値



### イベント主催者の考える価値



再来場意向上を高めたい  
他の人への推薦率を高めたい  
科学への関心・認知度の向上

# イベント空間：行動と気持ちの生成・観測の場

## イベント空間での実証実験



○行動を規定しながらかつ行動の変化を観測しやすい

○見たものに対して、「いいね！」が出やすい(気持ちの変化が明確に出る)

例

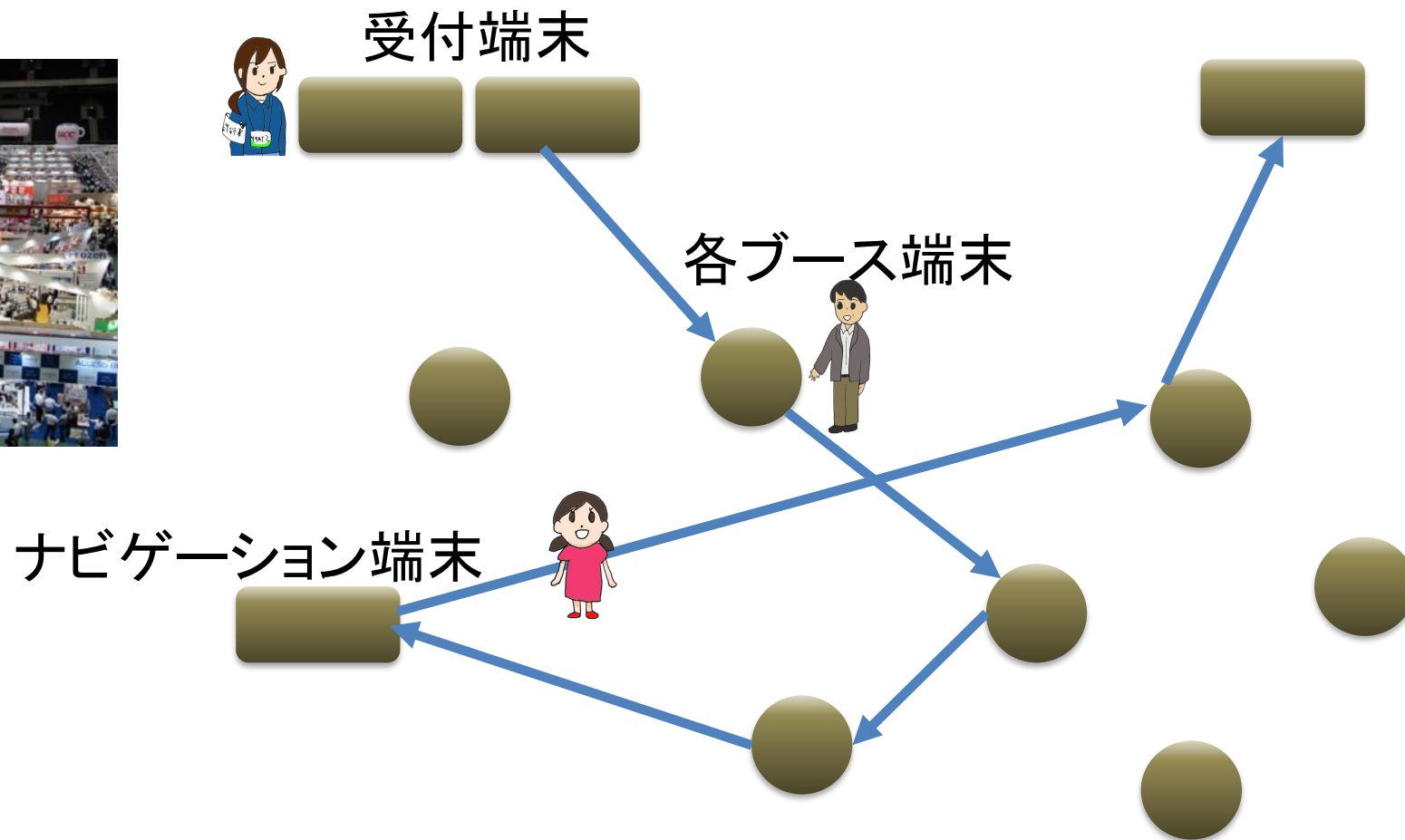
○サイエンスアゴラ@お台場  
(2016年～2019年)

- 科学と社会をつなぐ日本最大級の公開イベント(科学技術振興機構 主催)
- 例年3日間の出展団体数は100~200、来場者数は約4000人~1万人

○産総研・一般公開(2016年～2019年)

# イベント空間：行動と気持ちの生成・観測の場（続き）

対象イベント：複数の独立した展示ブースが集合し、来場者が自由に回遊することのできるイベント





# AIタッチラリー

Artificial Intelligence Research Center



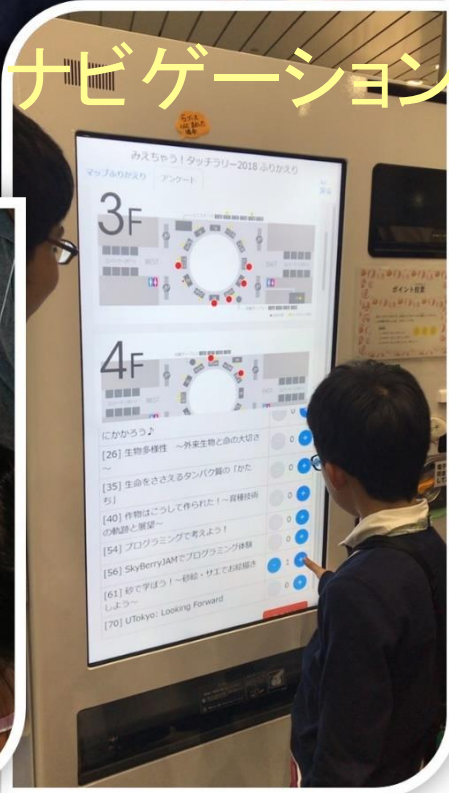
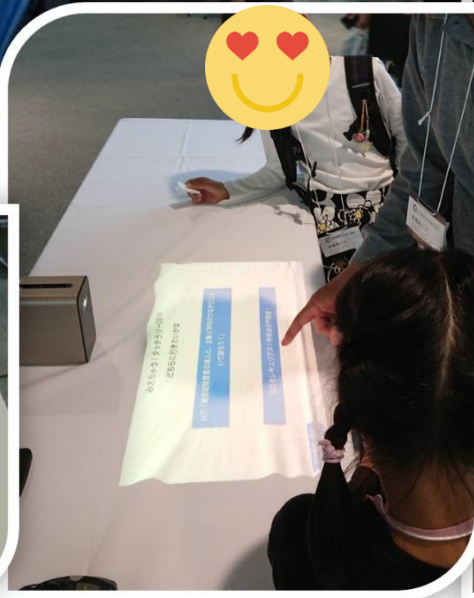
受付

今年で4年目

毎年、  
○サイエンスアゴラ  
○産総研・一般公開  
で継続的に  
実施→改良



ブース回遊



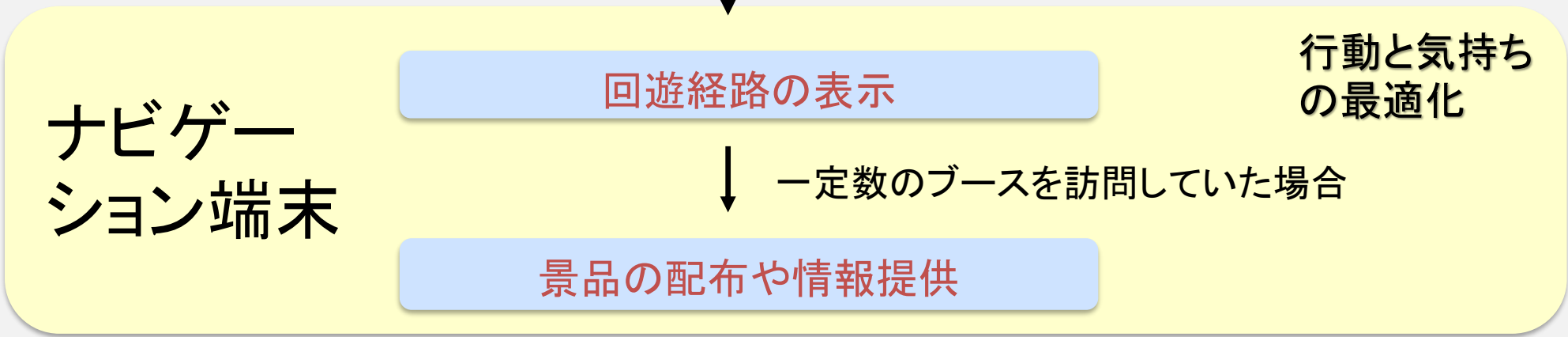
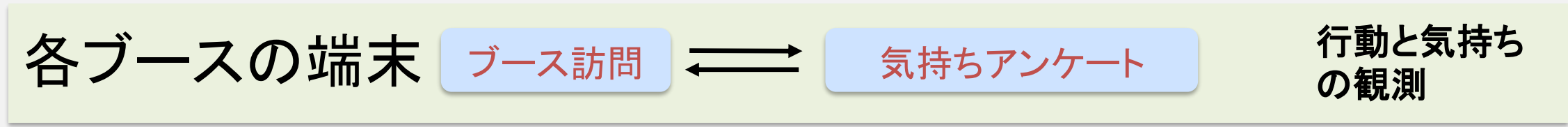
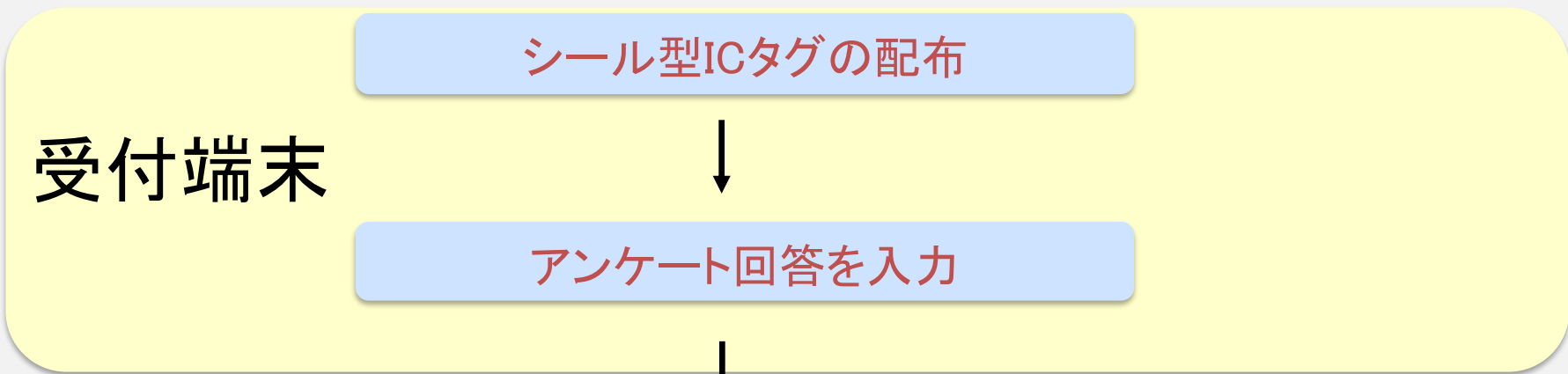
ナビゲーション

(近藤 等2016,近藤 等2017,近藤 等2018, 大和田 等2018,山下 等2019,大和田 等 2019)

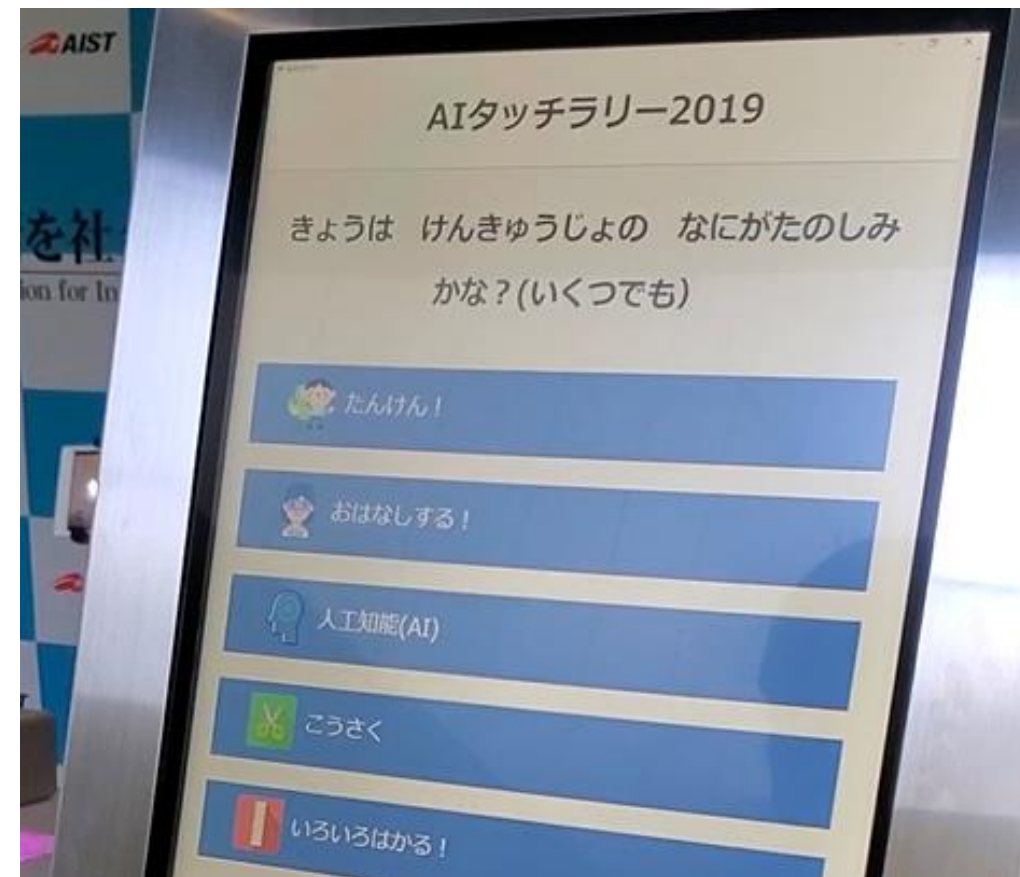
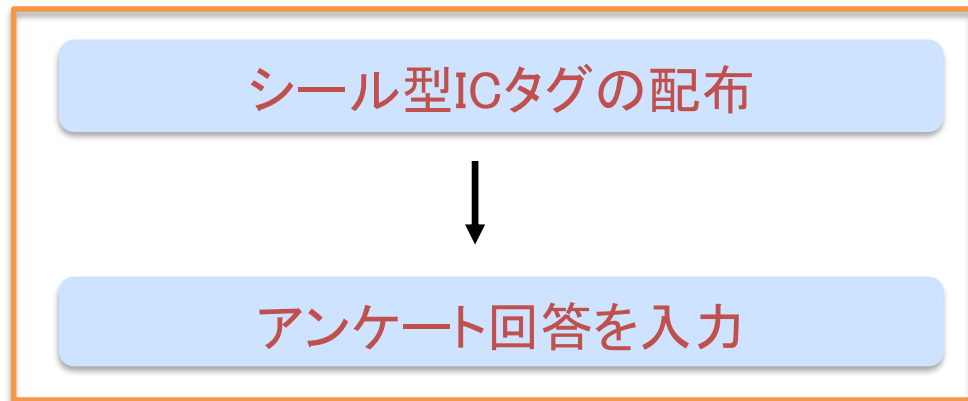
今年9月  
商業施設で

Xperia Hello! Xperia Touch

# AIタッチラリーの流れ

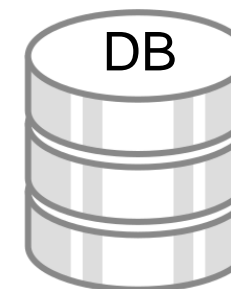
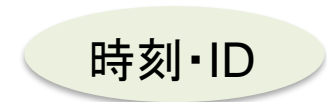
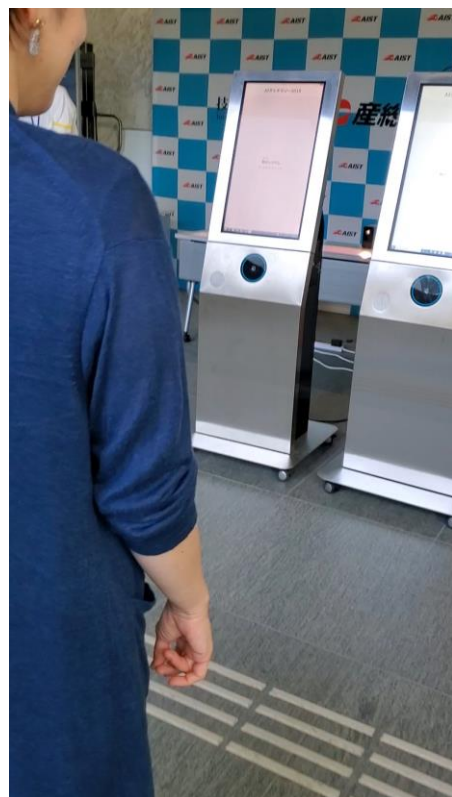


# AIタッチラリーの流れ(受付端末)



<https://gigazine.net/news/20120127-nfc-evernote-touchanote/>

ICカード  
をタッチ





# タッチラリーの流れ(各ブース端末)

ブース訪問



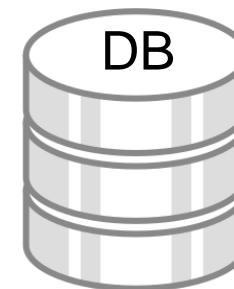
気持ちアンケート

ICカード  
をタッチ



- ・回遊データ
- ・気持ちデータ

時刻・ID





# AIタッチラリーの流れ(ナビゲーション端末)

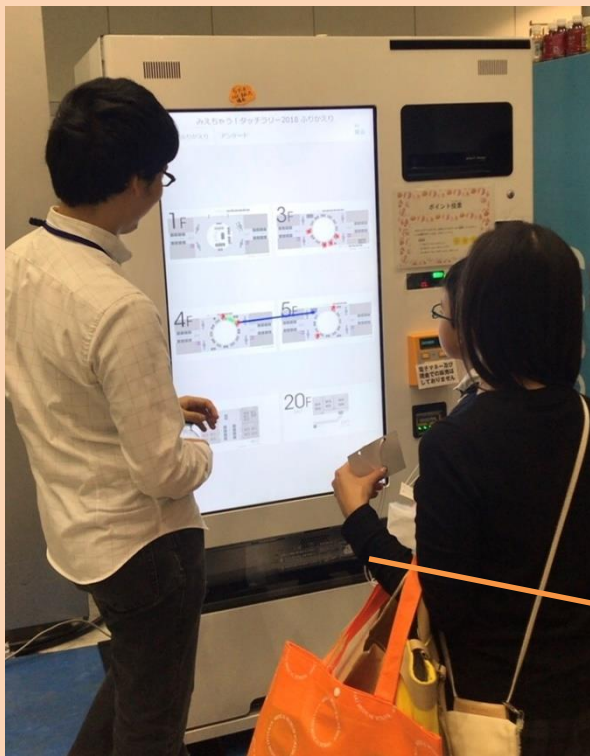
回遊経路の表示



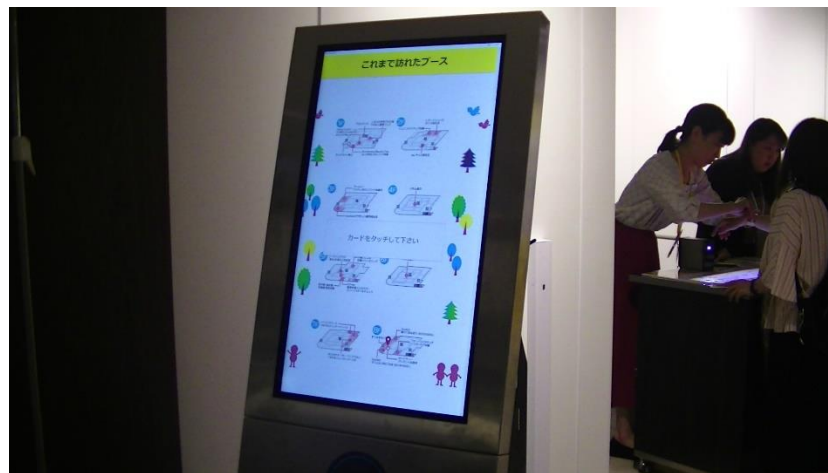
一定数のブースを訪問していた場合

景品の配布

自販機一体型サイネージ



ICカード  
をタッチ

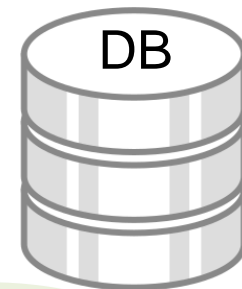


行動・気持ちを変  
える事が出来る  
チャンス

・ふりかえり  
マップ表示

・AIによる対話へ

- ・ふりかえりマップ表示
- ・おすすめ情報
- ・イベント回遊後アンケート

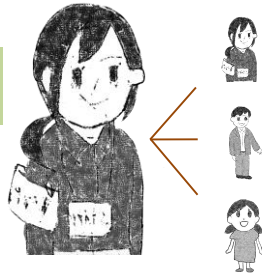


時刻・ID

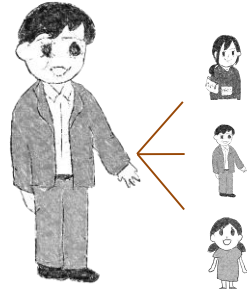
# サービスシステムにおける価値構造とその進化

潜在的な価値をステークホルダーごとにまとめる  
(ただし暗黙的なので気づきと進化が不可欠)

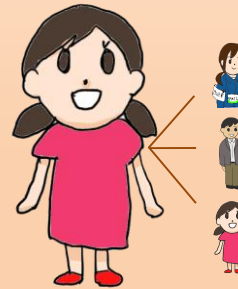
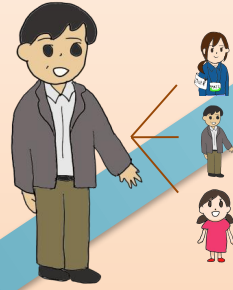
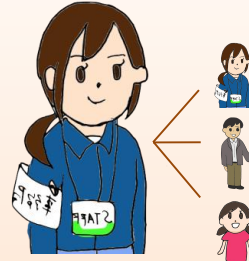
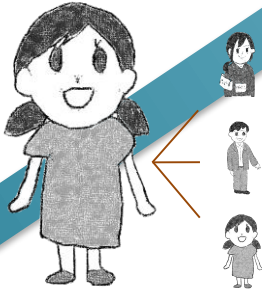
マネジメント層



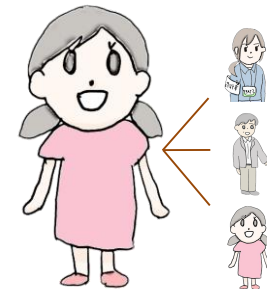
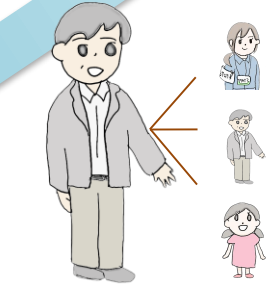
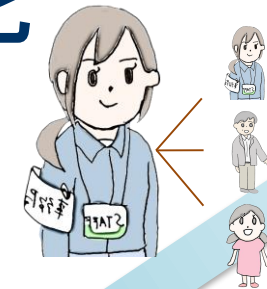
現場スタッフ



ユーザー



実際の結果



次回

サービスシステムが提供する価値が進化していく

事前の予想

# サービスシステムにおける情報システムの進化

Artificial Intelligence Research Center

今年うまくいった  
でも来年は..?

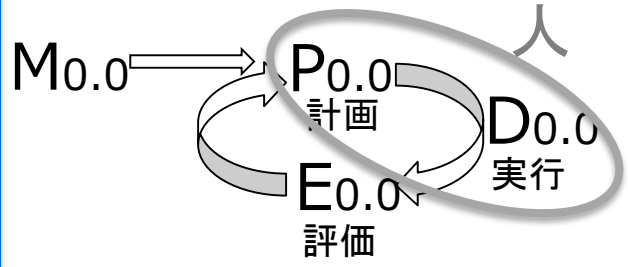


経験を積んだ  
イベント主催  
者Kさん

どうしてKさんは  
これまでこうし  
ていたの？  
今回も同じで良  
いのかな？

新しく挑戦する  
イベント主催者

経験と勘に支えられていて  
再現性が低い

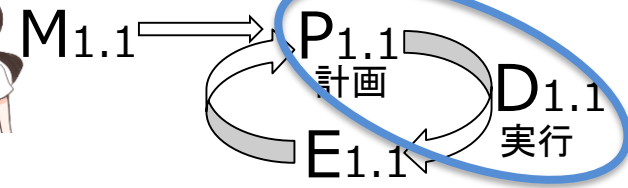


AIタッチラリー導入前

サイエンスアゴラ2015以前

滞在時間と立ち寄った  
ブースの**可視化**

AIタッチラリー

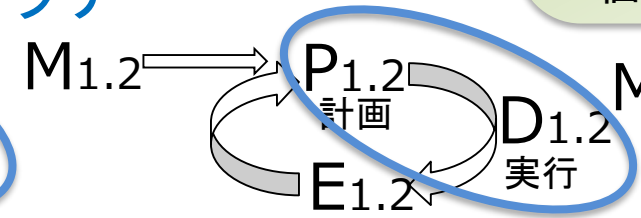


ブースの場  
所によって  
来場者数に  
偏り(不利な  
場所)がある  
ことが判明

AIタッチラリーで  
入力できる段階

サイエンスアゴラ2016,17

ブース場所による**偏り解消**



ブース場所による**偏り解消**

ナビゲーションで  
出力した段階

サイエンスアゴラ2018,19

円形会場にし、  
最初に上階への  
ナビゲーション

ナビ機能を拡充  
した段階

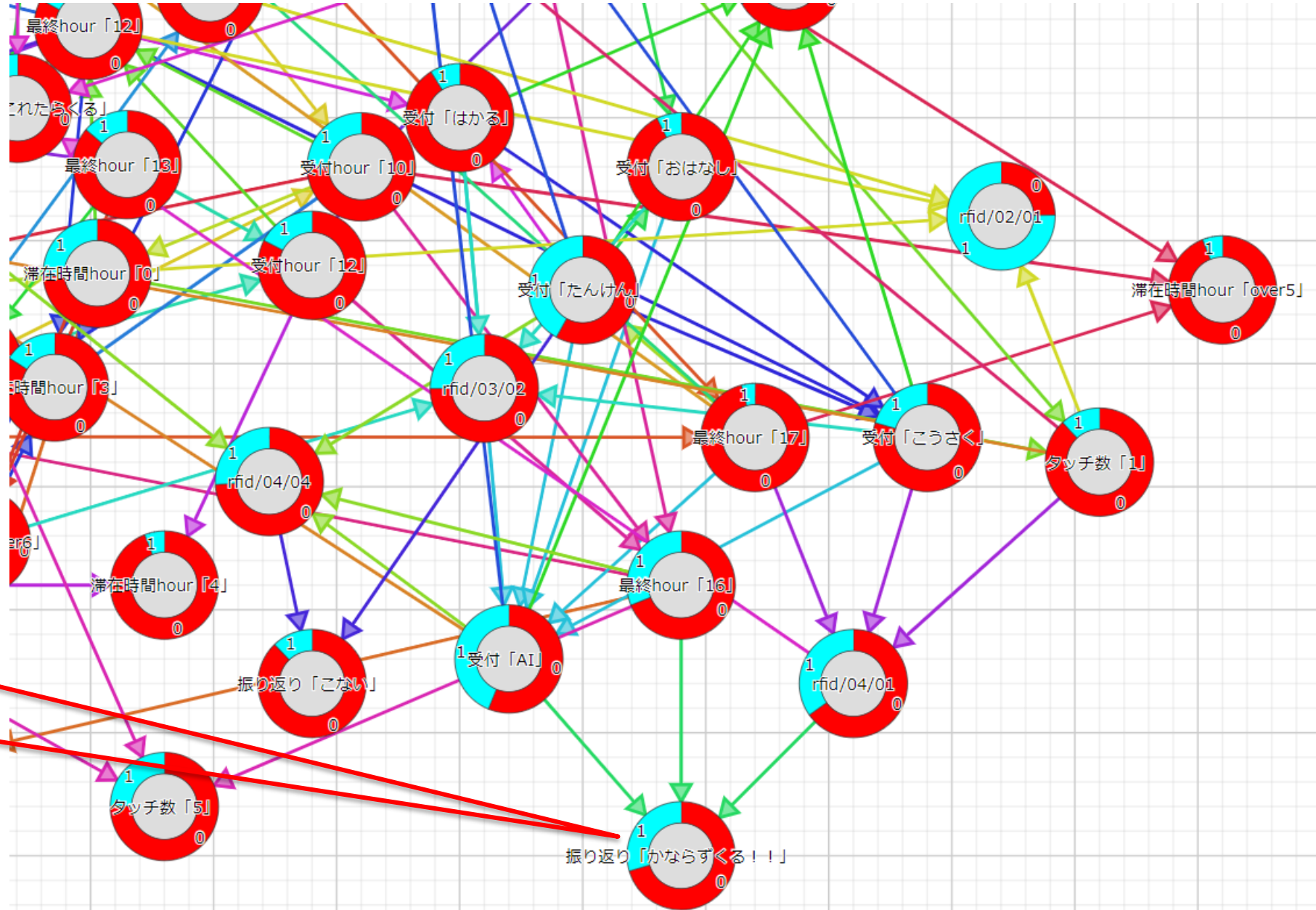
商業施設での例

- M: サービスシステム (情報システムと価値構造)
- P: 商業施設でのイベントの計画
- D: 実際のイベントの実行
- E: その結果をデータに基づいて評価する

情報システムの進化

# 実データを用いた一連のPLASMA利用・操作例

産総研・一般公開2019  
の確率的潜在意味構造  
モデリング

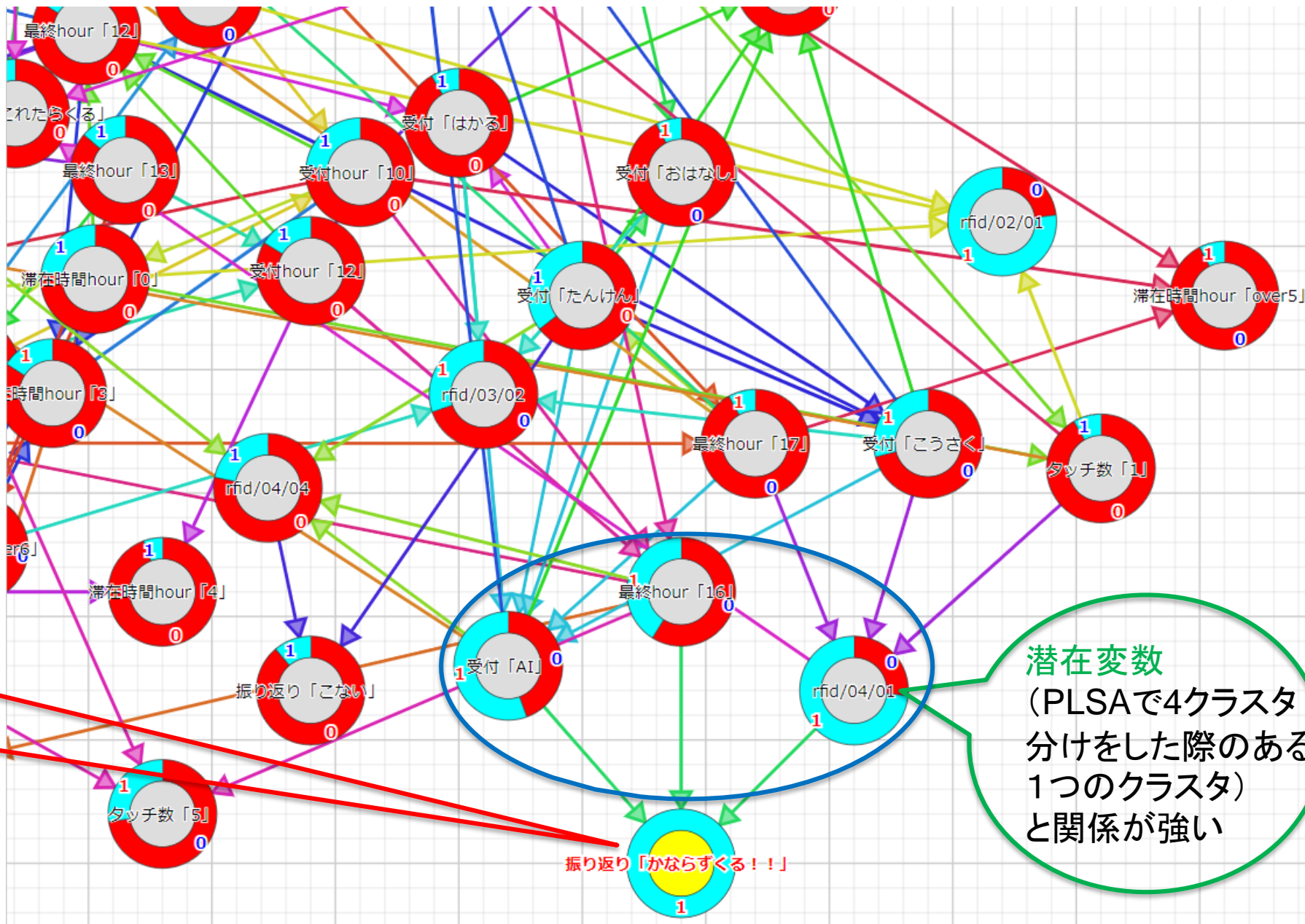


**再来場意向**  
(=帰り時に「また来てく  
れるかな」の設問に「かな  
らずくる!!」と答える)  
を**目的変数**



# 実データを用いた一連のPLASMA利用・操作例

産総研・一般公開2019  
の確率的潜在意味構造  
モデリング



## 再来場意向

(=帰り時に「また来てく  
れるかな」の設問に「かな  
らずくる!!」と答える)  
を**目的変数**

## 潜在変数

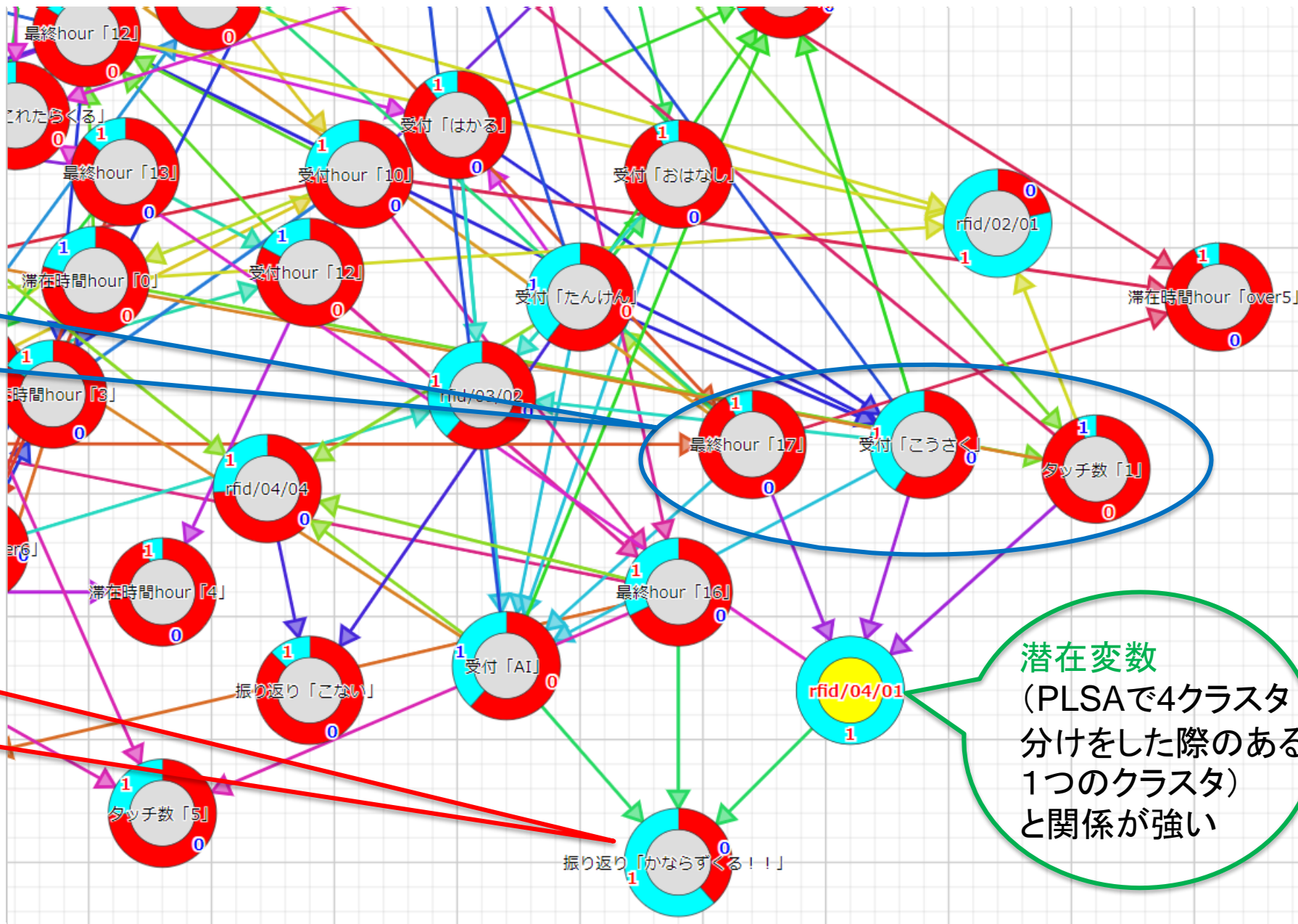
(PLSAで4クラスタ  
分けをした際のある  
1つのクラスタ)  
と関係が強い

# 実データを用いた一連のPLASMA利用・操作例

産総研・一般公開2019  
の確率的潜在意味構造  
モデリング

工作に興味があり  
イベント終了時間まで  
複数ブースを巡って  
いた人

再来場意向  
(=帰り時に「また来てく  
れるかな」の設問に「かな  
らずくる!!」と答える)  
を**目的変数**



潜在変数  
(PLSAで4クラスタ  
分けをした際のある  
1つのクラスタ)  
と関係が強い

# PLASMAの全体像

## ●モデル構築

### ○GUI

アプリケーション層。

→プログラミングなしでモデル構築。

### ○Scala API

Java API を Scala から使いやすくするためのラッパープログラム。

Scala はインタプリタ形式の実行環境を持ち、Java よりも記述量が少なく、試行錯誤に適している。

→データサイエンティストの試行錯誤をアルゴリズム化

## ●アプリケーション作成

### ○アプリ構築用ライブラリ

→手間をかけずアプリを作成

### ○Java API

PLSA 及びベイジアンネットワークを扱うための Java 言語によるクラスライブラリ

→細かい仕様まで自分で決めてアプリ作成

POSEIDON  
と連携

確率的潜在意味構造モデリングの社会実装をサポート



# PLASMA Scala

## データサイエンティストの試行錯誤をアルゴリズム化

The screenshot displays the IntelliJ IDEA IDE with a Scala project named 'plasma\_scala'. The main editor shows the file 'Sample2018.scala' with the following code:

```

7  //
8  package sample
9
10 import data._
11 import plsa._
12
13 /** 2018 年度開発 動作確認用サンプルスクリプト */
14 object Sample2018 extends App {
15
16 //
17 // 1. 最適クラスター探索 (手動)
18 {
19 // 学習データ
20 val data = Data.fromCSV( path = "data/POSデータ.csv", enc = "Shift_JIS")
21
22 // 学習データから共起行列を作成する
23 var cm = Comatrix(data, "顧客ID", "商品コード", "価格")
24
25 // 値が得られるように調整
26 // 何もしないと単調増加になるので集計値を減じている
27 cm = cm.multiply(0.01)
28
29 // 初期値 (生成シード) 4 個
30 val seeds = PlsaLoop.seed( num = 4)
31
32 // クラスター数 * 初期値数のモデルを構築する
33 var pl = PlsaLoop(List(5, 50, 100), seeds, cm)
34 pl.em( iterNum = 50)
35
36 // AIC による評価を取得する
37 val results = pl.aic()
38
39 // results に対する、その他の操作
40 results.view() // クラスター数と AIC (最小値) との関係
41 results.next() // 次に探索 k すべきクラスター数
42 results.best() // 最良モデルのクラスター数と初期値
43 results.asData().p() // 初期値 * クラスター数形式の評価値表
44
45 // 次のステップのモデル構築 (探索範囲を狭くしていく)
46 pl = PlsaLoop(List(27, 75), seeds, cm)
47 pl.em( iterNum = 50)
48 results.add(pl.aic()) // 結果を追加していく
49 results.view()
50 results.next()
51
52 // 次のステップのモデル構築
53 pl = PlsaLoop(List(38, 62), seeds, cm)

```

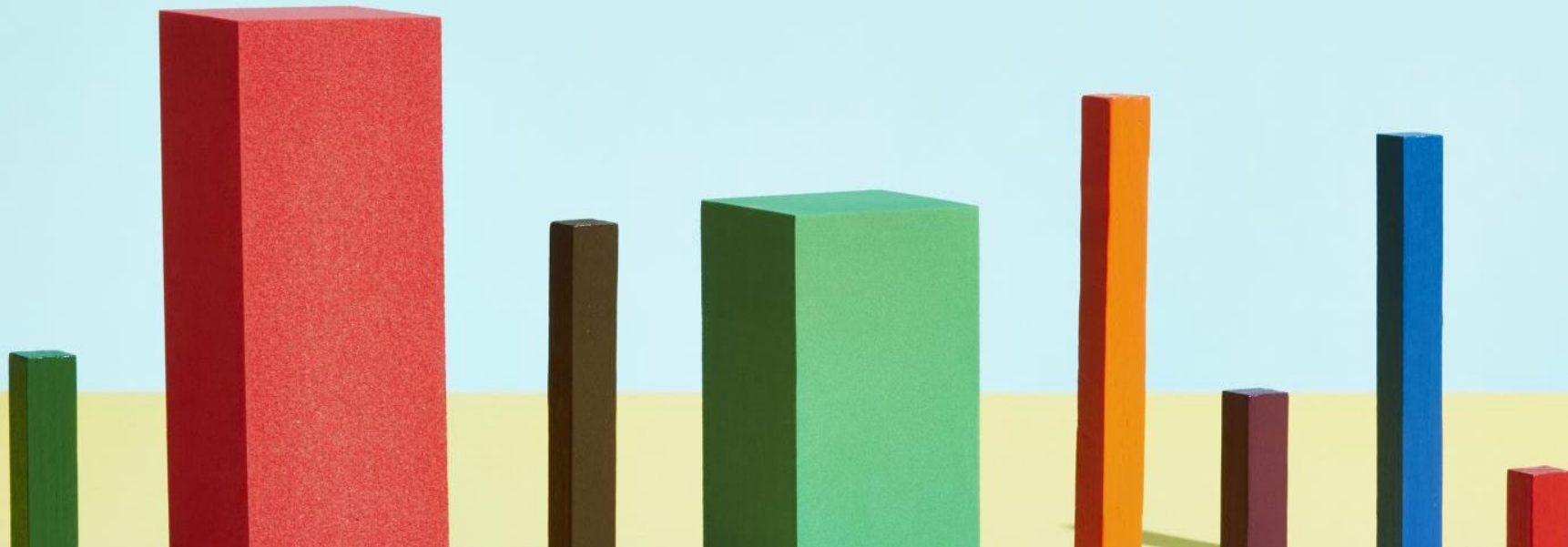
The Run window on the right shows the execution output:

```

Run: Scala Console x Scala Console x Sample x
"C:\Program Files\Java\jdk1.8.0_144\bin\java.exe" ...
Process finished with exit code -1

```

A 'Stop' dialog box is visible in the lower-left area of the IDE, indicating the execution has been manually stopped.



# PLASMA 操作について



データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

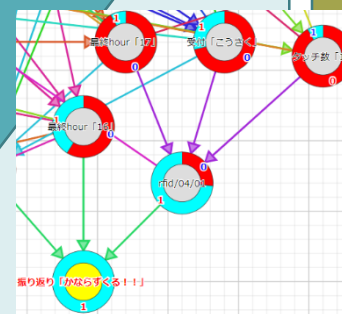
**GUI**

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論



データ  
加工

pLSA

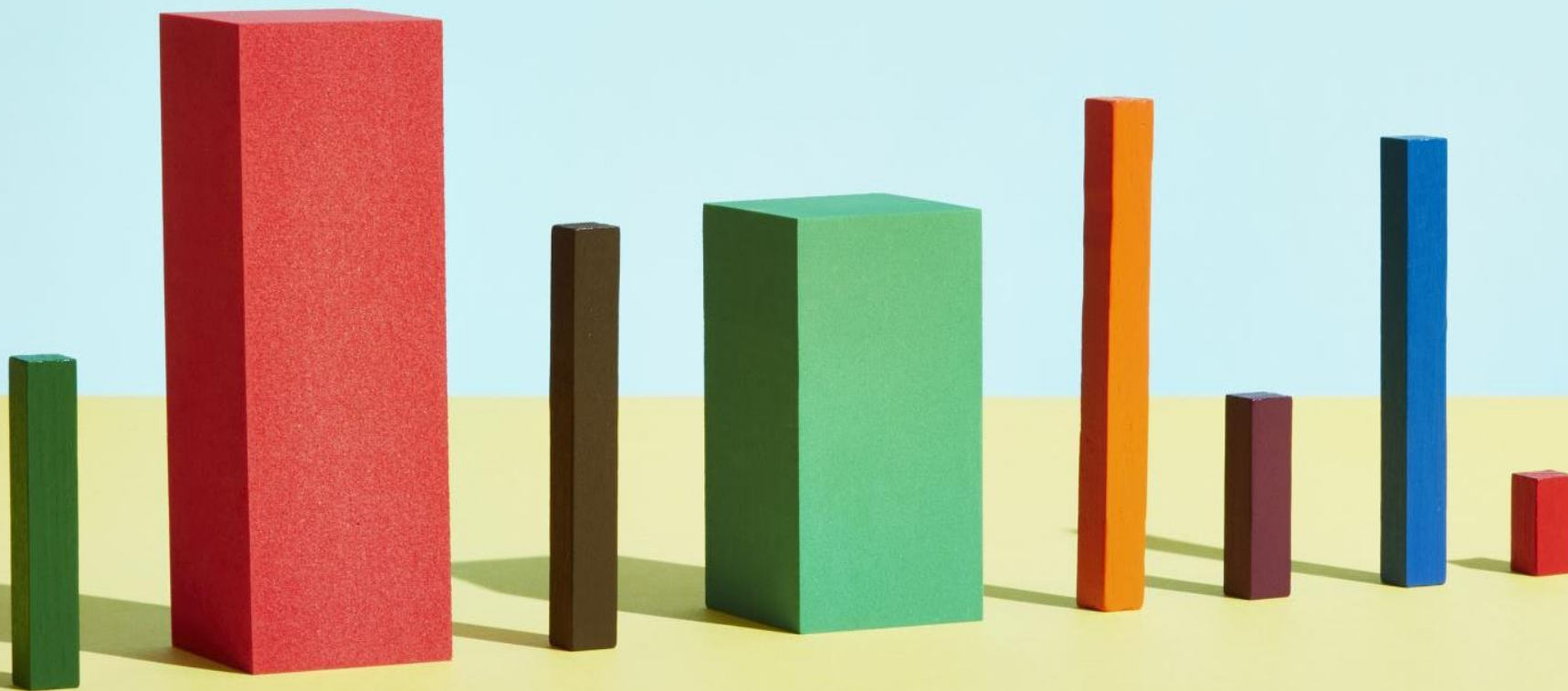
BN  
モデル  
構築

確率  
推論

コマンド

# PLASMA 操作解説ムービー

<https://www.youtube.com/watch?v=5c7bk2qitTc&list=PL5ddiPcj1tnX3VvLiZ4kewqlC7dWTzuOK>



具体的な操作の仕方は上記の動画プレイリストにありますので  
以下のスライドと併せてご参照ください



# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

6

0

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

★ 3

★ 4

5

★ 6

基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

6

0

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

6

0

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

6

0

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

6

0



# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

★  
3

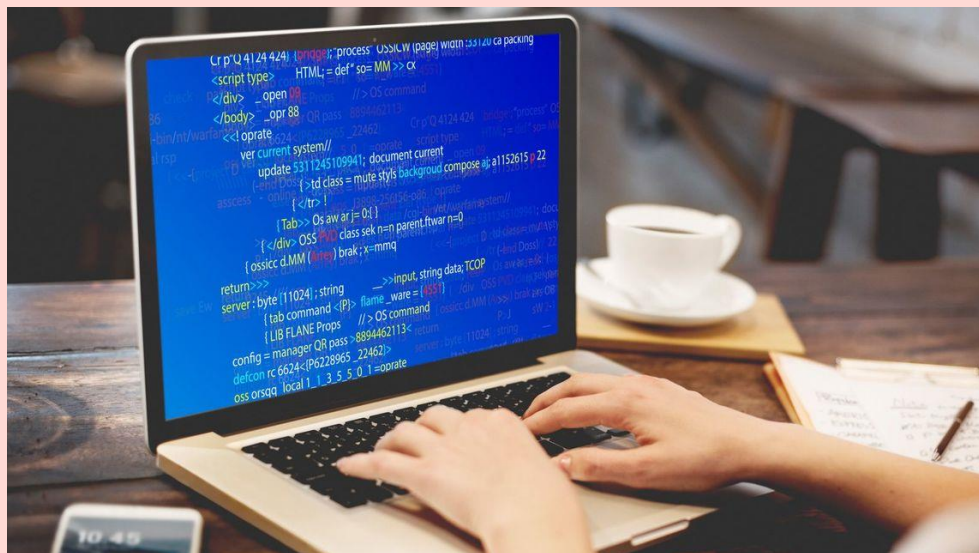
4

5

6

3

# 複雑なデータのpLSA



Before

PLASMA-N2

プロジェクト ウィザード バージョン情報

共起行列の生成

X(ROW) rfid

Y(COL) 受付hour,受付「たんかん」,受付「おはなし」,受付「AI」,受付「こたぐ」,受付「はかる」

カラム名をプレフィックスに利用する 結合文字

ウェイト無し

	A	B	C
1		Z001	Z002
2	受付hour_10	1	0
3	受付hour_11	0	1
4	受付hour_12	1	0
5	受付hour_13	1	0
6	受付hour_14	1	0
7	受付hour_15	1	0

After

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

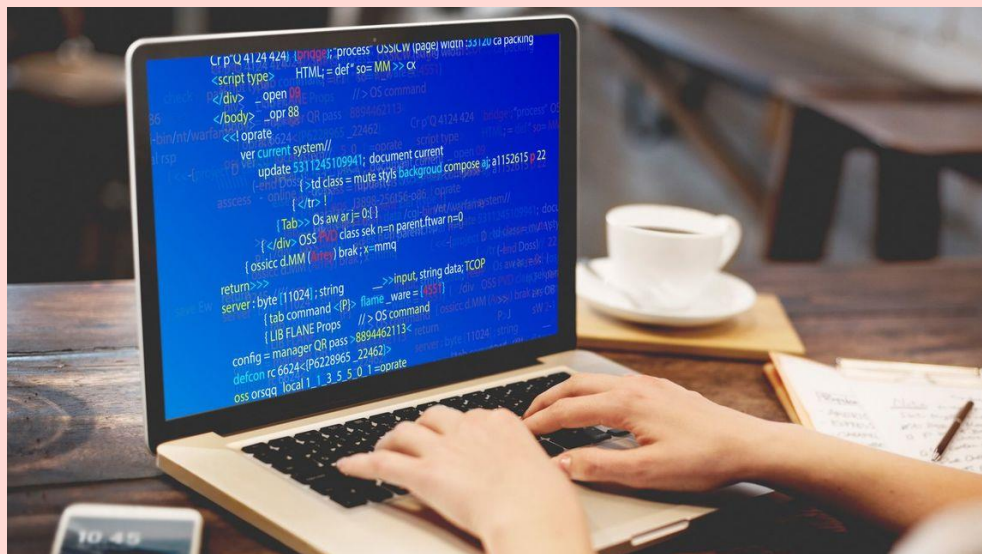
6

0

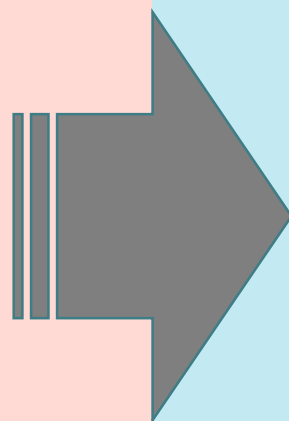




# PLSAマージ



Before



PLASMA-N2

プロジェクト ウィザード バージョン情報

PLSA Merge  
プロジェクト

変更する PLSA横軸複数 ... PLSAプロジェクト

モデル

最良のモデルを自動で選ぶ  
評価基準 AIC  必ず対象属性に最良のモデルを選択

自分で選ぶ  
モデル 02\_000000\_200213\_131307

データ生成・結合  
分析軸 X: rfid の結果を、カラム rfid

ハードクラスタ化  
 所属確率が最大  所属確率が閾値以上 閾値 0

カラム名  
 変更する  
[[ 軸名]]

区切りの文字 / プレビュー

前へ 次へ

[情報] ImportData : finished : 100 %

After

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

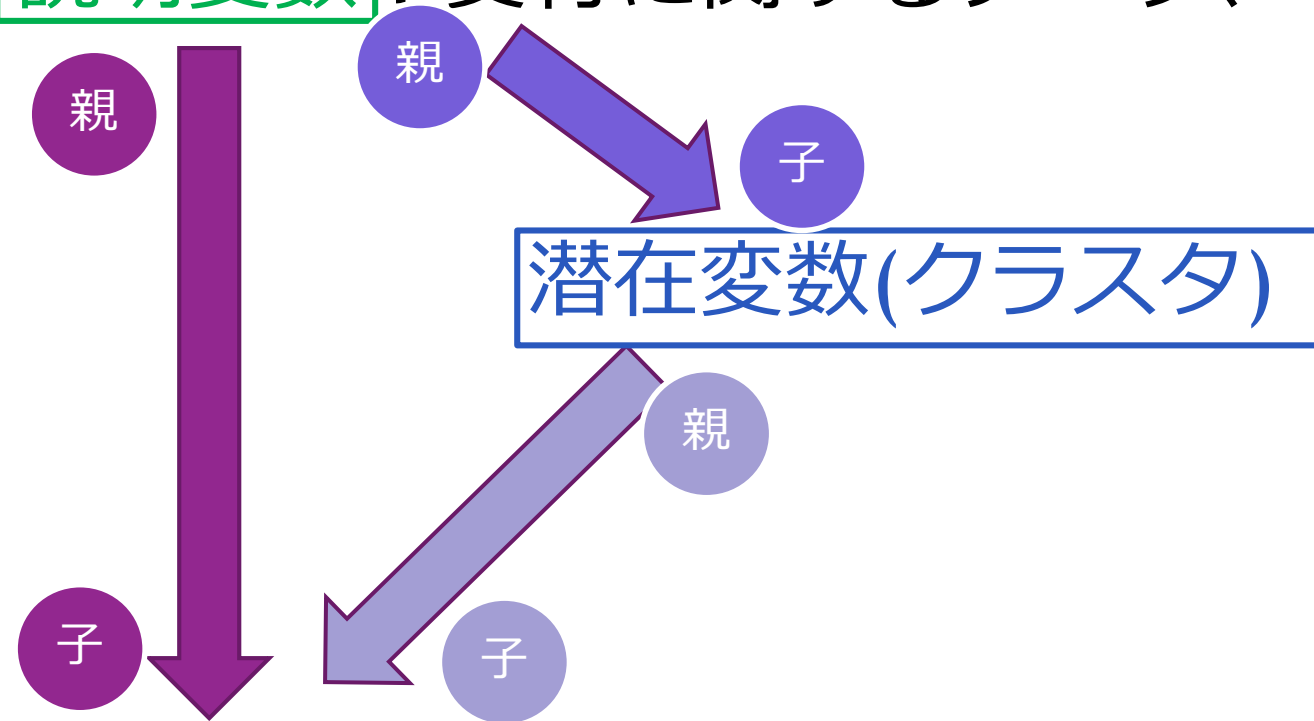
5

6



# 今回のイメージ...

**説明変数**：受付に関するデータ、時間、タッチ数



**目的変数**：振り返りデータ

# 基本操作

データ  
加工

pLSA

BN  
モデル  
構築

確率  
推論

1

2

3

4

5

★ 6

0

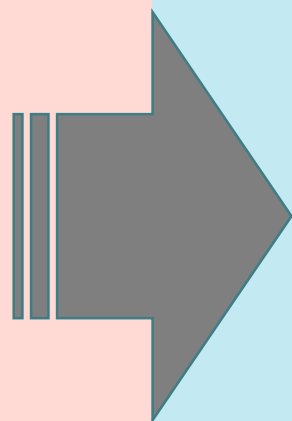
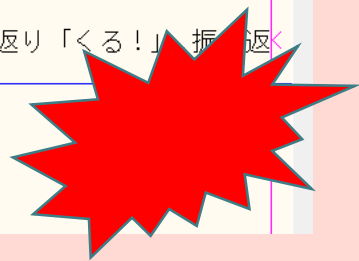


# BNINFER

```

1 # -*- coding:Shift_JIS -*-
2 target=振り返り「かならずくる!!」,振り返り「くる!」,振り返り
   「こない」,振り返り「これたらくる」
3 evidence=受付「たんけん」,受付「おはなし」,受付「AI」,受付「こう
   さく」,受付「はかる」,受付hour「10」,受付hour「11」,受付hour「12
   」,受付hour「13」,受付hour「14」,受付hour「15」,受付hour「16」,
   最終hour「10」,最終hour「11」,最終hour「12」,最終hour「13」,最終
   hour「14」,最終hour「15」,最終hour「16」,最終hour「17」,滞在時間
   hour「0」,滞在時間hour「1」,滞在時間hour「2」,滞在時間hour「3」,
   滞在時間hour「4」,滞在時間hour「over5」,タッチ数「0」,タッチ数「
   1」,タッチ数「2」,タッチ数「3」,タッチ数「4」,タッチ数「5」,タッ
   ち数「over6」
4 reference=振り返り「かならずくる!!」,振り返り「くる!」,振り返
   り「こない」,振り返り「これたらくる」
5 output.separator=.
6 output.format=%.%7f
7 output.map=true
8 output.map.postfix=MAP
[EOF]

```



```

win_x64\plasma_v3_win_x64\bin
plasma_v3_win_x64\bin>bnninfer validate -o property.prop
plasma_v3_win_x64\bin>bnninfer validate -o result.csv -p
plasma_v3_win_x64\bin>_

```

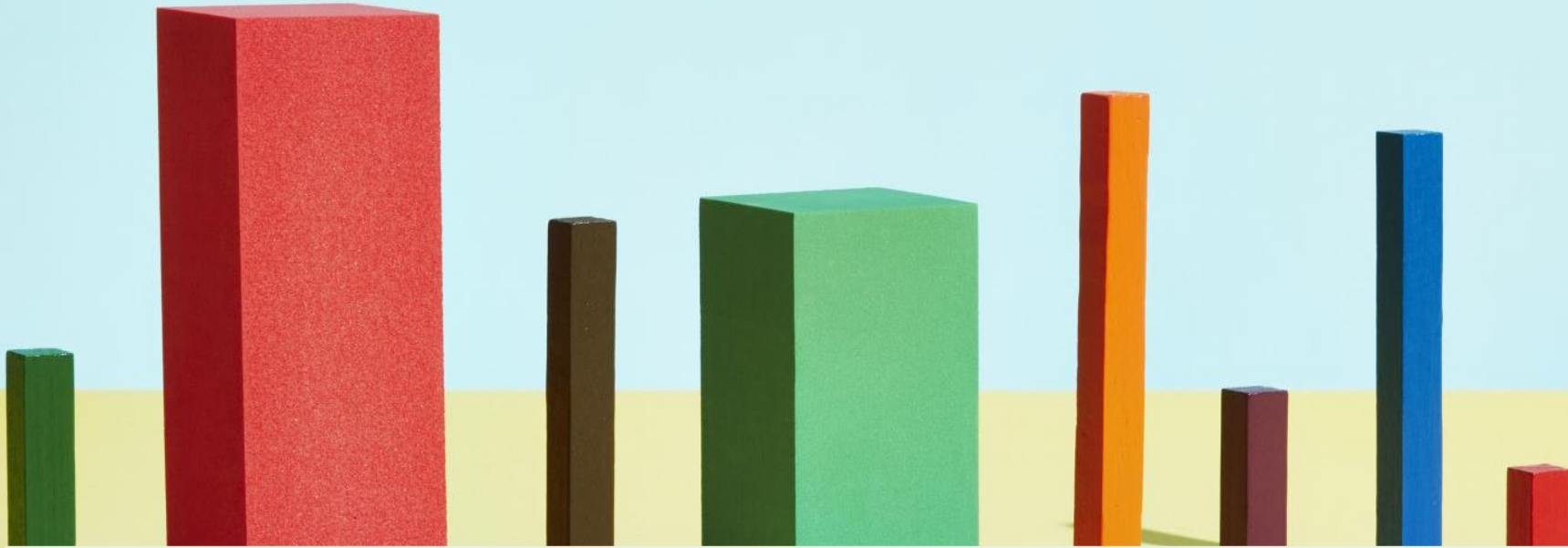


plasma\_n1\_v3\_20200220 > plasma\_v3\_win\_x64 > plasma\_v

名前	
result.csv	
property.prop	
evidence.csv	種類: PROP ファイル
model.bif	サイズ: 838 バイト
plasma_svr.exe	更新日時: 2020/02/27 14:00

Before

After



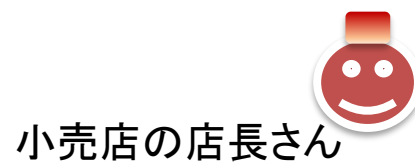
ありがとうございました

上記までが本編・下記はFAQでございます

# FAQ.連続値の予測は？



工場の良品の生産量予測したい

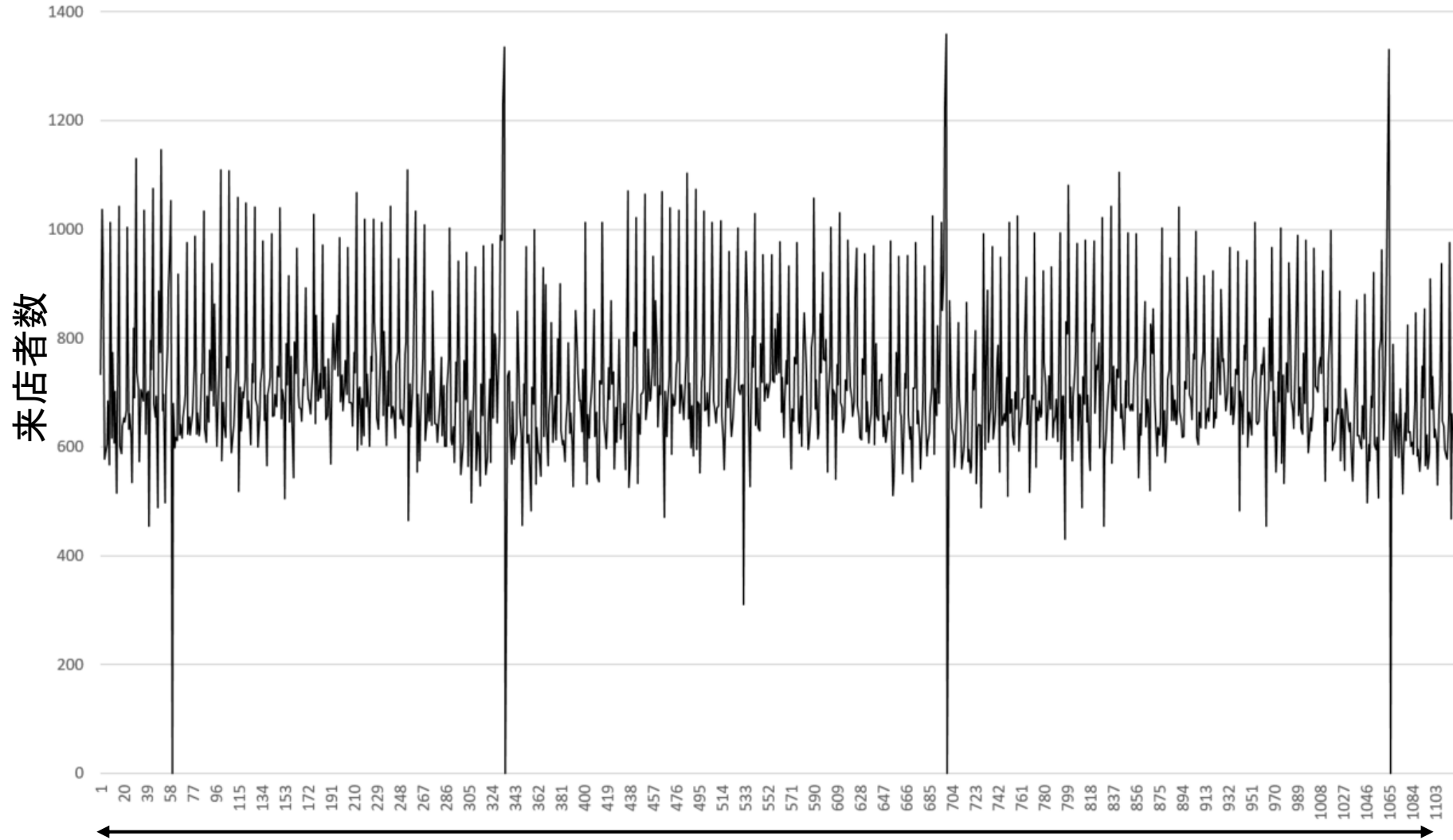


来店者数を予測したい



# 連続値の時系列

(ある店舗を訪れた来店者数を例に)



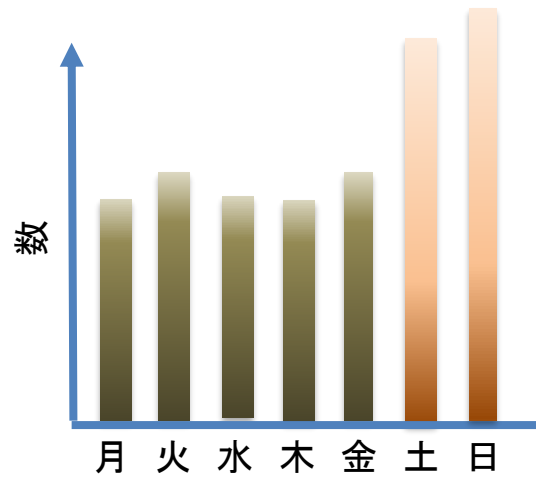
3年分

製造業

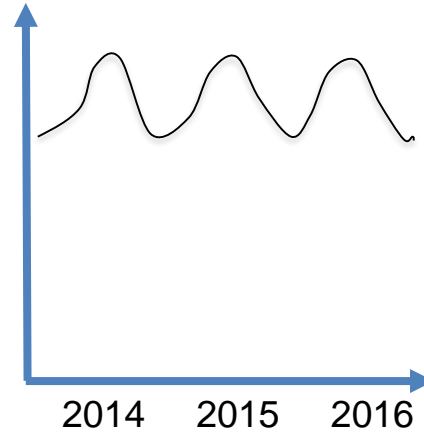


工場の良品の生産量  
予測したい

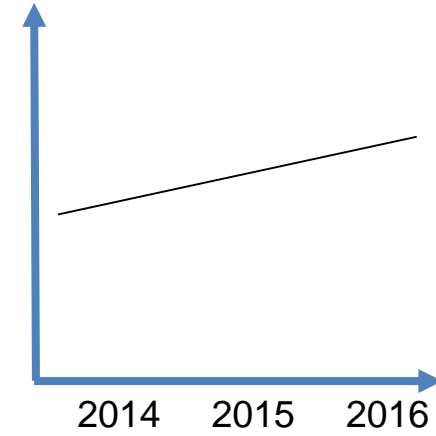
# ベースモデル



週変動



年変動



トレンド

ベースモデルに

# カスケードした形のモデル

## ベースモデル

(1次モデル)

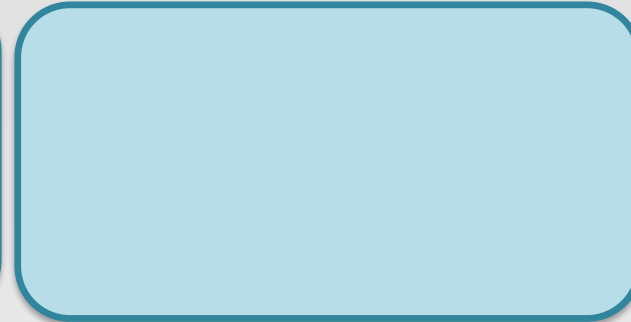
典型パターンを現す



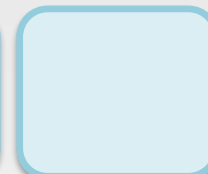
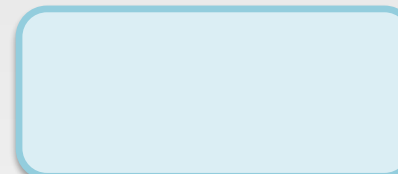
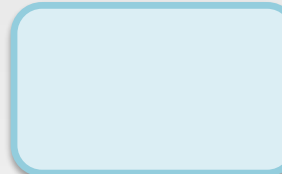
## 補正モデル

(2次モデル)

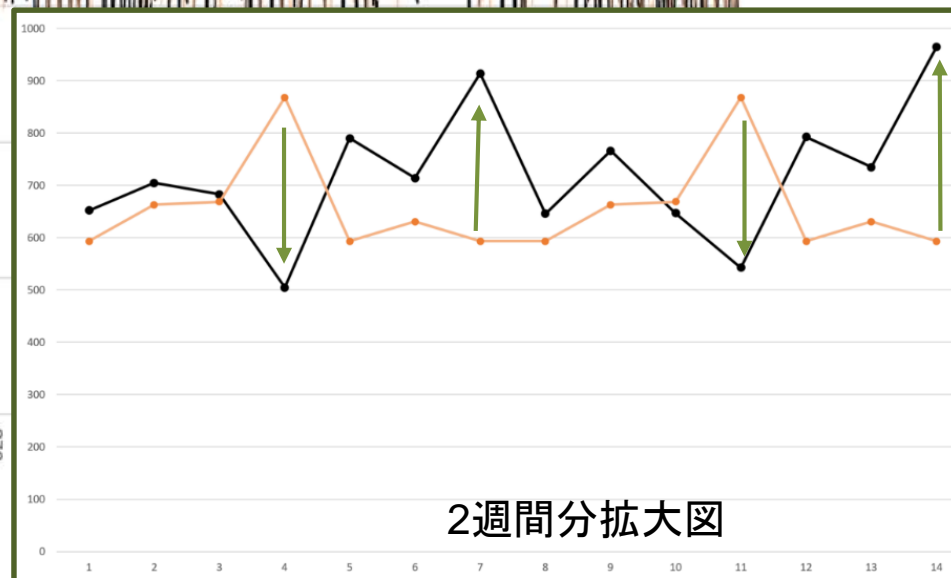
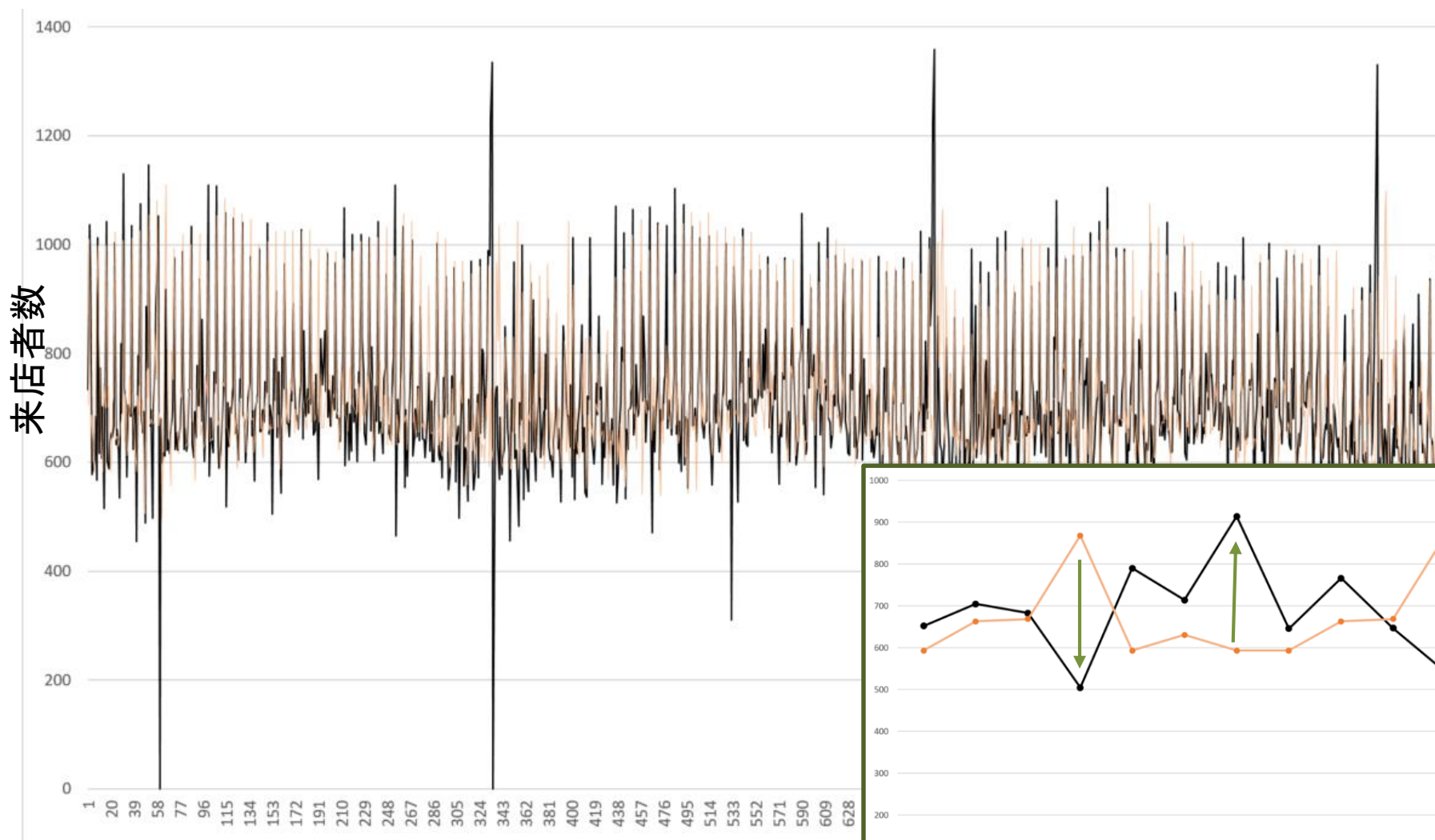
・ベイジアンネットワーク



## 3次モデル



# 注目の変数の時系列

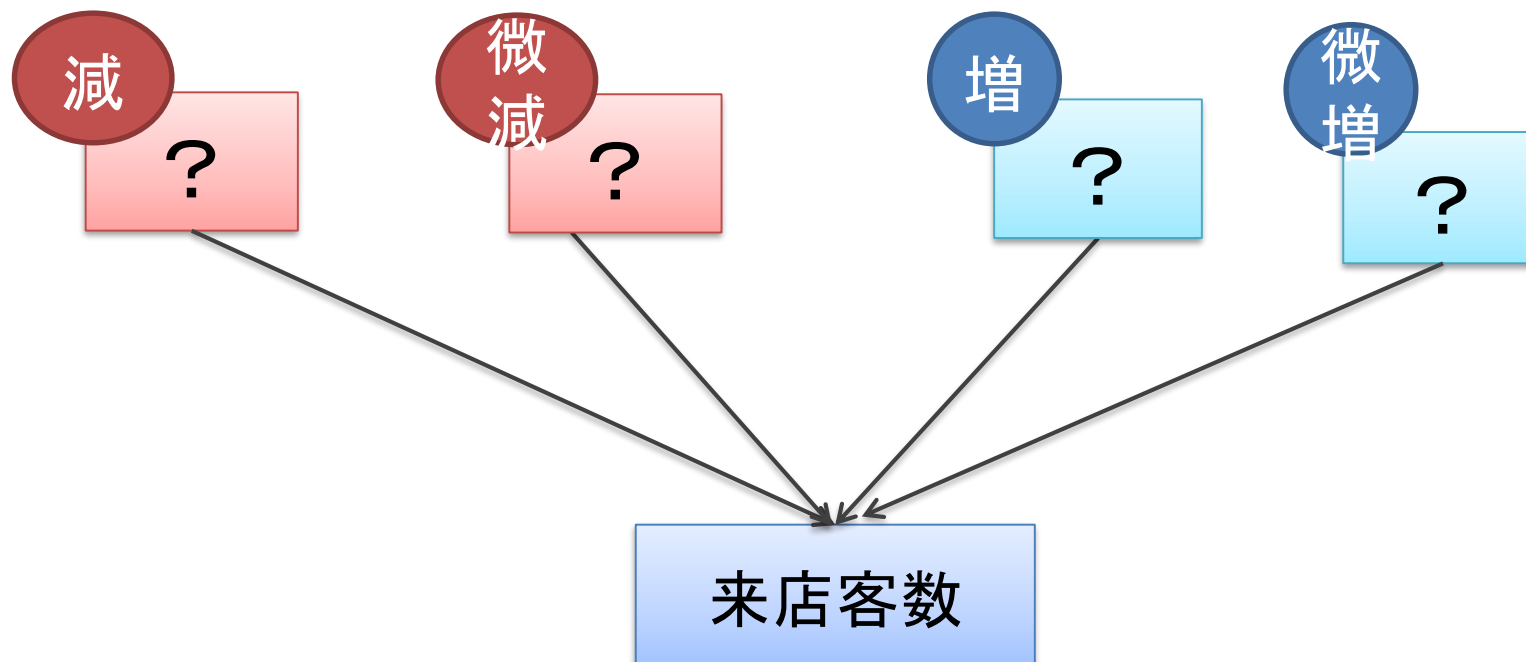


2週間分拡大図

—●— 実測  
—●— ベースモデル

# PLASMAによるベイジアンネットワーク構築

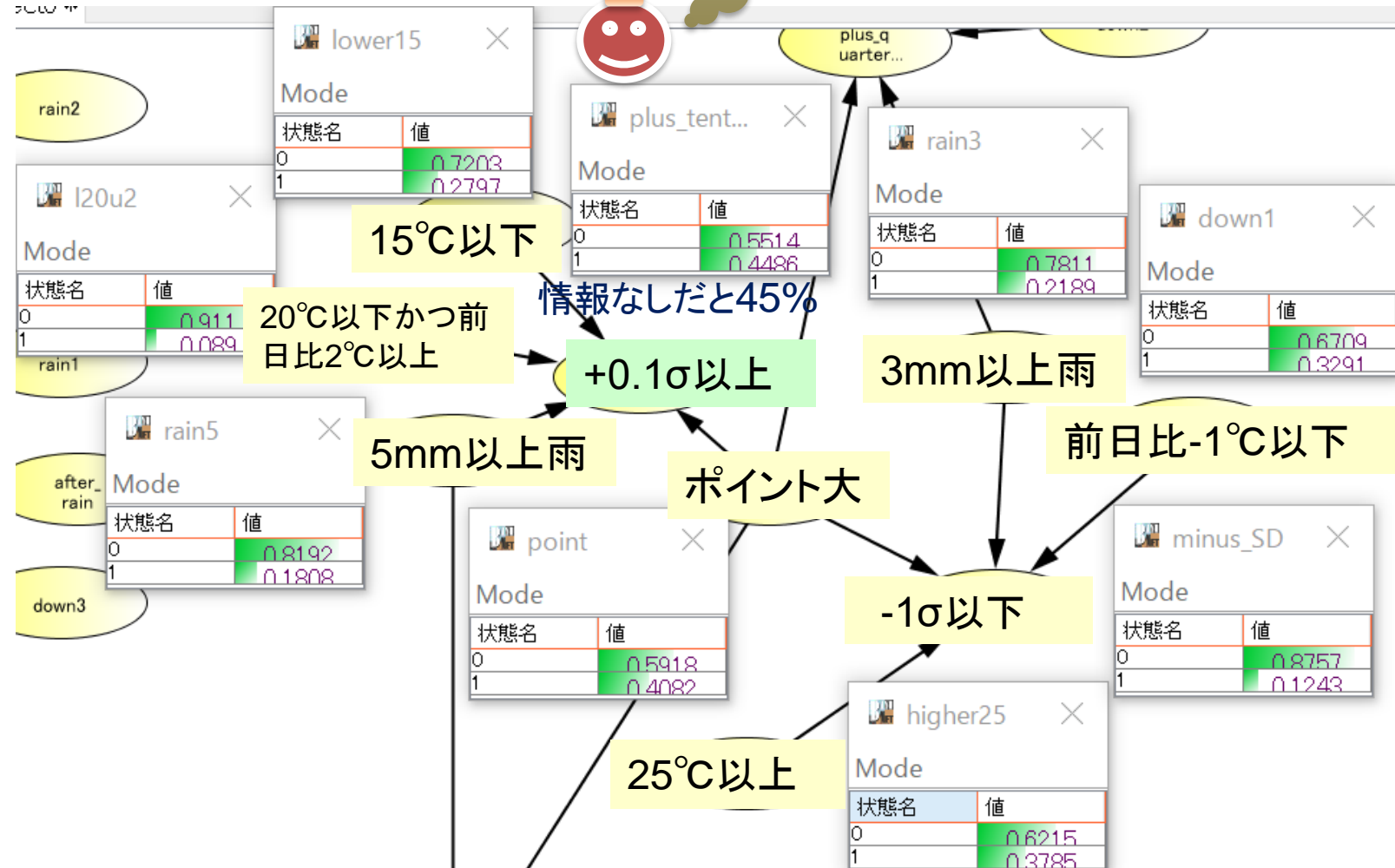
ベースモデル(1次モデル)で予測出来なかった増減を  
補正モデル(2次モデル)としてベイジアンネットワーク  
を構築する



# 来店者数モデル

7月3日、この時期の例年に比べて来客は増えるだろうか？

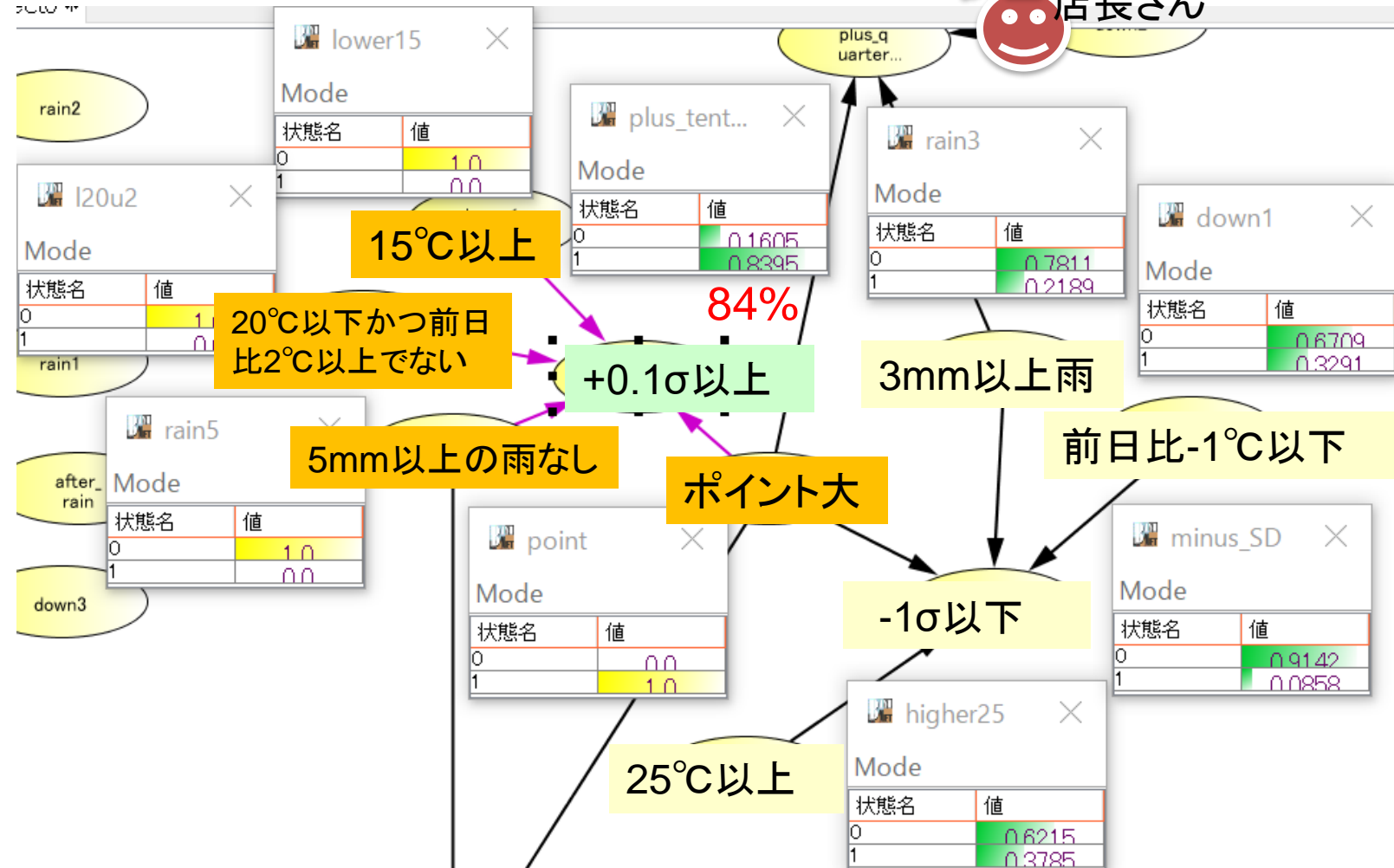
店長さん





# 確率推論(7名以上の来店増加予測)

ポイントおトクデーで、強い雨はない。増加しそうだ。



# モデルによる予測

平均: 654人  
標準偏差 $\sigma$ : 67人

		平均からの差 (単位: $\sigma$ )		モデル	
		実績	店長さん 手動予測	" $+0.1\sigma$ 以上の確率"	" $-\sigma$ 以下の確率"
2014/6/30	月	-0.17	-0.64	33%	1%
2014/7/1	火	+0.61	+0.09	84%	1%
2014/7/2	水	+0.25	-0.53	33%	5%
2014/7/3	木	-2.34	-0.05	3%	59%
2014/7/4	金	+1.86	+0.54	84%	1%

事前確率: 44%

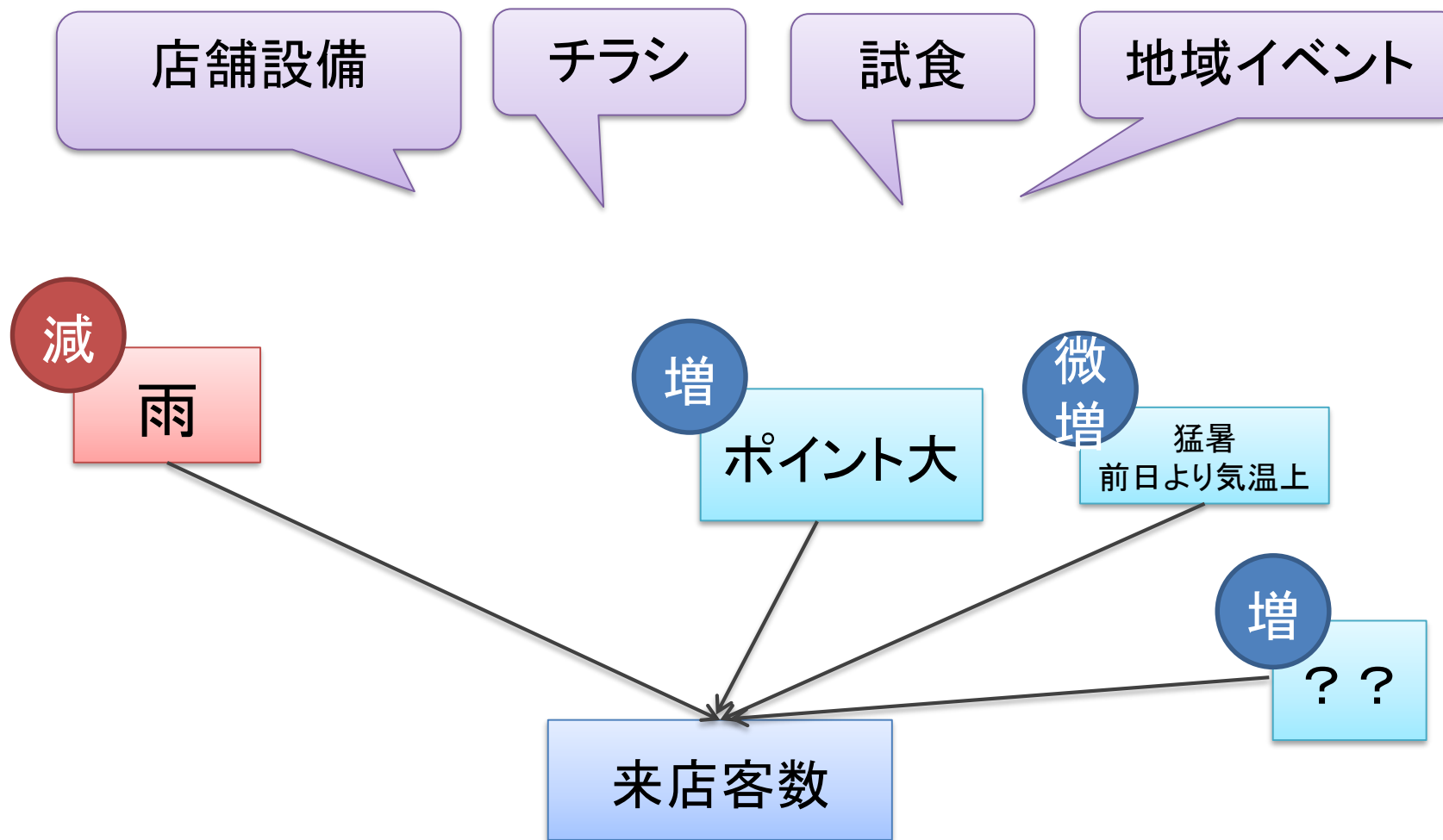
事前確率: 12%

1. 控えめ店長さんはもう少し攻められる (モデル→人)
2. 店長さんの知見をモデルに組み込める (人→モデル)
  - ・新たな変数の導入
  - ・モデルからのずれを予測するモデル
3. 新たな施策

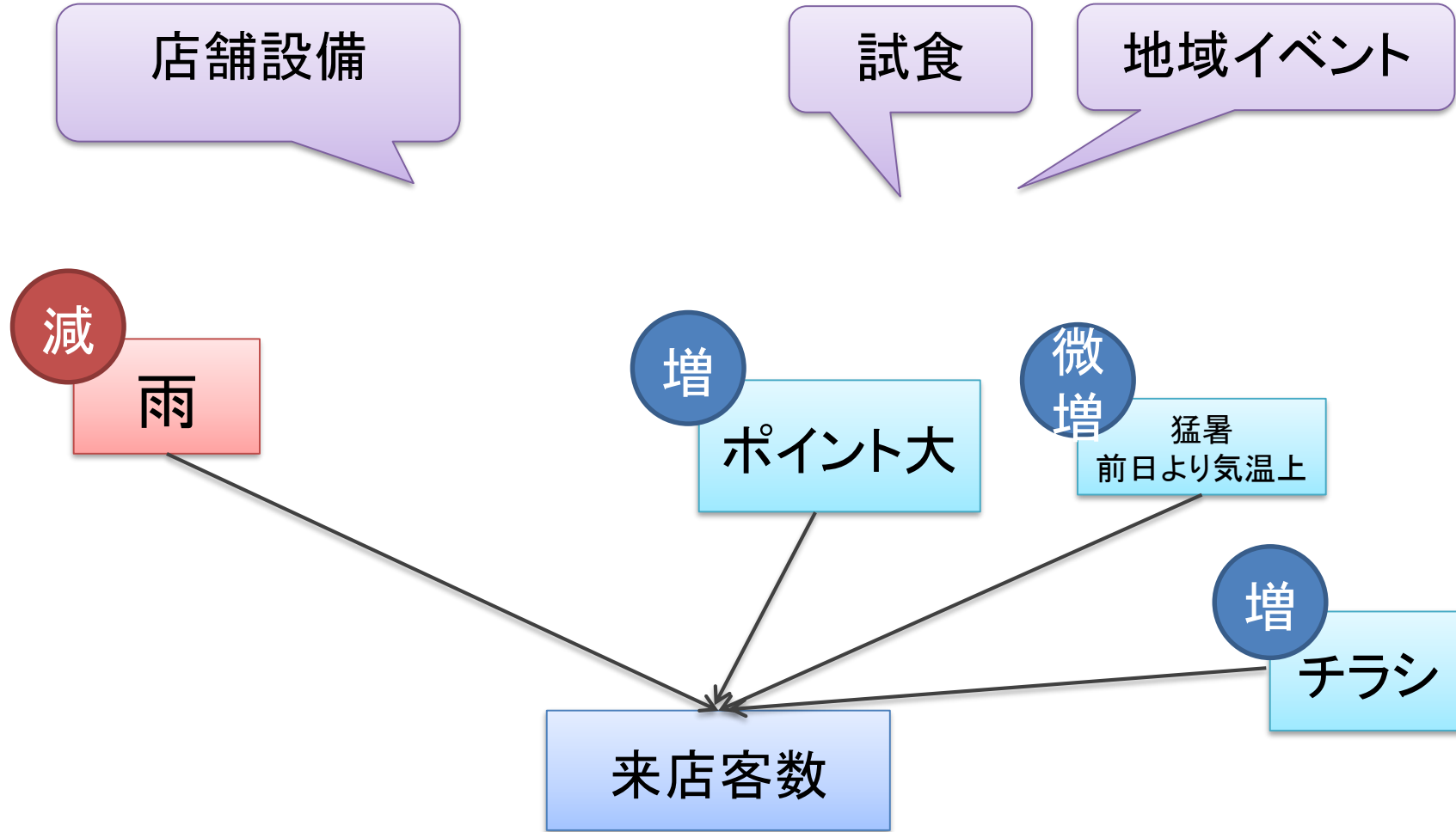
人と相互理解出来るAIによるデータ・知識融合

# 新たな変数

変数をつけ加えていける



# 新たな変数



# PLASMA 新機能説明会

ありがとうございました