

# 行動の多様化に対応する 新たな個人属性推定

Estimation of New Individual Attribute  
for diversity of behavior

---

Team I 山梨大学 University of Yamanashi

M2 澤田茜 Akane Sawada

M1 末木祐多 Yuta Sueki

M1 権女玲 Kwon Yeoryeong

B4 有働友哉 Tomoya Udo

B4 大竹浩太郎 Kotaro Otake

三菱重工業 佐野純平 Junpei Sano

# 背景 Background

- ✓働き方・暮らし方・考え方の多様化＝行動の多様化  
Diversity of Work Style・Life Style・Attitude

行動モデルを推定する際

個人属性（年齢・性別・職業...）を用いることが一般的  
Generally, when estimate the behavior model, using the individual attribute data (age/ sex/ job)



本当に個人属性で、個人の行動を推定できるの???  
Can you really estimate correct individual behavior with individual attribute???

- &Big Dataで様々な行動が得られるが、個人属性がないことが多い  
We can get various behavior data by using BIG DATA, but that don't have individual attribute data

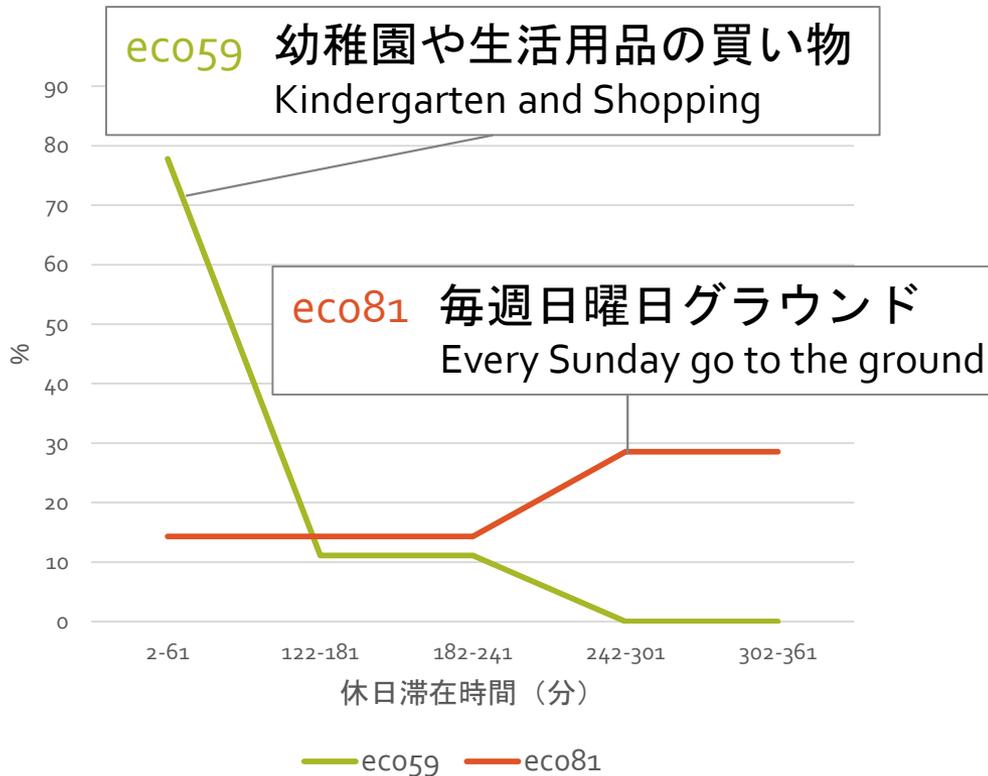
より個人を表現できる属性をつくりたい  
Our purpose is creation of individual attribute that can express individual better

# 基礎分析 Basic Analysis

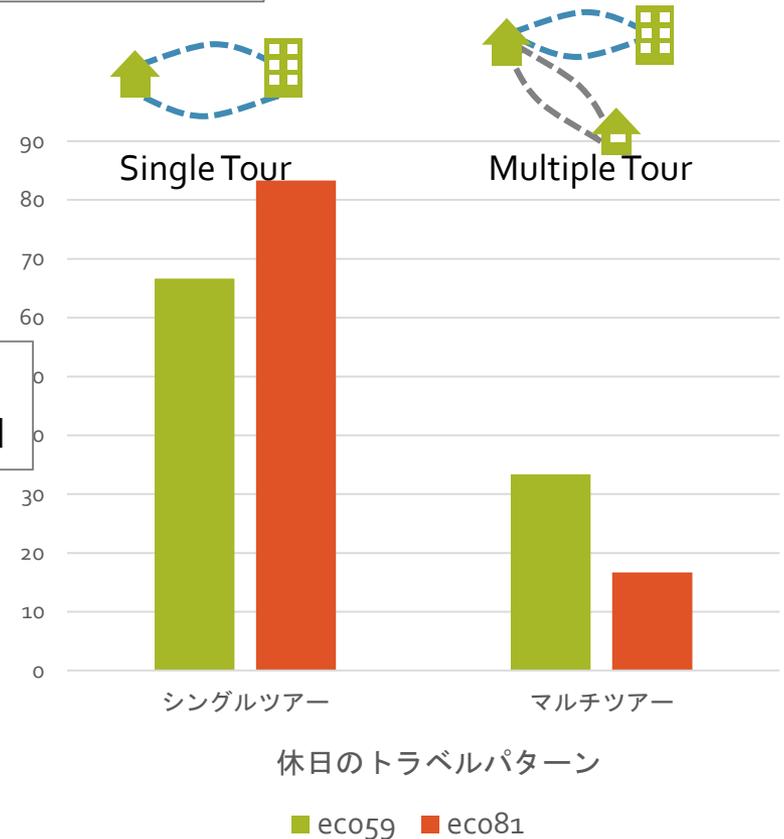
eco59 男性 33歳 会社員 (Man, 33 years old, Office worker)

eco81 男性 30歳 会社員 (Man, 30 years old, Office worker)

使用データ  
松山PP



Staying Time on holiday (minutes)



休日のトラベルパターン

Travel Pattern

# 新個人属性推定<sup>1</sup> Estimation of New Individual Attribute

✓同様の問題意識の研究（原 2017）では、GPSデータより滞在施設属性等を特定し、時間帯別の行動を集計

⇒これを主成分分析することで、個人の行動特性の抽出



新個人属性推定に私たちもチャレンジしたい！  
さらに個人の一日の行動に着目した形

個人のツアーごとの動きに着目 Focus on the behavior each individual tour

①主要ツアーと二次ツアーで分類 ②平日休日で分類

- ◆ 各ツアーの滞在時間
  - ◆ 各ツアーの出発時刻
  - ◆ 各ツアーのトリップ時間
1. Classify the tour by Primary tour and Secondary tour  
2. Classify the tour by Weekday and Holiday  
Staying / Departure / Trip time of each tour



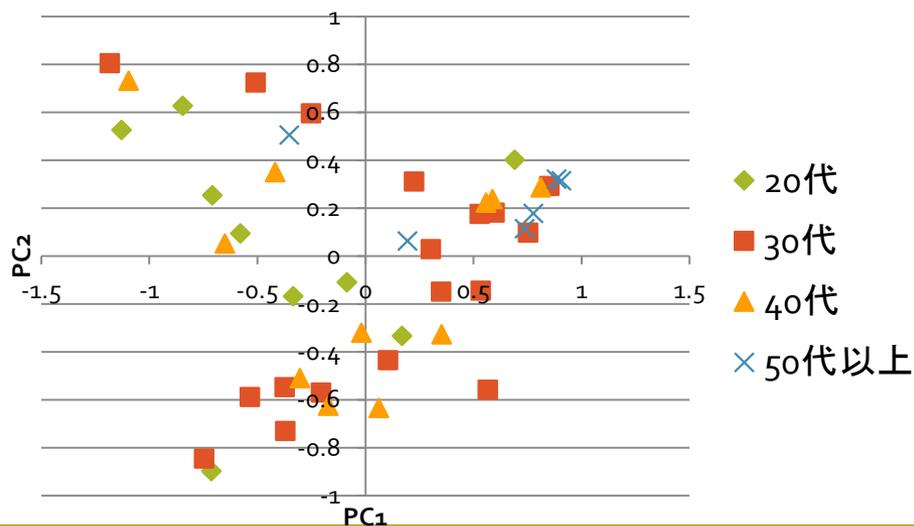
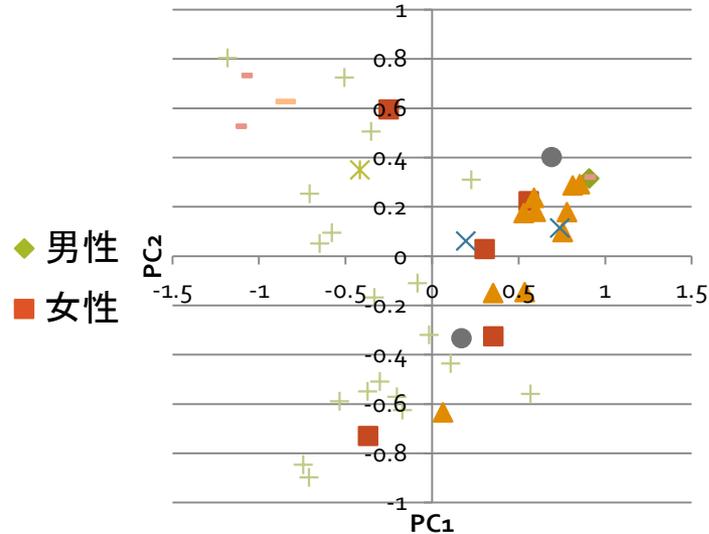
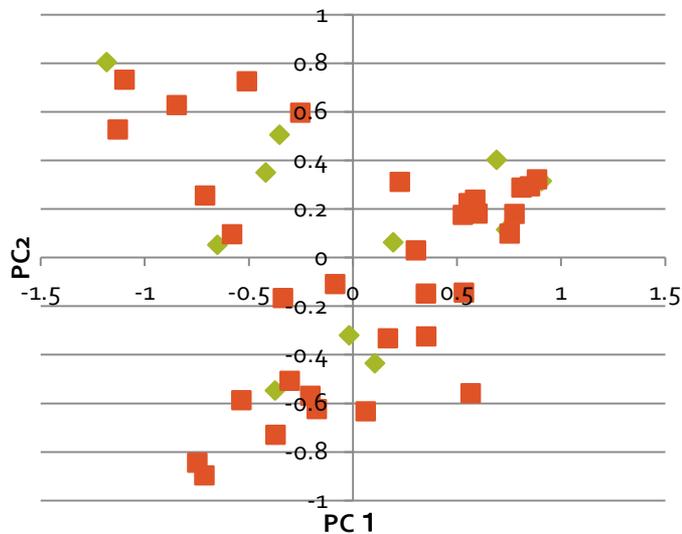
分類ごとに個人を集計・標準化し、主成分分析  
Counting individual for each classification, then using principal component analysis

# 主成分分析結果 Principal Component Analysis Results

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
滞在時間	シングル主要休日0.179	0.192865191	-0.014246177	0.238406735	-0.1284512	-0.129319527	-0.084310701	-0.028411631	0.231138635	-0.031
	シングル主要休日180.359	0.01262909	0.004822543	0.156686248	0.095080734	0.120325101	-0.106523677	-0.34310035	0.036913548	0.031
	シングル主要休日360.539	-0.05253655	0.210661797	-0.021526785	0.350193015	-0.135122918	-0.163609565	0.214446631	-0.109180899	-0.171
	シングル主要休日540.719	-0.099667696	-0.07159728	0.288793697	-0.078335533	0.164422669	-0.165810355	0.077008521	-0.080136671	-0.631
	シングル主要休日720.	-0.059604272	-0.087703467	-0.032928644	0.260107118	-0.132927053	-0.142801175	0.090386425	0.07066808	-0.411
	シングル主要平日0.179	0.316756169	-0.033330162	-0.051096966	-0.055963895	-0.008422155	-0.024062687	0.182470381	0.023631664	0.041
	シングル主要平日180.359	-0.014393708	-0.136575813	0.093808333	0.003774245	0.056177612	0.027354657	-0.26066678	-0.00457165	-0.111
	シングル主要平日360.539	-0.106950544	0.310566827	0.074830974	0.107101418	-0.126383107	0.2431156302	-0.059588023	0.073612984	-0.111
	シングル主要平日540.719	-0.0290783515	-0.034958394	-0.082788649	-0.108043643	0.10466071	0.035716221	-0.112952579	-0.073164855	0.111
	シングル主要平日720.	-0.100580761	-0.155422188	0.054375476	0.177090735	-0.157638301	-0.129293738	-0.060314765	0.057067626	0.261
	マルチ主要休日0.179	-0.070076861	-0.226980544	-0.006160916	0.152242485	-0.088107221	0.168914098	0.049382784	-0.047841043	0.011
	マルチ主要休日180.359	0.05635829	0.066548049	0.181548101	0.016342854	0.029398315	-0.02026639	-0.112845559	-0.014660643	-0.211
	マルチ主要休日360.539	-0.056425939	-0.06938884	-0.011894845	0.102782007	0.322902294	-0.126266593	0.256275275	0.023184976	0.091
	マルチ主要休日540.719	-0.034205354	-0.001624129	0.136530954	-0.047412233	0.087388845	0.10294187	0.345729704	0.217140704	0.071
	マルチ主要平日0.179	0.287457506	-0.007084373	0.116601881	-0.175469148	-0.00654241	-0.1653832823	-0.052200195	-0.008536075	-0.031
	マルチ主要平日180.359	0.183917861	-0.139174085	0.080459758	0.249004257	-0.130239511	0.062692172	0.071967334	0.2583131	-0.201
マルチ主要平日360.539	-0.103048706	0.354729655	-0.081222526	0.064301099	-0.126922586	0.124289993	-0.066510156	0.144784616	0.131	
マルチ主要平日540.719	-0.045231671	-0.015256709	-0.103976728	-0.229329701	0.188494129	-0.168333535	0.006885842	-0.169826597	0.021	
出発時間	平日.主要ツアー-0.5	0.089147324	0.004342656	0.107161902	0.007683206	0.105327567	-0.038160553	-0.14410877	-0.047073659	0.001
	平日.主要ツアー-6.7	-0.0298793716	-0.11335733	0.243084046	0.058148036	-0.125262472	-0.119902091	-0.024490275	-0.166387083	-0.011
	平日.主要ツアー-8.9	0.0551509052	0.030236365	0.230499867	-0.057946219	-0.008260089	0.138997155	-0.065816683	0.289305147	0.081
	平日.主要ツアー-10.11	0.0522860133	-0.013923915	0.02553929	-0.109535544	-0.022787435	0.02082445	-0.121779745	-0.046207018	-0.151
	平日.主要ツアー-12.13	0.284317272	-0.078508109	0.15163015	-0.1179484	-0.03521711	-0.09422923	0.040815542	-0.041478874	-0.011
	平日.主要ツアー-14.15	0.1228293242	0.014461087	0.38933857	-0.100636865	-0.062120468	0.029376408	0.277366848	-0.137037496	0.041
	平日.主要ツアー-16.17	0.182978096	0.0334789668	-0.050681159	0.154498587	-0.005783955	-0.085791963	0.14666442	-0.07880246	0.041
	平日.主要ツアー-18.19	0.193559393	-0.043598899	-0.048636221	0.183745776	0.338972211	0.073376177	-0.074407313	-0.036297585	0.131
	平日.主要ツアー-20.21	0.03564833	-0.089766853	-0.111678987	-0.115834817	-0.004105511	0.087821371	-0.023907584	-0.157753791	-0.041
	平日.主要ツアー-22.23	0.099114671	-0.121516489	-0.081004055	0.348056595	0.163087165	-0.014844381	-0.06276611	-0.06276611	0.911
	平日.2次ツアー-0.8	-0.047107927	-0.217565078	0.131207972	0.149889571	-0.58726875	0.022286866	-0.011316683	-0.083650553	0.031
	平日.2次ツアー-9.11	0.125635039	0.223170996	0.27201218	0.040064381	-0.123964195	-0.049533175	0.000816512	-0.24016261	0.011
	平日.2次ツアー-12.14	0.10636792	-0.230814978	0.025692003	0.236874125	0.103215019	0.084414768	-0.075783782	-0.033564265	0.201
	平日.2次ツアー-15.17	0.14450526	-0.007452799	-0.1817317124	-0.041882001	-0.137057805	-0.186028045	0.114849881	0.229599741	-0.041
	平日.2次ツアー-18.20	-0.056630314	0.057513901	-0.201524567	-0.1451902	0.079406826	0.03906102	-0.066884365	0.003477139	-0.091
	平日.2次ツアー-21.23	-0.103328986	-0.021328258	0.13245323	-0.100525147	0.039813	0.250381457	0.29244364	0.097773292	0.001
	休日.主要ツアー-0.5	-0.077273452	-0.009440208	0.262179382	0.133647587	0.092419287	0.006623256	0.150948905	0.246056001	-0.171
	休日.主要ツアー-6.7	-0.153481575	-0.165923438	-0.0005561739	0.179360433	0.031590137	-0.250699805	0.21454322	0.032844075	-0.151
	休日.主要ツアー-8.9	-0.096064722	-0.127821337	0.378054031	-0.074405873	0.021094281	-0.050870555	-0.047948975	-0.107562443	-0.031
	休日.主要ツアー-10.11	-0.021906563	0.024733972	0.075421983	0.128806239	-0.069821291	-0.1231128109	-0.033453519	0.172687259	-0.001
	休日.主要ツアー-12.13	0.089467988	0.021875439	0.16191964	-0.099393857	-0.132903647	-0.202593375	0.005402848	0.184960865	0.011
	休日.主要ツアー-14.15	0.107305558	-0.004390713	0.213252136	-0.187820212	-0.041464299	0.198639847	0.079515433	0.104986759	-0.031
	休日.主要ツアー-16.17	0.083280576	0.143587346	-0.158238403	0.076292952	-0.136737532	0.1424118612	0.027480419	-0.27717847	0.011
	休日.主要ツアー-18.19	0.010894169	-0.002200998	0.195883361	0.035300104	0.234405703	0.212331119	-0.10736998	-0.075107992	-0.071
休日.主要ツアー-20.21	0.0084750784	0.021732464	0.13434992	-0.13903713	0.090037789	0.243563829	0.282799319	0.101146672	-0.011	
休日.主要ツアー-22.23	-0.002056838	0.36047568	0.19245274	0.225324478	-0.130501958	-0.03304964	0.140950617	-0.091683066	0.101	
休日.2次ツアー-0.8	0.04024233	0.1321303	0.052882428	-0.089890565	-0.102532089	-0.14694256	-0.005761563	-0.076779186	0.271	
休日.2次ツアー-9.11	0.08292364	-0.102504763	-0.050863394	-0.080247548	-0.28691638	0.198499403	0.179000471	-0.253424232	-0.161	
休日.2次ツアー-12.14	-0.022385693	0.198030666	-0.044203718	0.196662811	-0.091926414	0.058346379	0.011813184	0.079192326	0.171	
休日.2次ツアー-15.17	0.05023709	0.122418595	-0.10479792	0.059781792	-0.064689931	-0.04657093	-0.32922525	0.315600727	-0.111	
休日.2次ツアー-18.20	0.05445797	0.045357162	0.186598023	0.021059471	0.292352071	0.111161709	-0.129771195	-0.041890008	-0.081	
休日.2次ツアー-21.23	-0.147890549	0.052171857	-0.124962825	-0.148746256	0.214822845	-0.205539024	0.074772782	-0.083175865	0.041	

# 個人属性比較<sup>1</sup> Compared Individual Attribute with New One

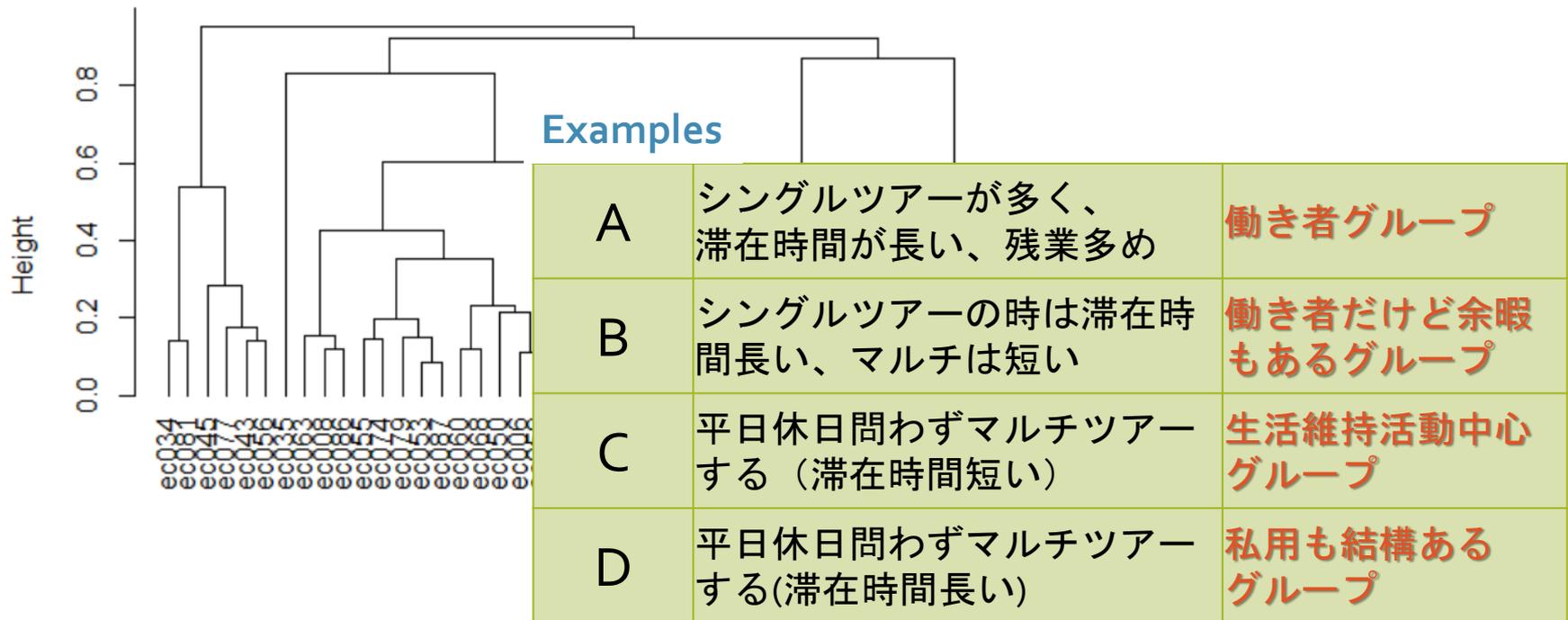
## 第一主成分・第二主成分と個人属性の比較



# 新個人属性推定2 Estimation of New Individual Attribute

個人ごと集計・標準化したデータを**クラスター分析**により分類  
...滞在時間 / 出発時刻 / トリップ時間

Cluster Dendrogram



# 個人属性比較 Compared Individual Attribute with New One

- ✓ビッグデータ(GPS etc.) は目的が分からないことが多い  
Big Data (GPS etc.) can't identify the purpose Sometimes combined with questionnaires
- アンケート調査と組み合わせて用いて分析することもある

ツアーごとの目的を当てにいきたい Construct the Purpose choice model

⇒個人属性と新個人属性の比較を行い、有効性を示したい

Compared the individual attribute with the New individual attribute

- ✓機械学習 (Random Forest) によりそれぞれの個人属性で目的を当てた

①個人属性 ②新個人属性<sub>1</sub> (主成分) ③新個人属性<sub>2</sub> (クラスタ)

《入力データ》

ツアー	開始時間	滞在時間	手段	平日 ダミー	個人属性...>	目的
1	8	562	鉄道	1		仕事
2	7	600	鉄道	1		仕事
3	12	25	徒歩	0		買い物
4	17	120	バス	0		その他
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.

# 個人属性比較結果 Results of Compared Individual Attribute

✓機械学習（Random Forest）によりそれぞれの個人属性で目的を当てた

## ①個人属性

## ②新個人属性<sub>1</sub>（主成分）

## ③新個人属性<sub>2</sub>（クラスタ）

個人属性	的中率	PCA	的中率	クラスタ	的中率
その他私用	61%	その他私用	69%	その他私用	98%
帰社・帰校	0%	帰社・帰校	40%	帰社・帰校	100%
業務	53%	業務	39%	業務	83%
娯楽	25%	娯楽	25%	娯楽	93%
出勤・登校	98%	出勤・登校	97%	出勤・登校	100%
食事	8%	食事	5%	食事	93%
買い物	59%	買い物	57%	買い物	98%
計	70%	計	70%	計	97%

クラスタ分析を用いた新個人属性がかなり当たる？！  
 （的中率97%...しかしクラスタリングが過学習の可能性あり）  
 ⇒適切なクラスタリングをするように今後考慮していきたい

# 都市間比較 Compared between the city

- ✓今回松山PPを用いて作成した新個人属性を横浜PPに付与し、他都市間で今回の新個人属性が共通なのかを検討したい



横浜PPのツアーサンプル数が少なく、計算が回らなかった

```
> ##### 主成分分析PC4 結果使用 #####
>
> data <- read.csv("~/yokohamaPCA.csv",head=T)
>
> ndata <- nrow(data)
> ridx <- sample(ndata,ndata*0.38)
> data.learn <- data[ridx,]
> data.test <- data[-ridx,]
>
> forest <- randomForest(object~.,data=data.learn,ntree=500)
Error in randomForest.default(m, y, ...) : Can't have empty classes in y.
> pred.forest <- predict(forest,newdata=data.test,type="class")
Error in eval(expr, envir, enclos) : object '平日ゴミ-' not found
> aaa <- table(pred.forest,data.test[,1])
Error in table(pred.forest, data.test[, 1]) :
 全ての引数は同じ長さでなければなりません
>
> write.table(aaa,"PCA比較20.csv")
>
> ##### 主成分分析PC4 結果使用 #####
>
> dataq <- read.csv("~/yokohamaPCat.csv",head=T)
>
> ndataq <- nrow(dataq)
> ridxq <- sample(ndata,ndataq*0.38)
> data.learnq <- data[ridxq,]
> data.testq <- data[-ridxq,]
>
> forestq <- randomForest(object~.,data=data.learnq,ntree=500)
Error in randomForest.default(m, y, ...) : Can't have empty classes in y.
> pred.forestq <- predict(forestq,newdata=data.testq,type="class")
> aaaa <- table(pred.forestq,data.testq[,1])
>
> write.table(aaaa,"PCA比較21.csv")
```

# 提案 Suggestion



同じ属性でも行動スタイルが多様になってきた今日

本当に個人属性で、個人の行動を推定できるの???



私たちは個人の行動履歴をクラスターに分割し  
個人属性にすることを提案します

## まとめ Conclusion

### 行動の多様化が進んでいる



- ✓ より詳細な個人属性の付与  
⇒ 目的推定の精度向上

より様々な行動モデルが当たりやすくなるのでは？！

### ビッグデータ活用も可能



- ビッグデータは個人属性・目的が分からないことが多い
- ✓ 行動パターンにより個人属性を付与可能
  - ✓ 目的推定も可能であることを示した

ビッグデータを行動モデルに活用できる