

2023年度 統計データ分析コンペティション

審査員奨励賞 [大学生・一般の部]

変数重要度に着目したクラスタリングによる
社会構造と睡眠時間の関係性の解析

薫田 千雪・ 高橋 希帆

(千葉工業大学先進工学部知能メディア工学科)

変数重要度に着目したクラスタリングによる 社会構造と睡眠時間の関係性の解析

薫田 千雪・高橋 希帆

(千葉工業大学先進工学部知能メディア工学科)

1. 研究のテーマと目的

今日経済の成長と競争の激化により、労働者は仕事に多くの時間を費やすケースが増加する傾向にあり、長時間労働がワークライフバランスに悪影響を与えることが懸念されている。ワークライフバランス憲章^[1]によれば、ワークライフバランスが実現された社会とは、「国民1人ひとりがやりがいや充実感を感じながら働き、仕事上の責任を果たすとともに、家庭や地域生活などにおいても、子育て期、中高年期といった人生の各段階に応じて多様な生き方が選択・実現できる社会」であるが、近年日本でワークライフバランスが声高に叫ばれている背景には、出生率の低下や、過労死や脳心臓疾患、精神疾患につながる過重労働の問題がある^[2]。ワークライフバランスを改善することで、「健康」「生産性」「家族時間」「趣味」「ストレス軽減」「モチベーションの維持」など様々な観点から問題を改善し、生活の質を向上させることが期待できる。実際に、余暇時間、食事や入浴などの生活時間と睡眠時間の組み合わせが睡眠の質やメンタルヘルスと関連することが報告されている^[3]。ワークライフバランスに影響する要因の一つに睡眠が挙げられる。良好な睡眠は仕事の安全性、健康の促進、職業生活の充実に密接に関連する。労働者に特化して国内での睡眠時間の変化をみると、睡眠が6時間未満である労働者が2012年では46.5%に至り、1992年前と比べて2.2倍になっている^[4]。十分な睡眠時間を確保することで、体調面のみならず心理的な健康の改善にも大きく改善できるといえる。

一方で、生活様式によって睡眠に割ける時間は大きく異なることから、社会構造や産業構造といった地域の在り方によって睡眠時間も変化すると考えられる。以上より、本研究では睡眠時間に着目し、睡眠時間に影響を及ぼす要因と産業構造の関連性を解析する。

2. 研究の方法と手順

本研究では、睡眠時間に重点を置きどの変数が影響を及ぼしているのか、また、社会構造が睡眠時間にどのような影響を及ぼしているのかを調査するにあたり、以下の手順で解析を行った。

手順1: 重要な変数を選択する

ここでは、まず睡眠時間に影響を及ぼす生活時間や職業を調べるために各都道府県における業種別の従業者数の割合データと、睡眠を除いた生活時間データを特徴量、睡眠時間を目的変数としてランダムフォレスト回帰により変数重要度を算出した。

手順2: 標準化したデータをクラスタリングする

睡眠時間と関係のある変数が、地域ごとにどのような差を示しているのかを調べるために手順1で得られた項目を対象として都道府県をクラスタリングする。少量のデータに対して高性能であり、クラスタ数を事前に決める必要のないAffinity Propagationを用いてクラスタリングを行う。

手順3：クラスタリング結果の可視化と解釈

ここでは、t-SNEによって低次元マッピングをし、Affinity Propagationによって各クラスタの代表点(exemplar)を検出し、手順1で算出した変数重要度から重要そうである変数を取り出し比較することで睡眠に影響する産業の特徴について議論する。

3. データセットの加工

本研究では、各都道府県の産業構造・睡眠時間と生活様式を調べるために教育用標準データセットの市区町村データと社会生活データを主に利用して分析を行った。SSDSE-Aの市区町村データを都道府県ごとに集計し、経済基盤と労働データを抽出した。また、従業者数を割合で表示するために全ての値を人口で割って表示している。SSDSE-Dの社会生活データからは、1日の生活の中で費やす時間を表す項目として生活時間のデータから総数のデータのみを抽出した。

手順1で示したように、各都道府県における業種別の従業者数の割合データと、睡眠を除いた生活時間データを特徴量、睡眠時間を目的変数としてランダムフォレスト回帰により変数重要度を算出した。最終的に抽出して使用した変数を以下の表に示す。

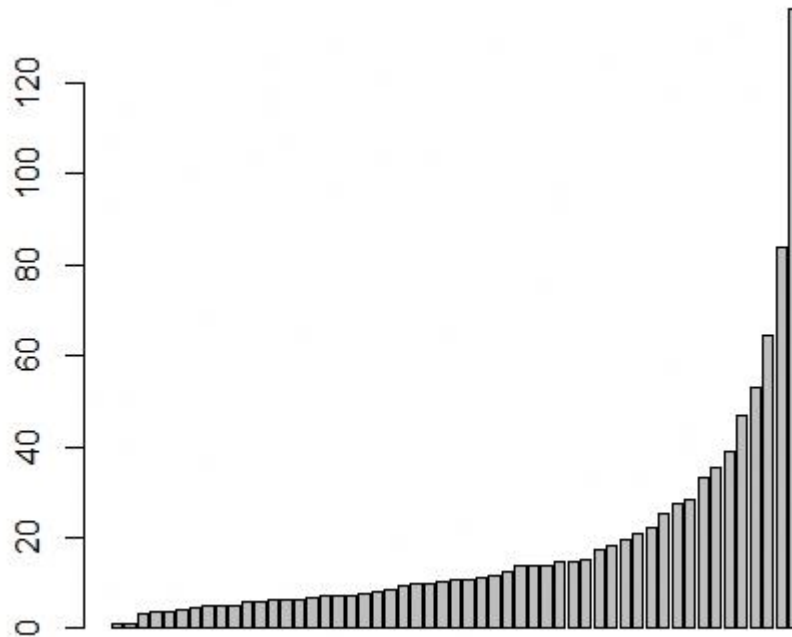
表1 本研究で使用したデータとその出典一覧

使用した変数	出典	調査年度
a 従業者数(民営)(農業, 林業)	総務省統計局, 経済産業省大臣官房 調査統計グループ「経済センサス-活動調査」	2016
b 従業者数(民営)(漁業)		
c 従業者数(民営)(建設業)		
d 従業者数(民営)(運輸業, 郵便業)		
e 従業者数(民営)(医療, 福祉)		
f 第1次産業就業者数	総務省統計局	2020
g 第2次産業就業者数	「国勢調査」就業状態等基本集計	
h 睡眠	総務省統計局「社会生活基本調査」	2021
i 通勤・通学		
j 仕事		
k 休養・くつろぎ		
l 趣味・娯楽		
m 交際・付き合い		

4. データ分析の結果

手順1のようにランダムフォレストによって計算された睡眠時間の予測における変数の重要度を図1に示す。縦軸は変数重要度であり、横軸は各変数名に対応している。ここでは、文字サイズの表示上の都合より変数名を削除している。上位の変数の中で特徴的なものを抽出し変数として利用した。使用したデータは表1に掲載している。

図1 手順1で計算された変数重要度



次に、各クラスターの重心と特徴量の関係性を表2に示す。Exemplarsの列はAffinity Propagationによって検出された各クラスターの重心となる都道府県を表している。ここでは、群馬、東京、富山、兵庫、愛媛、熊本、大分、沖縄が各クラスターの代表県として検出された。尚、各代表県に所属する県は表3に示す。目的変数は変数重要度を調べるために使った睡眠時間であり、重要と判断された特徴量をそれぞれ表示している。上の表は生活様式に関する特徴量に対応しており、一日の生活において割く活動の時間を表しており、下の表は各都道府県の産業の割合を表している。表は、各変数において値が上位に位置するセルを赤色、下位になるにつれて青色となるよう色付けしている。なおここに掲載している変数は、表1で示している手順1で選ばれた重要度が高い変数のみを表示している。

表2 検出されたクラスターと特徴量の関係性

都道府県	目的変数	生活様式の特徴量							
exemplars(重心)	睡眠	仕事	休養 くつろぎ	趣味、娯楽	交際、付き合い	通勤通学			
群馬	475	209	117	44	10	67			
東京	468	222	118	56	11	95			
富山	475	224	119	42	10	64			
兵庫	470	199	118	46	10	84			
愛媛	473	198	117	48	10	57			
熊本	476	205	120	40	9	68			
大分	475	201	120	46	10	66			
沖縄	481	215	111	38	12	65			
都道府県	目的変数	産業の特徴量							
exemplars(重心)	睡眠	従業者数. 民営..農業.林業	従業者数. 民営..漁業	従業者数. 民営..建設業	従業者数. 民営..運輸業.郵便業	従業者数. 民営..医療.福祉	第1次産業就業者数	第2次産業就業者数	
群馬	475	0.00351502	5.72E-05	0.03033196	0.02487997	0.05834687	0.02190902	0.1484841	
東京	468	0.00025321	5.69E-06	0.03251261	0.03147222	0.05711256	0.00156582	0.06216232	
富山	475	0.00619338	0.00070931	0.03855765	0.02269587	0.05612989	0.01491186	0.16630622	
兵庫	470	0.00122708	9.24E-05	0.02015315	0.0239193	0.0588331	0.00796615	0.10497489	
愛媛	473	0.00334796	0.00103533	0.03148839	0.02493256	0.06754363	0.03061488	0.10505221	
熊本	476	0.00464764	0.00055744	0.02893745	0.01753206	0.0754363	0.04128629	0.09777651	
大分	475	0.0066014	0.00099835	0.0352324	0.02023131	0.07440926	0.02839965	0.10476113	
沖縄	481	0.00222831	0.00014378	0.02728487	0.0187178	0.06396339	0.01585507	0.05407433	

表 3 代表クラスと所属する県

cluster1	cluster2	cluster3	cluster4	cluster5	cluster6	cluster7	cluster8
群馬	東京	富山	兵庫	愛媛	熊本	大分	沖縄
宮城		石川	埼玉	和歌山	鳥取	北海道	
福島		福井	千葉	山口	徳島	青森	
茨城		長野	神奈川	香川	長崎	岩手	
栃木		島根	滋賀	福岡	宮崎	秋田	
山梨		山形	京都		鹿児島	新潟	
岐阜			大阪			高知	
静岡			奈良				
愛知							
三重							
岡山							
広島							
佐賀							

図 2 は t-SNE によって表示された低次元マップを表しており、右の図は睡眠時間を把握しやすくするために都道府県名のサイズを睡眠時間に比例させて表示している。左の図は都道府県名の配置を確認するため全ての文字サイズを統一して同じ図を表示している。

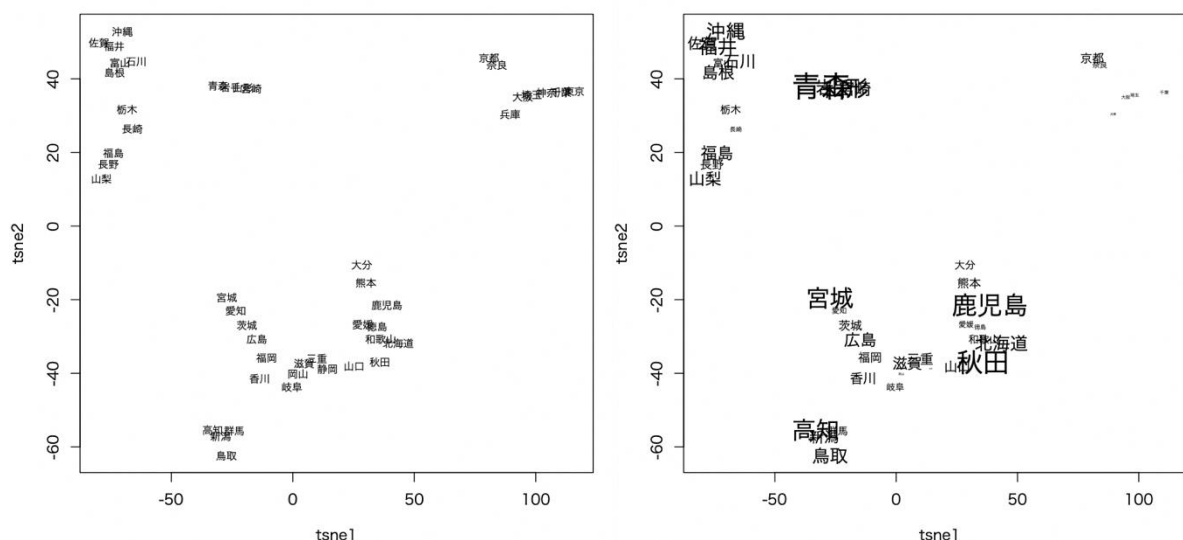


図 2 都道府県ごとの睡眠時間を表示させた低次元マップ

これらの結果から以下の内容が予想できる。

ランダムフォレスト回帰により変数重要度を算出した睡眠を除く 12 の変数のうち、睡眠との関係が認められそうな変数は、「休養・くつろぎ」「趣味・娯楽」「通勤・通学」「従業者数(民間)(農業, 林業)」「従業者数(民間)(建設業)」「業者数(民間)(運輸業, 郵便業)」「業者数(民間)(医療, 福祉)」「第 1 次産業従業者数」の 8 つの変数である。

表 2 より、生活様式と睡眠時間を比較すると、趣味・娯楽に費やす時間が多いと睡眠時間は少なく、逆に趣味、娯楽に費やす時間が少ないと、睡眠時間は多い傾向にあることが読み取れる。また、通勤通学の時間が多いと、睡眠時間は少ないことも読み取れる。一方で、仕事時間と睡眠時間には相関が認め

られないことが予想される。

産業構造と睡眠時間を比較すると、従業者数(民営)(農業、林業)が多いと睡眠時間は少なくはないことが予想できる。従業者数(民営)(建設業)が多いと睡眠時間は少ない傾向、従業者数(民営)(運輸業、郵便業)が多いと睡眠時間は少ない傾向にあり、逆に少ない場合は睡眠時間が長い傾向、従業者数(民営)(医療、福祉)が多いと睡眠時間も長く、少ないと睡眠時間も少ない傾向にある。また第1次産業就業者数が少ないと、睡眠時間も少ない傾向にあると予想できる。

5. 結果の解釈

生活様式と睡眠時間を比較した結果、睡眠時間は仕事に使う時間そのものに影響はないことが示唆された。一方で、通勤通学の時間が多いと、睡眠時間は少ないことが予想できた。また、趣味・娯楽に使う時間が多いほど睡眠時間は短く、趣味・娯楽の時間が短いほど睡眠時間が長い傾向にあることが示唆された。これらの結果から、生活様式において睡眠不足を補うための解決策に、通勤通学の時間を短くすることが考えられる。趣味、娯楽に関しては、睡眠時間とトレードオフの関係にあることから、ワークライフバランスにおいてどちらを優先すべきか検討する必要があると考える。

産業構造と睡眠時間を比較した結果、従業者数(民営)(農業、林業)が多いと睡眠時間は少なくはないことが予想できた。また、第1次産業就業者数が少ないと、睡眠時間も少ない傾向にあることが示唆された。これら二つの変数は日中、日光にあたる時間があり、体内のリズムが整うことで睡眠時間と正の相関があると予想する。仕事の時間そのものではなく、仕事内容が睡眠時間に影響を及ぼしていることが、これらの結果から予想できる。また、従業者数(民営)(建設業、運輸業、郵便業)が多いと、睡眠時間は少ない傾向にある。特に運輸業及び、郵便業は24時間運営が必要な産業であり、労働者は夜間勤務や長時間労働が一般的である。これにより、睡眠時間が不足し、労働者の健康問題や安全上のリスクが増加していることは明白である。よってこれらの産業に対して、労働環境の改善、労働時間の適正化、睡眠意識の向上などのアプローチを検討すべきである。具体的な策として、シフトの適切な調整や、過度な労働時間の削減、テクノロジーの導入により、労働者の睡眠管理支援などが考えられる。これらの提案を実施するためには、産業関係者、労働組合、政府などの協力が必要である。労働者の健康と安全を重視する姿勢を持ち、持続的な改善を図る体制を整備することが重要である。

本研究の試みによって、日本における産業や生活様式が睡眠に及ぼす影響について示唆を与えることができた。また、ワークライフバランスの改善を目指す上で、産業や生活様式といった社会構造も視野に入れることの重要性を指摘することができた。テレワークの推進や男性の育休取得、フレックス制度の促進などの柔軟な労働環境の促進を目指す上で働き方の改善策を提示するための支援が期待できる。

また、今回は男性と女性を分けずに睡眠時間と産業構造についての分析を行なったが、男性の場合は労働時間と生活時間でワークライフバランスについて考えることができるが、女性の場合は家事や育児などを主に行う人が多くいるため、性別ごとに分けて分析を行うことでより詳細な分析を行うことができると考える。

参考文献

- [1] 内閣府ホームページ「仕事と生活の調和」推進サイト ワーク・ライフ・バランスの実現に向けて仕事と生活の調和(ワーク・ライフ・バランス)憲章
- [2] 職場の同僚に着目したワークライフバランス支援制度の利用促進に関連する要因の検討ー地方自治体における実証分析ー
細見 正樹 (大阪大学大学院), 関口 倫紀 (大阪大学教授)
- [3] 勤務間インターバルと睡眠時間の組み合わせと職業性ストレス簡易調査票による高ストレス判定及び病欠欠勤の関連ー日本の日勤労働者を対象としたWEB横断調査ー
池田 大樹, 久保 智英, 井澤 修平, 中村 菜々子, 吉川 徹, 赤松 利恵
- [4] 特集:睡眠と健康ーライフステージとライフスタイルー. 睡眠と健康:成人期(勤労者)保健医療科学. 高橋正也. (2015). 2015 Vol. 64 No. 1 p. 23 独立行政法人労働安全衛生総合研究所作業条件適応研究グループ