

2015年6月17日
株式会社アイスファクトリー

オープンデータの可能性

医療費削減のヒントをオープンデータから考察

休養よりも自己啓発やアクティブな活動が健康に効果的！？

地方創生のカギは、地域特性に合った無理のない施策づくりに

15日、全国の病院ベッド数20万床の削減が政府調査会から発表された。2025年には団塊の世代が75歳以上となる「超高齢社会」を迎え、医療や介護のニーズも今後ますます高まっていくとみられる。それに伴い、社会保障給付は今後も急激な増加が見込まれ、税や社会保険料などの国民負担の増加は大きな課題である。

特に医療・介護分野における給付は、消費税収や現役世代の負担の伸びを上回って増加する見込みとなっており、財源の確保と同時に、国民一人一人の健康増進による医療費削減にも期待したい。

医療費の多い都道府県、医療費の抑制に成功している都道府県の特徴把握と、医療費を削減するために有効な示唆の発見を目的に、株式会社アイスファクトリー（東京都千代田区、代表取締役：大場智康、以下「アイスファクトリー」）では、サービス提供しているオープンデータ用自動解析ツール bodais を用いて各地域の医療費に関する分析を試みた。

■結果概略

今回、病気・食・住・労働・メディアに関するデータの分析と、時間の使い方に関するデータの分析の2種類の分析を行った。その結果は下記の通り。

1. 地域特性に合った、無理のない医療費削減施策の可能性発見

労働や食生活・健康状態、メディアとの関わりなどにおいて、特徴が似ている都道府県を明らかにすることによって、同じ特徴をもつ都道府県間で医療費を少なく抑えられている地域を参考に、医療費削減の対策を検討することができる可能性がある。

2. 知的好奇心とアクティブな活動が、医療費削減に影響！？

仕事以外の時間を、非活動的な「休養・くつろぎ」に充てる都道府県よりも、「子供などとのレジャー」や「読書・自己啓発」などの活動を行う都道府県の方が医療費が抑制されている傾向がみられた。

分析には、各地方公共団体や業界団体より発表されているオープンデータを用いた。

■利用した公開データ

以下の 164 個の変数

1. 病気

厚生労働省 特定健康診査・特定保健指導に関するデータより

平成 24 年度 特定健康診査（対象者数・受診者数・実施率）

平成 24 年度 特定保健指導（対象者数・終了者数・実施率）

平成 24 年度 メタボリックシンドローム（該当者数・予備群者数・割合）

2. 食

厚生労働省 国民健康・栄養調査より

平成 22 年 都道府県別の肥満及び主な生活習慣の状況

3. 住

総務省統計局 第 2 章 人口・世帯より

平成 22 年 都道府県別人口と人口増減率

平成 25 年 都道府県, 年齢階級別人口

平成 22 年 都道府県, 世帯人員別一般世帯数と世帯の種類別世帯人員

平成 22 年 都道府県別昼間人口と 15 歳以上の自宅外就業・通学者数

平成 25 年 都道府県別転出入者数

平成 25 年 都道府県別出生・死亡数と婚姻・離婚件数

4. 労働

総務省統計局 労働力調査（基本集計）都道府県別結果より

平成 26 年 完全失業率（年平均）

総務省統計局 第16章 労働・賃金より

平成22年 都道府県, 産業別就業者数

総務省統計局 第22章 教育より

平成25年 都道府県別進学率と就職率

平成23年度 都道府県別地方教育費

5. メディア

総務省統計局 通信利用動向調査より

平成25年 情報通信機器の保有状況

平成25年 パソコンの保有状況

平成25年 携帯電話（PHS含む）の保有状況

平成25年 スマートフォンの保有状況

平成25年 デジタルテレビ放送受信機器の保有状況

6. 生活

総務省統計局 社会生活基本調査より

平成23年 男女, スポーツの種類別平均行動日数

平成23年 男女, 趣味・娯楽の種類別平均行動日数

平成23年 男女, 学習・自己啓発・訓練の種類別平均行動日数

平成23年 男女, 旅行・行楽の種類, 頻度別行動者率

平成23年 曜日, 男女, 行動の種類別総平均時間, 行動者平均時間及び行動者率(15歳以上)(週全体)

7. 医療費

厚生労働省 医療費の地域差分析より

平成24年度 医療費の地域差

■分析概要

1. 病気・食・住・労働・メディアに関するデータの分析

bodais クラスタリングを用いて、病気・食・住・労働・メディアに関するデータから各都道府県のカテゴリ（教師なしクラスタリング*¹）を行った。分類されたそれぞれのクラスターに含まれる都道府県を、1人当たりの医療費の金額に応じて5段階に分類し考察した。

【病気・食・住・労働・メディアに関するデータの分析 結果】

クラスター名	クラスターの特徴	1人あたりの医療費				
		少ない	やや少ない	普通	やや多い	多い
(1) 総労働時間の少ない、都市圏・独身型	<ul style="list-style-type: none"> 自治体が教育費をかけている 総実労働時間数が少ない 1人世帯の割合が多い 高卒就職率が低い 第3次産業割合が高い 	埼玉 千葉 東京	神奈川	京都 大阪 兵庫	北海道 福岡	広島
(2) マイホームパパ型	<ul style="list-style-type: none"> デジタル放送チューナー保有率が高い 4人世帯の割合が高い 肥満者の割合が低い 飲酒習慣者の割合が低い 	茨城 栃木 群馬 愛知	山梨 長野 岐阜 静岡 滋賀 奈良	宮城 福井 三重	石川 岡山	香川
(3) 九州型	<ul style="list-style-type: none"> メタボ予備軍が多い メタボ対策を受けている人が多い パソコン保有率が低い スマホ保有率が低い ワンセグ放送が受信できる機器保有率が低い 	沖縄	-	宮崎	熊本	長崎 佐賀 鹿児島
(4) 大家族型	<ul style="list-style-type: none"> 5人以上世帯の割合が高い 総実労働時間が長い 離婚率が低い 出生率が低い 野菜摂取量が多い 	-	福島	岩手 山形 新潟 富山	-	-
(5) 過疎傾向型	<ul style="list-style-type: none"> 20~24歳の割合が低い 45~54歳の割合が低い 所定外労働時間数が少ない タブレット型端末の保有率が低い 	-	和歌山 青森	-	鳥取 愛媛 高知 秋田 徳島	山口 島根 大分

(1) 総労働時間の少ない、都市圏・独身型

独身一人暮らしが多く、教育費も多いことから子供も多いとみられる。労働時間が比較的短く、都市圏が多く含まれる。

医療費の少ない都道府県が多いクラスターであるため、医療費の多い傾向のある広島・北海道・福岡は、ワークライフバランスの改善が医療費の抑制に有効である可能性がある。

(2) マイホームパパ型

デジタル放送チューナーの保有が多く、4人世帯が多い。飲酒習慣者が少なく肥満者も少ないことから、健康的な生活習慣を持つクラスターといえる。外食や付き合いは程々にし、デジタル放送を楽しむなど家庭での時間を大切にしている可能性がある。

医療費が少ない地域が多く含まれており、医療費の比較的多い香川・石川・岡山は家庭での時間を意識して設けることで医療費の抑制ができる可能性がある。

(3) 九州型

このクラスターには、九州の多くの地域が含まれる。特徴としては、パソコンやスマートフォンの保有が少なく、ワンセグ放送受信機器の保有も少ない。沖縄以外は総じて医療費が平均より高く、メタボ対策を受けている人も多いことから、肥満者が多い傾向にあると思われる。

地域住民への医療費削減のための施策実行の際は、パソコンやスマートフォンの保有の少ない地域特性を踏まえて、伝わりやすい告知手段を考える必要があるかもしれない。

インターネット環境を活用した施策は、告知の即時性もあり双方向のコミュニケーションが可能になるという利点もあるので、インターネットの活用に興味を持ってもらえるよう働きかけることで、住民の生活に変化をもたらす医療費の改善が期待できる可能性がある。

(4) 大家族型

このクラスターの家族構成は5人以上の世帯が比較的多く、総実労働時間が長い。離婚率の低さに反して出生率は少ないクラスターであるが、データを確認すると他のクラスターに比べて若い人の割合が少ない。若い人が少なければ直接的に医療費を押し上げる要因になるが、このクラスターの地域の医療費はほぼ平均程度であることから、家族が多いことや野菜の摂取量が多い点は医療費を抑制する要因になっていると考えられる。

このクラスターの地域が医療費を抑制するためには、労働時間を短くするなど働き方の面から生活を見直すことも、効果的であると考えられる。

(5) 過疎傾向型

このクラスターでは、20～24歳、45～54歳といった若者や働き盛りの年代が他のクラスターよりも少なく、所定外労働時間も少ない。医療費の高い都道府県が多いと同時に、過疎化が進む地域の割合も多い。過疎化が進み高齢化が進む傾向にあるため、単に医療費がかさんでいる可能性が高い。

医療費削減のヒントとしては、地域産業の活性化や新たな魅力発掘など、過疎化に対する対策が医療費削減にも寄与すると考えられる。

2. 時間の使い方に関するデータの分析

bodais クラスタリングを用いて、時間の使い方に関するデータから各都道府県のカテゴリ（教師なしクラスタリング）を行った。分類されたそれぞれのクラスターに含まれる都道府県を、1人当たりの医療費の金額に応じて5段階に分類し考察した。

今回使用したデータだけでは、クラスターの特徴が明確に差別化できない場合もあったため、医療費との関係が伺える一部のクラスターから知見を得た。

【時間の使い方に関するデータの分析 結果】

クラスター	クラスターの特徴	1人あたりの医療費				
		少ない	やや少ない	普通	やや多い	多い
クラスター1	<ul style="list-style-type: none"> 詩・和歌・俳句・小説などの創作の平均行動日数が少ない 家政・家事の平均行動日数が少ない 生涯学習においてパソコンなどの情報処理を行う平均行動日数が少ない 休養・くつろぎに使う時間が多い スポーツの総平均時間が少ない 卓球に使う日数が多い 	-	福島	山形 新潟 富山 福井	秋田 鳥取	島根 山口 佐賀
クラスター2	<ul style="list-style-type: none"> 趣味娯楽における園芸・庭弄り・ガーデニングに使う平均行動日数が少ない 仕事に使う平均行動日数が多い 	茨城 栃木 群馬	山梨 長野 岐阜	岩手	岡山	-
クラスター3	<ul style="list-style-type: none"> 学習・自己啓発や、読書、遊園地や動物園に行く時間が多い 野球や柔道、日曜大工に使う時間は少ないが、ボウリングを行う人は多い 	埼玉 千葉 東京都	神奈川 奈良	京都 大阪	-	-
クラスター4	<ul style="list-style-type: none"> 水泳や編み物、洋舞・ダンス、邦舞・社交ダンスなどに使う時間が多い 将棋や和裁洋裁に使う時間が多い 	-	-	宮城	北海道 福岡	香川
クラスター5	<ul style="list-style-type: none"> 和裁・洋裁、コーラス・声楽、商業実務・ビジネス関係に使う時間が少ない 通勤通学を除いた移動に時間を使う人が多い 	愛知	滋賀	三重 兵庫	石川	広島
クラスター6	<ul style="list-style-type: none"> ゴルフ、英語、スポーツ観戦、サッカーに使う時間が多い 介護関係、受診・療養、家政・家事に使う時間が多い 	沖縄	-	-	熊本	長崎 鹿児島
クラスター7	<ul style="list-style-type: none"> テレビ・ラジオやTVゲームに使う時間が多い つり、ウォーキング、サイクリングに使う時間が多い 家事に使う時間が多い 	-	青森 静岡 和歌山	宮崎	徳島 愛媛 高知	大分

時間の使い方で、医療費との関係が見られたのはクラスター1とクラスター3であった。

(1) クラスター1

このクラスターは、労働時間が平均的であるのに対し、余暇の過ごし方はそれほど活動的ではない。スポーツを行う人が少なく、休みは休養・くつろぎに充てる傾向が見られた。

また、医療費の多い地域が多く含まれている。

(2) クラスター3

このクラスターは、趣味の読書や学習といった自己啓発や、おそらく子供と一緒にみられるレジャー（遊園地・動物園）に割く時間が多い。また、食事にかかる時間も多くとれている。クラスター3に含まれる全ての都道府県が、医療費が平均程度もしくは平均以下であることから、家族と過ごす時間を大切にしたり、自己投資を意識して行うことが健康を促進し医療費の削減にも関係がある可能性がある。

■結論

今回オープンデータのクラスタリングによって、医療費削減のヒントとなる様々な仮説が得られた。行政や企業による施策検討の際に、こういった仮説を確かなものにすることで、有効な施策が見えてくる。統計解析の技術を用いることで、利用可能な様々なデータから新たな施策が検討できる可能性があり、今後様々な社会課題での活用が期待できる。

今回のケースでは、クラスタリングの結果に医療費データを紐づけて考察することで、医療費削減の余地のある都道府県は、地域特性の似た都道府県から医療費削減のヒントを得ることができた。より実効性の高い施策を検討するために、オープンデータを活用した解析がスピーディーで有効である。

■オープンデータの背景

世界各国でオープンデータの試みが加速する中、国の単位のほかに、州・県レベルでの取組が着目され始めている。現在は渋滞情報やバス運行等のGISデータをアプリに取込む等、データをそのまま利活用する動きがある。

今後は、活用できるデータの選別をすることなく、可能性のあるデータを全て利用し相関や因果を確認し、利活用できるデータを絞り込むことで、より効率よく活用できる可能性がある。データからの発見が期待されるデータマイニングは、分析前のデータ項目選定、データのクレンジング処理過程に莫大な時間がかかる。この処理スピードアップがPDCAサイクルに直結してくる。

オープンデータを活用した地域のイノベーションでは、分析前の処理スピードアップが重要なファクターとなり、新たなデータ項目の追加なども積極的に行うことが肝となる。

■アイズファクトリーについて

アイズファクトリーは設立以来、数値データ、テキストデータの解析とそれを応用できるシステム開発に特化した解析専門企業です。各種データを客観的に解析できることに評価をいただき先端データ解析に取り組まれる企業をはじめ、新たにデータ解析に取り組まれる企業へ人材育成、解析運用、解析ツール提供など、ビッグデータ活用をワンストップでサポートする会社です。

*1 教師なしクラスタリングとは、『セグメント』や『ランク』など既存の分類がついていないデータを、そのデータの持つ特徴に応じて分類する分析です。

例えば各都道府県に、人口や年代構成比、労働状況などのデータを紐づけた上で教師なしクラスタリングを行うと、使用したデータにおいて類似した都道府県（人口が多い、30代が多い、正社員が少ない…など）がまとめられ、分類されます。

<p><input type="checkbox"/> 会社概要</p> <p>株式会社アイズファクトリー</p> <p>設立：2000年4月25日</p> <p>代表者：代表取締役 大場智康（理学博士）</p> <p>資本金：156,171千円</p>	<p><input type="checkbox"/> 本件に関するお問い合わせ</p> <p>TEL：03-5259-9004</p> <p>E-mail：mrk@isfactory.co.jp</p> <p>担当：マーケティング部 生田</p> <p>http://bodais.jp/company/</p>
---	--