

2020年3月25日 全5頁

AI と検索数を用いた新型コロナウイルスによる品薄問題の分析

グループごとの政策対応が求められる

経済調査部 エコノミスト 新田堯之

[要約]

- 新型コロナウイルス感染拡大に伴い、様々な品目が品薄状態に陥っている。こうした需給の逼迫状態をより広範囲から捉え、政策当局などが対応を検討する上で役立つ情報を探るべく、様々なキーワードの検索件数の活用を検討した。しかし、個々の品目のデータの特徴を一瞥しただけで、直感的に理解するのは困難であろう。そこで、複雑な状況から本質的な構造を抽出すべく、データをいくつかのまとまりに分類した上で、それぞれのグループの特徴を把握しようとした。分類の際には、高い精度を確保しつつ、恣意性を可能な限り排除するためにAI（機械学習）を活用した。
- その結果、①品切れが慢性的に発生している品目群、②品薄状態がより深刻な品目群、③一時的に需要が急増した品目群、④新型コロナウイルスの影響を比較的受けなかった品目群、と考えられる4グループに分類できた。政策当局には、①や②の商品群には、メーカーに向けた生産増の呼びかけ、③のような商品群には、発端がデマか否かを探り、前者の場合は人々やメーカー、小売店などに注意喚起を行うなどの対応が考えられる。他方、④に属する商品群に対しては、差し迫った対応をすべき必要性は低いだろう。
- もちろん使用上の留意点はあるものの、AI と検索数データの組み合わせは、人々が食料品や日用品などを購入できない「不安の度合い」を品目レベルで可視化し、政策当局等がより迅速に適切な対応を行う上で強力なツールになり得ると期待できる。

1. 検索データで品薄状態の全体像を把握する

新型コロナウイルス感染拡大に伴い、買いだめを含む消費者の需要が急増したことを受け、マスクをはじめ様々な品目が品薄状態に陥っている。また、足元では緩和・解消したものの、過去に品薄問題が発生した品目も存在する。

Google Trends の検索データを用いた品薄状態の把握とそのメリット

こうした需給の逼迫状態をより広範囲から捉え、政策当局などが対応を検討する上で役立つ情報はあるのだろうか。こうした問題意識の下、本レポートでは、様々なキーワードの検索件数の活用を検討した。その上で、日本における検索エンジン市場の約 75%¹を占める Google の検索データに基づき、Google Trends のデータを利用した。

検索データを用いた主な動機は、求める商品が入手できない場合、人々はインターネットを通じ、オンラインストアからの注文や入荷情報の収集等を試みられると思われるためである。さらに、ある程度の速報性を兼ね備えている（本レポートの執筆時点では3日前までのデータが取得できた）ことも、本データを用いる利点の一つであろう。今回の分析に際しては、各種メディアで品薄が指摘されたマスク、トイレットペーパーなど10品目の食料品および日用品の日本全国における検索数を取得した（日次ベース、期間は2020年1月1日～3月20日）。また、比較対象として、需給のバランスが安定的に推移したとみられる歯ブラシと髭剃りの検索数も取得した（地域、頻度、期間は同様）²。

2. AI と検索数データを活用し、12 品目を 4 グループに分類

データの特徴ごとに分類する際の難しさ

これらの品目のデータが持つ特徴を一度に理解するのは困難であろう。例えば、図表1の上図（次ページ）は、今回取り上げる12品目の検索数の推移を示したものだが、それぞれの品目が人間の目には複雑な変動パターンを示しているように見え、一瞥してその特徴を掴むのは容易ではない。さらに、各データの特徴を人手により無理に分類しようとすれば恣意性が入る余地もある。そのため、一見するとここからデータの傾向を掴み、意思決定に資する正確な情報を得るのは難しそうである。

AI を活用した分類のメリット

そこで、複雑な状況から本質的な構造を抽出すべく、客観的な基準を用いてデータをいくつか

¹ Statcounter GlobalStats ウェブサイトに基づく2020年2月時点の数値（パソコン・スマートフォン・タブレットのいずれの端末を含む）。

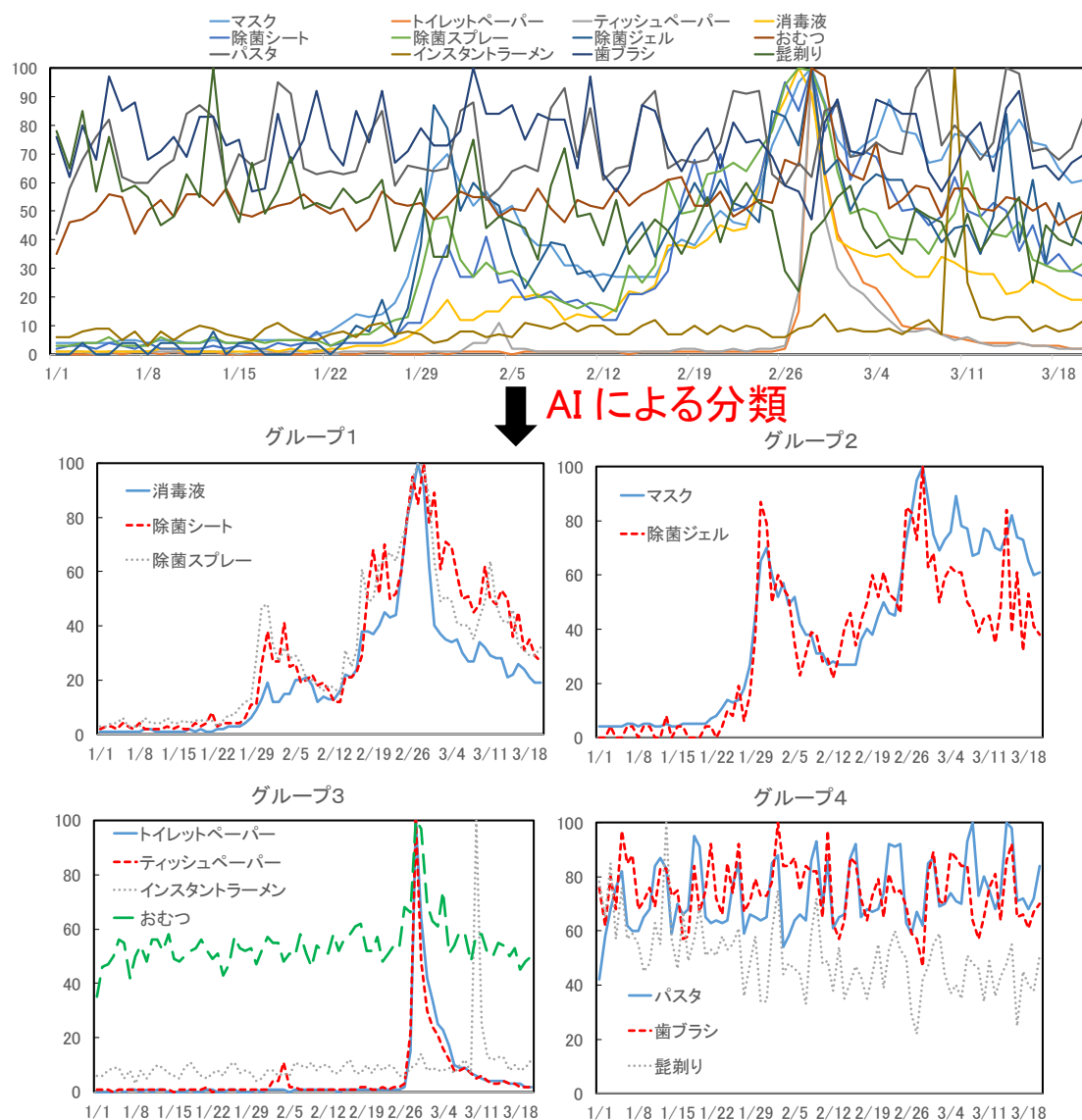
(URL : <https://gs.statcounter.com/search-engine-market-share/all/japan>)

² なお、「マスク 不足」、「マスク 品切れ」などより品薄をより想起させる検索語句も検討したが、①全体的にデータの振れが激しくなり、傾向を掴みにくくなった、②検索語句によってはデータが取得できなくなった（検索数が一定ラインを下回ったためと思われる。）、③より限定された地域で分析しようとする場合、②の問題がさらに顕在化する可能性が高い、ことなどを受け、品目名のみを検索対象とした。

のまとまりに分類した上で、それぞれのグループの特徴を把握しようとした。

ここでは分類の際に、高い精度を確保しつつ、恣意性を可能な限り排除するために AI（機械学習）を活用した。具体的には、教師なし学習の一種であり、かつ時系列データを分類（クラスタリング）する機械学習モデルの一つである k-Shape と呼ばれる手法³を用いた。分類されるグループ（クラスター）数は、エルボー法の結果に基づき四つと設定した。図表 1 の下図は、この手法で 12 品目を 4 グループに分類し、グループごとに各品目の検索数を示したものである。各グループは何らかの共通した特徴ごとに分類されたものと解釈できる。

図表 1 各品目をキーワードとした検索数の推移
(上図：AI による分類なし 下図：AI が分類したグループ（クラスター）ごと)



(注 1) 特定のキーワードの検索件数が最も多かった時期を 100 として、人気度の推移を確認したもの。

(注 2) 分析期間は 2020 年 1 月 1 日～3 月 20 日

(出所) Google Trends より大和総研作成

³ John Paparrizos and Luis Gravano (2015) “k-Shape: Efficient and Accurate Clustering of Time Series” SIGMOD 2015, pp. 1855-1870.

(URL : <http://www1.cs.columbia.edu/~jopa/Papers/PaparrizosSIGMOD2015.pdf>)

3. 各グループの特徴

以下では、AIによって分類された四つのグループごとに特徴を説明する。

品切れを示唆する2つのグループ

まず、消毒液や除菌シートなどが含まれるグループ1は、品切れが慢性的に発生している商品群だと思われる。データの動きを見ると、1月末頃に盛り上がり、その後はいったん落ち着いたものの、2月の後半には感染者数の増加ペースと合わせるように一気に加速した。3月に入るとやや落ち着きを取り戻したが、依然としてやや高い水準である。

グループ2はマスクと除菌ジェルが属する。基本的に、両者の検索数の動きはグループ1のそれと類似する。それにもかかわらず、AIが別のグループに分類した理由の一つは、グループ1に属する品目と比較して、3月以降も高止まりが続いているためであろう。すなわち、AIから見ると、マスクと除菌ジェルの品薄は頭一つ抜けて深刻な状況のようである。

一時的に需要が急増した商品のグループ

次に、グループ3は一時的に需要が急増したグループである。2月末頃に「紙製品が不足する」との情報が拡散したことをきっかけに、人々が小売店にトイレットペーパーやティッシュペーパーなどを我先に買い求めたことは記憶に新しい。この動きに合わせて、検索数は一時的に急増したが、その後は落ち着きを取り戻している。他方、インスタントラーメンの検索数が3月10日に一時的かつ急激に増加した主因は、品薄ではなく、同日にとあるテレビ番組でインスタント袋麺が特集されたためであろう。仮に、このイベントがなければ、インスタントラーメンはグループ4に分類された可能性が高い。

新型コロナウイルスの影響を比較的受けなかった商品のグループ

最後に、グループ4は新型コロナウイルスの影響を比較的受けなかった商品群だと思われる。もちろん、個々の店舗レベルや商品レベルでは品薄状態のケースも存在すると思われるものの、他のグループと比べると、小売店が保有する在庫は比較的豊富であり、消費者が入手しやすい商品群だと思われる。あるいは、品薄状態ではあるものの、入手手段をインターネットで検索するほどまで困っていない商品群と解釈できると思われる。

4. 各グループに求められる対応策

上記の結果を踏まえると、政策当局には、グループ1やグループ2のような品切れを示唆する商品群には、メーカーに向けた生産増の呼びかけ、グループ3のような需要が急増した商品群には、発端がデマか否かを探り、前者の場合は人々やメーカー、小売店などに注意喚起を行う利用方法が考えられる。他方、グループ4に属する商品群に対しては、差し迫った対応をすべき必要性は低いだろう。

AIと検索数データの組み合わせは「不安の度合い」を可視化する強力なツールに

このように、AIと検索数データの組み合わせは、人々が食料品や日用品などを購入できない

「不安の度合い」を品目レベルで可視化し、政策当局等がより迅速に適切な対応を行う強力なツールになり得るだろう。また、検索エンジン次第では地域単位の情報をも把握できる。例えば、本レポートで用いた Google Trends は一定以上のデータがあれば、都道府県単位のデータを取得できる。これをうまく活用すれば、地域ごとの実情を踏まえた、よりきめ細やかな施策を打てる可能性がある。

より正確な実態把握には関連キーワードの確認も必要

他方、最も注意すべき点の一つは、上記のインスタントラーメンのように、新商品の発売やメディア露出など、人々の不安とは関係のない部分でデータが大きく振れるリスクである。このような問題への対応策の一つとして、関連キーワードの確認が挙げられる。例えば、品薄が深刻なマスクを検索すると、「入荷予定」、「在庫」、「手作り」などの関連キーワードが得られる。一方、インスタントラーメンの関連キーワードは調理法や商品名が多く、品薄と関係がありそうなキーワードは発見できなかった。したがって、より正確な実態を捉えるためには、時には検索数データと関連キーワードの双方に目を配る必要があるだろう。加えて、本レポートで扱った時系列データの分類は、分析期間が変わればその結果も変わり得る⁴点も留意すべきである。

最後に、新型コロナウイルスは人々の不安を増幅させ、時にはパニックを引き起こす存在である。こうした時だからこそ、検索データを含めた幅広いデータとテクノロジーの活用を通じ、早い段階で現実を理解した上で、的確な判断を下す姿勢が求められるのではないかと。

以上

⁴ 実際、分析期間を本レポートで扱ったケースより1週間短い2020年1月1日～3月13日に設定したケースでは、除菌ジェルはグループ2ではなく、グループ1に属していた。除菌ジェルの検索数が3月14日に急増し、その後も高止まりが続いたことが両者の結果の違いを生み出したとみられる。