

背景

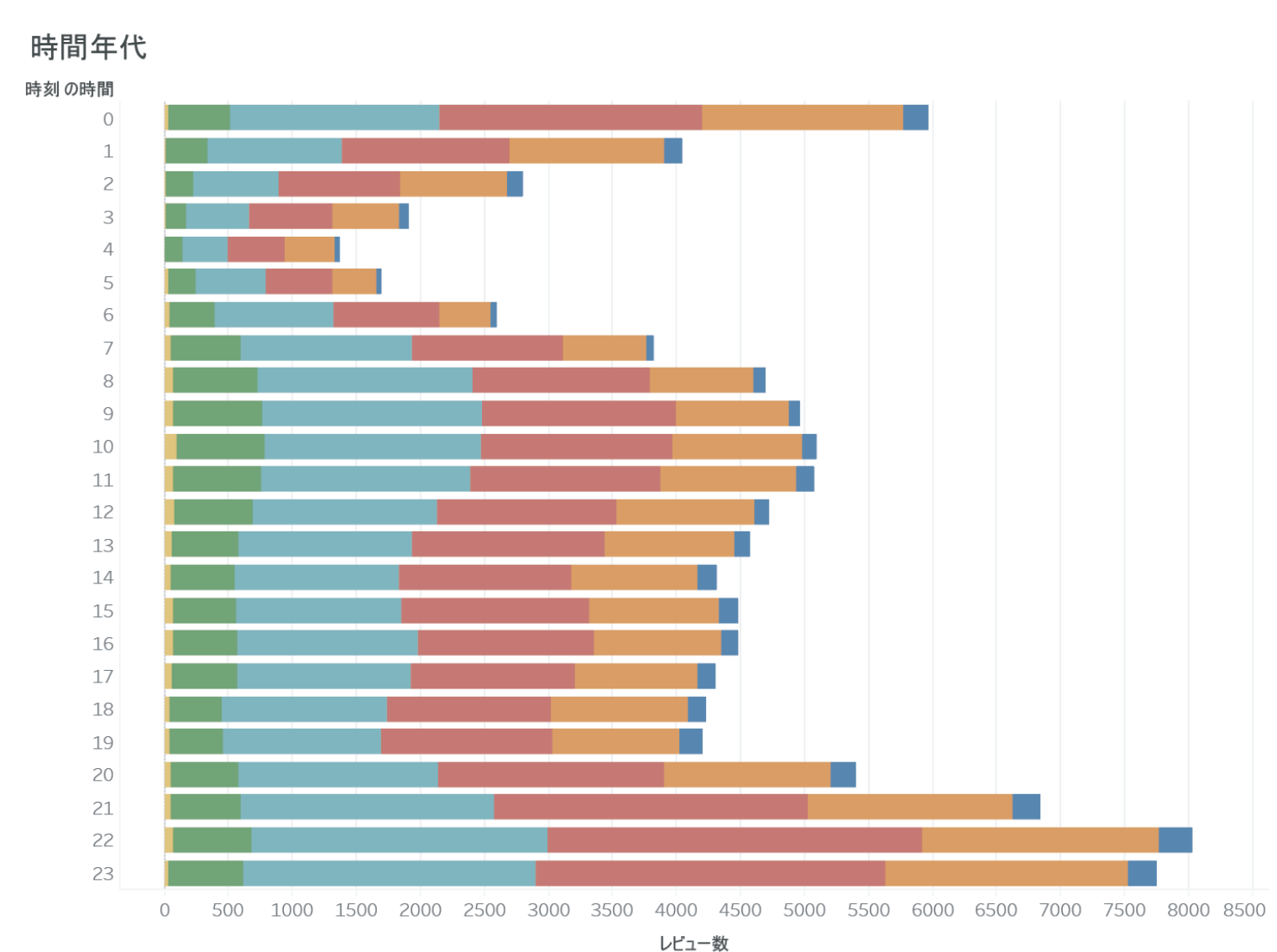
一般に EC サイトにおいては、ユーザーからの評価を受け付けるレビュー投稿によるフォーラムを備えている。今回の分析対象である「ダイエット口コミデータセット」の提供元である「ダイエットカフェ」を含め、それらの多くは星など数段階、あるいは点数などでの評価によるレーティング機能と、当該商品の感想を投稿できる口コミ機能を具備している。

ユーザーの商品およびブランドに対する感情は概ねこれらの口コミデータを分析することによって概ね導き出すことができると考えている。

個別の商品あるいは紐づけられたユーザーの年齢、性別、また当該データセットにおいては体重変化などの客観的なデータに応じてどのような変化が見られるか、特にその成否が社会的評価にもつながりうるが故に正直かつ感情的な文言が見られるであろうダイエットのデータにより探索しようと考えた。

分析

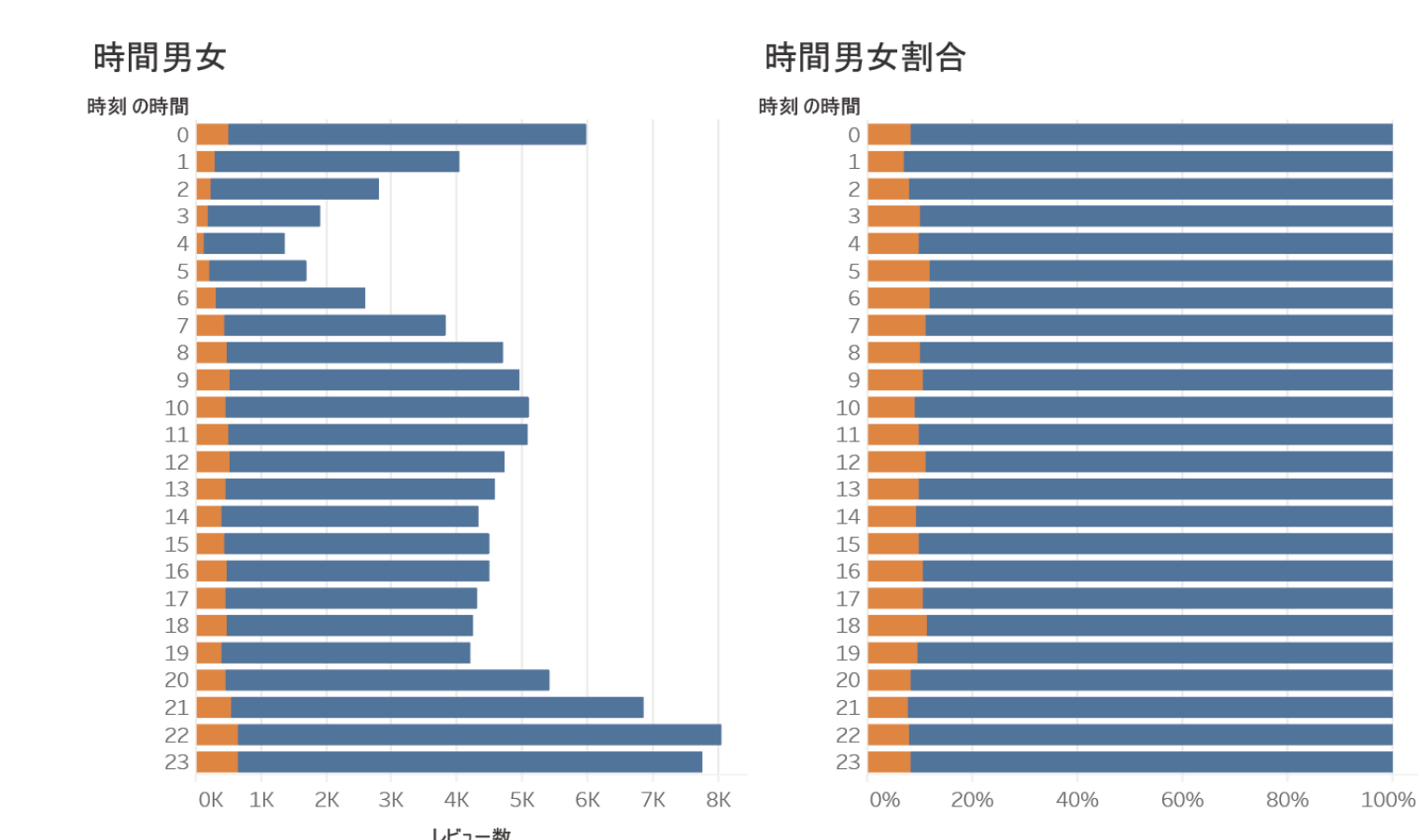
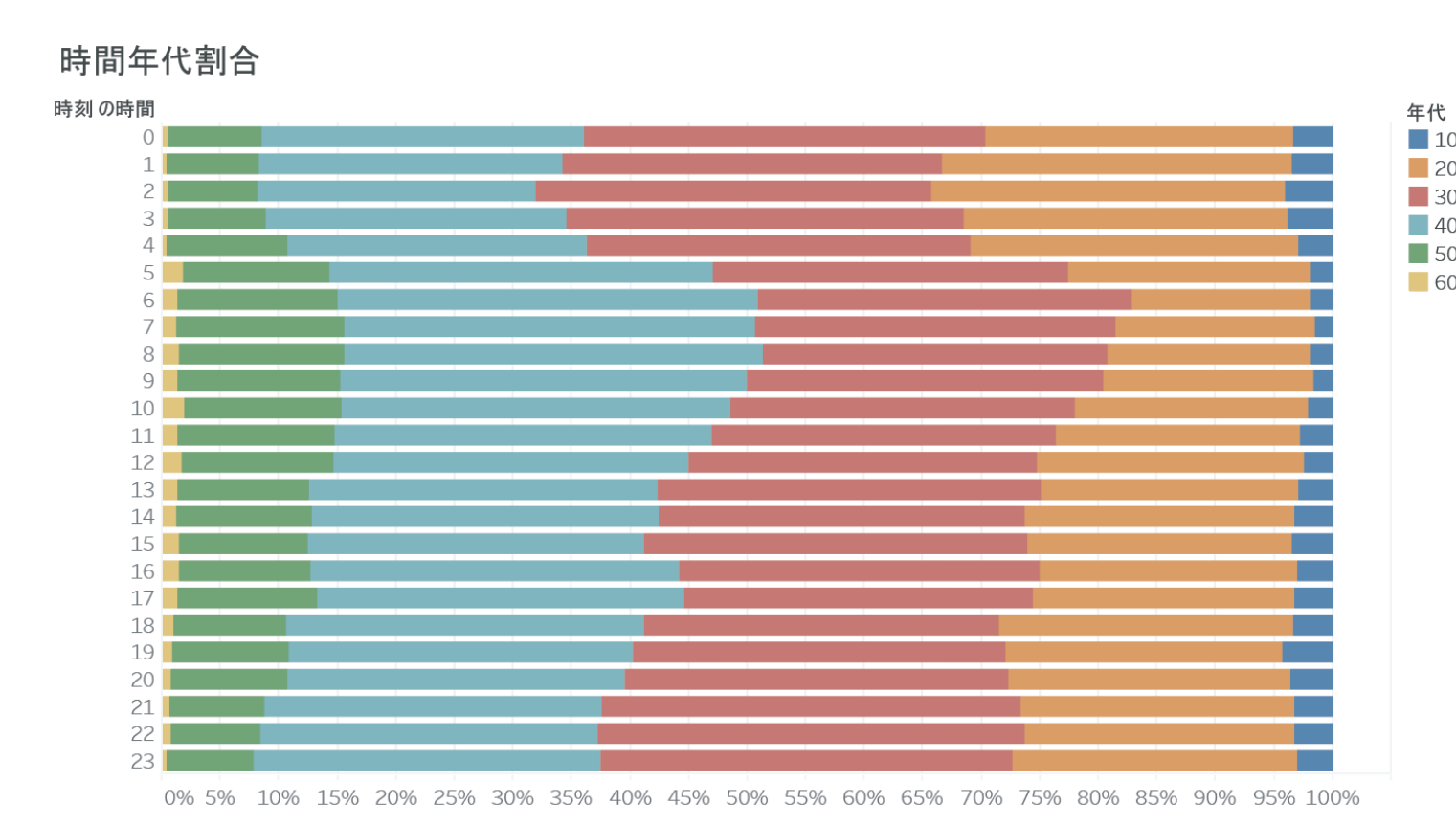
データの可視化には Tableau を使用した。



利用の時間帯と利用者の属性による分析

時間・利用者年齢層を軸にとって検証すると、午前 4 時が最も少なく、21~0 時が最も多い。また、日中 7~19 時においても 4000 件程度は集まっており、日中に私用で投稿などができる人々も多く利用していることがわかる。これを比率に落とし込むと主な利用者層は 20~40 代であり、また、全ての時間帯において女性の割合が圧倒的に多いことがわかる。

この結果から、ユーザーのセグメントとして、件数の一定割合を専業主婦あるいは夜勤者が占めている可能性がある。また、時間帯に関わらず女性の割合が多く、男性のダイエット商品に対する興味の薄さが伺える。



結果・今後の展望

ユーザーの時間帯別の属性からは、利用者層が日中にも相当数の投稿量があり、しかもその利用者は 20~40 代で、尚且つ女性が中心であることから、主たる利用者層は専業主婦である可能性がわかった。

また、投稿内容の対象について分析すると、特定時期の投稿数の変化が目立った商品も多く、口コミデータの内容の分析処理に加えて、その要因を行事等外部データも交えながら検証していきたい。

今回は数値化されたデータを中心に分析した結果を発表したが、自然言語処理を用いたセンチメント分析や、今回提示した分析結果の多変数化にも挑戦している。

目的

ターゲットに応じた商品ニーズおよび持ちうる感情の分析

利用データ

ダイエットカフェ株式会社様「ダイエット口コミデータベース」として、NII-IDR にて提供されたデータ。

提供データの内容としては次のとおり。

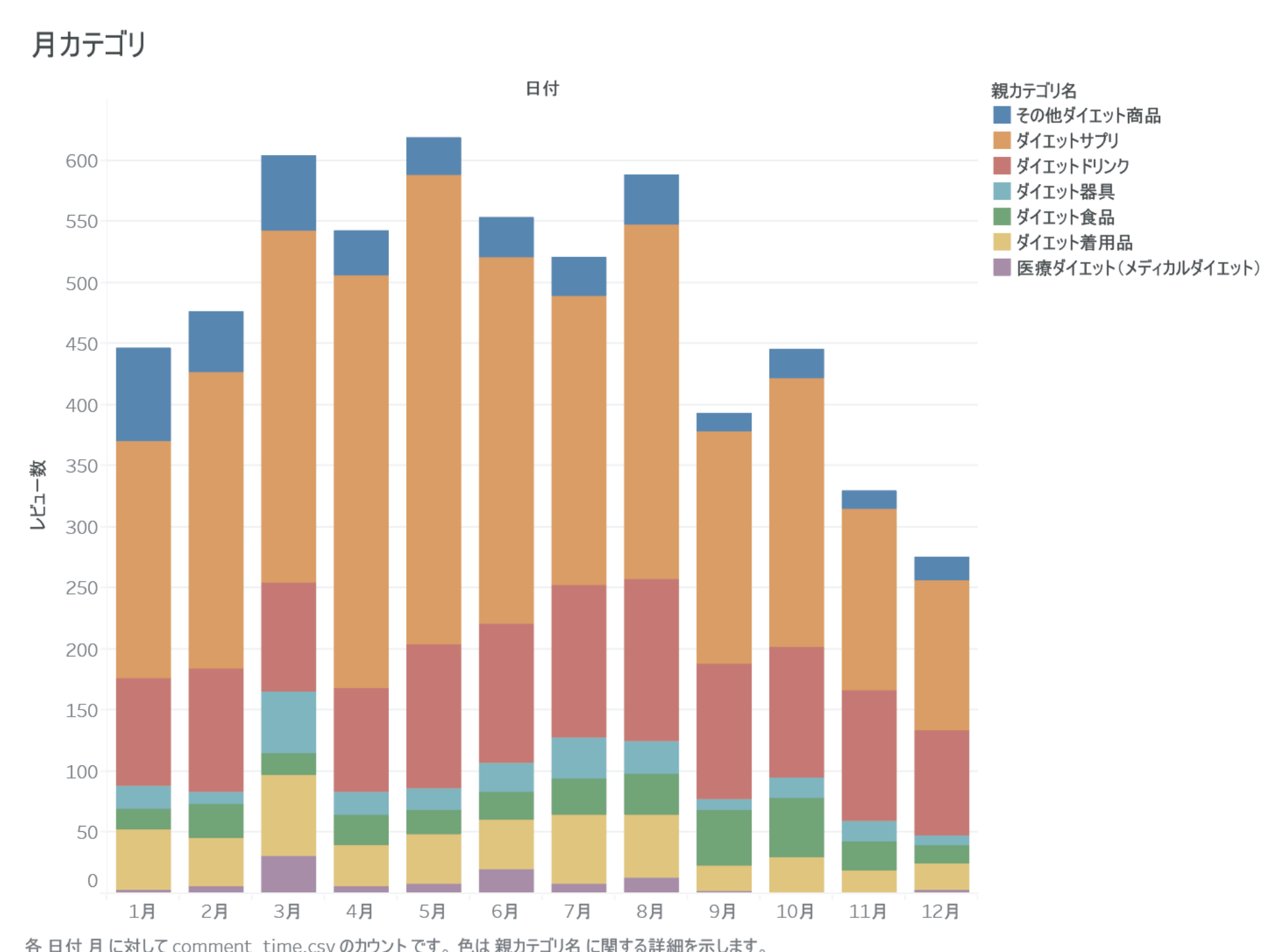
商品データ (識別子、商品名、カテゴリ、認証、成分、データ登録時期、評価点など口コミデータの簡易な集計)

口コミデータ (識別子、投稿者名、投稿日時、商品識別子、本文、評価、年齢、身長、使用前後の体重、使用期間、性別、タグ)

ダウンロードはこちら

<https://www.nii.ac.jp/dsc/idr/diet-review/>

投稿対象商品と時期の分析



月・カテゴリ別の投稿数について分析したこのグラフにおいては、

1. 投稿の多くをサプリメントについてのもので占めていること
2. 医療ダイエットやダイエット器具については特定の月に増減が激しいことが読み取れる。健康診断などのイベントが関連している可能性もある。

謝辞

データをご提供いただき、2023 年 11 月にはデータを大幅に追加していただきました。ダイエットカフェ株式会社に、深く感謝申し上げます。