

Hiddenによるデータ分析

食べログ お好み焼き百名店2019に関して

コンピュータグラフィックス 情報可視化

はじめに

グルメサイト・食べログにおいて、

「**お好み焼き百名店2019**」 に選ばれた
100店舗に対して情報可視化を行う。

百名店とは...

その名の通り、食べログの**総合点上位100店舗**のことである。料理ジャンルごとに100店舗が選出され、ラーメンやスイーツなどの部門がある。

その中で、個人的に一番好きな**お好み焼き**についてデータを分析することにした。

データ項目①

～お店の基本情報～

▶ A：都道府県

関東地方→1：東京、2：神奈川
近畿地方→6：大阪、7：兵庫、8：京都
四国地方→11：愛媛、12：徳島

中部地方→3：愛知、4：富山、5：石川
中国地方→9：広島、10：岡山

▶ B：最寄り駅からの距離(m)

▶ C：座席数

▶ D：昼の予算目安

▶ E：夜の予算目安

予算に関してはD,E共通して、以下の分類を用いた。この分類は食べログに準ずる。

0：なし、1：¥～999、2：¥1,000～1,999、3：¥2,000～2,999、4：¥3,000～3,999、
5：¥4,000～4,999、6：¥5,000～5,999、7：¥6,000～6,999、8：¥7,000～7,999、
9：¥8,000～9,999

データ項目②

～お好み焼きの情報～

▶ F：お好み焼きの種類

1：広島風、2：関西風、3：カキオコ（岡山）、4：三津浜焼（愛媛）、5：遠州焼（静岡）

何種類か提供している店に関しては、メインで提供している種類をその店のお好み焼きの種類とする。

▶ G：お好み焼きのメニュー数

モダン焼き、ねぎ焼き、べた焼きもお好み焼きの一種とする。

トッピングの種類もそれぞれ1とカウントする。また、そばとうどんの違いは考慮しない。

▶ H：お好み焼き以外の料理のメニュー数

焼きそば/うどん、一品料理、とんぺい焼き、鉄板焼きなど。

デザートは含み、ドリンクは除く。

データ項目③

～食べログ評価点～

▶ I：星

口コミ評価をもとに、食べログが独自につけている評価点。5点満点で小数点第2位までの値が出される。店名の下に表示されている。（図1参照）

▶ J：料理・味

▶ K：サービス

▶ L：雰囲気

▶ M：コストパフォーマンス

▶ N：酒・ドリンク

J～Nは、項目別評価の平均点。星と同様に、5点満点で小数点第2位までの値が出される。PC版の口コミ欄から閲覧ができる。（図2参照）



(図1)



(図2)

データ項目④

～評価数・保存数～

▶ 0：口コミ数

口コミを投稿した人の人数。星の右横に、吹き出しマークと数値が表示されている。（図1参照）

▶ P：保存数

店情報を保存している人の人数。口コミ投稿数の右横に、クリップマークと数値が表示されている。（図1参照）

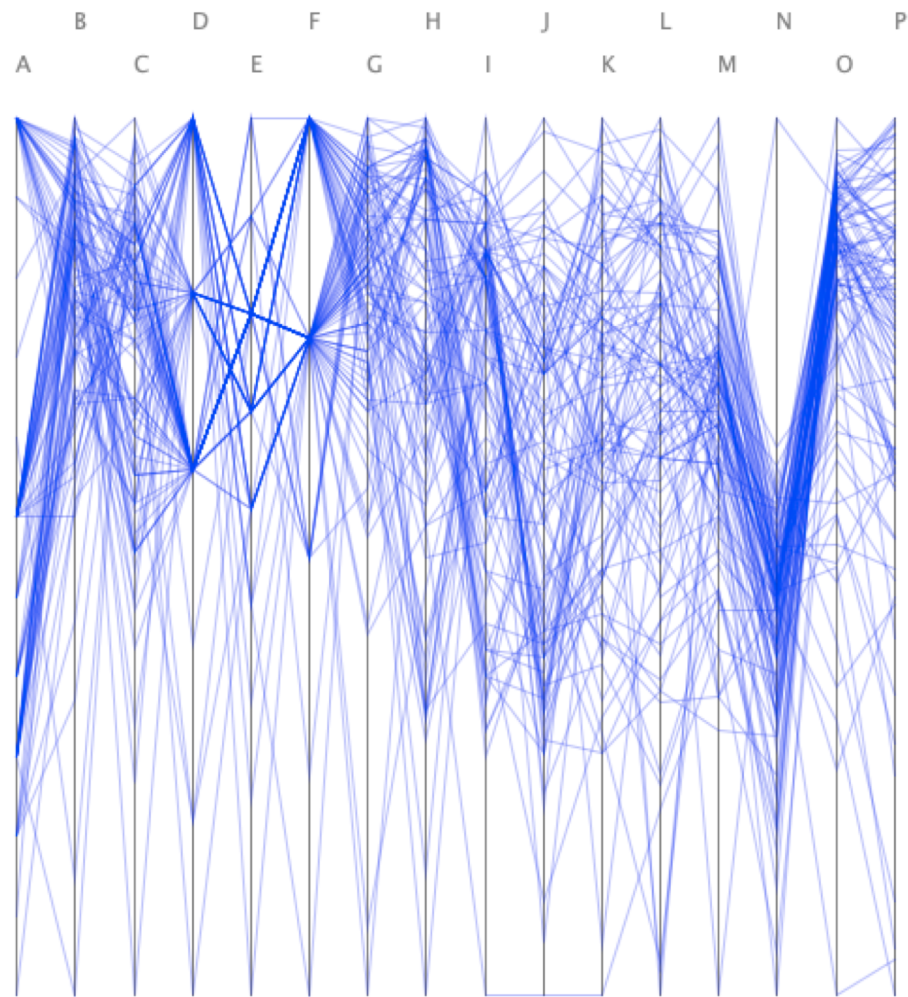
食べログの評価点は、口コミの星を単純に平均しているのではなく、評価が高いユーザーなどの口コミを優先したり、不正業者による口コミを防止したりして独自のアルゴリズムで点数化している。（また、評価3.8以上は年会費を払わなければ3.6に下げられるという疑惑もある...）

星などの評価点だけでは、店の価値を判断しかねると考え、上記のような評価された回数や保存された回数も評価軸として取り入れた。

ただ、今回のデータは食べログに載っている店舗の中でもランキング上位店であるので、どの店もとても評価の高いことは間違いない。

可視化結果

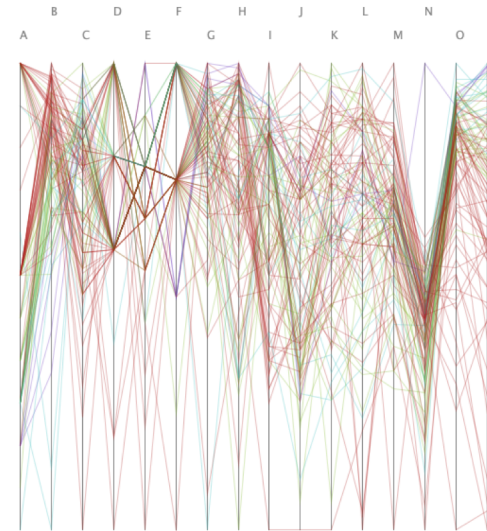
- ▶ 素の状態では、図3のようになつた。
- ▶ これから、各項目の相関関係について調べていく。



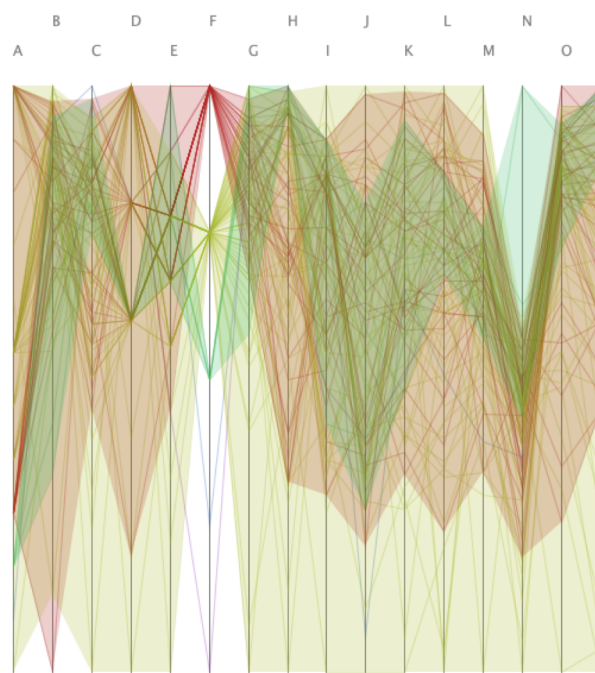
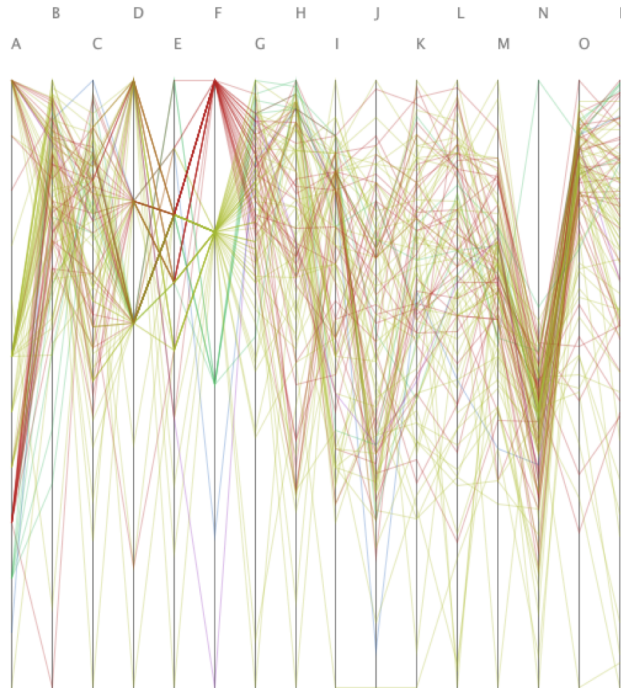
(図3)

色分け

- ▶ はじめにクラスタリングしてみたが、色分けが分かりづらかった。(図4参照)
- ▶ そこで、「Category」を使ってF（お好み焼きの種類）の数値ごとに色分けを行なった。



(図4)

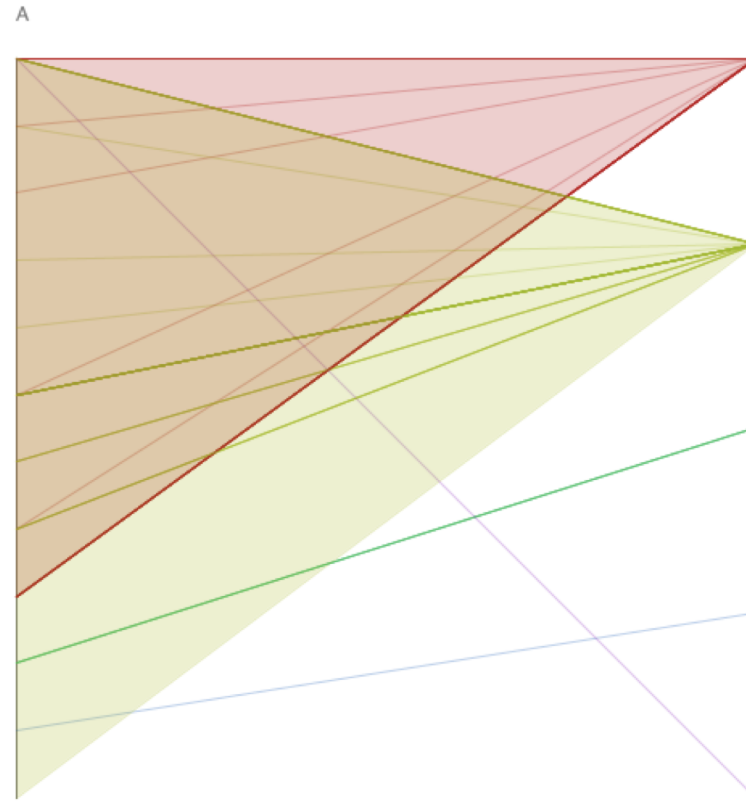


(図5(左))
Fの数値ごとの色分け
1:赤、2:黄緑、3:緑、
4:青、5:紫

(図6(右))
図5をPaintsしたもの

カリング (1)

- ▶ カリングを最大限してみると、PCPの軸をAとFのみにすることができた。
- ▶ このPCPから、A(都道府県)とF(お好み焼きの種類)の関係性に関して考察していく。
- ▶ 赤と黄緑の範囲の広がり具合から、**広島風**よりも**関西風**の方がより広い地域で人気であることが分かる。
- ▶ しかし、東京(1)や大阪(6)などで**広島風**と**関西風**どちらの店もある一方、広島(8)では**広島風**の店しかないことが読み取れる。

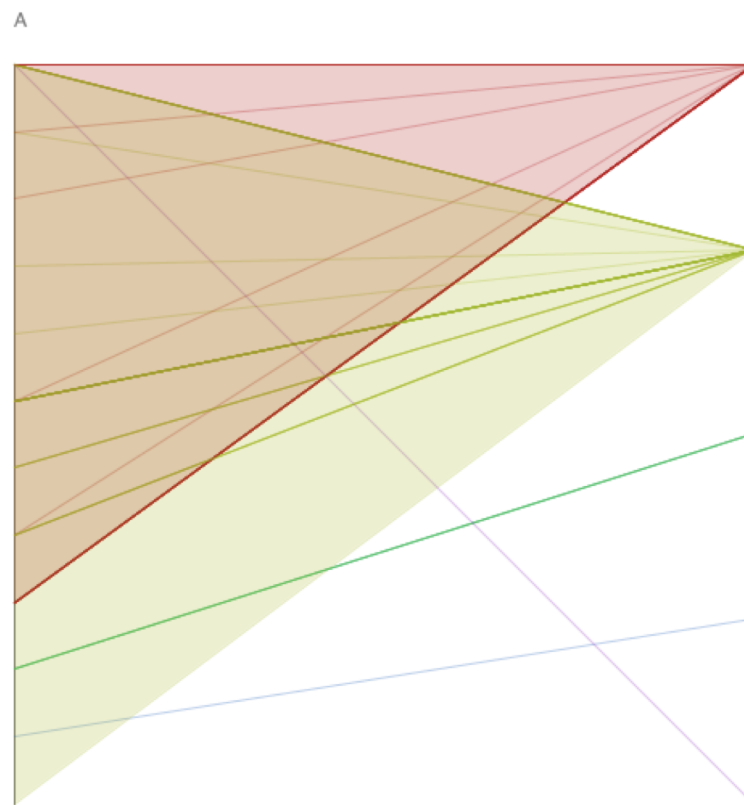


(図7)

参照 1:**広島風**、2:**関西風**、3 : **カキオコ**、4:**三津浜焼**、5:**遠州焼**

カリング (2)

- ▶ 緑の線が一本の太い線になっていることから、カキオコは岡山県内でのみ食べられており、根強い人気があることが読み取れる。
- ▶ 三津浜焼や遠州焼に関しては、1～3に分類しがたかったため一種としたが、他地域への広がりは見られなかった。

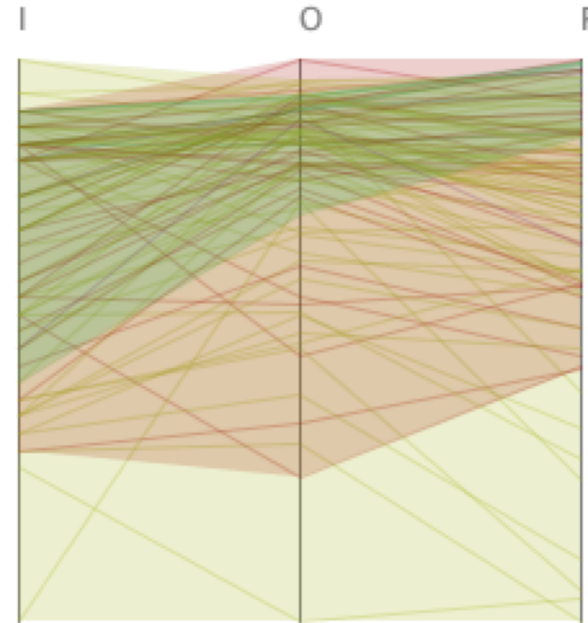


(図7)

参照 1:広島風、2:関西風、3:カキオコ、4:三津浜焼、5:遠州焼

右上スライダーをあげる (1)

- ▶ 図9より、I(星), O(口コミ数), P(保存数)に強い相関関係があることが読み取れる。
- ▶ I(星)とO(口コミ数)に負の相関
→ 評価が高いからと言って、口コミ数が多いわけではないことが分かる。口コミが多いほど人気も高いだろうという予想は外れた。
- ▶ I(星)とP(保存数)に正の相関
→ 評価点が高いほど、「今度行ってみたい」と思う人が多くなることが分かる。このことから、多くの人が店選びにおいて星を重要視していることが分かる。



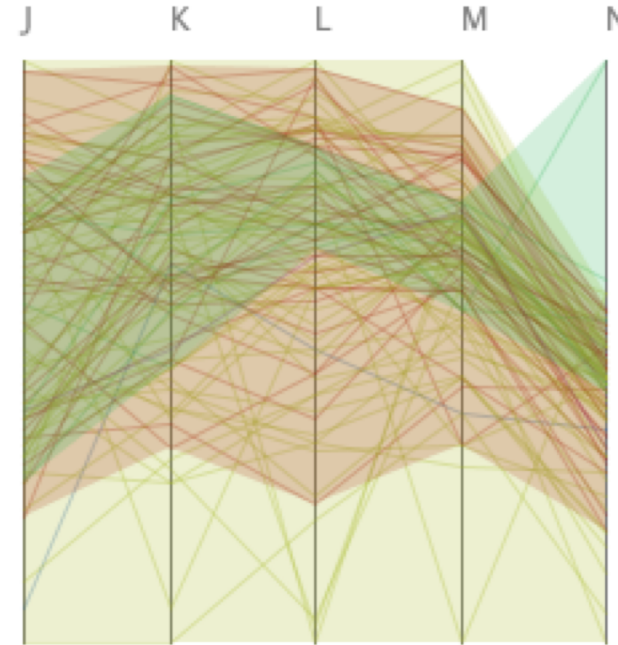
(図8)



(図9)

右上スライダーをあげる (2)

- ▶ 図11より、J (料理・味) , K (サービス) , L (雰囲気) , M (コストパフォーマンス) , N (酒・ドリンク) に相関関係があることが読み取れる。
- ▶ N (酒・ドリンク) とK (サービス) , L (雰囲気) に正の相関
→ ドリンクの品揃えが多いほど、サービス・雰囲気が良いという評価をしやすいくことが分かる。
- ▶ M (コストパフォーマンス) とK (サービス) , L (雰囲気) に負の相関
→ 逆に、サービス・雰囲気が良すぎるとコストパフォーマンス面では評価が低くなることが分かる。



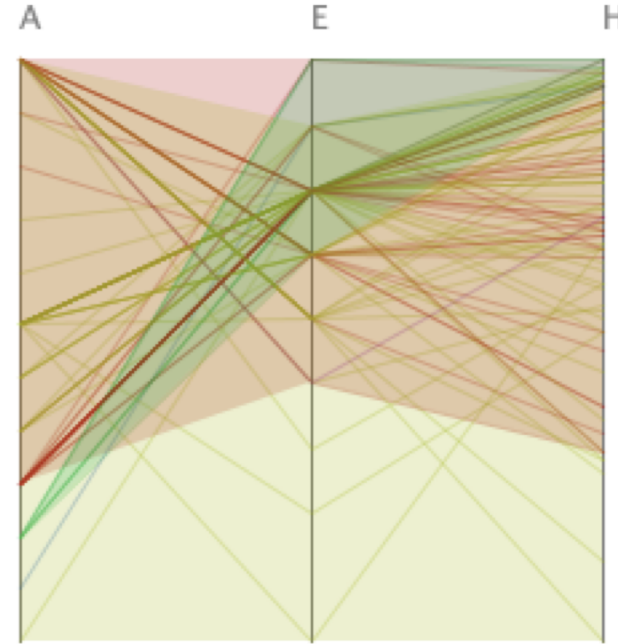
(図10)



(図11)

右上スライダーをあげる (3)

- ▶ A（都道府県）とE（夜の予算目安）に負の相関
→東京から西の都道府県に行くほど、リーズナブルであることが分かる。
- ▶ E（夜の予算目安）とH（お好み焼き以外のメニュー数）に負の相関
→やきそばや一品料理など手軽に食べられる品目が多いほど、予算を安くおさえられる。



(図12)

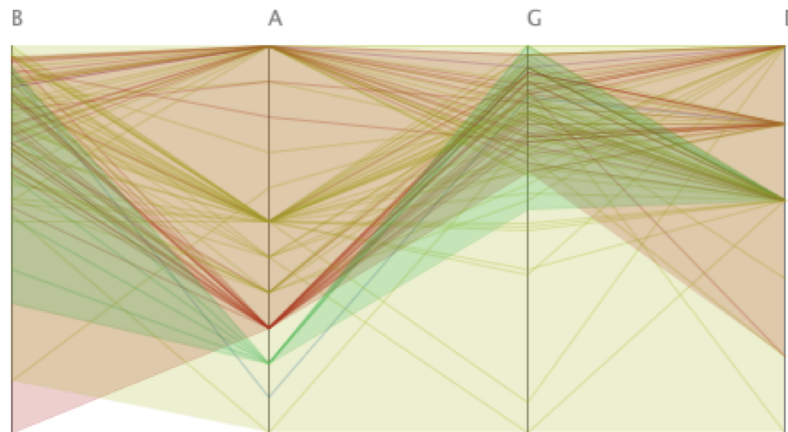


(図13)

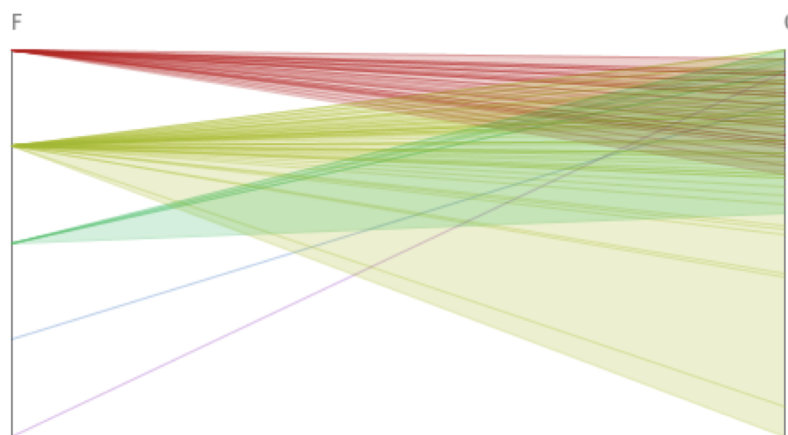
カリングして、右上スライダーをあげる

- ▶ A（都道府県）とB（最寄りからの距離）に正の相関

→西の都道府県ほど、駅から店までの距離が遠い。



- ▶ 図14の下図より、F（お好み焼きの種類）とG（メニュー数）に関して見てみると、メニューの多さは**広島風**、**カキオコ**、**関西風**の順で多いことが分かる。特に**広島風**に関しては、**カキオコ**よりも少なく、オリジナルの味勝負で経営している店が多いことが分かる。



(図14)

参照 1:**広島風**、2:**関西風**、3 : **カキオコ**、4:**三津浜焼**、5:**遠州焼**

まとめと感想

～まとめ～

- ▶ お好み焼き百名店のデータを可視化してみた結果、広い地域・豊富なメニューのバリエーションで勝負する関西風と、広島を筆頭する地域で根強い人気を誇る広島風お好み焼きという違いを知ることができた。
- ▶ また、口コミの投稿数が少なくても食べログの評価（星）が高い店があることや、ユーザーがいかに星の点数を重視しているかがわかった。

～感想～

- ▶ 今回、大好きなお好み焼きのデータを集めてみて、初めて知るお好み焼きの種類（カキオコなど）があった。データ集めは大変だったが、新しい発見のおかげでとても楽しめた。
- ▶ 反省としては、店の基本情報などの項目の中であまり必要なかったものもあった（最寄りからの距離や座席数など）ので、こういったデータを集めるか決める段階で注意する必要があると感じた。

参考文献

- ▶ 食べログ/百名店 お好み焼き2019

<https://award.tabelog.com/hyakumeiten/okonomiyaki>

- ▶ 食べログ/点数・ランキングに関して

<https://tabelog.com/help/score/>

- ▶ 食べログ3.8問題を検証

https://clean-copy-of-onenote.hatenablog.com/entry/tabelog38_problem

- ▶ 福太郎 本店

<https://tabelog.com/osaka/A2701/A270202/27002665/dtlratings/>