

# 群馬県富岡市周辺施設の 評判分析

清水嵩文

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

1	(1)	北海道	54.2
2	(2)	京都府	46.3
3	(3)	東京都	35.1
4	(4)	沖縄県	32.7
5	(6)	神奈川県	30.2
6	(7)	奈良県	28.7
7	(9)	大阪府	24.1
8	(10)	長野県	24.0
9	(11)	石川県	23.3
10	(8)	福岡県	21.4
11	(5)	長崎県	20.0
12	(19)	静岡県	17.6
13	(17)	愛知県	17.2
14	(18)	宮城県	16.8
15	(32)	大分県	16.5
16	(13)	兵庫県	16.3
17	(25)	青森県	15.8
18	(30)	富山県	15.4
19	(25)	秋田県	14.9
20	(14)	千葉県	14.4
21	(15)	熊本県	14.2
21	(20)	鹿児島県	14.2

25	(38)	愛媛県	13.1
26	(21)	三重県	12.6
27	(41)	島根県	12.4
28	(25)	和歌山県	12.0
29	(24)	高知県	11.9
30	(36)	山梨県	11.7
31	(33)	岡山県	11.4
32	(43)	山口県	11.3
33	(41)	滋賀県	10.9
34	(36)	香川県	10.8
35	(23)	新潟県	10.5
36	(28)	岩手県	10.3
37	(21)	山形県	10.0
38	(45)	佐賀県	9.4
39	(40)	福島県	9.2
39	(46)	埼玉県	9.2
41	(31)	福井県	9.1
42	(29)	岐阜県	9.0
42	(39)	鳥取県	9.0
44	(34)	徳島県	8.6
45	(44)	群馬県	8.5
46	(35)	栃木県	7.8



## 世間から浴びる群馬への冷たい視線

群馬県は本来素晴らしい場所であるにも関わらず、世間では人気のない・魅力のない県であることが今回の課題選択のモチベーション



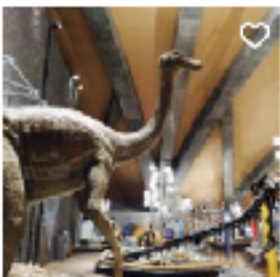
# 観光地としてのポテンシャル

---

群馬県には世界遺産「富岡製糸場」もあるので、富岡製糸場を軸に評判分析をすることで、観光地としての群馬の強みをより活かせると考えた。



富岡市  
Tomioka City



群馬県立自然史博物館

51件の口コミ  
富岡市の観光スポットで18件中1位  
専門美術館

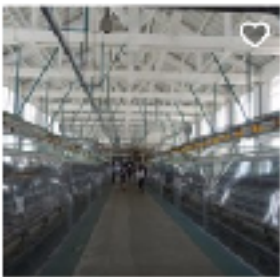
詳細



群馬サファリパーク

125件の口コミ  
富岡市の観光スポットで18件中2位  
動物園

詳細



富岡製糸場

518件の口コミ  
富岡市の観光スポットで18件中3位  
史跡、歴史博物館

詳細



妙義神社

52件の口コミ  
富岡市の観光スポットで18件中4位  
神社/寺院/教会/モニュメントなど

詳細



一之宮貫前神社

39件の口コミ  
富岡市の観光スポットで18件中5位  
神社/寺院/教会/モニュメントなど

詳細



もみじ平総合公園

6件の口コミ  
富岡市の観光スポットで18件中6位  
公園、遊び場

詳細

# 富岡周辺観光地

- 富岡製糸場
- こんにゃくパーク
- 道の駅甘楽
- もみじ平総合公園
- 長岡今朝記念ギャラリー
- 群馬県立自然史博物館
- 群馬サファリパーク
- 妙義神社
- 一ノ宮貫前神社
- もみじ平総合公園
- 富岡諏訪神社 etc.

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

# データ収集サイト：トリップアドバイザー

---

トリップアドバイザーを選択した理由は2つ。

- 1 広告収益がメインなので中立性が保たれている可能性が高い
- 2 知名度も高く、ある程度口コミのボリュームもあるはず

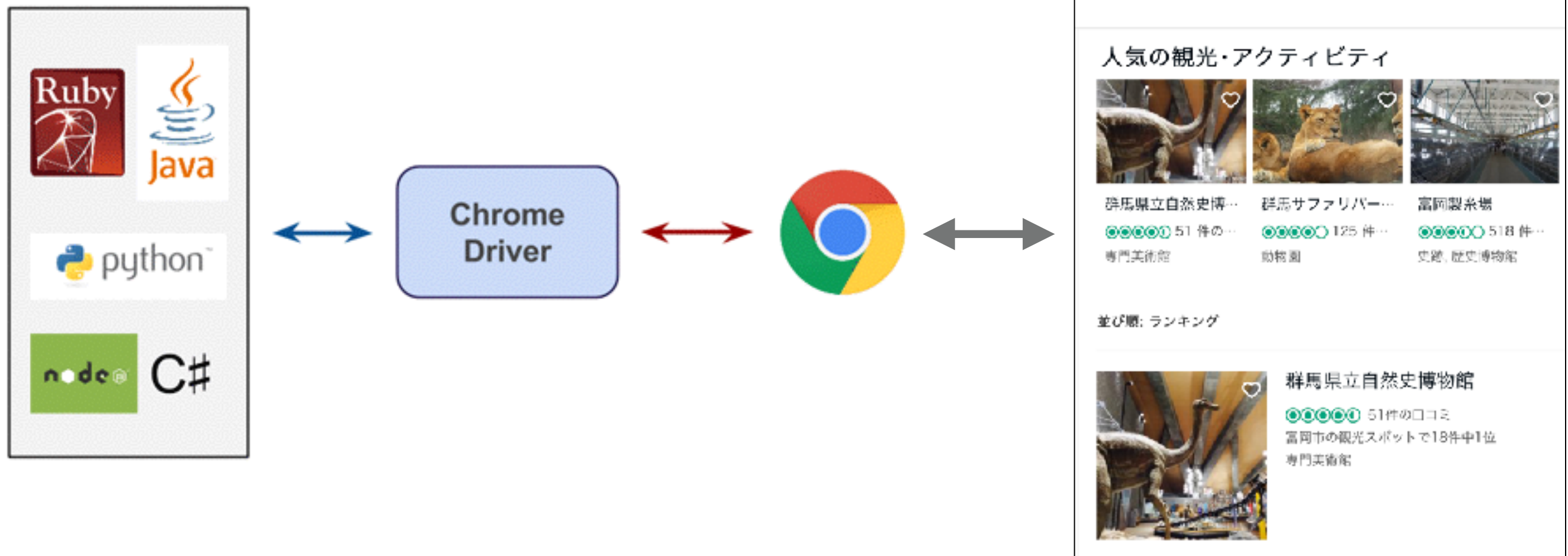




# データ収集方法：SELENIUM WEBDRIVERの活用

.....

pythonでGoogleChromeを操作し、トリップアドバイザーの「富岡 観光」ページよりデータ取得



# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

この口コミ情報はTripAdvisor LLCのものではなく、トリップアドバイザー会員の主観的な意見です。



静百丸 茨城県



口コミを投稿：2018/01/22

## 近代史の一部を垣間見る

世界遺産というよりは明治以降の近代経済の一部を垣間見るという感想。当時の工場は女工を始め多くの従業員が集団生活し、製糸場が一つの町になっていたと思われる。15歳から25歳の女子が集団で寝食を共にし毎日製糸作業に打ち込む姿に日本の近代史を見る。数百年前の○○殿とかいう歴史... [さらに表示](#)

[富岡製糸場について質問する](#)

役に立った



この口コミ情報はTripAdvisor LLCのものではなく、トリップアドバイザー会員の主観的な意見です。



青々283 坂東市, 茨城県



口コミを投稿：2018/01/14

## 混雑なしだけど、高い

世界遺産に登録されてすぐには混雑していたようだが、冬季だったためか空いていた。しかし、思っていたより見学者に対しての十分なサービスがなされておらず、入場料1000円を高く感じてしまった。ガイドツアーにも参加したが、よく理解できたかと言えば自分でも疑問。また行こうとは思わ... [さらに表示](#)

[富岡製糸場について質問する](#)

役に立った



この口コミ情報はTripAdvisor LLCのものではなく、トリップアドバイザー会員の主観的な意見です。



mickato



口コミを投稿：2018/01/08

## 近代史の重要な遺構

世界遺産に絶景を期待する方には不評かもしれませんが、日本の近現代史が好きな私にとっては大変興味深い場所でした。人もまばらで、ゆっくりと見学できました。あまり大人数で行くところではないかもしれません。

[富岡製糸場について質問する](#)

役に立った 2



この口コミ情報はTripAdvisor LLCのものではなく、トリップアドバイザー会員の主観的な意見です。

[前へ](#)

1

2

3

4

5

6

...

51

[次へ](#)

# 各観光地における口コミ投稿

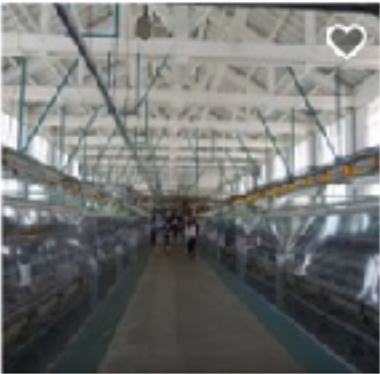
.....  
複数ページに渡る口コミを分析しようと試みる。

そのためにクローリングを行った。

だがしかし、、、

# 度重なる試練の山

Javascriptで変化するページはどうやらクロールすることが難しいらしい。。。




### 富岡製糸場

518件の口コミ  
富岡市の観光スポットで中3位  
史跡, 歴史博物館

明らかに情報が取れてない事件

	すめ。美術館は広々として落ち着けます。よい季節に行けば紅葉と...	4.0
	われている駅です。特に見るところはありません。名前が珍しいだ...	3.0
218	上信電鉄の駅、名前のウケるので降りてみました。古びた木造駅舎はいい感じ、駅名表はひらがなで、...	3.0
219	日本でも屈指の歴史を誇る上信電鉄の終点下仁田にほど近い小駅。でもその名前とローカル風情、そ...	4.0

220 rows x 3 columns



codena1667 港区, 東京都

5口コミを投稿：2017/12/19

## 話のタネに

世界遺産ということで一度は行こうと思  
歩きました。歴史的な背景はともかくです  
りというか。保存や防災上難しい面も多い  
も良いと思いました。ガ... [さらに表示](#)

「さらに表示」が  
押せない事件

民駐車場から5分くらい  
のわりにはちょっとがっか  
は思いですが、もう少し見せ方に工夫があって



# OUTLINE

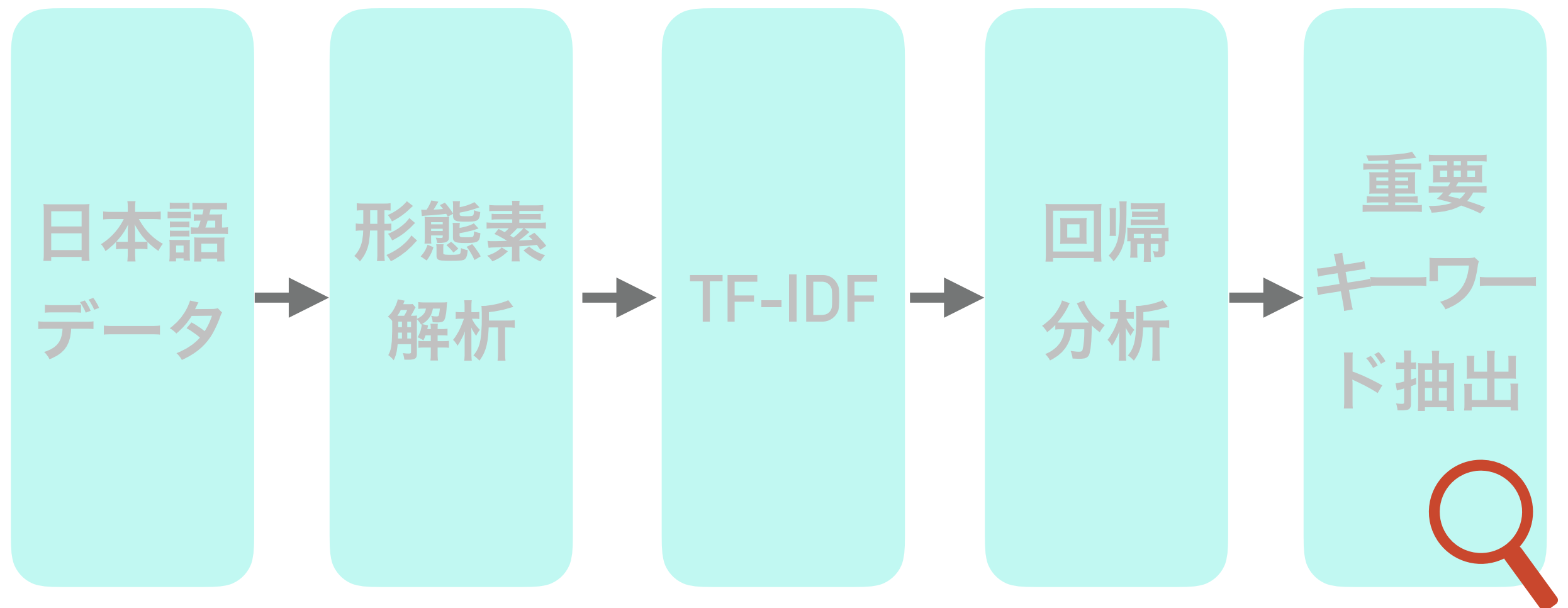
---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

# GOAL：回帰分析で重要なキーワード抽出

---

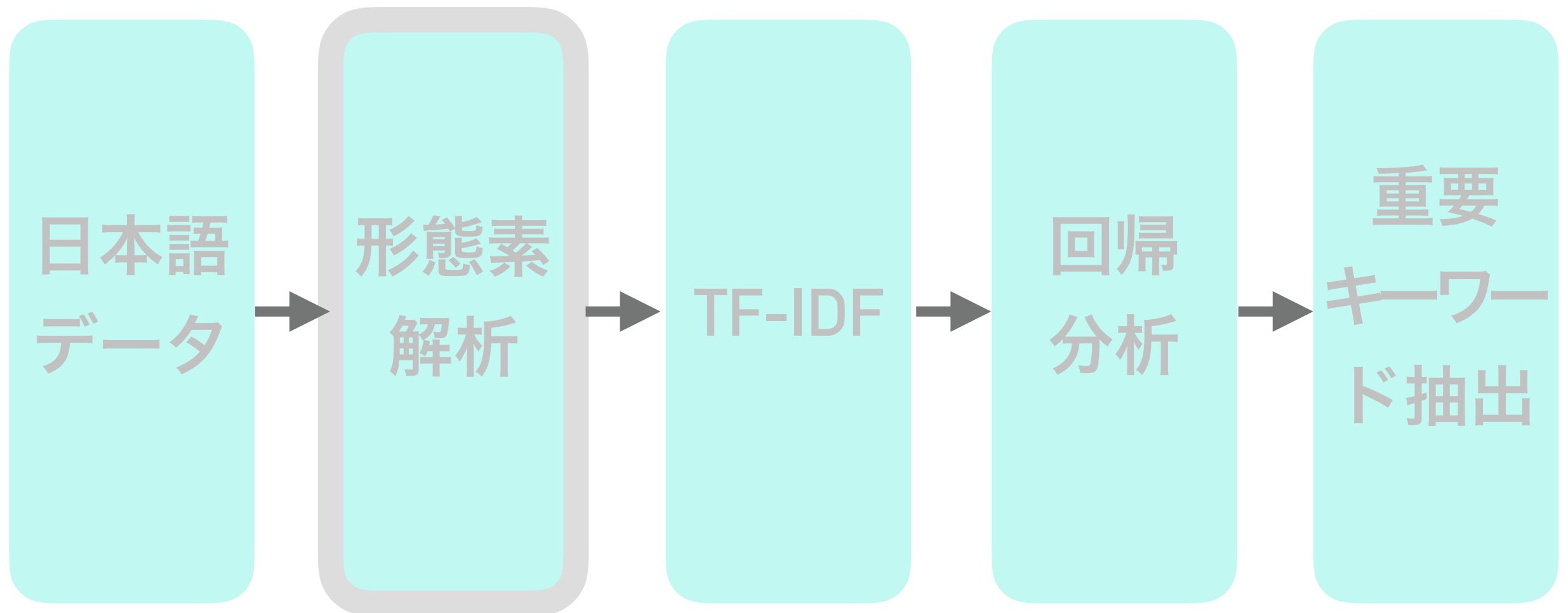
回帰分析を行うことで、高評価に繋がるキーワードを知り、富岡市の観光の強みを把握できるのではないか？



# 形態素解析とは？

---

日本語では「すもももももももものうち」のようにどこで文字区切りしたらいいのかわからないので、「すもも も もも も もも の うち」という形に分ける。（今回はMecabを使用）

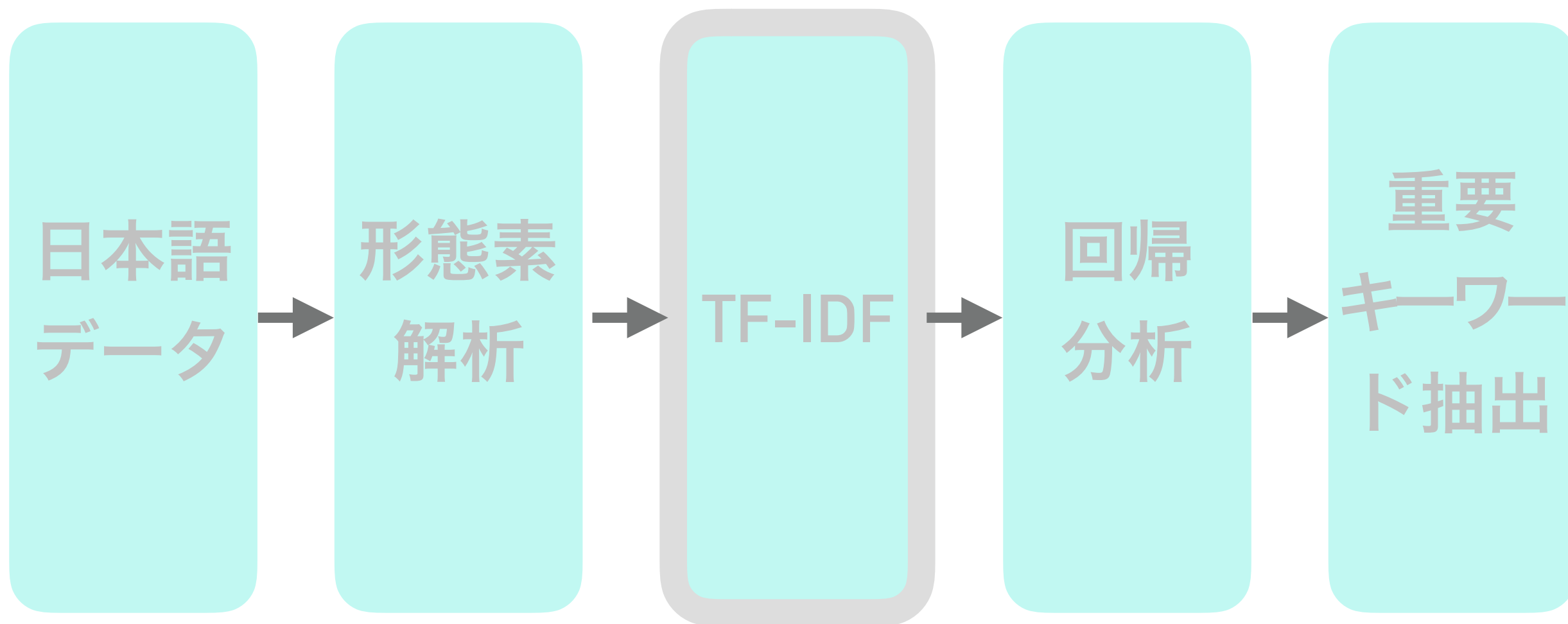


# TF-IDFとは？

---

ある単語の頻度に対して、その単語が出現する文書の割合で重み付けを行う手法。

(例：「私」という単語は頻出するが、「私」という文字は、文章に対して大きな意味を持たないので、重要度を下げるという。)

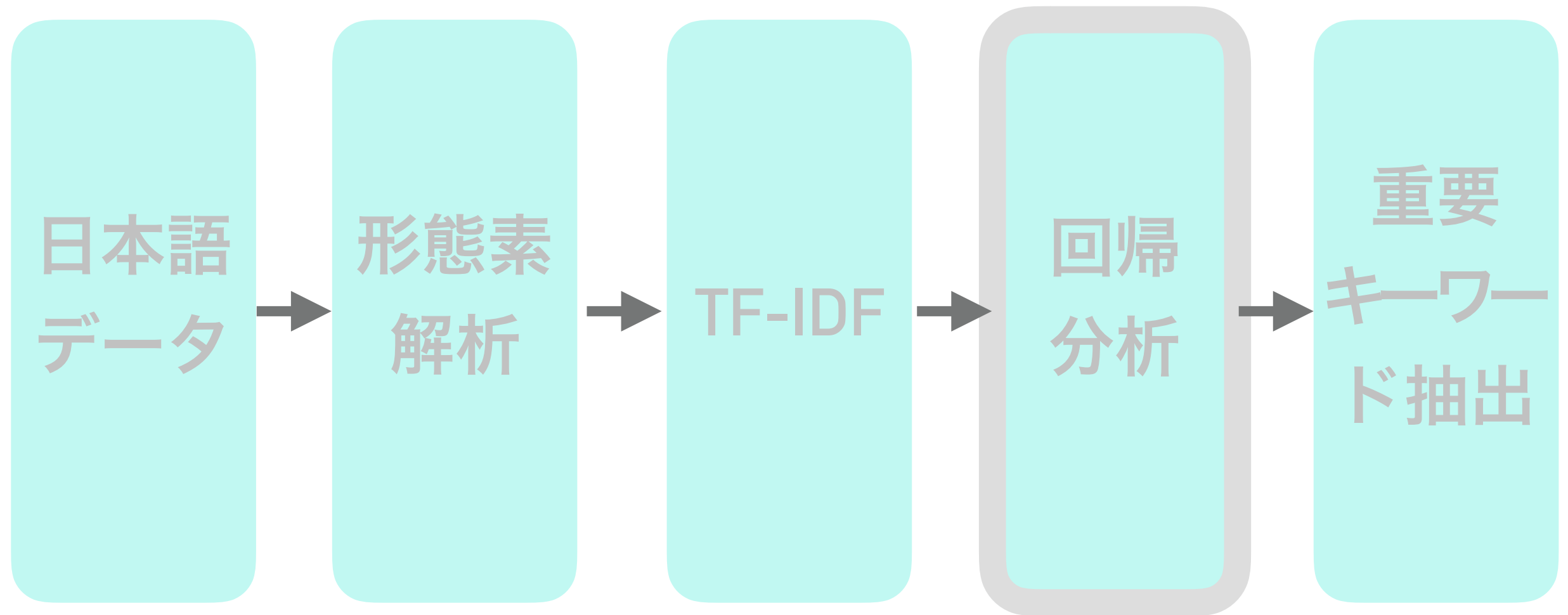




# 回帰分析について

---

今回はLasso回帰/Ridge回帰/ElasticNetを利用し、その結果を比較検討し、一番結果が良さそうなモデルを選択する。



# LASSO回帰/RIDGE回帰/ELASTICNETとは？

$$\min_{\beta_0, \beta} \left( \frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 + \lambda P_\alpha(\beta) \right),$$

ここで、

$$P_\alpha(\beta) = \frac{(1 - \alpha)}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1 = \sum_{j=1}^p \left( \frac{(1 - \alpha)}{2} \beta_j^2 + \alpha |\beta_j| \right).$$

ペナルティをかけた多重共線性の問題を避けながら、回帰係数を算出する方法。

	特徴	用途
Lasso回帰	$\alpha=1$	特定の回帰係数を0にし、変数選択したい場合
Rasso回帰	$\alpha=0$	過学習を起こす恐れがある場合
ElasticNet	$0 < \alpha < 1$	Lasso回帰のデメリットを補いたい時

# LASSO回帰のデメリットとは？

---

デメリット 1：

特徴量( $p$ ) > データ量( $n$ )の時、たかだか $n$ 個の変数までしか選択できない。

デメリット 2：

相関の高い変数群がある場合、Lasso回帰はその変数群のうちから一つしか変数として選択できない。（それ以外は係数0）



Lasso回帰のデメリットを補いたい時は、ElasticNetを活用

\*今回は3つのモデルを試して評価します。

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと



# 単語抽出に関して

---

キーワードになり得る単語は、一部の品詞のみ。

したがって今回は一般名詞・固有名詞・サ変接続・形容詞のみを抽出し、重要になりそうなキーワード抽出を行った。

## 一般名詞例

夏休み

カフェ

神社

・  
・  
・

## 固有名詞例

妙義山

甘楽

群馬

・  
・  
・

## サ変接続例

紅葉

利用

観光

・  
・  
・

## 形容詞例

青い

美しい

良い

・  
・  
・

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

# クローリング

「さらに表示」を読み取れなかったり、「次へ」ボタンの設定が各ページで違っていたため、全ての情報を集めるのにかなりの時間を要した。



codena1667 港区, 東京都



口コミを投稿：2017/12/10

## 話のタネに

世界遺産ということで一度は行こうと思いましたが、市民駐車場から5分くらい歩きました。歴史的な背景はともかくても、1000円の入場料のわりにはちょっとがっかりというか。保存や防災上難しい面も多いとは思いますが、もう少し見せ方に工夫があっても良いと思いました。ガ... [さらに表示](#)

「さらに表示」が  
押せない事件



Neomi T

197  
178

## 見所いっぱいな世界遺産

朝9時オープンです。オープンしたての時間に行きましたので、場内は大変空いていました。まずその広い敷地に驚きました。さすが世界遺産に認定されるだけあって、

役に立った

この口コミ情報はTripAdvisor LLCのものではなく、トリップアドバイザー会員の主観的な意見です。

「次へ」が押せ  
ない事件

前へ

1 2 3 4 5 6 ... 10

次へ

# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと



# 分析結果

Ridge回帰/Lasso回帰/ElasticNetの中でもっとも平均二乗誤差が少ない  
Ridge回帰の結果を元に、係数が大きいキーワードを抽出し、分析・考察  
する。

	平均 2 乗誤差
Lasso回帰	0.6066749910233624
Rasso回帰	<u>0.6045927890648886</u>
ElasticNet	0.6499036116811716

最も高評価に繋がる（係数最大）キーワードは！？

.....



「道」

・・・道ってなんだ???



# 評価が高い「道」がつく文章を探してみると、、、

---



TVの中山秀ちゃんのコマーシャルを見て群馬に行ったら見に行こう  
思っていました。 近くに道の駅も...

道の駅で、ピザを焼いているとは思いませんでした。やや小規模の道  
の駅ですが甘楽町物産センターを...



群馬サファリパークついでに寄ってみました。 大きな道の駅ではな  
いですが、食事も物販も地元色た...

軽井沢からの帰りに利用。道の駅らしく地元の野菜・名産品や弁当・  
食堂もある。帰りに道の駅で野菜...



・・・「道の駅」の満足度はかなり高いらしいから、「道」だったのかな

# その他高評価に繋がるキーワード

---



郵便局前でバスを降り、徒歩5分  
入場すると、ボランティアガイド  
申し込みに中庭へ。 毎00時・...

製糸工場を見学する事も初めてだっ  
たけど、とても楽しかった。時間が  
無かったのでガイドをお願い...



恐竜の骨やロボットはありますが  
そのほかの展示物もいろいろあっ  
てよかったです。

巨大な竜脚類3兄弟や、動きが地味  
めな恐竜ロボットなど、目玉はもち  
ろん恐竜関連なのかもしれませんが、...



自分の車で回って、自分達のペース  
で見学できたので良かったです。間  
近で動物を見れて、想像以...

明治の建築物として立派な建物です。  
骨組みは木造です。日本の技術の高  
さを誇りに思う事ができます...



たまたま近くに行ったので入って  
みましたが、無料施設と考えると  
満足です。コンニャクビュッフェ

# 高評価に繋がるキーワードを多く持っていた観光施設

---

甘楽道の駅

道

こんにゃくパーク

満足

想像

富岡製糸場

ボランティア

明治

お願い

(ガイドのお願いという文脈がほとんど)

県立自然史博物館

展示

想像

恐竜



# 甘楽道の駅の考察

.....

- ▶ 甘楽道の駅への訪問自体が満足度が高い傾向にある？



- ▶ 富岡市に訪れた人が甘楽道の駅にも訪れたくなる施策を考え、富岡市での観光トータルでの満足度を高める。



# 富岡製糸場の考察

---

- ▶ 「明治」と、「ボランティア」・「お願い」が満足に寄与している。



- ▶ 施設のバックグラウンドを理解している人々は満足度が高い。



- ▶ ビジュアル的なインパクトが少ない施設なので、より文化理解を促す仕組みを作るといいのではないかな？





## こんにゃくパークの考察

.....

- 「想像」と「満足」が好評価に寄与している。



- 無料施設だけど「想像」以上のような文脈が多く見受けられた。



- 施設内容の表現がうまくできていないから、「無料＝陳腐な施設」というイメージを抱かせてしまっているのではないかな？ よって今の打ち出し方を改める必要がある？





# 県立自然史博物館の考察

- .....
- ▶ 「想像」・「展示」・「恐竜」  
が好評価に寄与している。



- ▶ 「恐竜」が「想像」以上とい  
う内容が多かった。



- ▶ 「恐竜」というアピールポイ  
ントを、対外的に打ち出せて  
いるのか？  
今回を機に「県立恐竜博物館」  
への改名を勧める。

# OUTLINE

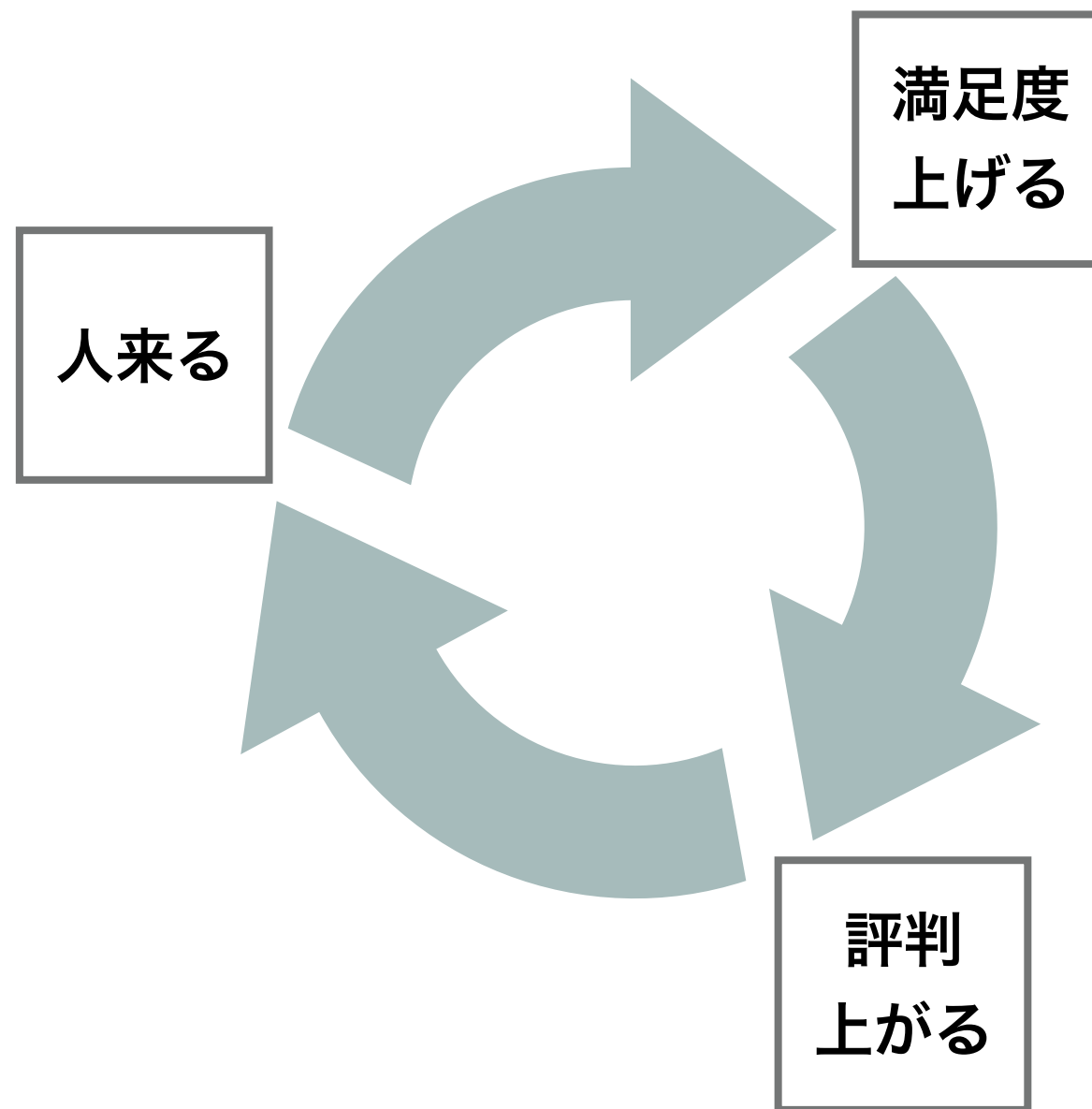
---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと

# 満足度が上がれば観光が盛り上がる、はず

---

満足度に繋がるキーワードから、「より満足度を向上させる体験を作る」・「より満足度の高い施設への誘導を行う」ことにより、富岡の評判も良くなり、より多くの人が観光にきてくれるのではないか！？



# OUTLINE

---

- 課題選択の理由
- データ収集方法
- 使用したデータについて
- 初期仮説・分析アプローチ
- 工夫したこと
- 難しかったこと
- 分析結果・考察
- ビジネスインパクト
- 今後さらに実施したいこと



# 地域を拡大させた分析

---

評判分析を行うことによって、世間の感覚を定量的に理解することができたと思う。

地域を拡大させて群馬県その他地域の分析を行いたい。

また満足度に寄与するキーワードの文脈も重要なので、その分析も行えるようにしたい。

