

エンジニア 採用課題

# UGC分析と マーケティング戦略に ついての企画書

2023年10月1日

石原正喜

---

## TABLE OF CONTENTS

---

01	企画の概要	3
02	デモの概要	4
03	Youtube分析のデモの内容	5
04	Google Map分析のデモの内容	7
05	デモの利点	9
06	デモのさらなる改良案	10
06	リスクと課題	11

---

# 01

## 企画の概要

---

本企画は、ユーザー生成コンテンツ（UGC）を活用し、より観光客の視点に立った観光分析フレームワークの構築を目的としています。デモンストレーションでは、歴史的な街並みと食文化で人気の高い川越を対象に、YouTube や Google Maps などに投稿された情報を収集・統合し、実際の観光体験や来訪者の嗜好を分析し、AIとデータ解析を通じて、地域の特性を踏まえた効果的なプロモーション戦略の立案を目指しました。

観光に関する動画・レビュー・コメントをクラウド上で一元管理し、GeminiをはじめとするAIや分析ツールで解析・可視化することで、観光客が地域をどのように捉え、何に価値を感じているかをデータに基づいて明らかにします。これにより、地域の魅力を可視化し、戦略的な観光施策の立案に活用することが可能となります。

今後は、このフレームワークを他地域や他SNSプラットフォームへ拡張し、地域間比較や国際的な応用を可能にしていきます。さらに、データ処理の自動化、感情分析や属性別傾向分析などの高度な手法を取り入れることで、より多面的で精度の高い観光分析を実現し、最終的には多様な地域が活用できる人間中心型の観光戦略支援プラットフォームを構築することを目指します。

---

# 02

## デモの概要

---

### [デモの概要]

今回のデモを作成するにあたって任意の地域「川越」を選定しました。川越は小江戸の街並みや歴史的建造物、食べ歩き文化などで観光需要が高く、SNSや動画サイト上でも多くの情報が発信されています。本プロジェクトでは、YouTube と Google Maps という二つの主要なユーザー生成コンテンツ（UGC）のソースを活用し、川越に関する観光情報やレビュー、コメントを体系的に収集・分析します。

YouTube Data API で観光動画と字幕を収集し、Google Maps Places API で観光地や飲食店の詳細とレビューを取得します。これらを GCP 上で整形・保存し、Gemini による要約や分析で観光体験や訪問者の嗜好を抽出します。さらに BigQuery と Looker Studio で可視化することで、トレンド把握やプロモーション戦略の立案に活用します。

本プロジェクト全体を通じて、UGC を活用した観光分析の実証的なフレームワークを構築することを目指しています。YouTube と Google Maps という異なる情報源を統合し、Gemini による戦略生成と BigQuery・Looker Studio による可視化を組み合わせることで、観光地プロモーションのための高度な分析が可能になります。

---

# 03

## Youtube分析のデモの内容

---

### 01.動画の一覧データを取得 (youtube\_seach.py)

- YouTube Data API v3 の Search list を利用
- 「川越」「Kawagoe」などのキーワードで検索し動画一覧を取得
- 一回のリクエストで取得できる件数に上限があるため四半期ごとに収集
- キーワードごとに JSON ファイルに保存

### 02.動画の詳細データを取得 (youtube\_enricher.py)

- NGワードを YAML ファイルで定義し、タイトルや概要欄に含まれる動画を除外
- YouTube Data API v3 の Videos list を利用
- 除外済みの動画IDリストをもとに統計情報やメタ情報を取得
- 取得する情報にはタイトル、概要欄、公開日、タグ、カテゴリ、再生回数、いいね数、コメント数、動画の長さなどが含まれる
- 情報を Pandas データフレームに整形し Parquet 形式で保存

### 03.動画の字幕を取得 (youtube\_captions.py)

- youtube-transcript-api を利用
- Parquet ファイルから再生回数の多い動画を選び字幕を取得
- GCP 上では大量リクエストによるIPブロックの可能性があるのでローカル環境で実行

---

# 03

## Youtube分析のデモの内容

---

### 04.Gemini による観光戦略の生成 (youtube\_strategy.py)

- 字幕データを基に Gemini 2.5 Flash で各動画を要約
- 字幕の要約に再生回数や高評価数、公開日などの指標を組み合わせて入力データを構築
- Gemini 2.5 Pro に入力し、観光コンテンツの傾向やターゲット層を踏まえたプロモーション戦略を生成
- 出力内容には観光体験のトレンド分析、訪問者の属性別嗜好、季節やターゲットに応じた施策提案を含める

### 05.BigQueryにデータを導入して可視化 (youtube\_video\_features.sql)

- 動画の詳細データ（統計値・カテゴリ・言語・動画の長さなど）を BigQuery に導入
- SQL で再生回数の区分やエンゲージメント比率（いいね率、コメント率など）を計算
- 公開月や曜日、ショート動画フラグ、日本語動画かどうかといった特徴量を追加
- 加工済みのデータを Looker Studio と接続し、観光動画のトレンドや人気要素を可視化
- 可視化結果をもとに、観光戦略や施策の改善に活用できるダッシュボードを構築

---

# 04

## Google Map分析のデモの内容

---

### 01. Google Mapの情報を取得 (places\_search.py)

- Places API を利用し、川越の中心地から半径4kmの範囲にある観光地や飲食店を検索
- 検索対象は「観光スポット」「レストラン」などのカテゴリを含む
- レビュー数が多い場所や評価が高い場所を優先し、上位50件を抽出
- Places API (New) を用いて、各場所の詳細情報（名称、住所、営業時間、料金帯、カテゴリなど）やレビューコメントを取得
- 結果を JSON ファイルとして保存

### 02. データの整形と保存 (places\_cleaner.py)

- 取得した JSON ファイルを読み込み、不要なフィールドを整理
- 場所ごとの基本情報（名称、住所、緯度経度、カテゴリ、評価値、要約など）を抽出して Parquet ファイルに変換
- レビューコメントは別のデータフレームに分割し、場所IDと紐づけて保存
- 出力として「gmap\_places.parquet」と「gmap\_reviews.parquet」を生成

---

# 04

## Google Map分析のデモの内容

---

### 04.Gemini による観光戦略の生成 (places\_strategy)

- 各場所の詳細情報とレビューコメントを結合し、入力データを構築
- Gemini 2.5 Pro に入力し、観光コンテンツの傾向やターゲット層を踏まえたプロモーション戦略を生成
- 出力内容には観光体験のトレンド分析、訪問者の属性別嗜好、季節やターゲットに応じた施策提案を含める

### 05.BigQueryにデータを導入して可視化 (gmap\_place\_features.sql)

- 整形済みの Parquet ファイルを BigQuery に取り込み、分析用テーブルを作成
- SQL により新たな特徴量を追加
- 可視化基盤として Looker Studio に接続し、観光地の分布や人気傾向を地図やグラフで表示
- ダッシュボード化することで、観光戦略や施策改善に活用可能なインサイトを提供



# 05

## デモの強み

### GCP 上で統合管理でき安価

YouTube と Google Maps はどちらも GCP 経由で API が提供されており、他の有料データソースと比べても安価に利用できる。

データの取得から保存・分析までを一元的に管理でき、運用コストも抑えられる。

### 観光客のリアルな視点を反映

実際の動画コンテンツや Google Maps のレビューコメントを直接利用することで、観光客の体験や評価に基づいた一次情報を収集できる。

これは自治体や観光事業者の公式データでは得られないリアルなインサイトとなる。

### 戦略立案に直結する実用性

Gemini による戦略生成と BigQuery・Looker Studio による可視化を組み合わせることで、観光トレンドの把握から具体的な施策提案まで一連のプロセスを自動化できる。

実際のプロモーション戦略や施策改善にすぐ活用できる実用性が高い。

### 国内・訪日観光客のバランス

#### 国内・訪日観光客のバランス

YouTube と Google Maps は多様な言語・国籍のユーザーから観光地に関する情報が集まる。

YouTube では海外観光客による旅行動画や多言語の発信が多く、Google Maps でも国内外の観光客が投稿したレビューを豊富に取得でき、国内・訪日観光客の関心や行動傾向をより正確に比較・分析できる。

# 06

## デモのさらなる改良案

### 他の観光地にも対応

- 現在は川越を対象としているが、キーワード選定・NGワード設定・ジャンル分類などを Embedding モデルや Gemini により自動化することで、他地域への展開が容易になる。
- 自動化により手作業を減らし、全国・海外の観光地に対しても同じフレームワークでデータ収集・分析が可能となる。
- 複数の観光地を同条件で分析できるため、地域間比較や競合観光地との差別化分析にも応用できる。

### 他のSNSに応用

- Instagram API から投稿のキャプション・コメント・ハッシュタグを取得して分析することで、ビジュアルやトレンド志向の高い若年層データも活用可能。
- Reddit や Tripadvisor など、海外で広く利用されているレビュー・コミュニティサイトも統合すれば、インバウンド観光客の意見をさらに反映できる。
- SNSごとの利用層・表現傾向の違いを分析することで、媒体別のプロモーション戦略を立てることも可能になる。

### 新たな機能の追加

- YouTube のコメント欄を Gemini に学習させる対象に加えることで、動画視聴者の感想や反応を分析でき、より「消費者のリアルな声」に基づいた戦略立案が可能となる。
- 投稿データを国・地域・言語ごとに分類する機能を追加することで、国別・エリア別のニーズに最適化された戦略提案ができる。
- さらに、公式観光サイトや自治体の発信情報と自動照合する仕組みを組み込むことで、誤情報や古いデータに基づく誤分析を防止できる。
- コメントや字幕を感情分析モデルで分類し、「満足・不満・驚き・感動」などの感情マップを作成。より人間中心の観光評価指標を構築する。