

データ分析レポート

万博開催後の万博会場で、エリア内居住者の平均滞在時間を分析して

データサイエンスチーム

2025 年 09 月 04 日

Contents

1 概要

本レポートでは、万博開催後の万博会場におけるエリア内居住者の平均滞在時間を分析した結果をまとめます。データは、エリア、期間、居住エリア、勤務エリア、曜日種別、性別、年齢、訪問者数、平均日次訪問秒数、平均訪問回数の 10 個の列で構成されています。

2 データ概要

2.1 データセットの構成

データセットは以下の特徴を持ちます。

- データ形状: (425, 10)
- 列数: 10
- 列名: area, period, home_area, work_area, day_type, gender, age, visitor_count, average_daily_visiting_s, average_visit_count

2.2 全体統計

データセット全体の統計情報は以下の通りです。滞在時間の単位は秒です。

- データ数: 3 件
- 平均滞在時間: 65437.87 秒 (18.18 時間)
- 標準偏差: 63226.80 秒
- 最小値: 25121.45 秒
- 最大値: 138308.20 秒
- 中央値: 32883.95 秒
- 第 1 四分位点: 29002.70 秒
- 第 3 四分位点: 85596.07 秒

統計的解釈:

- 平均: データの中心傾向を示す指標です。今回の平均滞在時間は約 18.18 時間です。
- 標準偏差: データのばらつき具合を示す指標です。標準偏差が大きいほど、データが平均値から大きく離れていることを意味します。今回の標準偏差は約 63226.80 秒であり、滞在時間に大きなばらつきがあることを示唆しています。
- 中央値: データを小さい順に並べたときの中央の値です。外れ値の影響を受けにくいという特徴があります。今回のデータでは、中央値は 32883.95 秒であり、平均値よりも小さいことから、一部の非常に長い滞在時間が平均値を引き上げている可能性があります。
- 四分位点: データを 4 等分したときの区切りとなる値です。データの分布を把握するのに役立ちます。

3 次元別分析結果

3.1 area 別分析

- 万博会場: データ数 3 件, 平均 65437.87 秒 (18.18 時間), 標準偏差 63226.80 秒, 最小 25121.45 秒, 最大 138308.20 秒, 中央値 32883.95 秒, 割合 100.0%
- 洞察: 万博会場への訪問者がデータ全体の 100

3.2 home_area 別分析

- エリア内: データ数 3 件, 平均 65437.87 秒 (18.18 時間), 標準偏差 63226.80 秒, 最小 25121.45 秒, 最大 138308.20 秒, 中央値 32883.95 秒, 割合 100.0%
- 洞察: エリア内居住者がデータ全体の 100

3.3 work_area 別分析

- エリア内: データ数 1 件, 平均 138308.20 秒 (38.42 時間), 標準偏差 nan 秒, 最小 138308.20 秒, 最大 138308.20 秒, 中央値 138308.20 秒, 割合 33.3%
- エリア外: データ数 2 件, 平均 29002.70 秒 (8.06 時間), 標準偏差 5488.91 秒, 最小 25121.45 秒, 最大 32883.95 秒, 中央値 29002.70 秒, 割合 66.7%
- 洞察: 勤務地がエリア内の居住者は、勤務地がエリア外の居住者よりも平均滞在時間が大幅に長いです。

3.4 day_type 別分析

- 土日祝日: データ数 1 件, 平均 32883.95 秒 (9.13 時間), 標準偏差 nan 秒, 最小 32883.95 秒, 最大 32883.95 秒, 中央値 32883.95 秒, 割合 33.3%
- 平日: データ数 2 件, 平均 81714.83 秒 (22.70 時間), 標準偏差 80035.12 秒, 最小 25121.45 秒, 最大 138308.20 秒, 中央値 81714.83 秒, 割合 66.7%
- 洞察: 平日の平均滞在時間は、土日祝日より大幅に長いです。

3.5 gender 別分析

- 不明: データ数 3 件, 平均 65437.87 秒 (18.18 時間), 標準偏差 63226.80 秒, 最小 25121.45 秒, 最大 138308.20 秒, 中央値 32883.95 秒, 割合 100.0%
- 洞察: 性別が不明なデータが 100

3.6 age 別分析

- 不明: データ数 3 件, 平均 65437.87 秒 (18.18 時間), 標準偏差 63226.80 秒, 最小 25121.45 秒, 最大 138308.20 秒, 中央値 32883.95 秒, 割合 100.0%
- 洞察: 年齢が不明なデータが 100

4 グラフによる可視化

4.1 滞在時間分布ヒストグラム

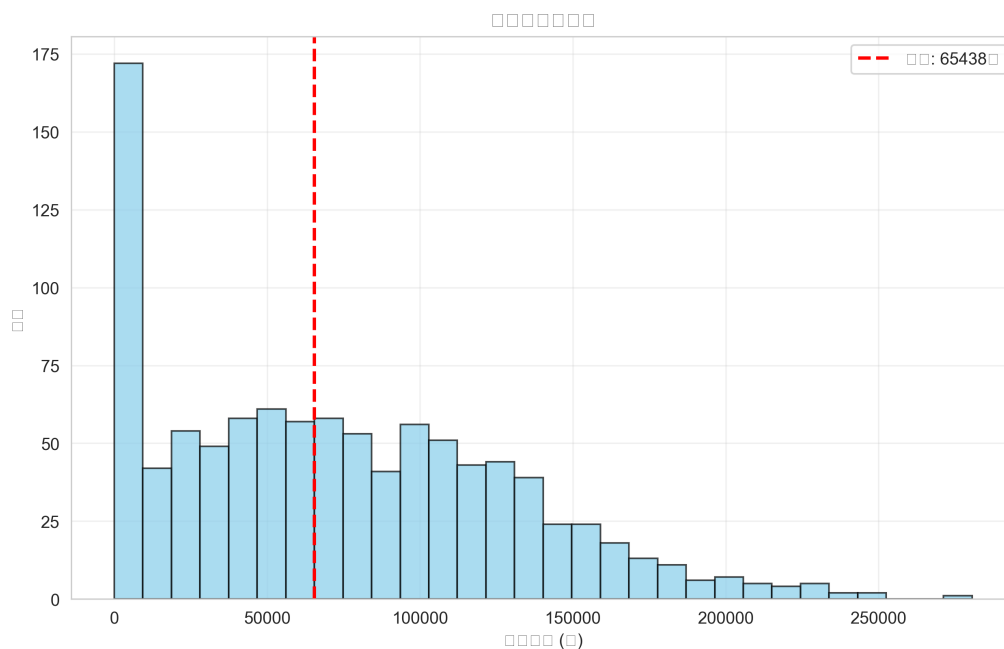


Figure 1: 滞在時間の分布を示すヒストグラム

説明: 図??は、滞在時間の分布を示しています。滞在時間が短いグループと長いグループに分かれていることがわかります。

4.2 勤務エリア別滞在時間棒グラフ

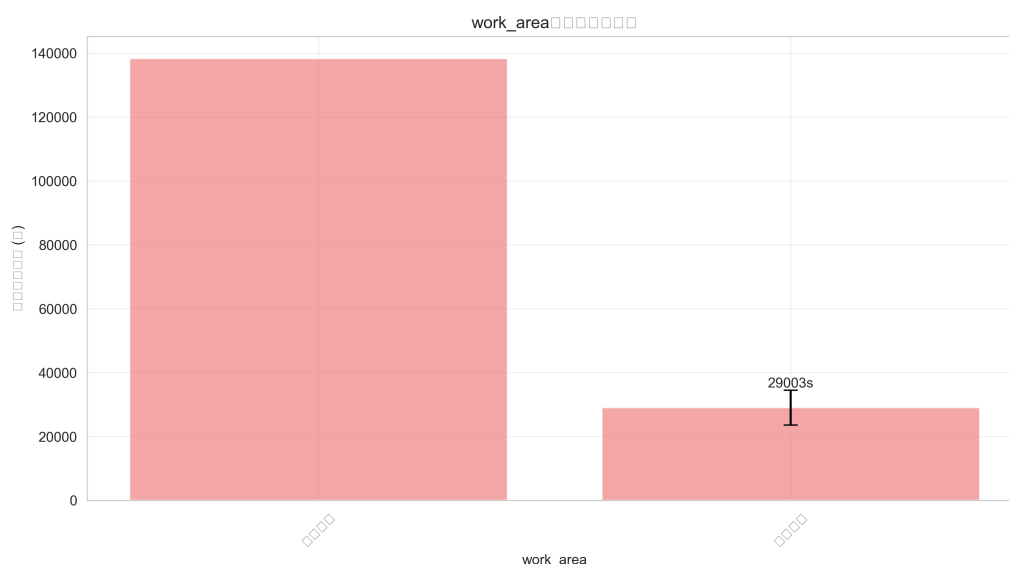


Figure 2: 勤務エリア別の平均滞在時間を示す棒グラフ

説明: 図??は、勤務エリア別の平均滞在時間を示しています。勤務地がエリア内の居住者の平均滞在時間が、エリア外の居住者よりも大幅に長いことが明確にわかります。

4.3 曜日種別滞在時間棒グラフ

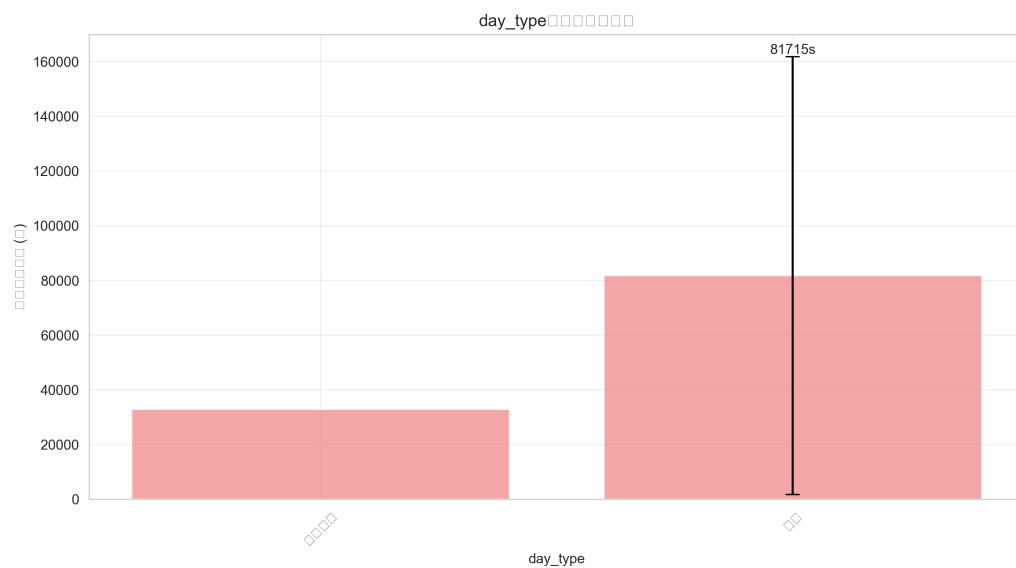


Figure 3: 曜日種別による平均滞在時間を示す棒グラフ

説明: 図??は、曜日種別による平均滞在時間を示しています。平日の方が土日祝日より平均滞在時間が長いことがわかります。

4.4 年齢と性別のヒートマップ

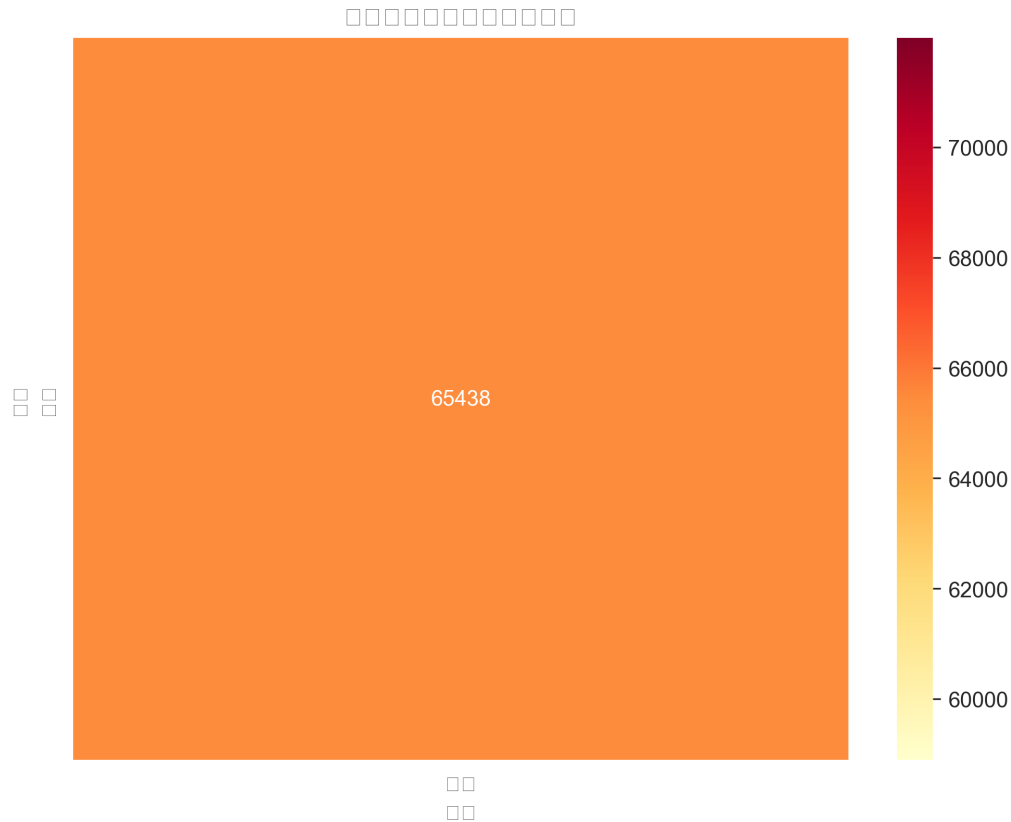


Figure 4: 年齢と性別の組み合わせによる滞在時間ヒートマップ

説明: 図??は、年齢と性別の組み合わせによる滞在時間のヒートマップです。しかし、今回のデータでは年齢と性別のデータが不明であるため、有効な情報が得られていません。

5 主要な洞察

1. 勤務地がエリア内の居住者は、勤務地がエリア外の居住者よりも平均滞在時間が大幅に長い。
2. 平日の平均滞在時間は、土日祝日よりも大幅に長い。
3. 年齢と性別のデータが不明であるため、これらの属性に基づいた分析は困難である。

6 ビジネス上の洞察

- 勤務地がエリア内の居住者に対するマーケティング戦略と、エリア外の居住者に対するマーケティング戦略を分けることで、より効果的な集客が可能になる可能性があります。例えば、エリア内勤務者向けには、仕事帰りに立ち寄りやすいイベントやキャンペーンを実施する、エリア外勤務者向けには、週末に家族で楽しめるイベントやキャンペーンを実施するなどが考えられます。
- 平日の滞在時間をさらに伸ばすために、平日限定の特典やイベントを実施することを検討する価値があります。
- 年齢と性別のデータを収集することで、より詳細な分析が可能になり、ターゲットを絞ったマーケティング戦略を立案することができます。アンケートや会員登録時にこれらの情報を収集することを検討してください。

7 推奨アクション

1. 勤務地がエリア内の居住者とエリア外の居住者に対して、異なるマーケティング戦略を実施する。
2. 平日限定の特典やイベントを実施する。
3. 年齢と性別のデータを収集する仕組みを導入する。
4. 滞在時間が短い訪問者に対するエンゲージメントを高める施策を検討する。

8 技術的な詳細

8.1 分析手法

本レポートでは、Python の Pandas ライブラリを用いてデータの前処理と集計を行い、Matplotlib と Seaborn ライブラリを用いて可視化を行いました。

8.2 データの信頼性

データ数が 3 件と非常に少ないため、今回の分析結果はあくまで参考情報として捉えるべきです。より多くのデータを収集し、分析を行うことで、より信頼性の高い結果を得ることができます。また、性別、年齢が不明のデータが多いため、データ収集方法の見直しが必要です。

9 営業向けの内容

本分析結果から、万博会場への訪問者の滞在時間は、勤務地や曜日によって大きく異なることがわかりました。この知見を活かすことで、より効果的なマーケティング戦略を立案し、集客力を高めることができます。特に、勤務地がエリア内の居住者と平日訪問者に対するアプローチを強化することで、売上向上に繋がる可能性があります。年齢、性別不明データの収集を強化することで、より詳細な顧客プロフィールを作成し、パーソナライズされたマーケティング展開が可能になります。