

Analysing Immigration Trends Using Facebook Advertising Data

Ma Xinyi 2025.9.29

Motivation:

Traditional data (censuses, surveys) offer official migrant estimates but lag 1-2 years with incomplete coverage of origins and moves (Yildiz et al., 2024; Rampazzo et al., 2021). Facebook data provides real-time proxies via user demographics, tracked by Monthly Active Users (Zagheni et al., 2017). Corrected for biases, it boosts precision and policy support (Yildiz et al., 2024), though limited by private access constraints.

```
import requests
import pandas as pd
import json
# query arguments
args = {
    "token": "8db25b1327d531554c6d81f8059a0733",
    "platform": "facebook",
    "country": "JP",
    "gender": "0",
    "date_start": "2020-01-10",
}

response = requests.get(url='http://digitrace-datahub.
ndph.ox.ac.uk/api/v1/query', params=args)

# format response as dictionary
response = response.json()

# check status
print(response.get('status'))
print(response.get('message'))

# extract data as pandas dataframe
if response.get('status') == 200:
    data = pd.DataFrame(json.loads(response.get('data')))
    data.to_csv('JAPAN.csv', index=False)
```

Our contribution:

ESCAP's diverse region, with 40% of global migrants, needs timely data. We use Facebook data to analyze trends, filling a gap in Europe-focused studies (Zagheni et al., 2017), offering real-time insights for policy amid regional shifts (Yildiz et al., 2024).

```
import pandas as pd

# Read CSV file
df = pd.read_csv("/Users/machine/Desktop/research/GUY ABEL/51_US_US_F.csv")

filtered_13 = df[
    (df['age_min'] == 13) &
    (df['age_max'] == 999) &
    (df['geo_locations'].str.contains('countries', case=False, na=False)) &
    (df['geo_locations'].str.contains('recent', case=False, na=False))
]

filtered_18 = df[
    (df['age_min'] == 18) &
    (df['age_max'] == 999) &
    (df['geo_locations'].str.contains('countries', case=False, na=False)) &
    (df['geo_locations'].str.contains('recent', case=False, na=False))
]

filtered_expats_13 = df[
    (df['age_min'] == 13) &
    (df['age_max'] == 999) &
    (df['geo_locations'].str.contains('countries', case=False, na=False)) &
    (~df['geo_locations'].str.contains('recent', case=False, na=False))
]

filtered_expats_18 = df[
    (df['age_min'] == 18) &
    (df['age_max'] == 999) &
    (df['geo_locations'].str.contains('countries', case=False, na=False)) &
    (~df['geo_locations'].str.contains('recent', case=False, na=False))
]

with pd.ExcelWriter("US_F_filtered.xlsx") as writer:
    filtered_13.to_excel(writer, sheet_name="male_local_13", index=False)
    filtered_18.to_excel(writer, sheet_name="male_local_18", index=False)
    filtered_expats_13.to_excel(writer, sheet_name="male_expats_13", index=False)
    filtered_expats_18.to_excel(writer, sheet_name="male_expats_18", index=False)
```

Limitations:

Facebook data biases favor younger, urban groups, missing older/rural migrants. API reliance risks disruptions (e.g., fbRads issues), varying updates affect consistency, and lacking sub-national data limits insights. Privacy and ethical concerns also restrict applicability.

Future direction:

This research will deepen Facebook data analysis for ESCAP immigration trends, refining models for regional patterns like Southeast Asia labor or South Asia refugees. We plan real-time tracking with API tools, machine learning predictions, and partnerships to expand data on understudied nations.

1_Hongkong	17_Iran-no data	33_Pakistan	49_Tuvalu
2_Afghanistan	18_Japan-no data	34_Palau	50_UK(Great Britain and Northern Ireland)
3_Armenia	19_Kazakhstan	35_Papua New Guinea	51_US
4_Australia	20_Kiribati-no data	36_Philippines	52_Uzbekistan
5_Azerbaijan	21_Kyrgyzstan	37_Korea-no data	53_Vanuatu
6_Bangladesh	22_Lao	38_Russian Federation-no data	54_Viet Nam
7_Bhutan	23_Malaysia	39_Samoa	55_American Samoa
8_Brunei Darussalam	24_Maldives	40_Singapore	56_Cook Islands
9_Cambodia-no data	25_Marshall Islands	41_Solomon Islands	57_French Polynesia
10_China	26_Micronesia	42_Sri Lanka	58_Guam
11_Korea(Democratic)-no data	27_Mongolia	43_Tajikistan	59_Macao
12_Fiji	28_Myanmar	44_Thailand	60_New Caledonia
13_France	29_Nauru	45_Timor-Leste	61_Niue-no data
14_Georgia	30_Nepal	46_Tonga	62_Northern Mariana Islands
15_India	31_Netherlands	47_Turkey	download_data.py
16_Indonesia	32_New Zealand	48_Turkmenistan	filter_data.py

- Alexander, M., Polimis, K., & Zagheni, E. (2022). Combining social media and survey data to nowcast migrant stocks in the United States. *Population Research and Policy Review*, 41(1), 1-28.
- Rampazzo, F., Bijak, J., Vitali, A., Weber, I., & Zagheni, E. (2021). A framework for estimating migrant stocks using digital traces and survey data: An application in the United Kingdom. *Demography*, 58(6), 2193-2218.
- Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., & Rango, M. (2019). Quantifying international human mobility patterns using Facebook Network data. *PLoS one*, 14(10), e0224134.
- Yildiz, D., Wiśniowski, A., Abel, G. J., Weber, I., Zagheni, E., Gendronneau, C., & Hoorens, S. (2024). Integrating Traditional and Social Media Data to Predict Bilateral Migrant Stocks in the European Union. *International Migration Review*, 59(1), 90-118. <https://doi.org/10.1177/01979183241249969> (Original work published 2025)
- Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review*, 721-734.