

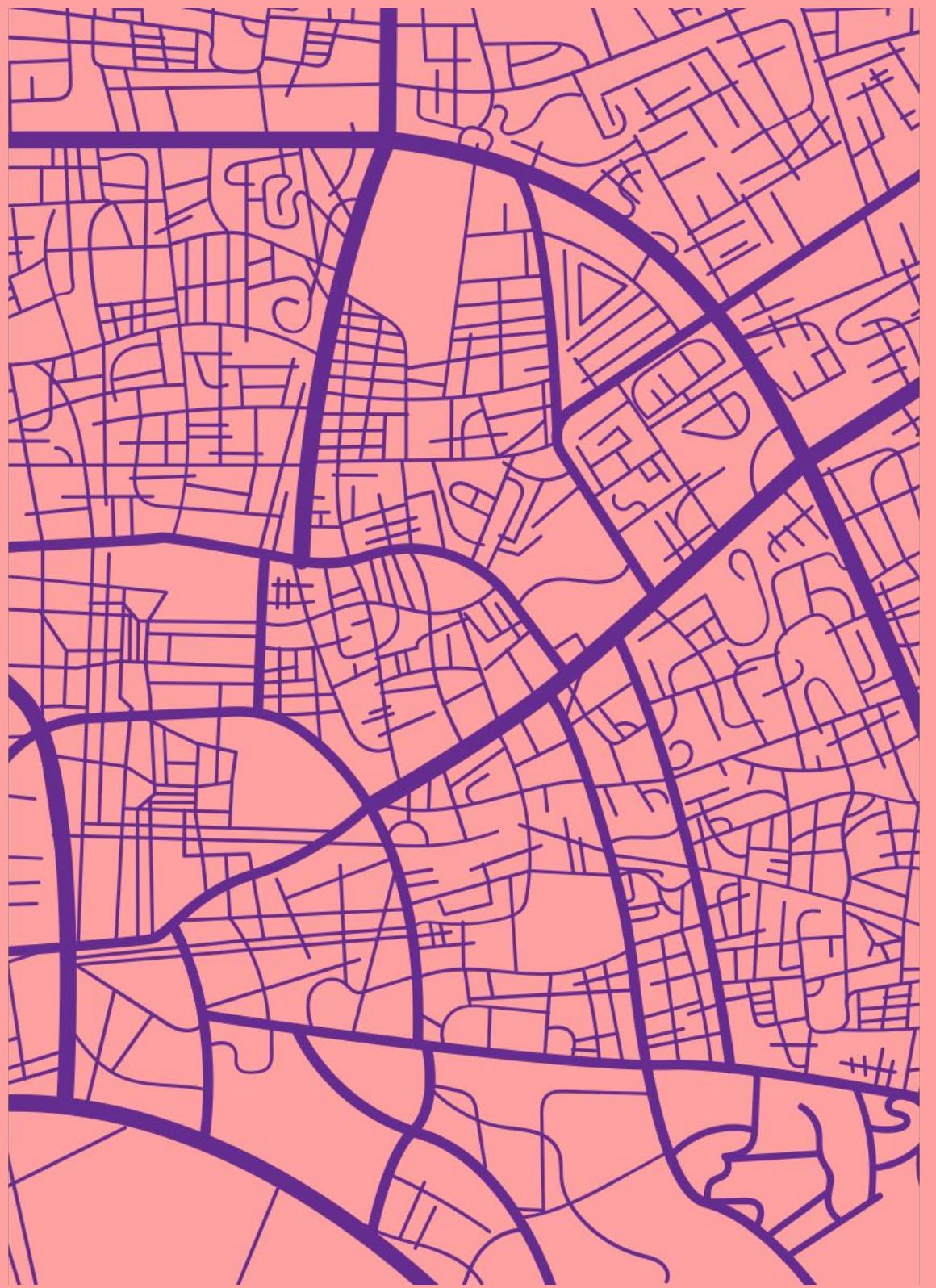


PowerThruData

2023

Predicting and Visualizing the Progression
of Marginal Villages: A Case Study of
Japanese Municipalities

Chloe T., Kaede H., Brook L., Angel P.



Introduction

Introducing the background and aims of our chosen topic



Implementation

Implementing our data-driven solution



Evaluation

Evaluating the strengths, limitations, and implications of our work for future research



Introduction

Background & Problem

Background

Japan's Ageing Population

- 1/10 aged 80 or over (Ng, 2023)
- World's highest % of people aged 65+ (Ng, 2023)
- By 2040, 34.8% will be aged 65+ (National Institute of Population and Social Security Research, 2023)
- Other nations, such as Italy (2nd)and Finland (3rd), are also facing an ageing populations (Washington, DC: Population Reference Bureau, 2019)
 - local problem that requires widely applicable solutions

Causes

1. Low birth rates
2. Increasing cost of living
3. Long working hours
4. Long life expectancy (e.g., centenarians)

Problem & Solution

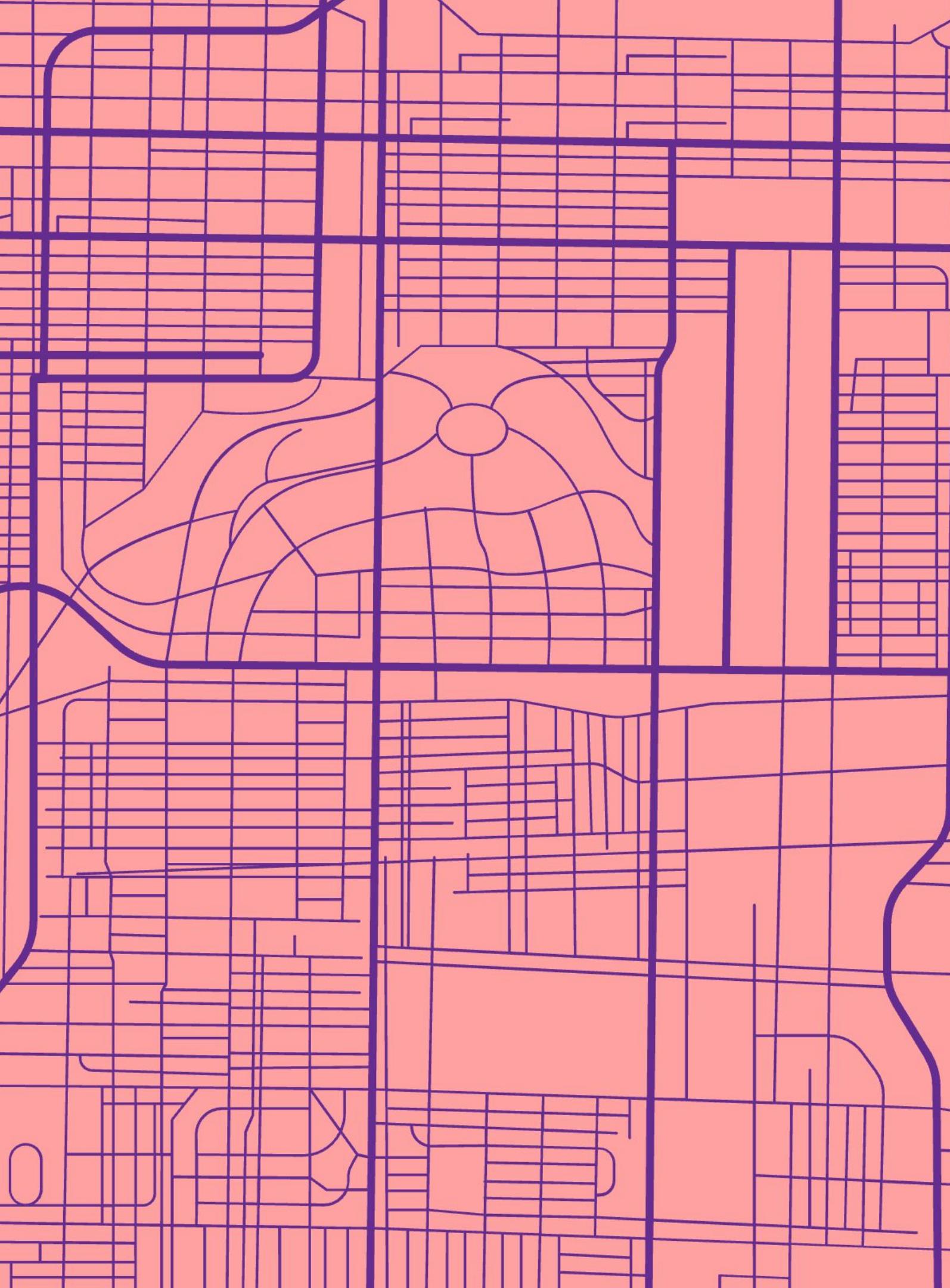
Localised problem: marginal villages (Sato, 2012)

- Lack of village management affects quality of life
- Arises from having 50%+ of local population being 65+ years old

Consequences (Niimura, 2009)

- Poor quality of life
 - Environmental issues (e.g., land and waste management)
 - Lack of security
 - Decreased economic growth = limited employment
- Connects to **Jobs and Social Protection**

Few good data visualisation or predictive tools on the progression of marginal villages (i.e., municipalities)



Solution

- Visualisation → PowerBI
- Prediction → Python linear regression model



Figure 1 (Tanaka, 2019)
Percentage of 65+ population over time

Predictive model provides insight into contribution of variables to ageing population
→ allows for recommendations on high-priority factors

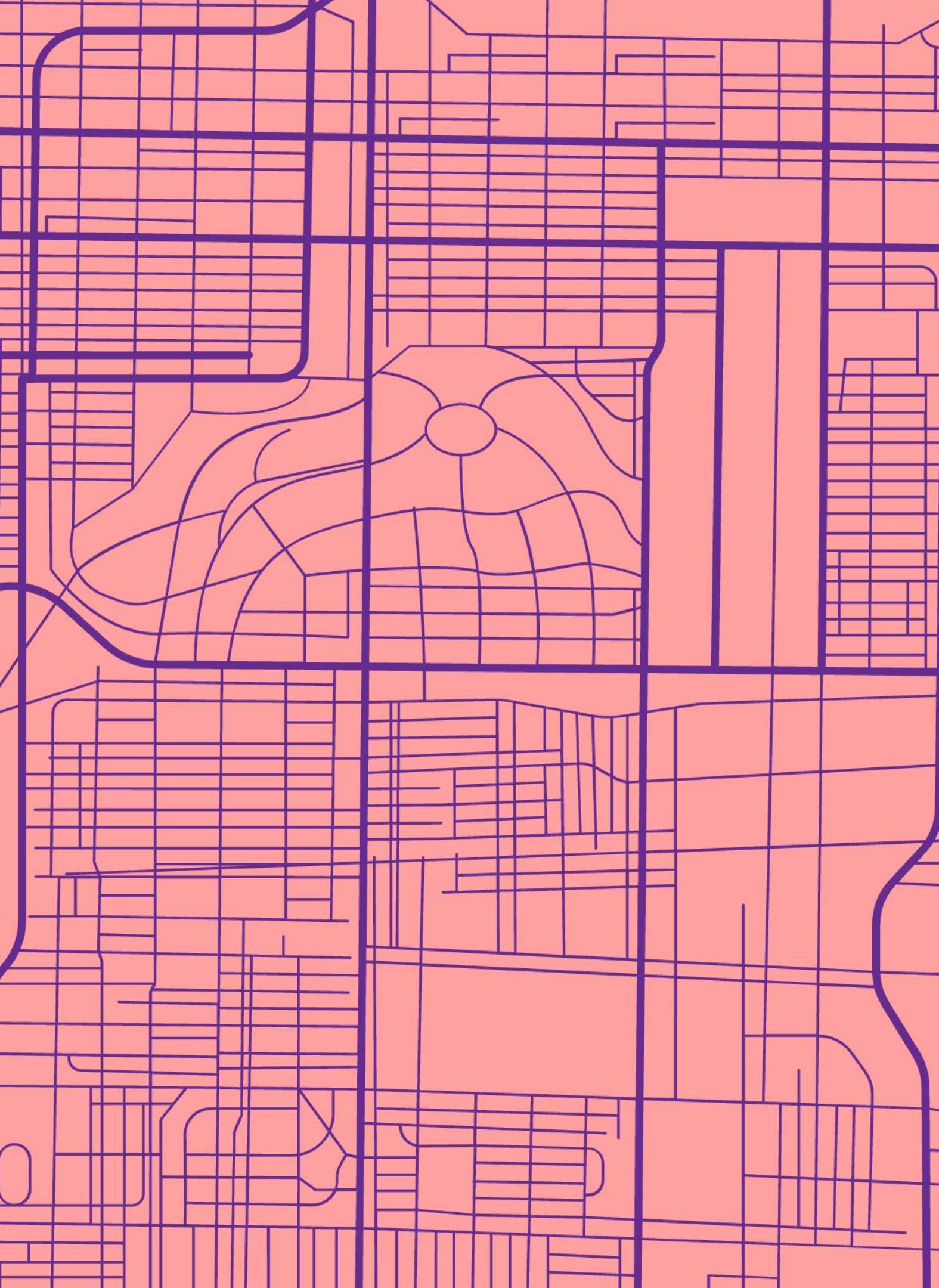
Methods

Data Wrangling



Portal Site of Official Statistics of Japan
→ Open-access governmental statistics

- Selected variables for all municipalities from 2000 to 2020 in 5-year intervals
 - Original dataset: 1,916 municipalities
 - Excluded 262 municipalities with missing or miscoded data across any year
 - Final dataset: **N=1,654 municipalities**



Variables

Based on Ministry of Finance (2011) and Niinuma (2009)

Demographic IVs

- % of employed persons
- Birth rate per 1,000
- Mortality rate per 1,000

No. of births/deaths*1,000

Total municipal population

Area IVs

- % of inhabitable area over total municipal area (ha)
- Number of educational infrastructure (i.e., kindergarten, elementary, and upper/lower secondary infrastructure)

DV

- % of total municipal population over 65 years old

Total over 65

Total municipal population

Data Analysis

Tableau Visualisation

- % of population over 65 per municipality
- Summary per prefecture
- Correlations between predictor and outcome variables
- Filters (e.g., by year, region, % of population over 65)

Linear Regression Model

- Run within municipalities

% of population over 65 ~ year + % of employed persons +
birth rate + mortality rate + % of inhabitable land + no. of
educational infrastructure



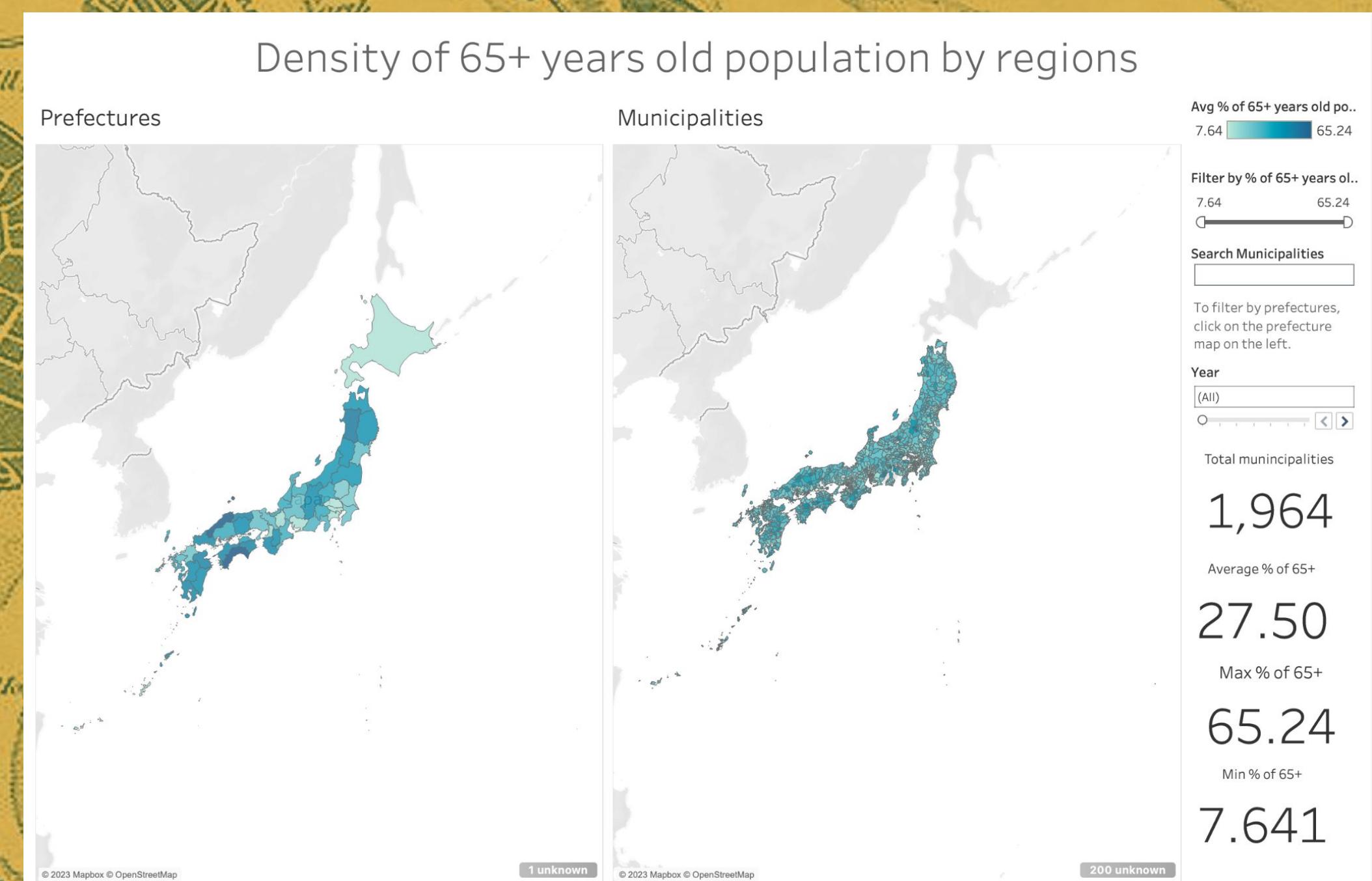
Implementation

Descriptive Statistics

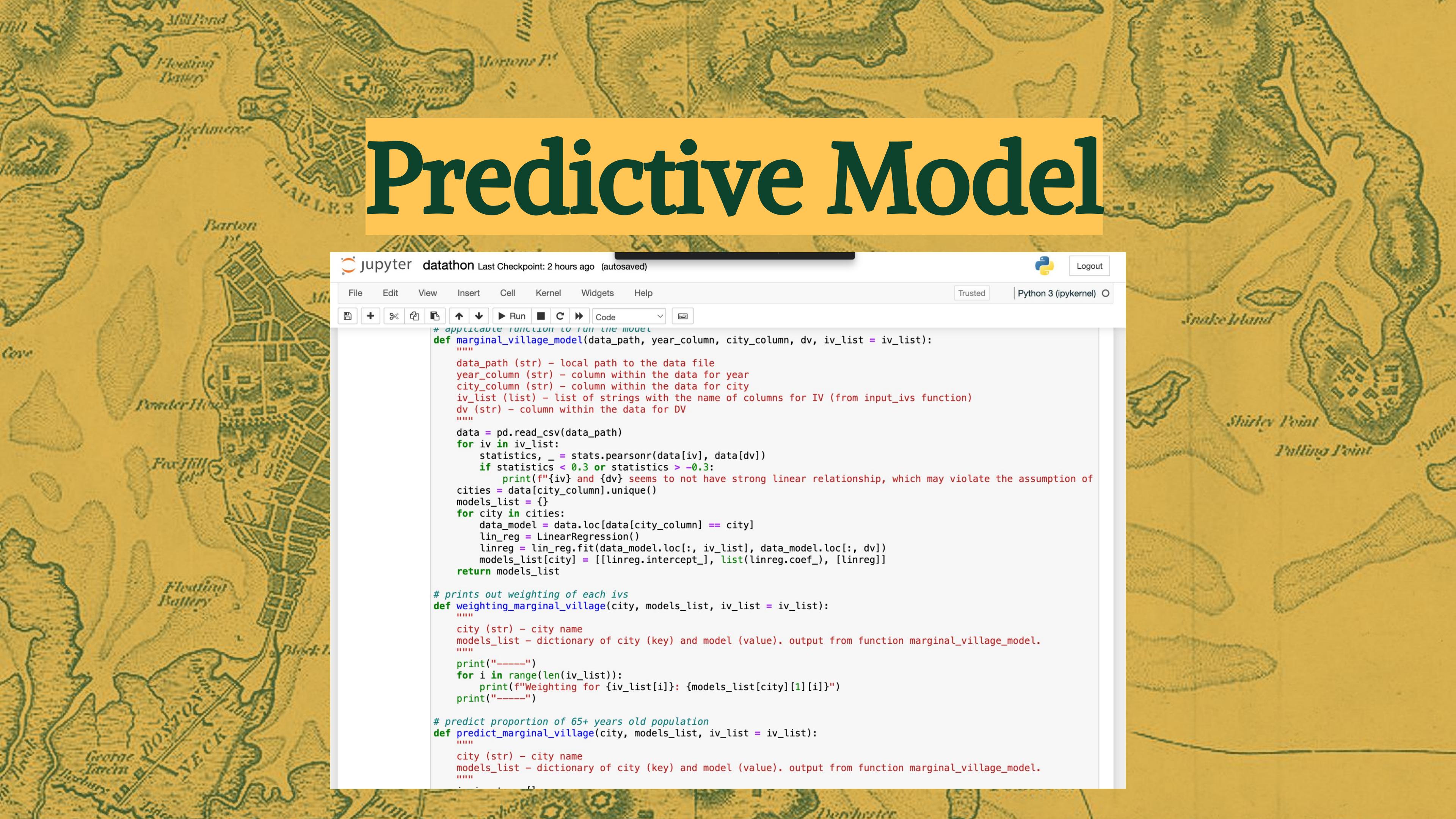
	2000	2005	2010	2015	2020
Municipalities	1654.00	1654.00	1654.00	1654.00	1654.00
M Total Population	83276.04	84099.16	84620.18	84330.74	84023.97
M % of Population > 65	21.46	24.35	27.08	30.89	33.74
M % of Employed Persons	49.96	48.90	47.29	48.06	47.96
M Birth Rate	8.66	7.62	7.49	7.07	5.85
M Mortality Rate	9.22	10.49	11.86	13.08	14.04
M % of Inhabitable Land	52.57	52.57	52.39	52.59	52.70
No of Educational Infrastructure	34.42	33.47	32.25	30.15	28.14

Note. Descriptive statistics were averaged across municipalities. M = mean.

Tableau Visualisation



Predictive Model



A historical map of Boston, Massachusetts, serves as the background for this image. The map shows the city's layout from an earlier era, with streets, landmarks, and bodies of water labeled. Key features include the Charles River, Fort Hill, and several points along the coastline.

```
jupyter datathon Last Checkpoint: 2 hours ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help
Trusted Python 3 (ipykernel) ○

# applicative function to run the model
def marginal_village_model(data_path, year_column, city_column, dv, iv_list = iv_list):
    """
    data_path (str) - local path to the data file
    year_column (str) - column within the data for year
    city_column (str) - column within the data for city
    iv_list (list) - list of strings with the name of columns for IV (from input_ivs function)
    dv (str) - column within the data for DV
    """
    data = pd.read_csv(data_path)
    for iv in iv_list:
        statistics, _ = stats.pearsonr(data[iv], data[dv])
        if statistics < 0.3 or statistics > -0.3:
            print(f'{iv} and {dv} seems to not have strong linear relationship, which may violate the assumption of')
    cities = data[city_column].unique()
    models_list = {}
    for city in cities:
        data_model = data.loc[data[city_column] == city]
        lin_reg = LinearRegression()
        linreg = lin_reg.fit(data_model.loc[:, iv_list], data_model.loc[:, dv])
        models_list[city] = [[linreg.intercept_, list(linreg.coef_), [linreg]]]
    return models_list

# prints out weighting of each ivs
def weighting_marginal_village(city, models_list, iv_list = iv_list):
    """
    city (str) - city name
    models_list - dictionary of city (key) and model (value). output from function marginal_village_model.
    """
    print("-----")
    for i in range(len(iv_list)):
        print(f"Weighting for {iv_list[i]}: {models_list[city][1][i]}")
    print("-----")

# predict proportion of 65+ years old population
def predict_marginal_village(city, models_list, iv_list = iv_list):
    """
    city (str) - city name
    models_list - dictionary of city (key) and model (value). output from function marginal_village_model.
    """
```



Evaluation

Summary of Implementation

Power BI dashboard: allows for visualisation of marginal village progression per municipalities/prefectures

Linear regression model: predicts progression of marginal villages per municipality and estimates contribution of significant factors



Strengths

- **Novel:** marginal villages are an important local issue but not well-addressed with data
 - Clear visualisation
- **Local:** municipality-specific predictive models
- **Applicable:** predictive function can be adapted to other regions with an ageing population



Limitations

- **Data:** limited high-quality data of Japanese municipalities
- **Variables:** choice of predictors and exclusion criteria must be more systematic and evidence-based
- **Model evaluation**
 - Doesn't account for acceleration of effect
 - Missing assumption checks and appropriate variable transformation
 - Missing variance explained for models

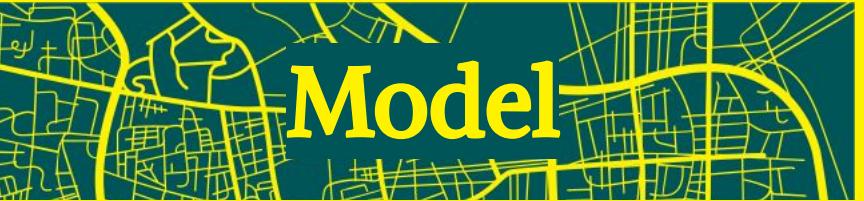


Future Directions



Improve data wrangling and analysis

- Use evidence-base on ageing population and marginal villages to select variables of interests



Improve model accuracy and evaluation

- Incorporate evaluation into Python function
 - Assumption checks
 - Variance explained
- Polynomial regression model to account for acceleration
- Machine learning
 - Training and testing sets



Improve flexibility of visualisation and predictive model

- Visualise predictions for ageing population
- Allow for different types and numbers of predictors/outcomes
- Evaluate applicability to other regions facing ageing population



A dense, abstract map of a city street network in blue on a light orange background. The map features a complex web of streets, with some major roads highlighted in thicker blue lines. Street names are visible in various orientations across the grid.

Thank you!

Contact Us

✉ chloe.turcan.20@ucl.ac.uk

✉ kaede.hasegawa.20@ucl.ac.uk



References

- D'AMBROGIO, Enrico. Japan's Ageing Society. EPRS: European Parliamentary Research Service, 2020.
- Ministry of Finance. 過疎地域等における集落の状況に関する現況把握調査. 総務省, 2011.
- . 過疎地域における集落の状況に関する現況把握 調査最終報告.
- National Institute of Population and Social Security Research. "日本の将来推計人口(令和5年推計) | 国立社会保障・人口問題研究所." www.ipss.go.jp, 26 Apr. 2023, www.ipss.go.jp/pp-zenkoku/j/zenkoku2023/pp_zenkoku2023.asp.
- Ng, Kelly. "Japan Population: One in 10 People Now Aged 80 or Older." BBC News, 19 Sept. 2023, www.bbc.co.uk/news/world-asia-66850943.
- Niinuma, Shiori. "「限界集落」における集落機能の維持と住民生活の持続可能性に関する考察—東京都西多摩郡檜原村 M 集落の事例から—." E-Journal GEO, vol. 4, 2009, pp. 21-36.
- Official Statistics of Japan. " Municipality Data | System of Social and Demographic Statistics(SSDS) | Search by Areas." Portal Site of Official Statistics of Japan, 2023, www.e-stat.go.jp/en/regional-statistics/ssdsvview/municipality. Accessed 5 Nov. 2023.
- Sato, Makoto. 「限界集落」論と北海道の農村社会. 北海学園大学開発研究所, 2012.
- Suzuki, Yusuke. IMPACT of JAPAN'S AGING POPULATION in 2030 – FOCUSING on the EFFECT on the SOCIAL SECURITY SYSTEM and LOCAL ECONOMIES – . Mitsui & Co., July 2023.
- Tanaka, Yusuke. 【限界集落】ふるさとの消滅は何を意味するのか." NewsPicks, 21 June 2019, newspicks.com/news/3950529/body/. Accessed 6 Nov. 2023.