

[Content](#)

Content

<u>Software/Tools Needed</u>	1
<u>Libraries Needed</u>	1
<u>Potential Problems</u>	1
<u>Possible Output Files</u>	1
<u>Topics Considered</u>	1
<u>Topic 1</u>	2
Key Factors for Explanation	2
Datasets Needed	2
Datasets Available on OpenDOSM	3
Layout Concept Draft	3

Content

Software/Tools needed:

- MS Visual Studio Code
- IBM Watson Studio
- IBM Cognos Analytics
- Terminal

Libraries needed:

- Pandas
- Matplotlib
- Folium
- Plotly express
- Dash (html, dcc)
- *Seaborn
- *NumPy
- *BeautifulSoup (for web scraping)

- Use GitHub to host the project

Potential Problem(s):

- hosting issue (if cannot host a website then try on Cognos Analytics to create dashboard)
- unable to access datasets

Possible Output Files:

- HTML file for the dashboard
- Excel worksheet for data tables, analysis, visuals
- Python notebook file for data wrangling, analysis, visuals
- Python coding file for dashboard

Topic Considered:

- Johor Population (total, male vs female, newborn by years, etc.)
- OR
- Malaysia Population/Crime Record/Immigrants etc.

TOPIC 1: Newborn Trend in Malaysia (Overall)

Purpose: Display the newborn trend in Malaysia over the years and some key factors/correlated factors (not necessary causation)

Descriptive and diagnostic factors (from the perspective of social and economy)

Descriptive	Diagnostic
<ul style="list-style-type: none">• Ratio of newborn to number of people in the age group suitable to give birth compared by years (<i>area chart</i>)• Ratio of married to single (<i>area chart</i>)• Marriage percentage for age 25 – 40/by race• Number of people in the age group suitable to give birth compared by years (<i>line chart</i>)	<ul style="list-style-type: none">• Income/annual salary by years (median/ mean, <i>line chart</i>)• GDP• Living cost by years (median/average, <i>line chart</i>)• Inflation• Unemployment rate• Tax• House price• Consumer goods price• Health indices like BMI by years (<i>line chart</i>)• Poverty by years (<i>line chart</i>)• Income inequality by years (<i>line chart</i>)• reasons of not getting marriage/birth (<i>wordcloud</i>)

Datasets required:

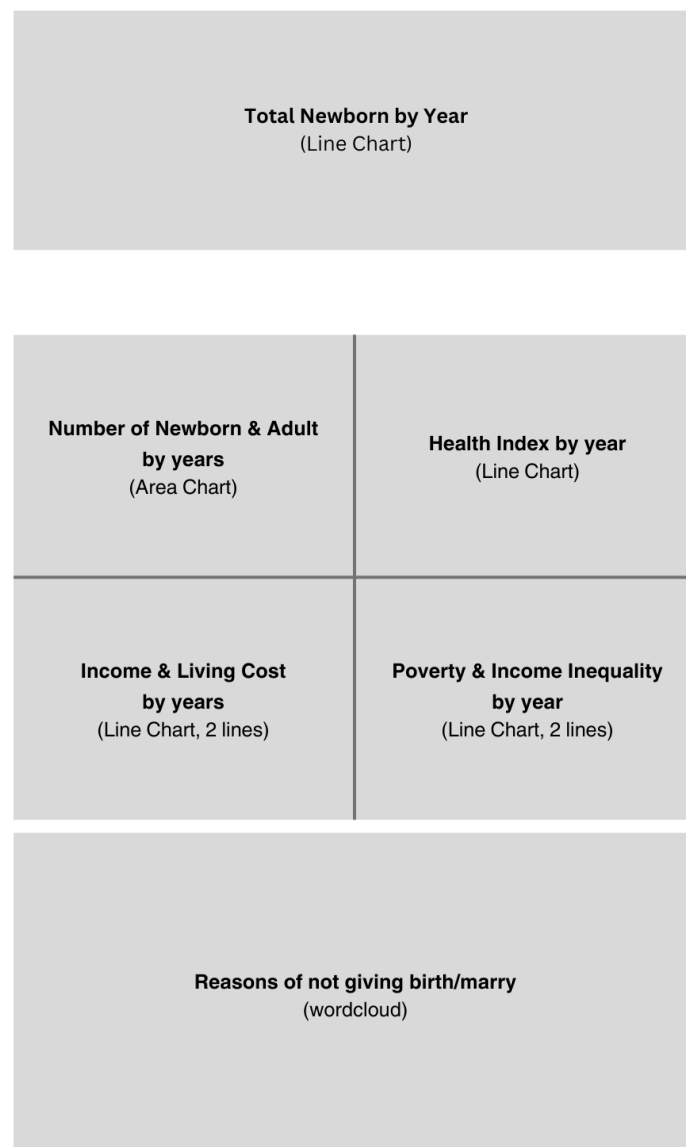
- Population by age group
- Newborn/birth by year, sex and ethnicity
- income/annual salary
- GDP
- living cost
- inflation
- unemployment rate
- tax
- Poverty
- Income inequality
- marriage percentage for age 25 – 40/by race
- survey data of reasons of not getting married/pregnant
- health indices like BMI

Datasets available on OpenDOSM:

Population	by year, state, age, sex, ethnicity
Birth	by year, state, sex, ethnicity
Income	mean and median by year, state
Expenditure	income (mean & median), expenditure and poverty by year, state
Poverty	by year, state
Income inequality	by year, state
GDP	
Consumer prices	

Dashboard layout concept draft:

Newborn Trend in Malaysia



Data source:

for personal learning purposes only