IDEAS

1. General attributes

a. Crude Death Rate by State

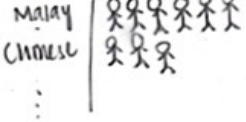
i) choropleth map

ii) proportional symbol map

b. Crude Death Rate by Ethnic Group

i) donut chart

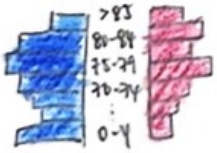
ii) isotype



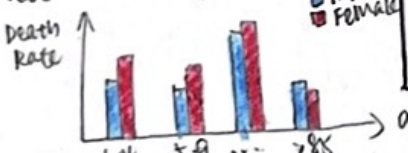
iii) pie chart

c. Crude Death Rate by Gender & Age

i) population pyramid

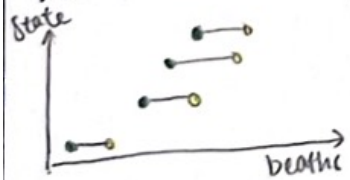


ii) Grouped Bar chart

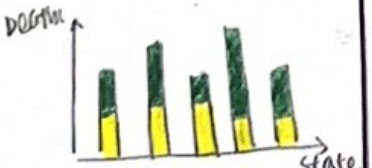


d. Deaths caused by medically & non-medically certified

i) dumbbell chart

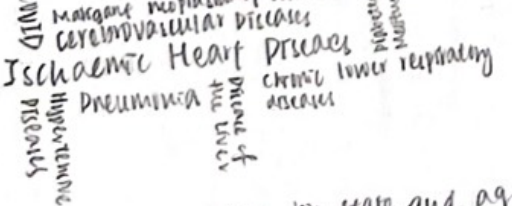


ii) Stacked Bar chart



e. Principal cause of death

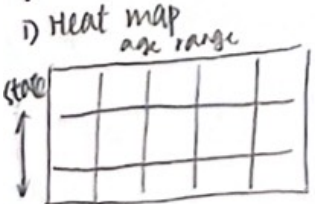
i) Word Cloud



ii) Pie Chart

f. Deaths of population by state and age

i) Heat map

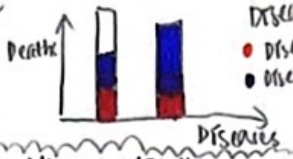


ii) dot map

g. Premature Mortality rate by state

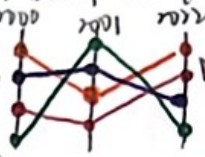
i) dot matrix

ii) Stacked Bar graph

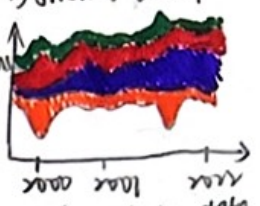


h. Time series of deaths (2000 - 2022)

i) Bump Chart



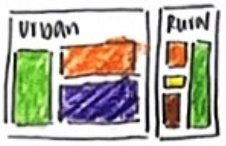
ii) Stream Chart



2. Medically & Non-medically attributes

a. Ten Principal of deaths by Stratum

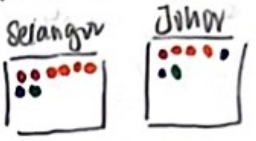
i) Tree Map



ii) Donut Chart

b. Ten principal cause of deaths by state

i) dot matrix



FILTER

1.a.ii) Proportional Symbol map

The pie chart might be too small & unclear

Better shown in 1.b.ii) !!

1.e.ii) Pie Chart

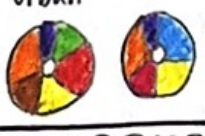
Too many groups, unclear

Better for ≤ 6 groups

1.f.ii) choropleth map

Have to create multiple maps for every age ranges, easy to compare between diff. states but hard to compare between diff. age ranges

2.a.ii) Donut Chart



This two charts are not in same scale

Misleading when comparing two stratums!

CATEGORIZE

Cause of death in Malaysia

- General
  - Crude Death Rate (1.a-c & f)
  - Cause of deaths (1.e & g & h)
- Compare between medically & non-medically
  - Cause of Deaths in each category (2.a-b)
  - Comparison of deaths (1.d)

COMBINE & REFINE

- Crude Death Rate

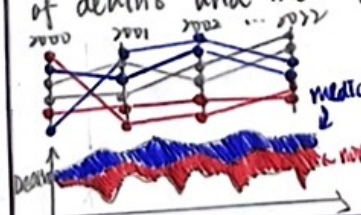
1.a.i) can be used to show total crude death rate in each state, while 1.f.i) shows the death rate of different age ranges of the state

1.b.ii) can be selected over 1.b.i) & 1.b.iii) as it clearly shows the data no matter how many races are there and also the isotype symbol can also enhance understanding

1.c.i) can be selected over 1.c.ii) as the population pyramid provides extra info

- Cause of Deaths

1.h.i) & 1.h.ii) can be combined to show the time series of deaths and the major cause of deaths



- Comparison of death

1.d.i) can be selected over 1.d.ii) as it better shows the difference.

1.g.ii) can be selected over 1.g.i) as it better compare the rate

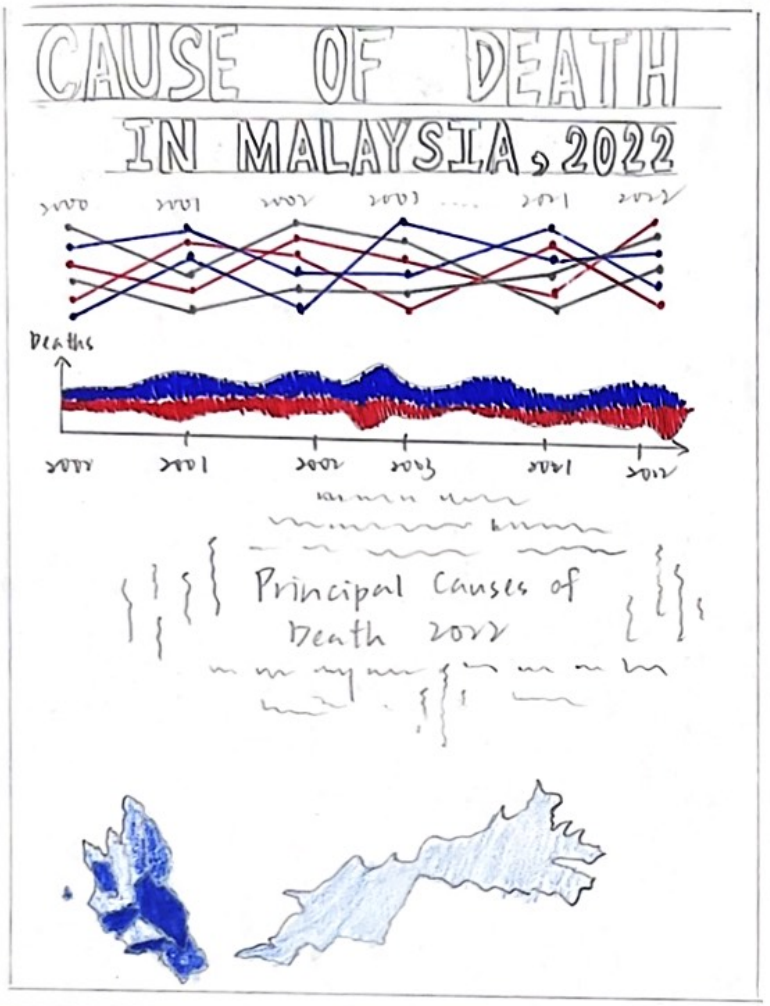
- Cause of deaths in each category

2.a.i) & 2.b.i) can be combined

QUESTIONS

1. Is too complicated to show all causes of deaths? Shall I focus on the top 5 / top 10? use filter function?
2. Has the data been normalised for certain graphs, such as choropleth map, Heat map and etc.?
3. Is it possible to draw all graphs using vega-lite? Any restrictions?

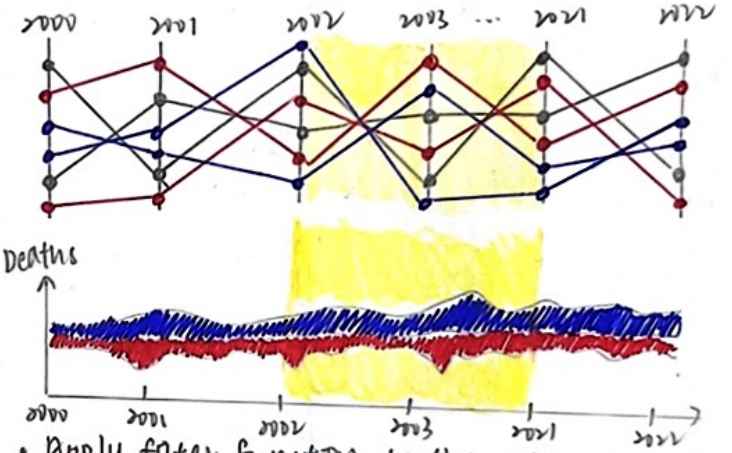
LAYOUT



Title: Partitioned poster
Author: Chai Ke Rou
Date: 29/9/2024
Sheet: 2
Task: Design a story-telling poster

OPERATION

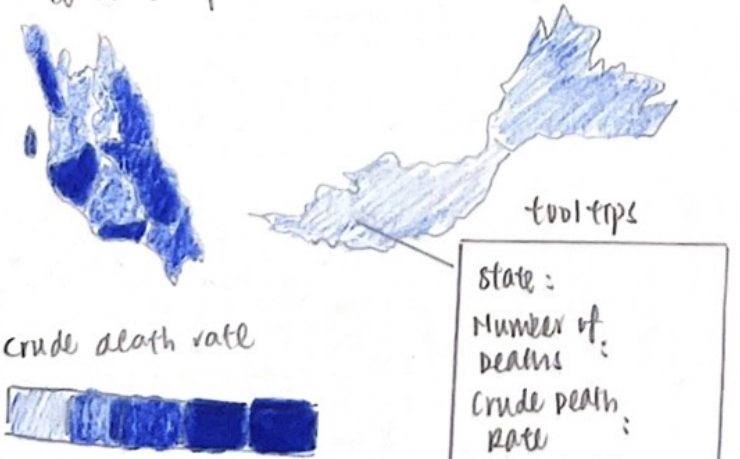
- Apply blushing function to the stream graph use stream graph to highlight certain time period.



- Apply filter function to the category in stream graph, which filter out relevant data in bump chart & word cloud..

FOCUS

- No major focus in this data visualisation
- All idioms presented show key components of the visualisation
- The tooltips in choropleth map is important, it provide details about the number of deaths, crude death race & the top 3 major causes of death in the state. This helps to come up with useful insights.



DISCUSSION

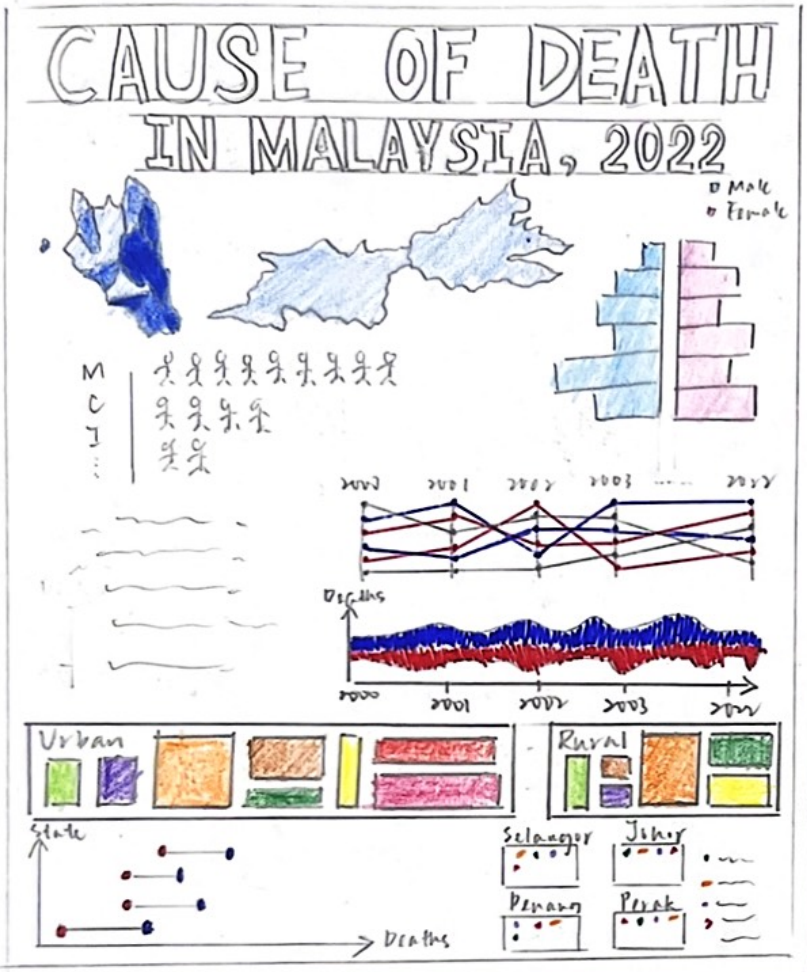
Advantages

- Well-structured, simple and clear layout
- Interative function (brushing function) is used to make the user easier to read and compare

Disadvantages

- The sequence of the idioms are mixed up, should guide the user from reading 2022 year relevant data to comparison charts.
- Lack of useful information, such as crude death rate by age ranges & gender, comparison idioms between medically & non-medically causes of death and etc.
- Too simple data visualisation, doesn't provide further information for reader to understand the domain and main topic.

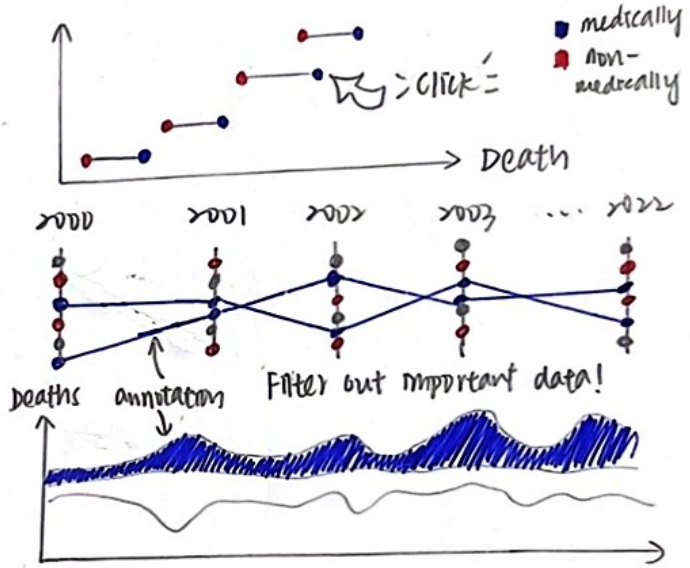
LAYOUT



Title: Partitioned Poster
Author: Chai Ke Ron
Date: 29/9/2024
Sheet: 2
Task: Design an infographic poster

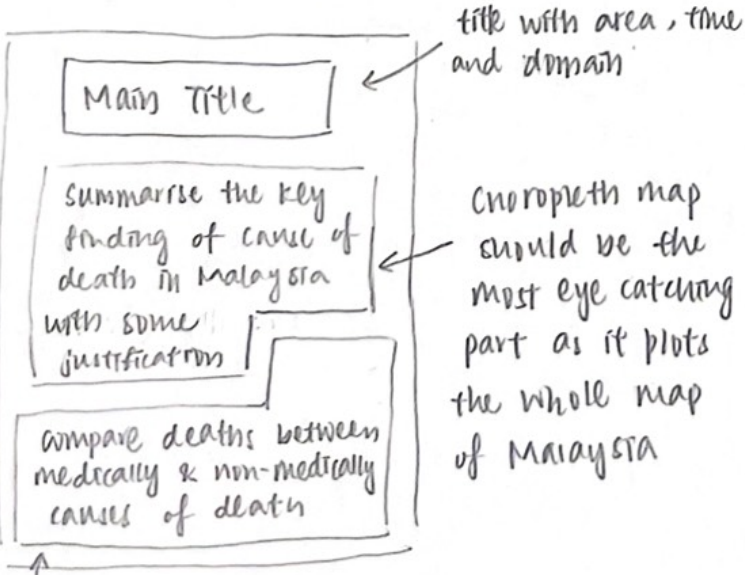
OPERATION

- Click on the medically / non-medically causes of death with filter out relevant data in bump chart & line graph



FOCUS

- No main focus in this data visualization.
- All columns are equally important.



title with area, time and domain

Choropleth map should be the most eye catching part as it plots the whole map of Malaysia

compare 2 different category of causes of death. This provides some important findings to investigate the major cause of deaths in 2022.

DISCUSSION

Advantages

- Provide detail information about the cause of death in Malaysia, clear comparison charts & annotations to show the difference between 2022 and previous years.
- Colour used across the columns are consistent.
- Interative operation is applied to provide better understanding to the readers.

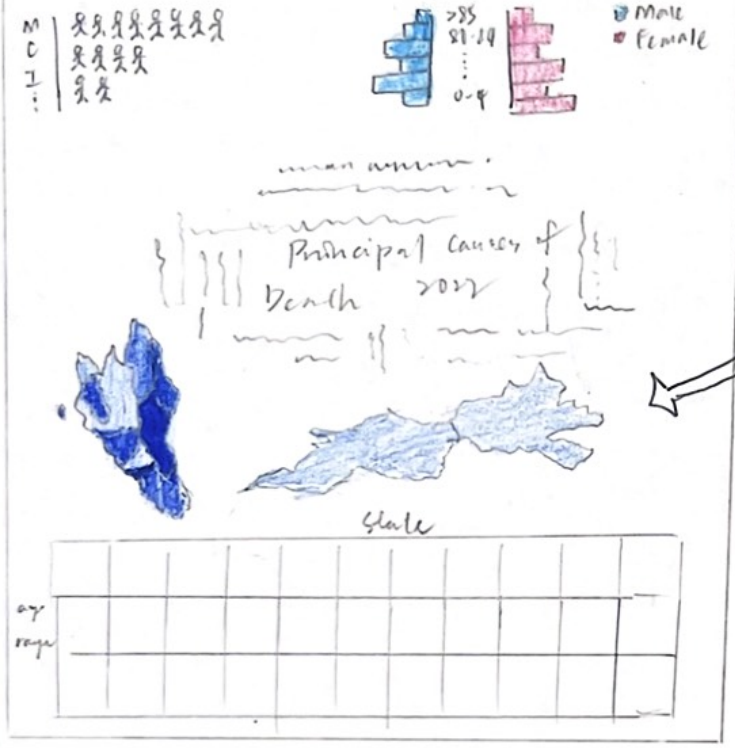
Disadvantages

- Messy and unclear layout, too many sight lines have been used.
- No main focus in the graph, hard to read, doesn't guide user to read (no start and end point)
- Avoid long paragraph, instead short and simple justification should be used.

LAYOUT

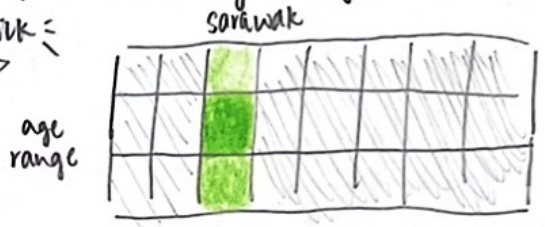
Title: Partitioned Poster
Author: Chai Ke Rou
Date: 29/9/2024
Sheet: 4
Task: Design a scrollable poster

CAUSE OF DEATH



OPERATION

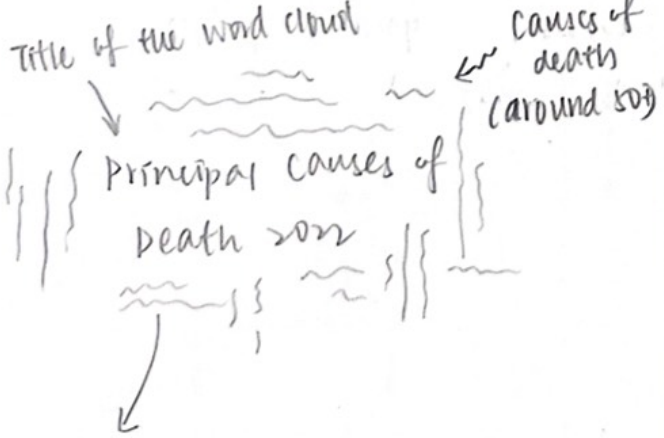
- Click on the state in choropleth map to highlight specific column in the heat map, easily compare the deaths of different age range in a state



- Click on different age range in population pyramid to highlight specific column in heat map, easily compare between different states.

FOCUS

Main focus: Word cloud (principal causes of death 2022)



Tool-tips

DISCUSSION

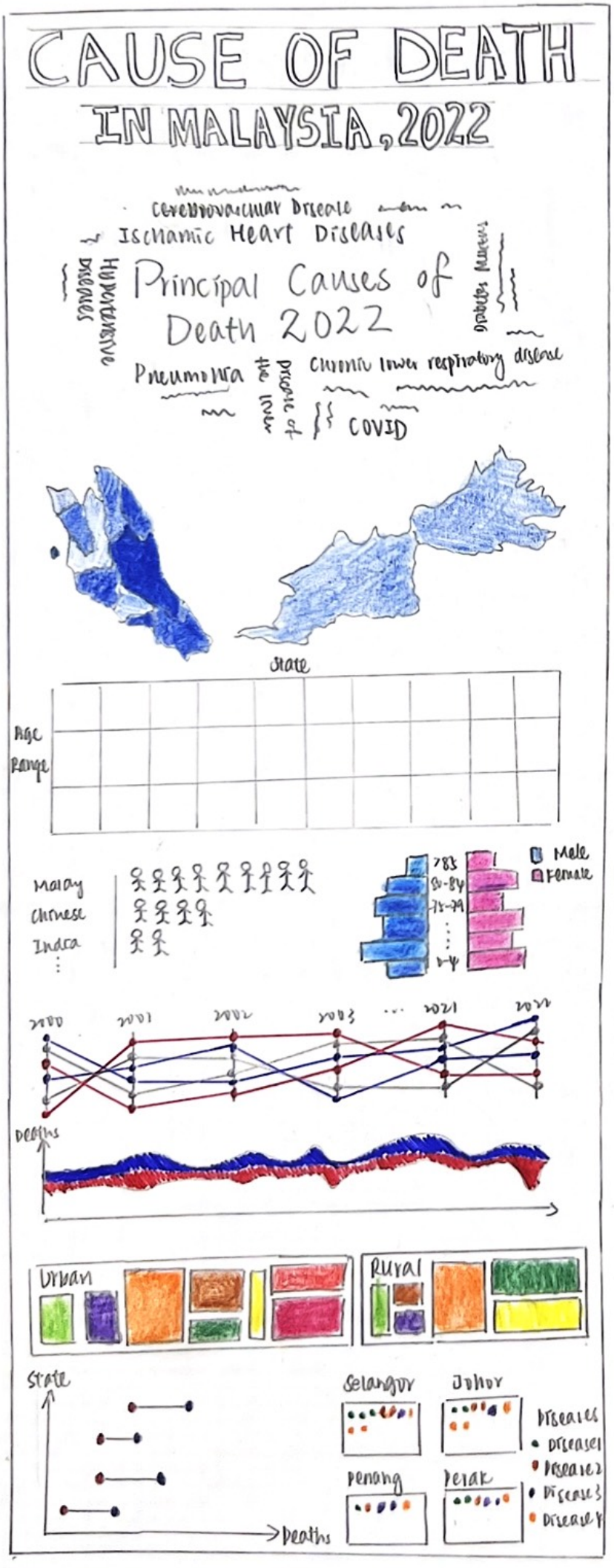
Advantages

- Key component is placed slightly above the geographic centre
- provide simple and clear data visualisation to show the distribution of causes of death
- Lack of area and time info in the main title of the data visualisation

Disadvantages

- Relevant graphs are placed far away from each other. such as population pyramid and heat map (same age ranges)
- Doesn't provide comparison between medically & non-medically causes of death
- Missing information about the time series of deaths, doesn't compare deaths in 2022 with previous data.

LAYOUT



Title: Five Design Sheet
Author: Chai Ke Rou
Date: 29/9/2024
Sheet: 5
Task: Design the final layout

FOCUS

- Main Focus: The word cloud and the choropleth map as they provide an overview of cause of deaths and the death rates of Malaysia in 2022
- Second Focus: The bump chart and stream graph as this shown the time series of deaths and imporeant cause of death from 2020 to 2022.
- Other graphs are equally important as they help to explain the distribution of causes of death more details.
- The tooltips and annotations are important as well as they provide useful information.

OPERATION

- Apply filter function for diseases across the website, consistent colour used for the disease in all idioms.
- Click on the states in choropleth map to filter out the details of the distribution of causes of deaths in that particular state.
- Apply brushing interaction for bump chart & stream graph to highlight a time period
- Click on specific age range in the population pyramid to highlight relevant data in the heat map
- Add tooltips in word cloud to show the percentage, number of deaths, category under
- Link bump chart, stream graph & dumbbell chart by medically & non-medically
- Add annotations to stream graph, bump chart, choropleth map and etc to indicate key findings.

DETAILS

- Dependencies: Excel [data preprocessing], vega-lite [data visualization], html [create website]
- Estimated time & effort: [1 day for final report]
  1 day for data preprocessing
  2 days for creating idioms
  5-6 days for final design & structure layout
- Specific requirements: can all the idioms be created using vega-lite? are they too many?