# Group project structure

## Project Idea: \*\*Comprehensive Analysis of Heart Failure Data with Integration of Socioeconomic Data from the Australian Census\*\*

### 1. \*\*Introduction and Background\*\*

- \*\*Objective\*\*: The goal of this project is to provide a comprehensive analysis of heart failure patients using the Zigong heart failure dataset, enhanced with socioeconomic data from Zigong Census and Australian heart failure dataset and Australian Census. The analysis will explore potential correlations between heart failure outcomes and socioeconomic factors such as income, education, and access to healthcare.

- \*\*Datasets\*\*:

- \*\*Already have Dataset\*\*: Zigong heart failure data.

- \*\*Extra Dataset\*\*: Australian heart failure dataset; Zigong and Australian Census data (particularly data related to socioeconomic factors such as income, education, employment, and access to healthcare facilities).

### 2. \*\*Data Exploration and Preprocessing\*\*

- \*\*Zigong Heart Failure Data\*\*: Begin with a detailed exploration of the heart failure data, focusing on key variables such as LVEF, NYHA classification, comorbidities, and demographic data (age, gender).

- \*\*Australian Census Data\*\*: Identify and extract relevant socioeconomic data at a geographic level (SA1, SA2) that can be linked or compared to the heart failure data.

- \*\*Data Integration\*\*: Develop a strategy to integrate the socioeconomic data with the Zigong dataset. This might involve mapping geographic areas or using similar variables for comparison.

### 3. \*\*Analysis Framework\*\*

- \*\*Socioeconomic Influence on Heart Failure\*\*:

- \*\*Objective\*\*: Analyze the influence of socioeconomic factors on heart failure outcomes, such as the type of heart failure, severity, and mortality rates.

- \*\*Methodology\*\*: Use regression models, clustering, or other statistical methods to identify correlations or patterns between heart failure outcomes and socioeconomic data.

- \*\*Visualizations\*\*: Interactive dashboards using Shiny that allow users to explore these relationships by filtering different socioeconomic and clinical variables.

### 3. \*\*[Optional] Geographic Analysis\*\*:

- \*\*Objective\*\*: Explore the geographic distribution of heart failure cases in relation to socioeconomic factors.

- \*\*Methodology\*\*: Use geographic information system (GIS) tools or similar packages in R (like `sf` or `tmap`) to visualize the distribution of heart failure cases and overlay it with socioeconomic data.

- \*\*Visualizations\*\*: Create detailed maps showing the distribution of heart failure cases across different socioeconomic strata, potentially using DataShader for ultra-detailed mapping.

### 4. \*\*Australian Census Data Connection\*\*

- The Census data offers a wealth of information on the social determinants of health, such as income, education, and access to services, which are crucial factors in understanding health outcomes, including heart failure.

- \*\*Potential Linkage\*\*:

- Although direct linking might be challenging due to the lack of geographic specificity in the Zigong data, you can draw comparisons between the heart failure data and broader Australian patterns in health outcomes and socioeconomic factors.

- Use synthesized individual-level data for comparative analysis, or simulate a similar population based on Australian Census data, then compare with the Zigong dataset.

### 5. \*\*Visualization and Scrollytelling\*\*

- \*\*Interactive Dashboards\*\*:

- Implement Shiny apps for interactive analysis, allowing users to filter and explore the data based on different parameters, such as age, gender, socioeconomic status, etc.

- \*\*Scrollytelling\*\*:

- Consider using the `parallaxr` package or other scrollytelling tools in R to create a narrative around the data. This could involve walking the viewer through the story of heart failure in the context of both Zigong data and Australian socioeconomics.

## Project Output:

### 1 \*\*Interactive Shiny Dashboards\*\*:

For dynamic data exploration and analysis.

### 2 \*\*Static Visualizations\*\*:

High-quality visualizations for geographic and statistical analysis.

### 3 \*\*Scrollytelling Narrative\*\*:

A webpage or presentation that narrates the findings in an engaging manner.

## Available Australian Heart Failure Data:

Australian Census Data

AIHW (Australian Institute of Health and Welfare) - various datasets and reports on heart failure, including hospitalizations and mortality data?

Australian Cardiovascular Disease Statistics - Reports and data are available from the Heart Foundation and AIHW

PhysioNet Public Data

Select a suitable Australian census data for comparison with the Zigong dataset

Description

The patient unique ID

Destination of hospital discharge, recorded after discharge

first admission ward

possible ways of admission are Emergency vs. non-emergency

occupation

the patient discharged from

the number of hospital admissions before this hospitalization

gender: Male, Female

body temperature in degrees celsius

pulse rate (beats per minute)

respiratory rate (breaths/min)

systolic blood pressure (mmHg)

diastolic blood pressure (mmHg)

mean arterial pressure (mmHg)

weight (kg)

height (m)

BMI (kg/m^2)

type of heart failure (left, right, both)

NYHA.cardiac.function.classification

Killip.grade: Class 1 No rales, no 3rd heart sound; Class 2 Rales in <1???2 lung field or presence of a 3rd heart sound; Class 3 Rales in >1???2 lung field?€?pulmonary edema;Class 4 Cardiogenic shock?€?determined clinically

myocardial.infarction

congestive heart failure (e.g., A minority of patients were not coded as having the diagnosis of ?€?congestive heart failure?€? in the comorbidity list because they did not have past history of congestive heart failure on admission. They were diagnosed with HF for the first time in this index hospitalization.)