BEHAVIOURAL FACTORS LEADING TO CARDIAC ISCHAEMIA:

Getting to the Heart of Heart Disease with Predictive Analytics

Chalamalasetti Sree Vaishnavi

Teo De Xuan Justin

Siah Wee Hung

Lim Zi Xiang

Nagammai Senthil Kumar

PRESENTATION OUTLINE

01 Introduction

02 Problem Statement

03 Project Objective

04 Data & Methodology

05 Modelling

06 Proposed Business Solution

Chalamalasetti Sree Vaishnavi

PREVALENCE OF HEART DISEASE

Cardiovascular diseases 2nd leading cause of death in Singapore – **nearly 1 in 3 deaths** in 2018

Chalamalasetti Sree Vaishnavi

PREVALENCE OF HEART DISEASE

Strain on healthcare resources → Overwhelm healthcare system

(CNA, 2022)

(Straits Times, 2022)

Chalamalasetti Sree Vaishnavi

PROBLEM STATEMENT

01 Time-consuming Diagnostic Tests

02 High Costs of Tests

03 Manpower shortage

04 Scant Medical Records

Chalamalasetti Sree Vaishnavi

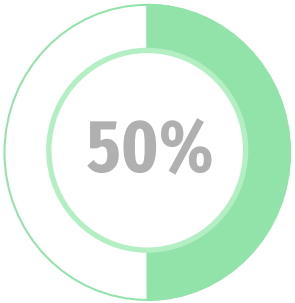
TIME-CONSUMING DIAGNOSTICS TESTS

Ambulatory Electrocardiogram

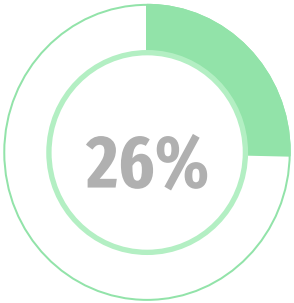
24-48 hours

Chalamalasetti Sree Vaishnavi

TIME-CONSUMING DIAGNOSTICS TESTS



Potentially salvageable heart muscle is lost within an hour



Heart attack patients managed to get treatment in time

Chalamalasetti Sree Vaishnavi

HIGH COSTS OF TESTS

Diagnostic Test	Cost
Consultation	$180-$300
Electrocardiogram	$60-$80
Echocardiogram	$450-$600
Treadmill Test	$350-$550

Deters high-risk patients from taking health screenings or seeking treatment →

- Further strain of healthcare resources
- Higher mortality rate

Chalamalasetti Sree Vaishnavi

MANPOWER SHORTAGE & SCANT MEDICAL RECORDS

Diversion of critical resources → Could have been better optimised

Initial assessment based on their medical history

Chalamalasetti Sree Vaishnavi

GAPS IN EXISTING MODELS

01 High false negative rates of existing models

02 Usefulness of non-medical factors

03 Limitations of use of complex data in existing models

INCOMPLETENESS OF EXISTING MODELS

- High false negative rates

Naïve Bayes and weighted approach	2 SVM's and XGBoost	SVM and DO	XGBoost
44.5%	47.5%	43.4%	45.7%

- Life threatening: Heart disease goes undetected

Chalamalasetti Sree Vaishnavi

USEFULNESS OF NON-MEDICAL FACTORS

❖ Behavioural Factors

(Global case-study on modifiable behavioural risk factors association with heart disease)

90%

❖ Psychological Factors

(Case-study on depression's association with heart disease)

3x

Chalamalasetti Sree Vaishnavi

LIMITATIONS OF USE OF COMPLEX DATA IN EXISTING MODELS

- Complex data needed for existing models for prediction

Reluctance

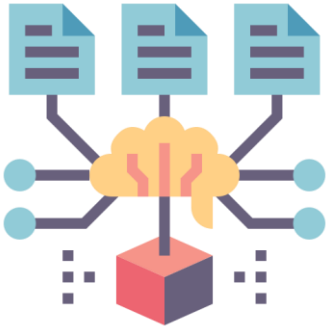
Data unavailable

Chalamalasetti Sree Vaishnavi

PROJECT OBJECTIVE

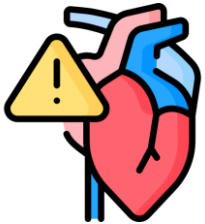
Behavioural Demographic

Inputs



Predictive analytics model

Output



Risk of Heart disease

- Easily accessible
- Time-efficient
- Affordable
- Accurate (Low false negatives)

Alleviate stress on NHCS through pre-diagnosis with non-medical data

Chalamalasetti Sree Vaishnavi

DATA SOURCING

- **Aim:** Construct a model that predicts likelihood of CHD based on non-medical data
- **Need:** Wide range of behavioural, psychological & demographic factors that lead to CHD

- Public health surveys
- Behavioural, Psychological & Demographic factors

DATA CLEANING

Teo De Xuan Justin

INITIAL DATASET

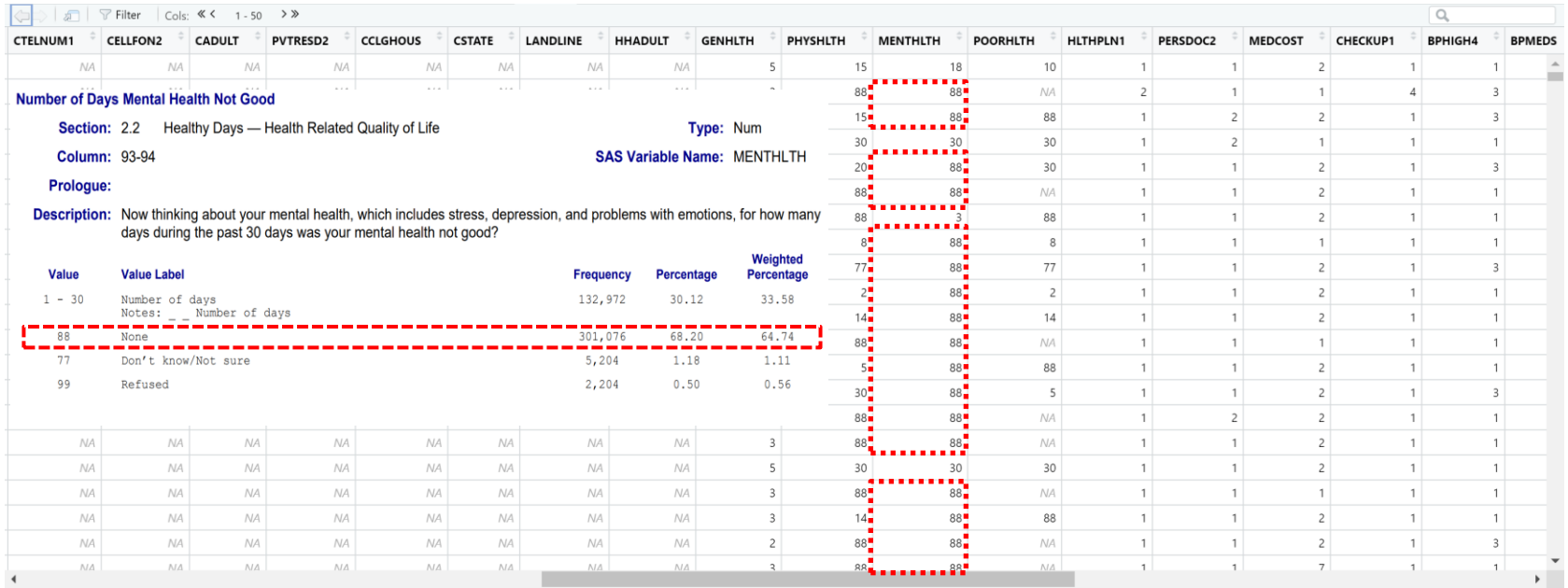
Data

dt 441456 obs. of 330 variables

	_STATE	FMONTH	IDATE	IMONTH	IDAY	IYEAR	DISPCODE	SEQNO	_PSU	CTELENUM	PVTRESD1	COLGHOUS	STATERES	CELLFON3	LADULT	NUMADULT	NUMMEN	NUMWOMEN	CTELNUM
1	1	1	b'01292015'	b'01'	b'29'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	3	1.000000e+00	2.000000e+00	
2	1	1	b'01202015'	b'01'	b'20'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
3	1	1	b'02012015'	b'02'	b'01'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
4	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	3	1.000000e+00	2.000000e+00	
5	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
6	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
7	1	1	b'01052015'	b'01'	b'05'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
8	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
9	1	1	b'01132015'	b'01'	b'13'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
10	1	1	b'01302015'	b'01'	b'30'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
11	1	1	b'01222015'	b'01'	b'22'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
12	1	1	b'01162015'	b'01'	b'16'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
13	1	1	b'01202015'	b'01'	b'20'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
14	1	1	b'01202015'	b'01'	b'20'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
15	1	1	b'01202015'	b'01'	b'20'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
16	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
17	1	1	b'01042015'	b'01'	b'04'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
18	1	1	b'01202015'	b'01'	b'20'	b'2015'	1200	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
19	1	1	b'01272015'	b'01'	b'27'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	5.397605e-79	1.000000e+00	
20	1	1	b'01062015'	b'01'	b'06'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	2	1.000000e+00	1.000000e+00	
21	1	1	b'01142015'	b'01'	b'14'	b'2015'	1100	2.015e+09	2.015e+09	1	1	NA	1	2	NA	1	1.000000e+00	5.397605e-79	

Teo De Xuan Justin

I. RECODE SPECIFIED VALUES



Filter | Cols: « < 1 - 50 > »

CTELNUM1	CELLFON2	CADULT	PVTRESD2	CCLGHOUS	CSTATE	LANDLINE	HHADULT	GENHLTH	PHYSHLTH	MENTHLTH	POORHLTH	HLTHPLN1	PERSDOC2	MEDCOST	CHECKUP1	BPHIGH4	BPMEDS
NA	NA	NA	NA	NA	NA	NA	NA	5	15	18	10	1	1	2	1	1	
									88	88	NA	2	1	1	4	3	
									15	88	88	1	2	2	1	3	
									30	30	30	1	2	1	1	1	
									20	88	30	1	1	2	1	3	
									88	88	NA	1	1	2	1	1	
									88	3	88	1	1	2	1	1	
									8	88	8	1	1	1	1	1	
									77	88	77	1	1	2	1	3	
									2	88	2	1	1	2	1	1	
									14	88	14	1	1	2	1	1	
									88	88	NA	1	1	1	1	1	
									5	88	88	1	1	2	1	1	
									30	88	5	1	1	2	1	3	
									88	88	NA	1	2	2	1	1	
NA	NA	NA	NA	NA	NA	NA	NA	3	88	88	NA	1	1	2	1	1	
NA	NA	NA	NA	NA	NA	NA	NA	5	30	30	30	1	1	2	1	1	
NA	NA	NA	NA	NA	NA	NA	NA	3	88	88	NA	1	1	1	1	1	
NA	NA	NA	NA	NA	NA	NA	NA	3	14	88	88	1	1	2	1	1	
NA	NA	NA	NA	NA	NA	NA	NA	2	88	88	NA	1	1	2	1	3	
NA	NA	NA	NA	NA	NA	NA	NA	3	88	88	NA	1	1	7	1	1	

Number of Days Mental Health Not Good

Section: 2.2 Healthy Days — Health Related Quality of Life **Type:** Num

Column: 93-94 **SAS Variable Name:** MENTHLTH

Prologue:

Description: Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?

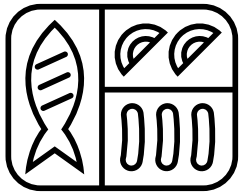
Value	Value Label	Frequency	Percentage	Weighted Percentage
1 - 30	Number of days Notes: _ _ Number of days	132,972	30.12	33.58
88	None	301,076	68.20	64.74
77	Don't know/Not sure	5,204	1.18	1.11
99	Refused	2,204	0.50	0.56

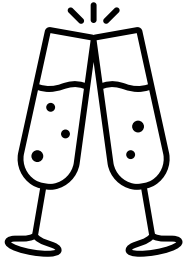
II. REMOVE DUPLICATES

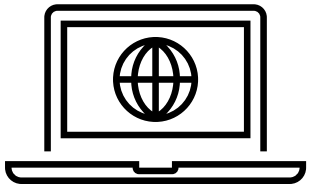
THE DATASET HAD NO DUPLICATES.

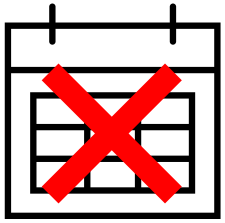
Teo De Xuan Justin

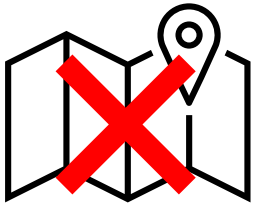
III. COLUMN PRESERVATION











Teo De Xuan Justin

IV. CHOOSE 'Y' – MICHD

Ever had CHD or MI

CalculatedVariables: 6.1 Calculated Variables **Type:** Num

Column: 1899 **SAS Variable Name:** _MICHD

Prologue:

Description: Respondents that have ever reported having coronary heart disease (CHD) or myocardial infarction (MI)

Value	Value Label	Frequency	Percentage	Weighted Percentage
1	Reported having MI or CHD Notes: CVDINFR4=1 OR CVDCRHD4=1	38,633	8.83	6.42
2	Did not report having MI or CHD Notes: CVDINFR4=2 AND CVDCRHD4=2	398,881	91.17	93.58
BLANK	Not asked or Missing Notes: CVDINFR4=7, 9 OR MISSING OR CVDCRHD4=7, 9, OR MISSING	3,942		

Teo De Xuan Justin

V. CORRECT NUMERIC DATA

Computed Weight in Kilograms

CalculatedVariables: 7.17 Calculated Variables **Type:** Num

Column: 1983-1987 **SAS Variable Name:** WTKG3

Prologue:

Description: Reported weight in kilograms

Value	Value Label	Frequency	Percentage	Weighted Percentage
2300 - 29500	Weight in kilograms [2 implied decimal places] Notes: 0001 <= WEIGHT2 <= 650 or 9023 <= WEIGHT2 <= 9295 (non-metric WEIGHT2 value divided by 2.2046)	410,535	93.00	93.29
99999	Don't know/Refused/Missing Notes: WEIGHT2 = 7777 or 9999 or Missing or not in accepted values	30,921	7.00	6.71

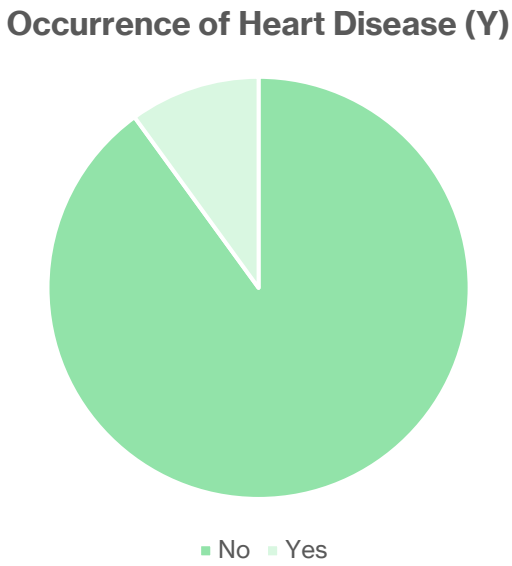
WTKG3, BMI5, GRENDAY, FRUITDA1, BEANDAY : * 0.01

VI. REMOVAL OF MISSING DATA

- Removed rows with **missing 'Y'**
- Removed rows with missing values in columns with **>20%** data missing (SCNTWRK1, PA1VIGM, PA1MIN)
  - ➢ No need to resort to imputation – enough data (441,456 obs.)
- Removed rows with **>20%** missing values

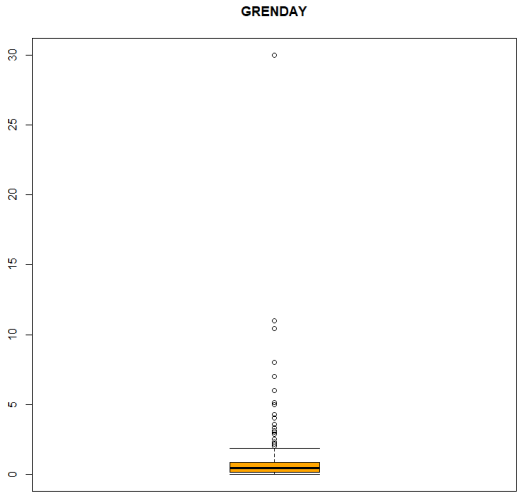
VII. BALANCE DATA

- 1:9 [Disease : No Disease] – bias towards predicting no disease
- **Undersampling**
- 1:9 → **1:1**

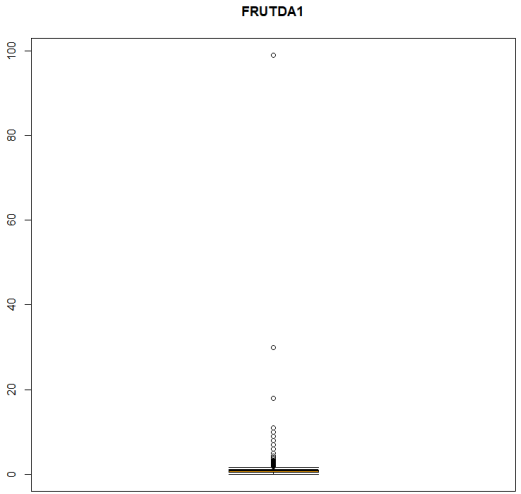


Teo De Xuan Justin

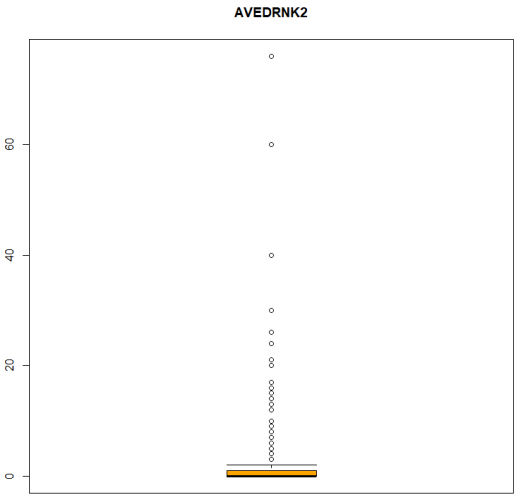
VIII. REMOVAL OF OUTLIERS



30 servings/day



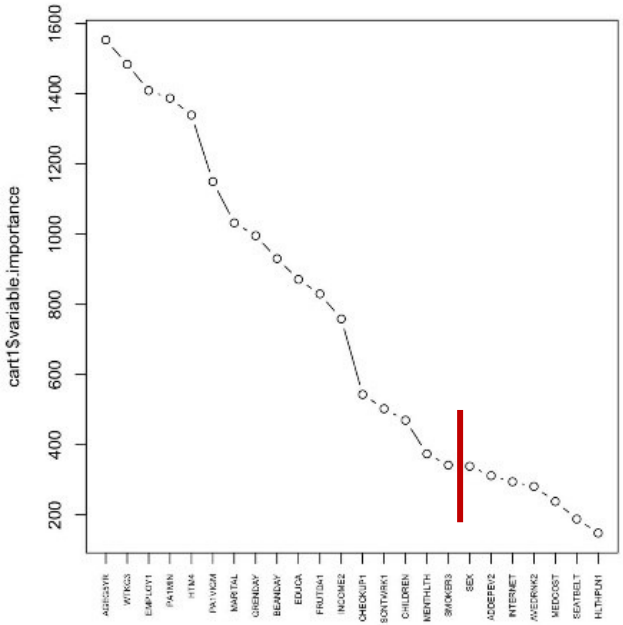
99 servings/day



72 servings/day

Teo De Xuan Justin

IX. REDUCE DIMENSION WITH CART



variable.importance : 27 → 18 Columns

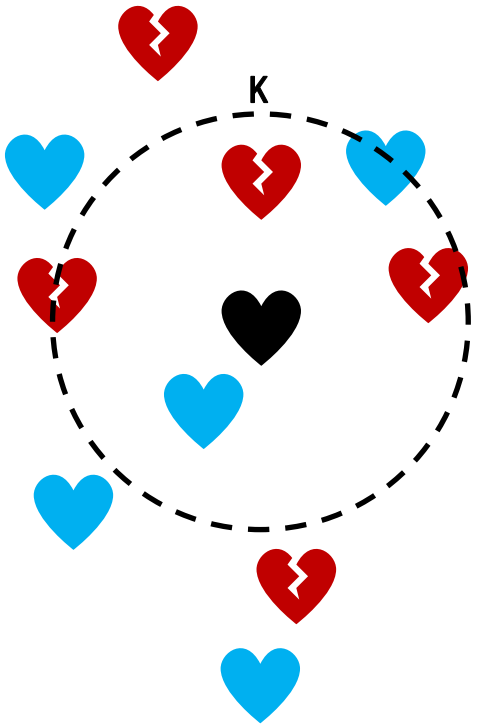
POST-CLEANING

- Dimensions: **21,214 rows, 18 columns**
- Impute data to fill missing values

Teo De Xuan Justin

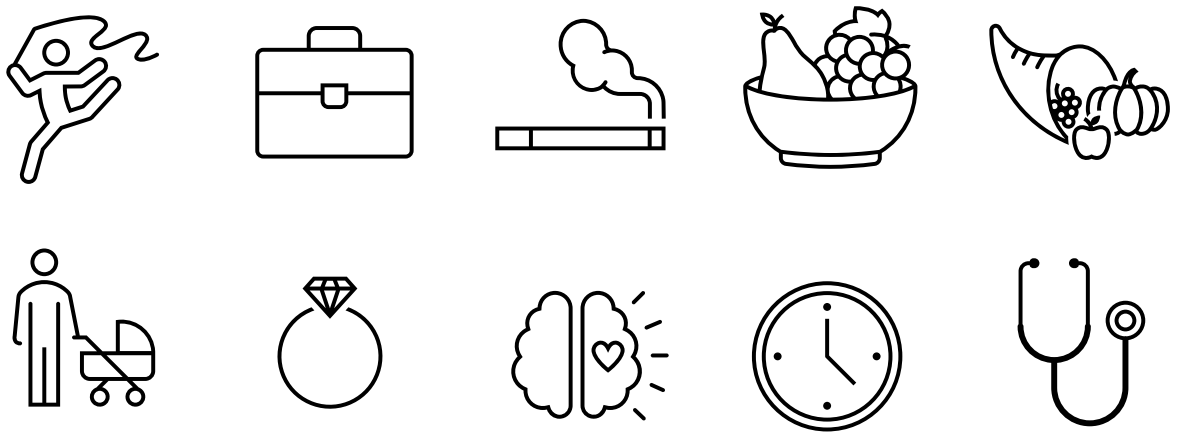
DATA IMPUTATION

- **K-nearest neighbours (kNN)** algorithm
- More robust than regression models
  - Can impute **both numerical, categorical** variables
  - Distance-based algorithm – can impute even with missing values by using alternative columns



FINAL DATASET

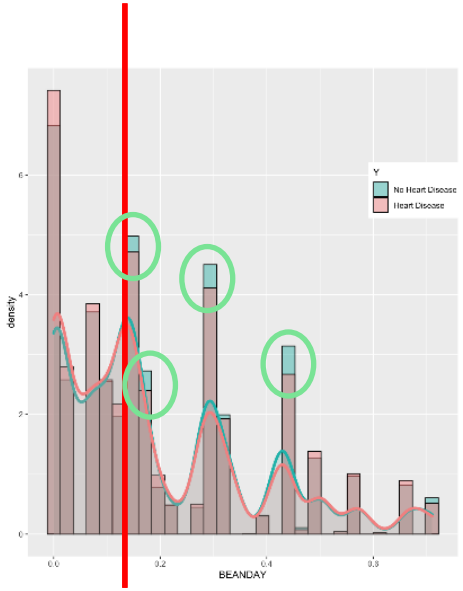
- **21,214 rows, 18 columns**
- Data includes:



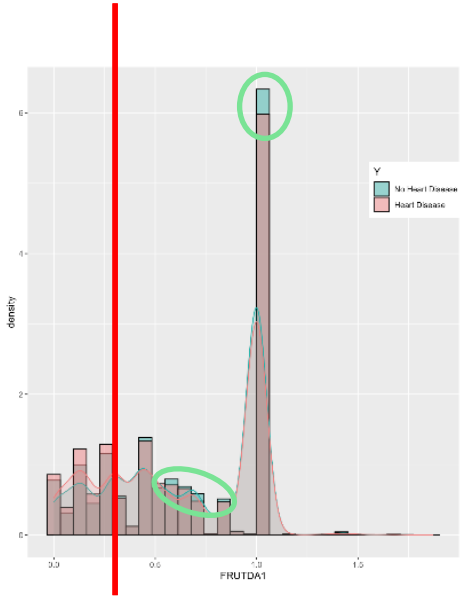

DATA EXPLORATION

Siah Wee Hung

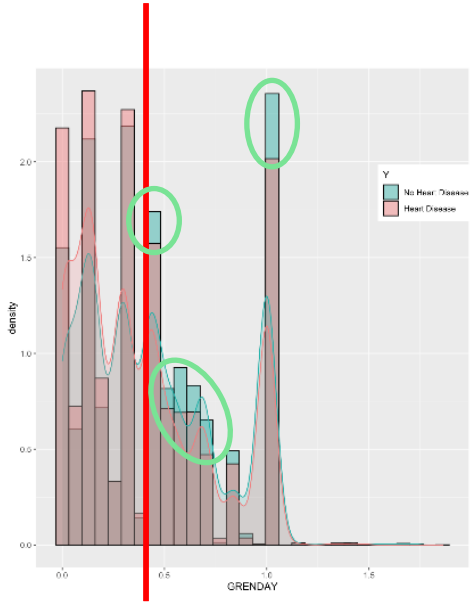
NUTRITION



DAILY BEAN INTAKE



DAILY FRUIT INTAKE

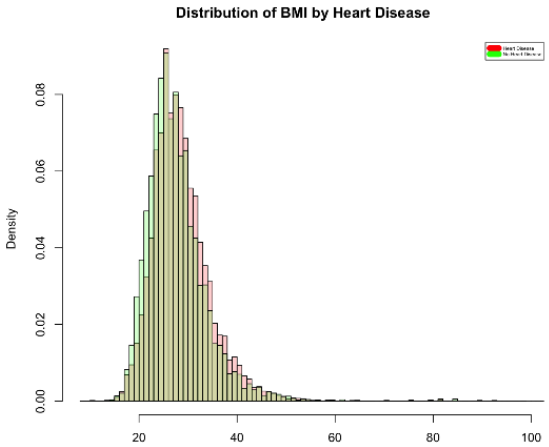


DAILY VEGETABLE INTAKE

HIGHER NUTRITION OFFERS SOME PROTECTIVE EFFECT

Siah Wee Hung

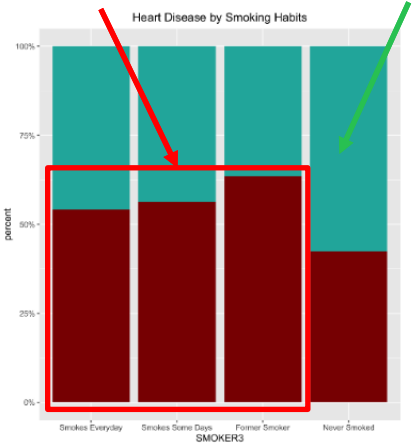
COMMON HEALTH FACTORS



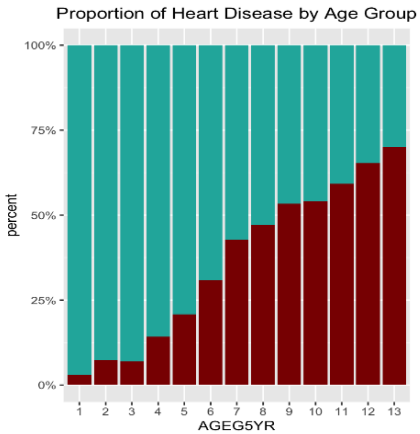
BMI

SMOKERS

NON SMOKERS



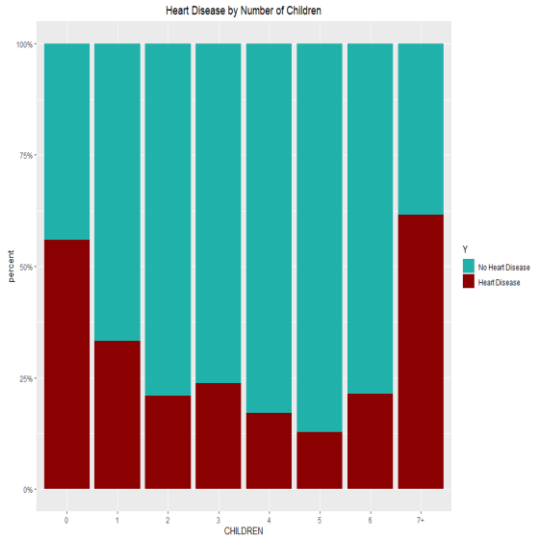
SMOKING HABITS



AGE

Siah Wee Hung

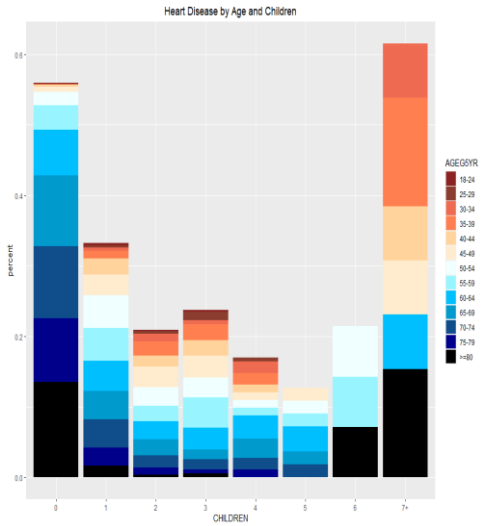
NO. CHILDREN



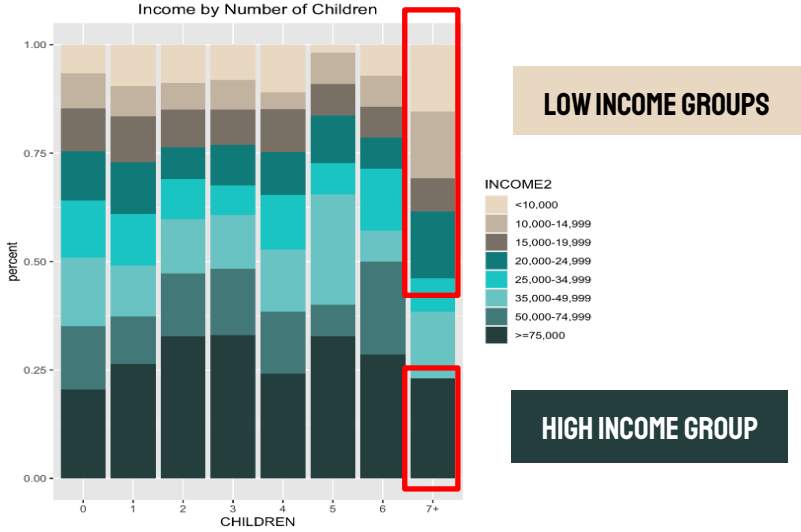
NO. CHILDREN VS HEART DISEASE

Siah Wee Hung

NO. CHILDREN



NO. CHILDREN
VS HEART DISEASE



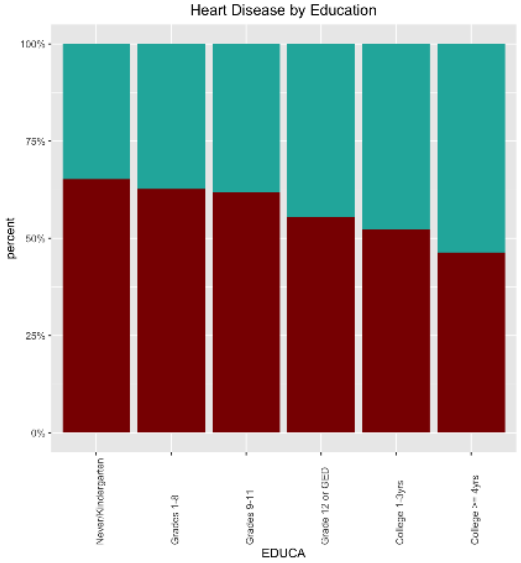
LOW INCOME GROUPS

HIGH INCOME GROUP

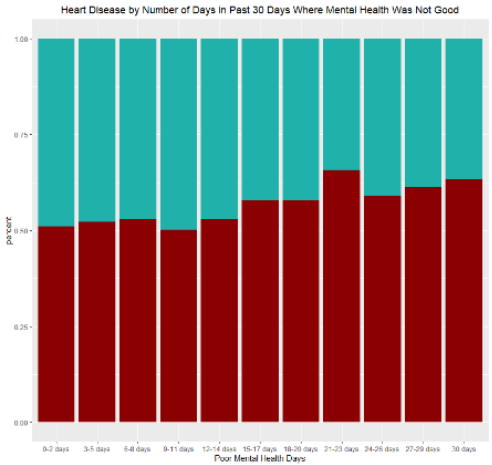
INCOME VS NO. CHILDREN

Siah Wee Hung

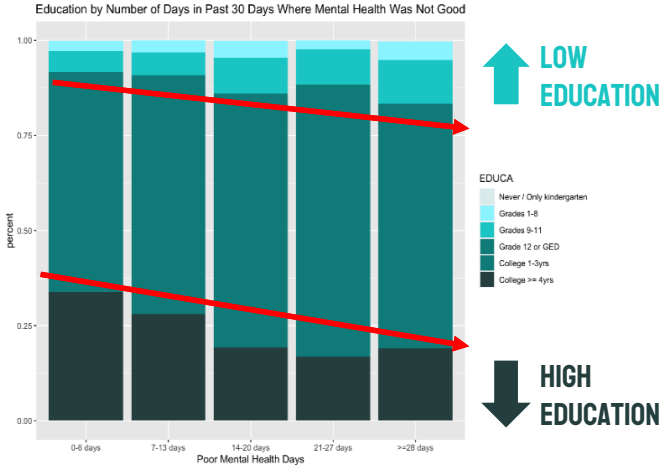
EDUCATION & MENTAL HEALTH



EDUCATION



MENTAL HEALTH



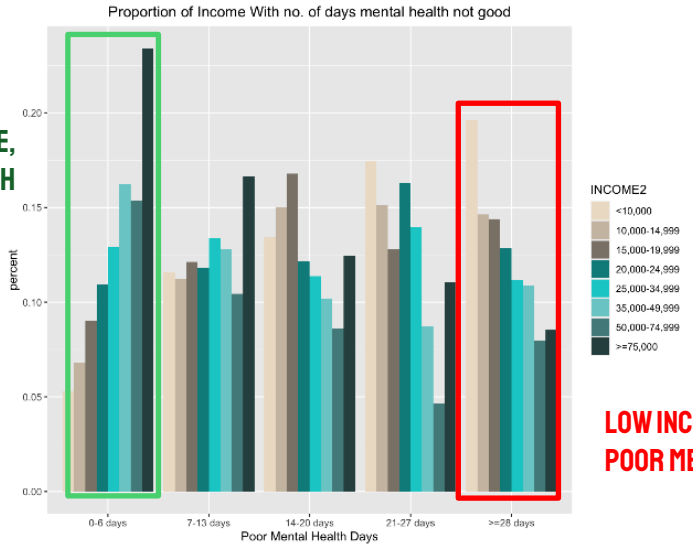
EDUCATION VS MENTAL HEALTH

"NUMBER OF DAYS IN PAST 30 DAYS WHERE MENTAL HEALTH IS NOT GOOD"

Siah Wee Hung

INCOME & MENTAL HEALTH

HIGH INCOME,
GOOD MENTAL HEALTH

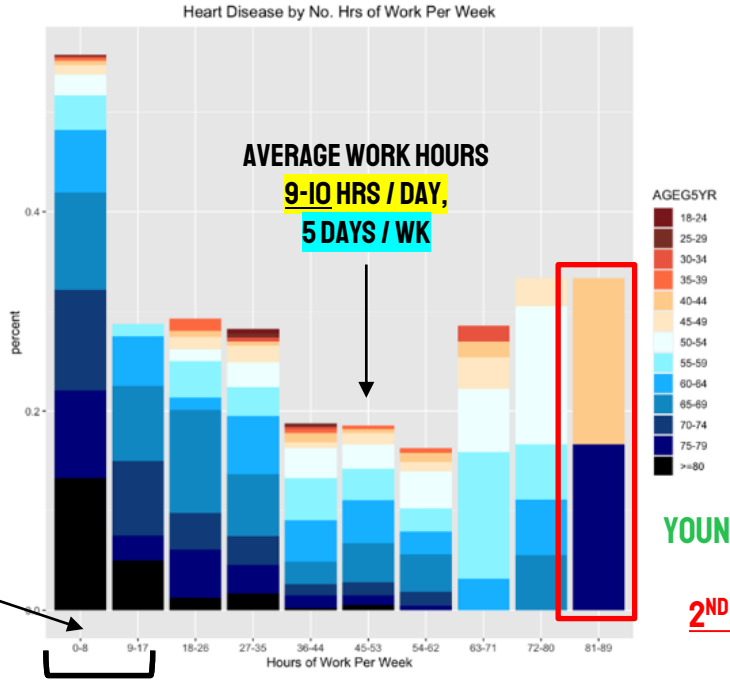


LOW INCOME,
POOR MENTAL HEALTH

INCOME VS MENTAL HEALTH

Siah Wee Hung

WEEKLY HOURS OF WORK



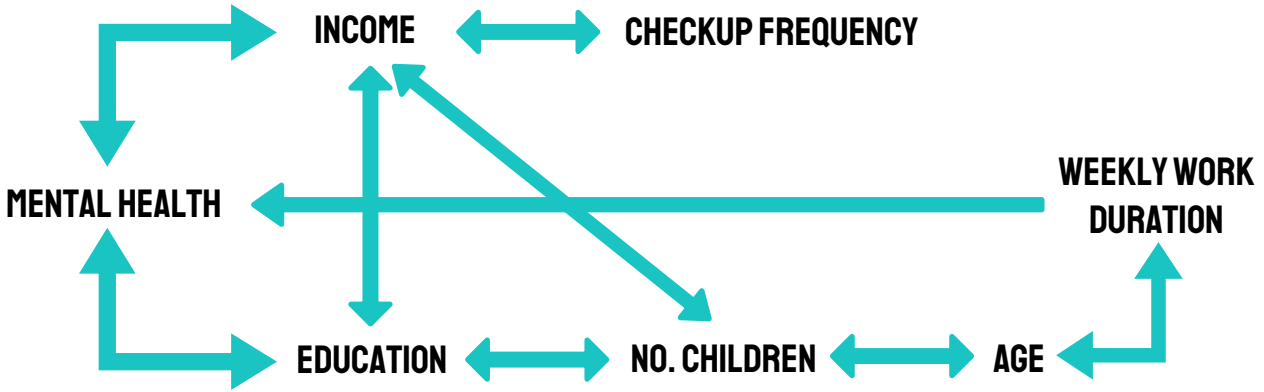
INACTIVE
(0 HRS OF WORK)

RETIREES

YOUNGEST AGE GROUP
BUT
2ND HIGHEST RISK!

Siah Wee Hung

INTERCONNECTEDNESS OF VARIABLES



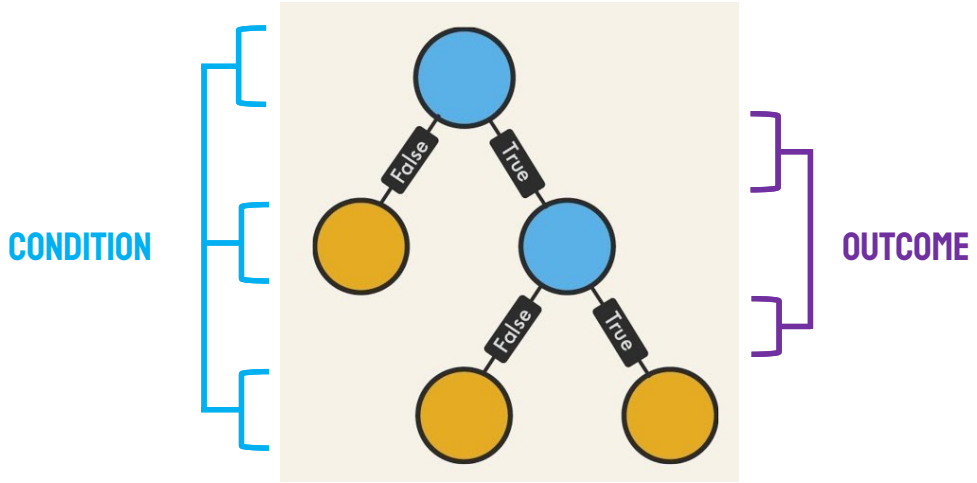
1. CORRELATION ≠ CAUSALITY

2. INTERDEPENDENT VARIABLES = MULTICOLLINEARITY?

Siah Wee Hung

MODELLING

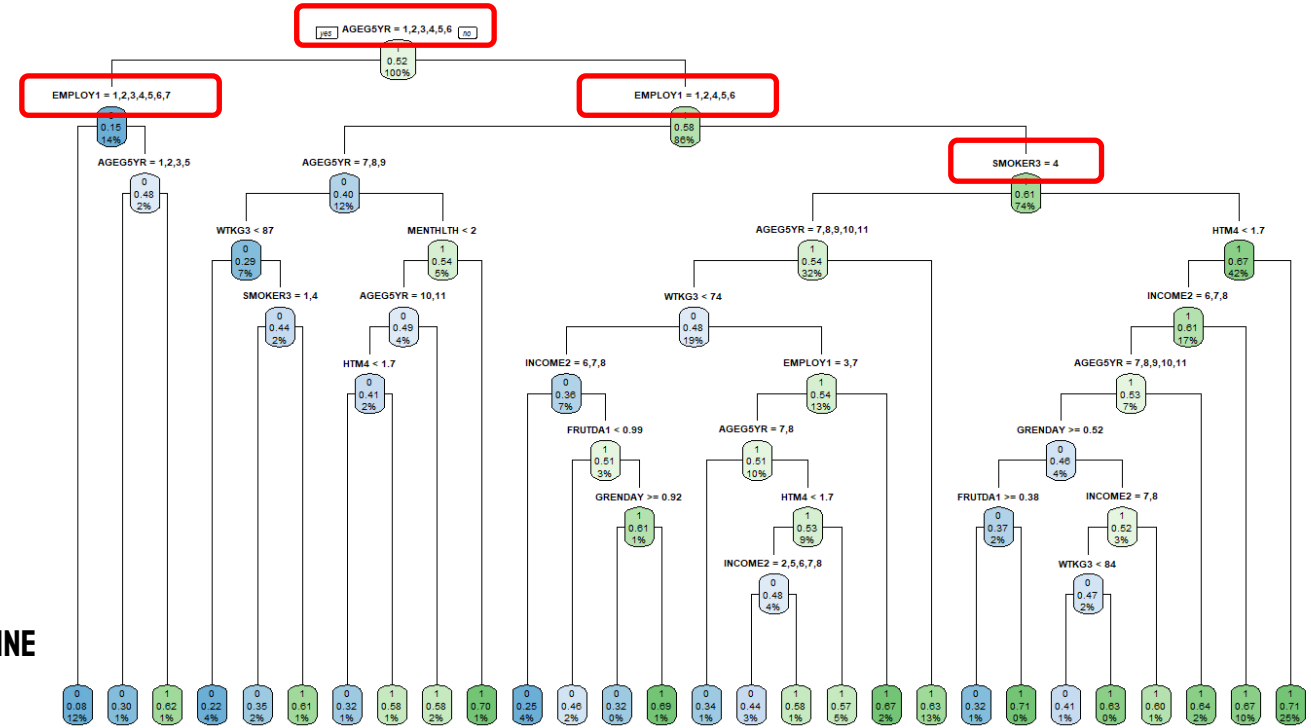
- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE



Siah Wee Hung

MODELLING

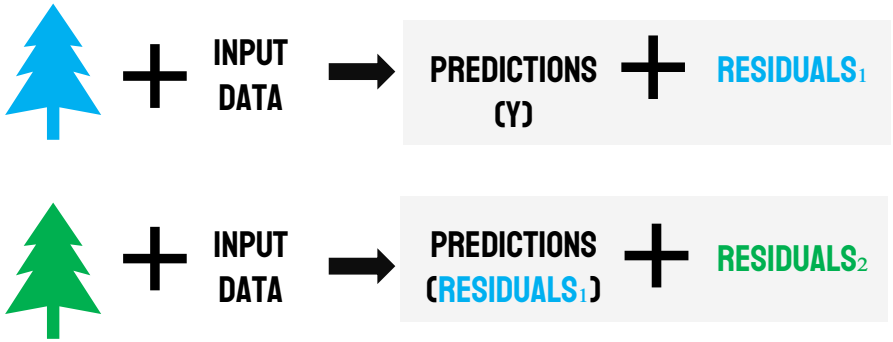
- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE



Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

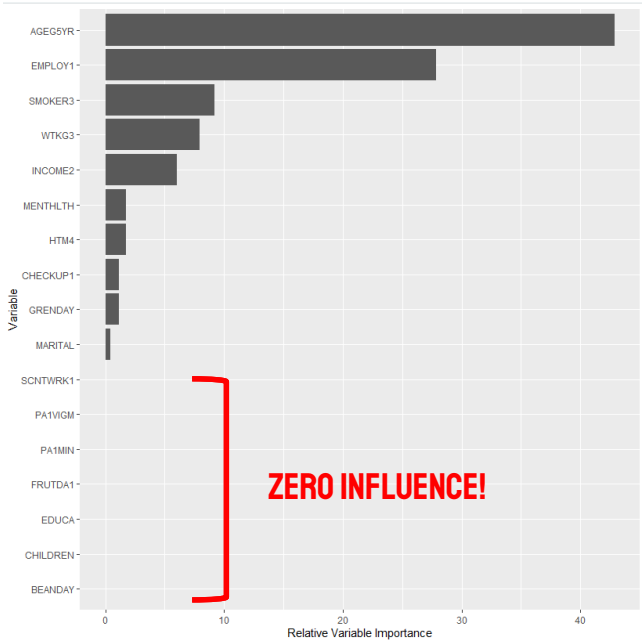


- COMBINE WEAK LEARNERS
- NEW TREE CORRECTS THE ERRORS OF PREVIOUS TREE

Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

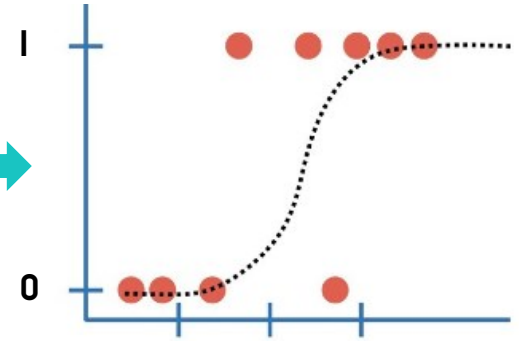


Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

$$Y_i = \beta_0 + \beta_1 X_i$$



LOGISTIC FUNCTION

Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

Backward Stepwise Logistic Regression

- Variables removed from the original model step-by-step
- Prevent overfitting
- Reduce multicollinearity

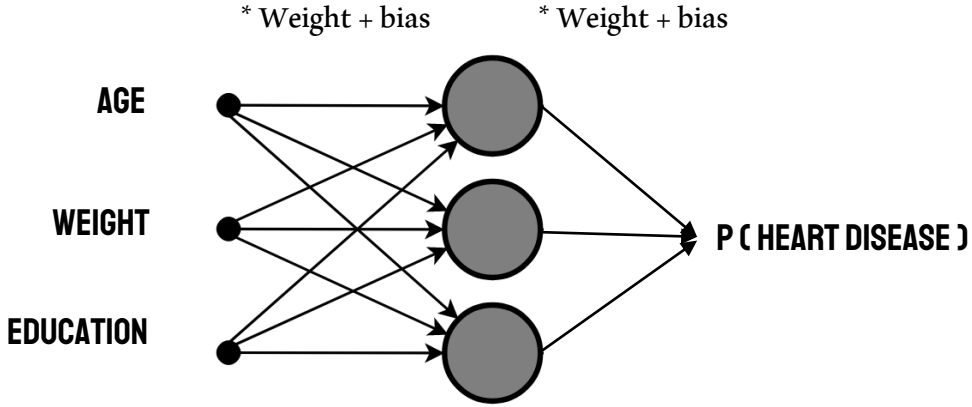
WTKG3
HTM4
AGEG5YR
EMPLOY1
GRENDAY
MARITAL
INCOME2
PA1VIGM
CHECKUP1
MENTHLTH
SMOKER3

~~PA1MIN~~
~~BEANDAY~~
~~FRUTDA1~~
~~EDUCA~~
~~CHILDREN~~
~~SCNTWRK1~~

Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

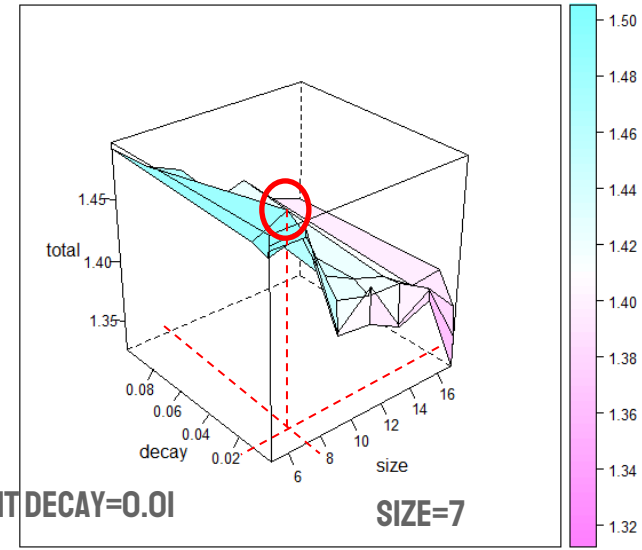


Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

GRID SEARCH



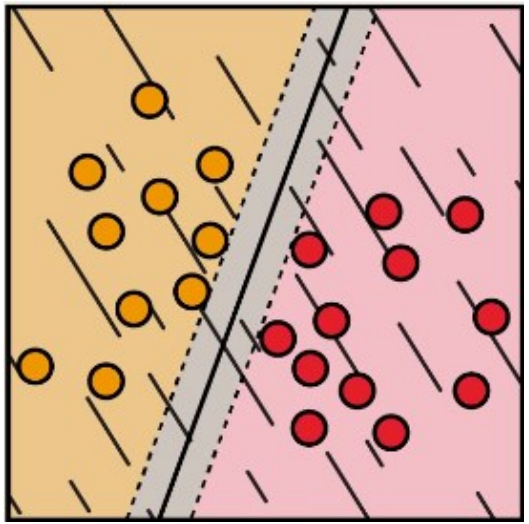
WEIGHT DECAY=0.01

SIZE=7

Siah Wee Hung

MODELLING

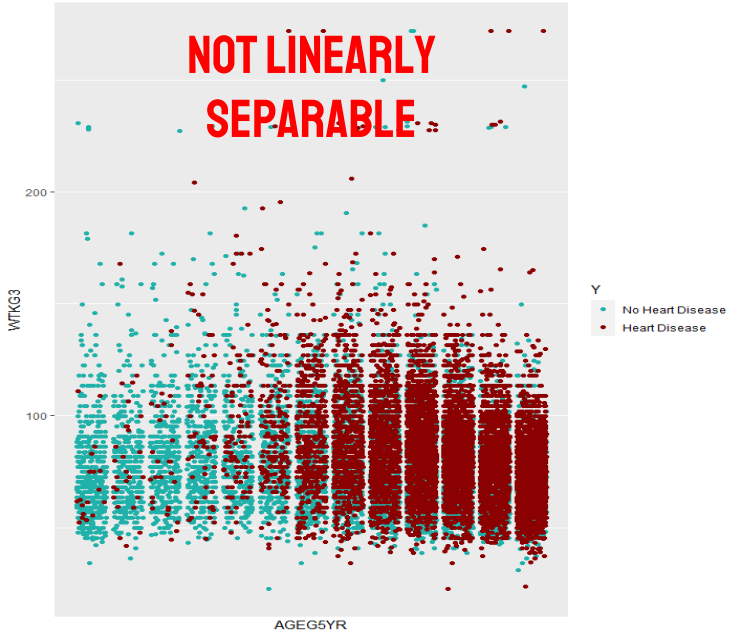
- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE



Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE



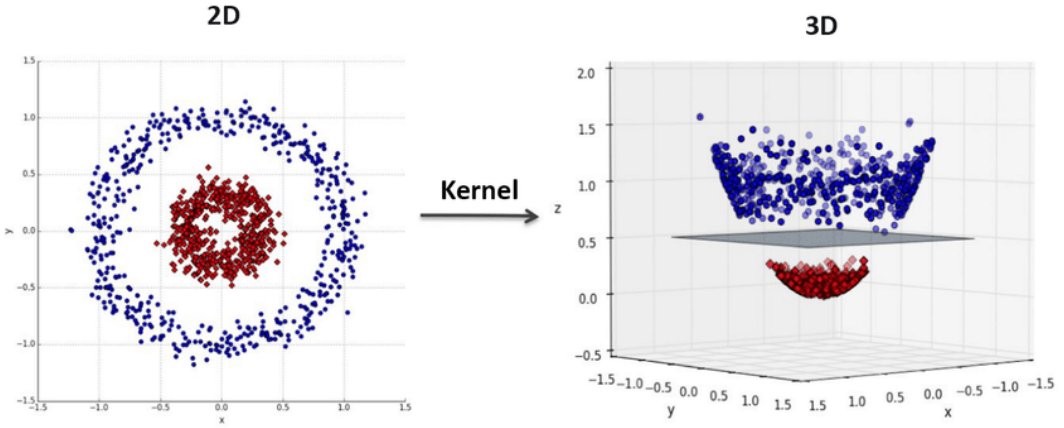
HEART DISEASE BY WEIGHT & AGE

Siah Wee Hung

MODELLING

- CLASSIFICATION AND REGRESSION TREE
- GRADIENT BOOSTING MACHINE
- LOGISTIC REGRESSION
- NEURAL NETWORK
- SUPPORT VECTOR MACHINE

KERNEL TRICK

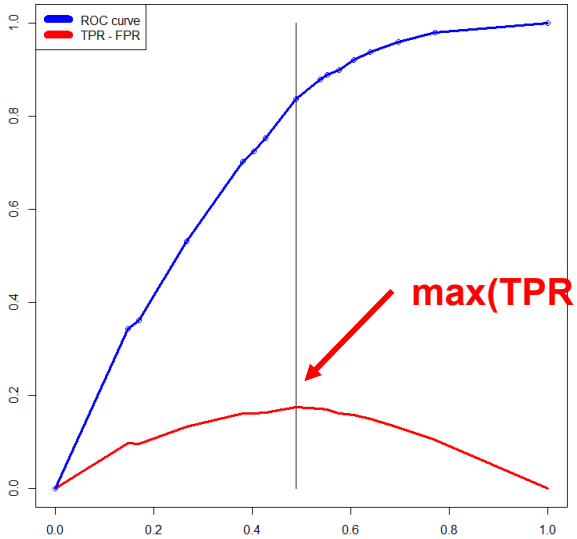


Siah Wee Hung

THRESHOLD MOVING

Receiver Operating Characteristic (ROC) Curve Method

TRUE POSITIVE RATE

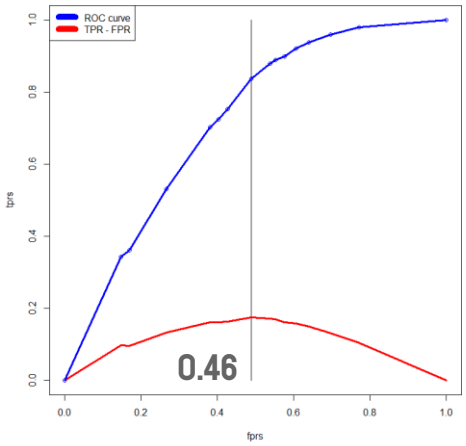


FALSE POSITIVE RATE

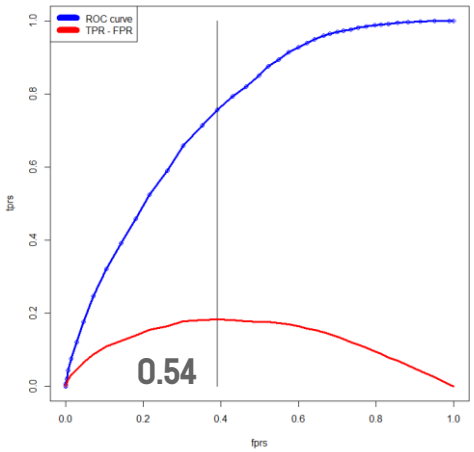
Siah Wee Hung

THRESHOLD MOVING

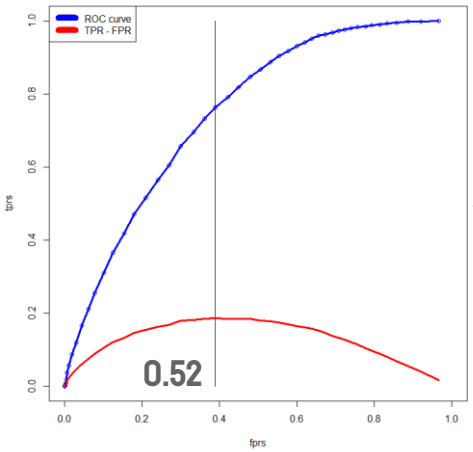
CART



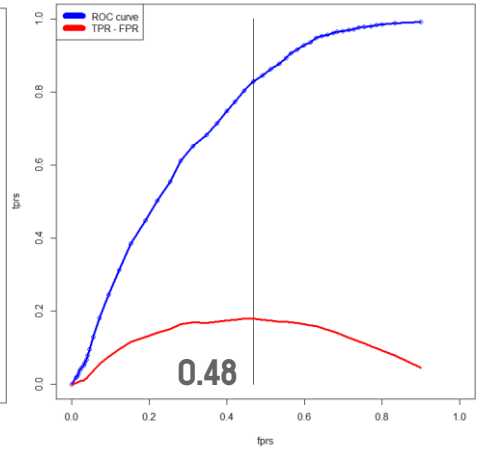
GBM



LOGISTIC REGRESSION



NEURAL NETWORK



Siah Wee Hung

MODEL EVALUATION

Validation Accuracy

$$\frac{\text{TRUE POSITIVES}}{\text{ACTUAL POSITIVES}}$$ Recall

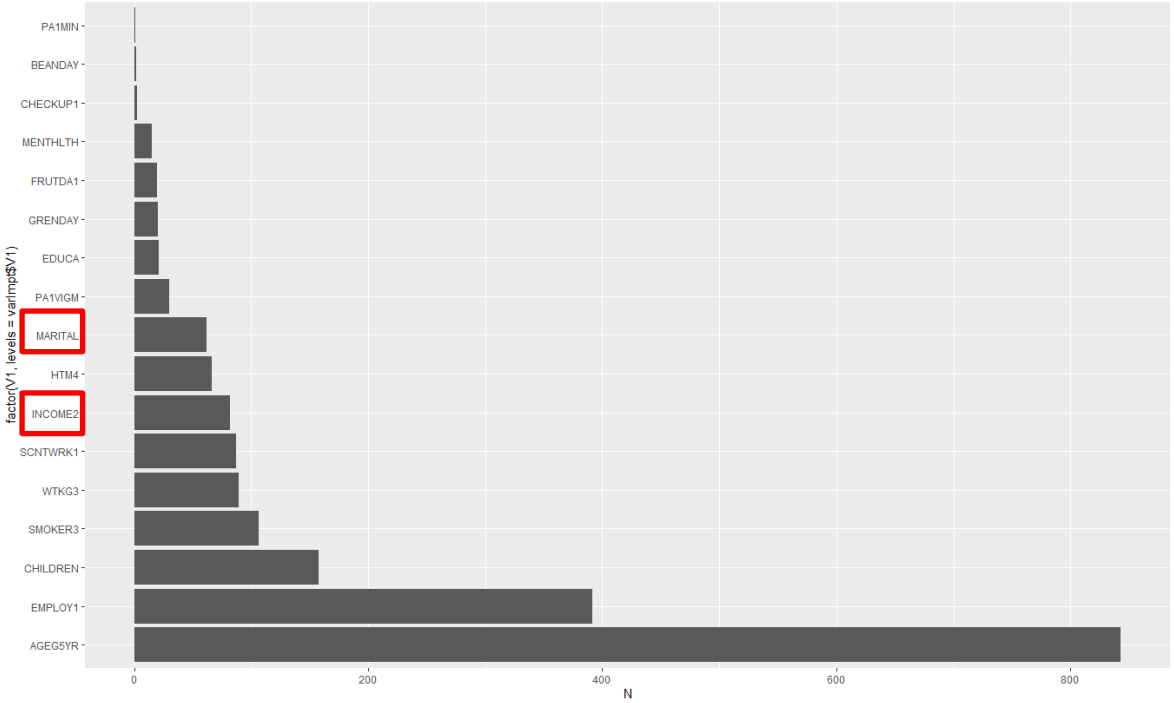
$$\frac{\text{TRUE POSITIVES}}{\text{POSITIVE PREDICTIONS}}$$ Precision

HARMONIC MEAN OF RECALL & PRECISION F1 Score

Lim Zi Xiang

MODEL INSIGHTS

CART'S VARIABLE IMPORTANCE



Lim Zi Xiang

MODEL INSIGHTS

CHI-SQUARE HYPOTHESIS

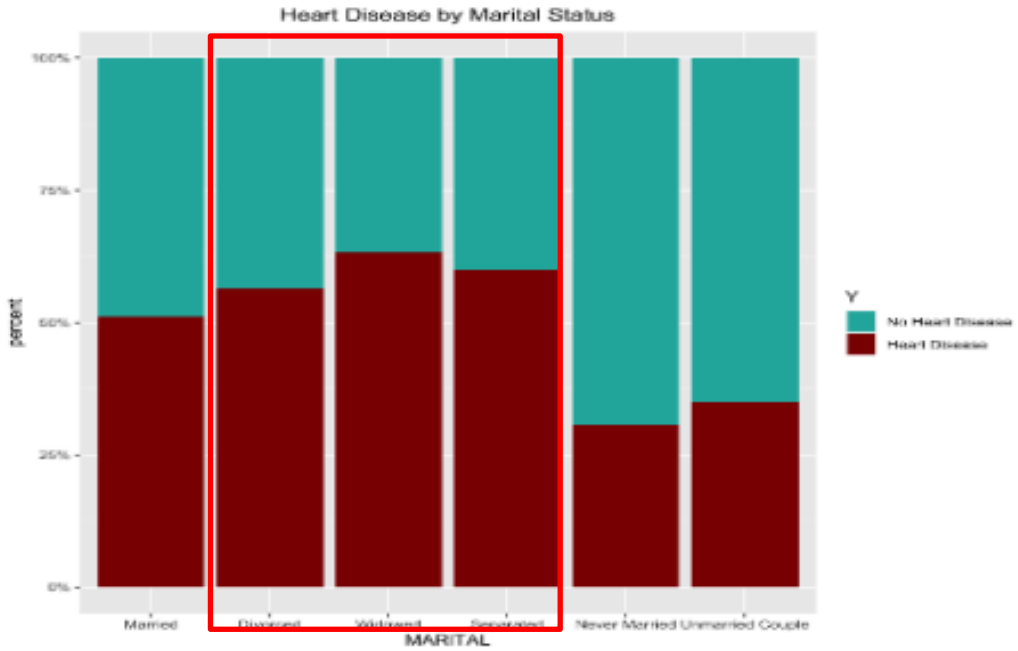
```
     Chi-square test of categorical association

Variables:  Y, MARITAL

Hypotheses:
   null:        variables are independent of one another
   alternative: some contingency exists between variables

Observed contingency table:
   MARITAL
Y      1    2    3    4    5    6
  0 5231 1320 1679  153 1519  224
  1 5469 1712 2888  229  670  120

Expected contingency table under the null hypothesis:
   MARITAL
Y      1    2    3   4    5   6
  0 5107 1447 2180 182 1045 164
  1 5593 1585 2387 200 1144 180

Test results:
   X-squared statistic:  709.708
   degrees of freedom:  5
   p-value:  <.001

Other information:
   estimated effect size (Cramer's v):  0.183
```

```
     Chi-square test of categorical association

Variables:  Y, INCOME2

Hypotheses:
   null:        variables are independent of one another
   alternative: some contingency exists between variables

Observed contingency table:
   INCOME2
Y      1    2    3    4    5    6    7    8
  0  660  643  821  974 1181 1517 1516 2814
  1  805 1033 1270 1400 1531 1755 1518 1776

Expected contingency table under the null hypothesis:
   INCOME2
Y     1   2    3    4    5    6    7    8
  0 699 800  998 1133 1295 1562 1448 2191
  1 766 876 1093 1241 1417 1710 1586 2399

Test results:
   X-squared statistic:  532.651
   degrees of freedom:  7
   p-value:  <.001

Other information:
   estimated effect size (Cramer's v):  0.158
```

Lim Zi Xiang

MODEL INSIGHTS

MARITAL

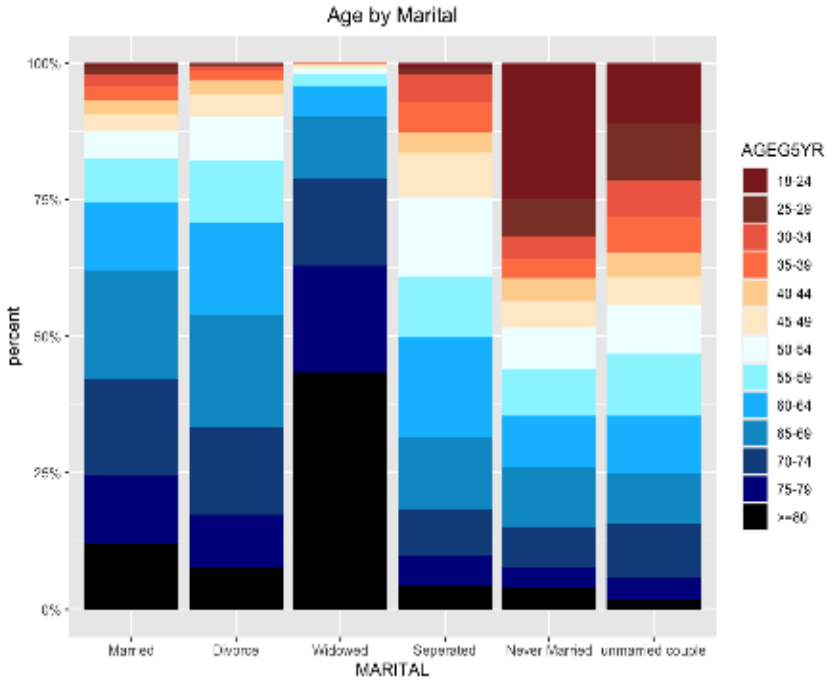


- Higher percentage of **"Divorced", "Widowed", "Separated"** who contracted heart disease

Lim Zi Xiang

MODEL INSIGHTS

MARITAL BY AGE

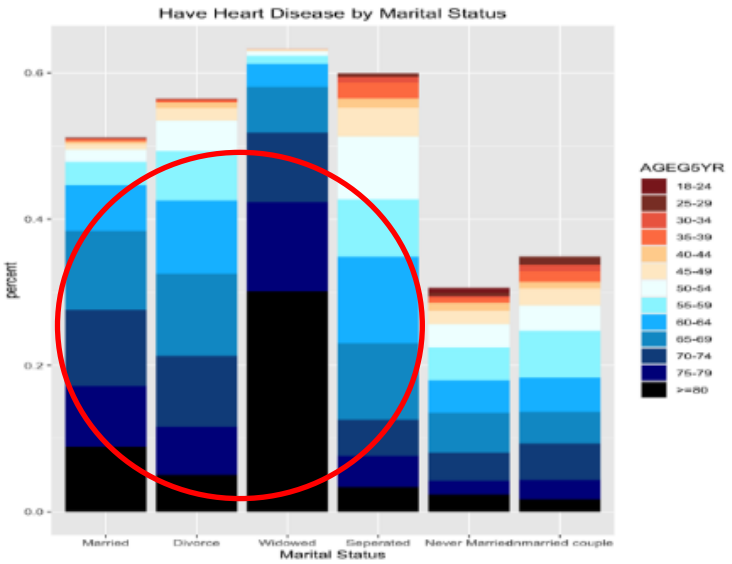


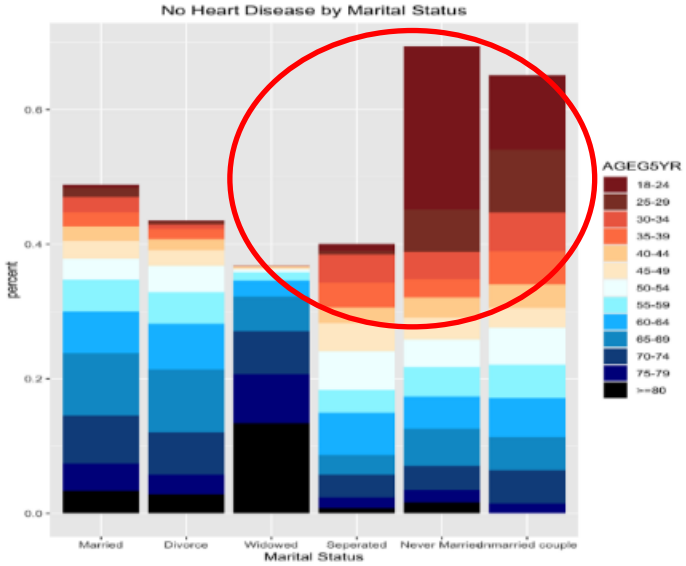
- **Respondents under 40** tend to fall under “non-married” and “unmarried couple”
- **Widowers** have the largest percentage of seniors

Lim Zi Xiang

MODEL INSIGHTS

MARITAL BY AGE



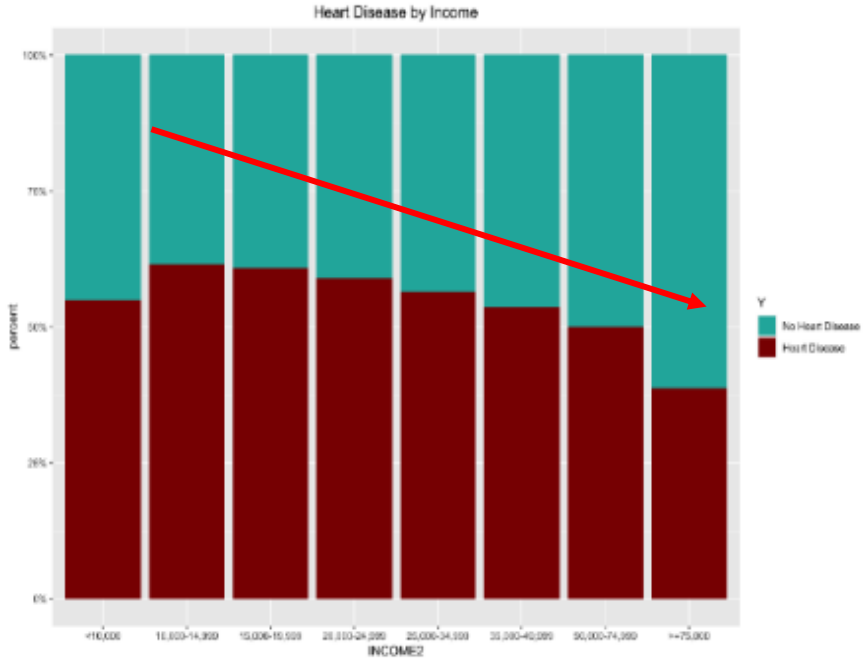


- **Age** factor in Marital Status may be the underlying factor in the prediction of heart disease

Lim Zi Xiang

MODEL INSIGHTS

INCOME

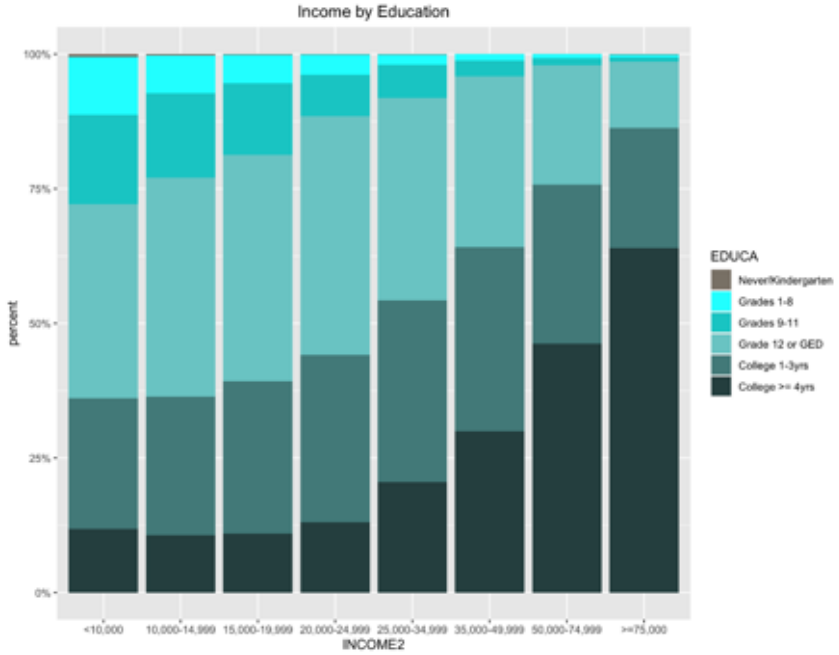


- As income increases, the risk of contracting heart disease decreases

Lim Zi Xiang

MODEL INSIGHTS

INCOME BY EDUCATION

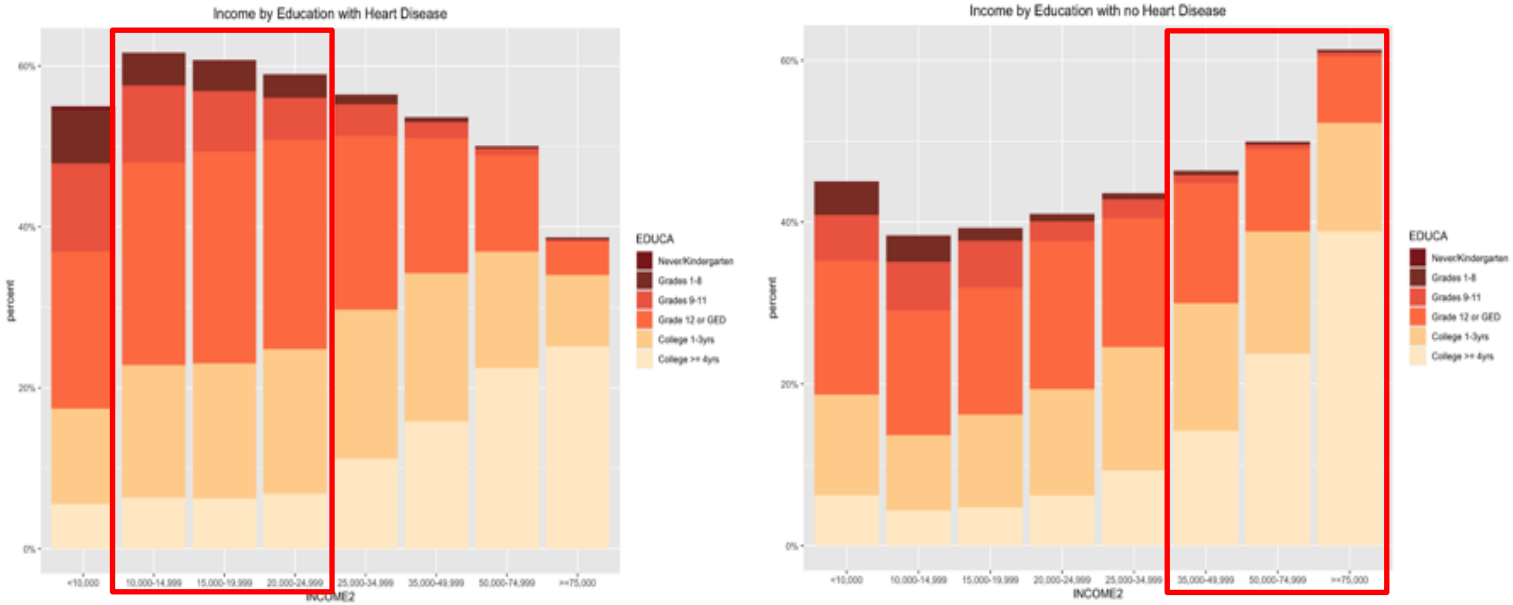


- Income has a positive correlation to the **level of education** one attains

Lim Zi Xiang

MODEL INSIGHTS

INCOME BY EDUCATION



- Correlation between Education -> Income -> Heart Disease

Lim Zi Xiang

MODEL INSIGHTS

MENTAL HEALTH

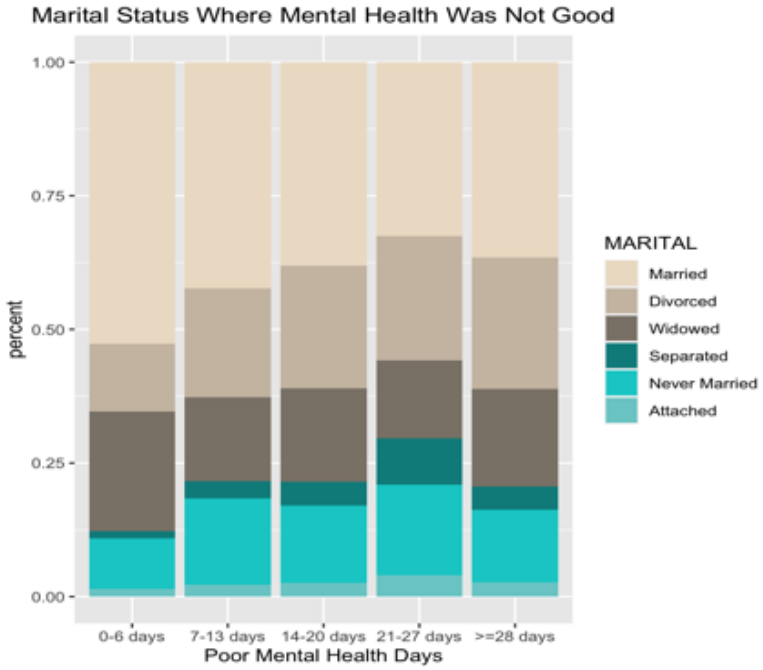
- Relatedness exist between marital status and heart disease through a syndrome named **“broken heart”** (AHA, 2022)

- Studies have shown that **low income and work stress** are researched to be linked to heart disease (ECS, 2019)

Lim Zi Xiang

MODEL INSIGHTS

MARITAL BY MENTAL HEALTH

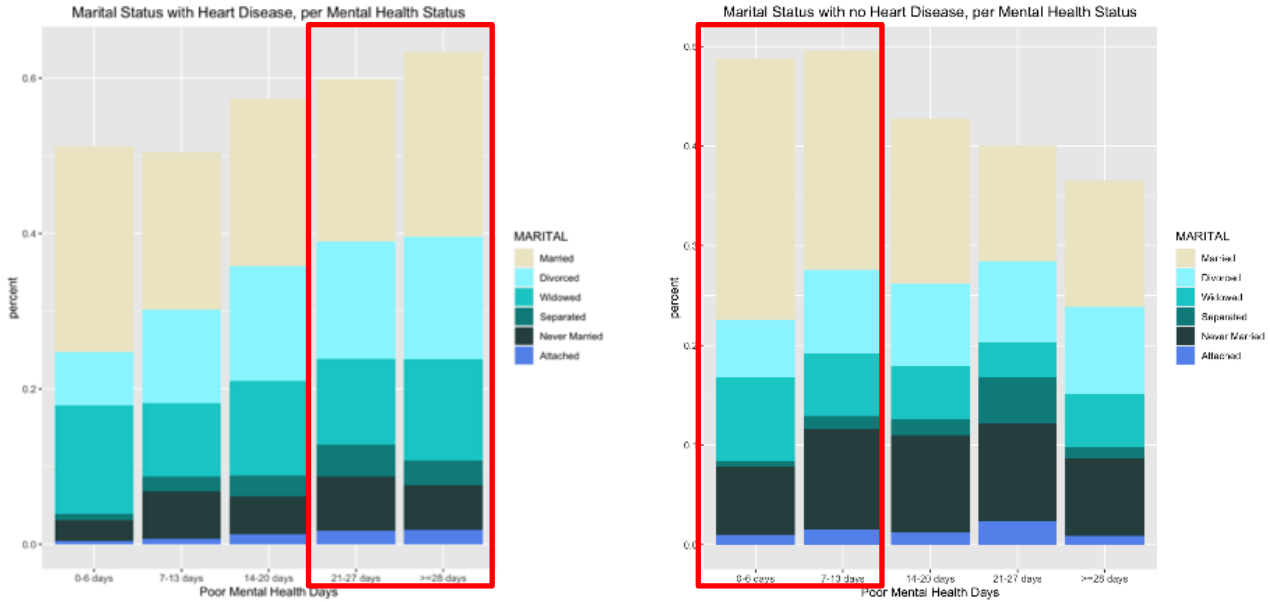


- As no. of poor mental health days decrease, ratio of married to divorce and separate increase

Lim Zi Xiang

MODEL INSIGHTS

MARITAL BY MENTAL HEALTH

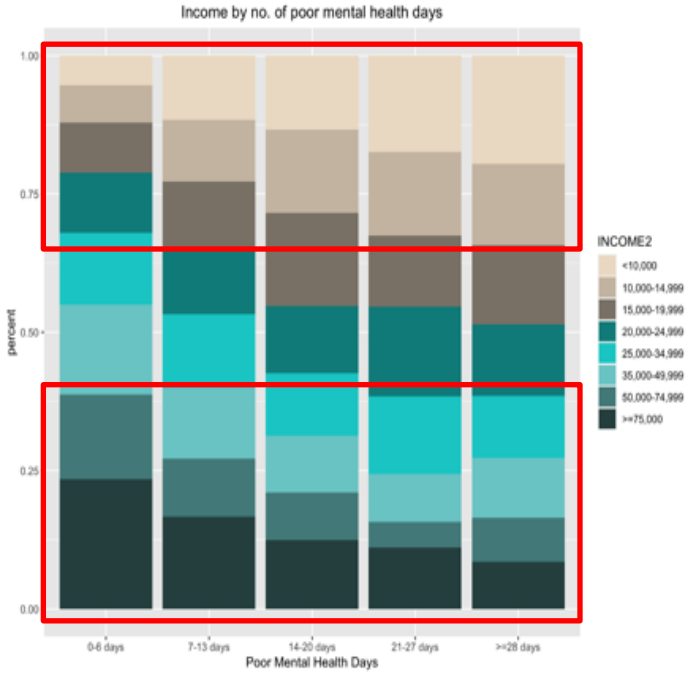


- Correlation between Marital Status -> Mental Health -> Heart Disease

Lim Zi Xiang

MODEL INSIGHTS

INCOME BY MENTAL HEALTH

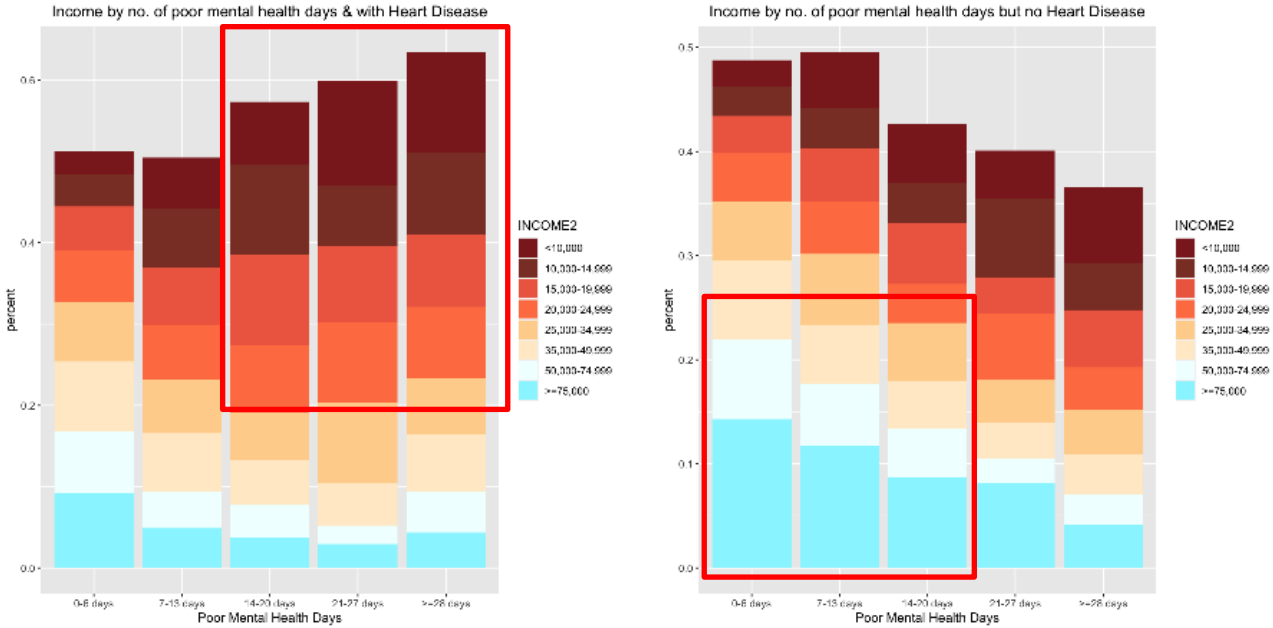


- As income increases, the no. of poor mental health days drop

Lim Zi Xiang

MODEL INSIGHTS

INCOME BY MENTAL HEALTH



- Correlation between Income-> Mental Health -> Heart Disease

Lim Zi Xiang

MODEL INSIGHTS

MENTAL HEALTH

- Correlation exists between
  - Income/Marital status
  - Mental Health
  - Heart disease
- More research has to be done
- More emphasis must be placed on mental health

Lim Zi Xiang

HUMAN TENDENCY

4 COMMON BELIEFS

"I am healthy, so health screening is unnecessary"

"Going for health screening is a hassle."

"I have done my health screening 5 years ago, why must I do it again?"

"Health screening is expensive."

(Raffles Health, 2018)

Lim Zi Xiang

PROPOSED BUSINESS SOLUTION

Prevention is **KEY** as first-time events may be fatal

Lim Zi Xiang

IS THIS APP A REPLACEMENT TO THE EXISTING MEDICAL TESTS?

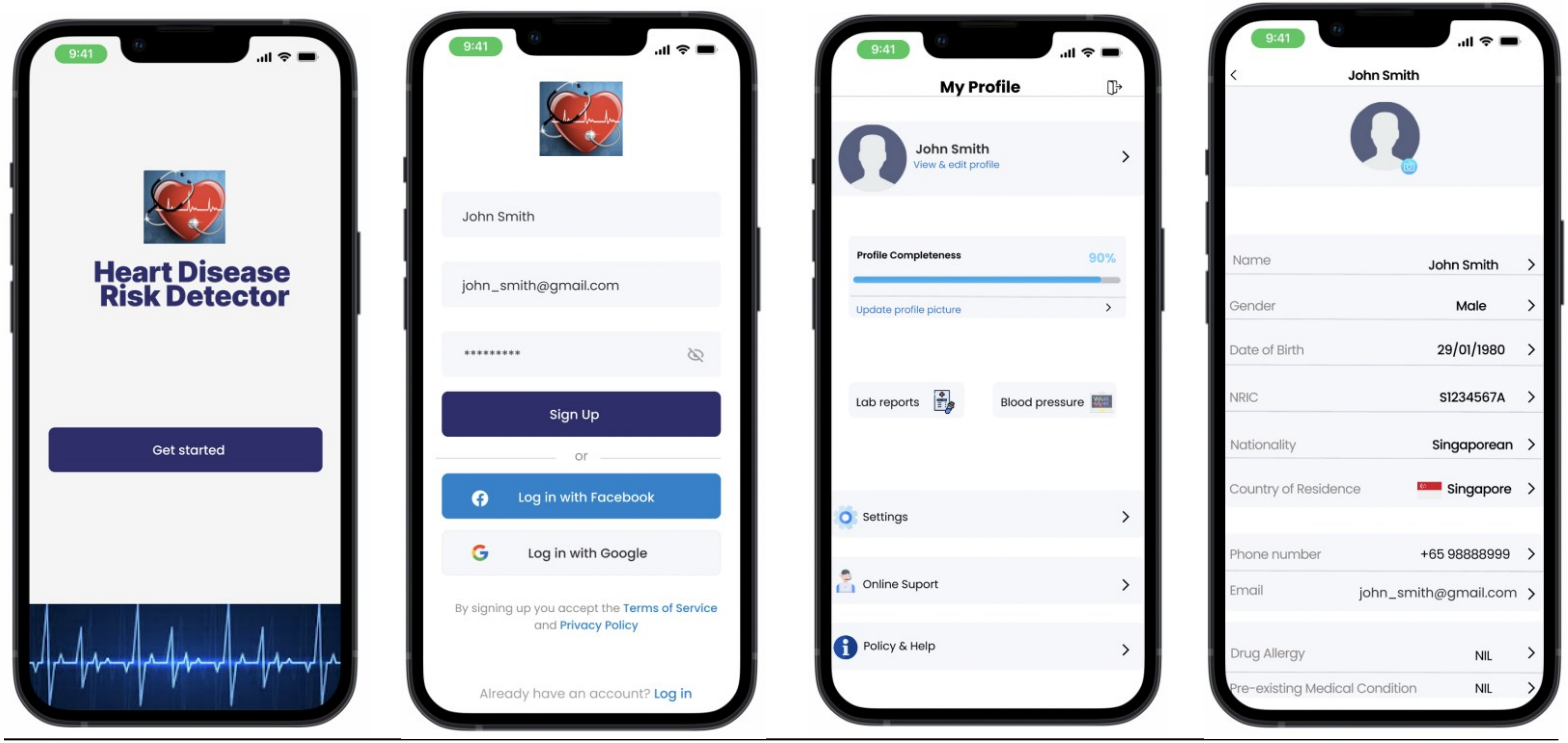
NO

The heart disease Risk Detector app is a supplement to the existing medical tests

Nagammai Senthil Kumar

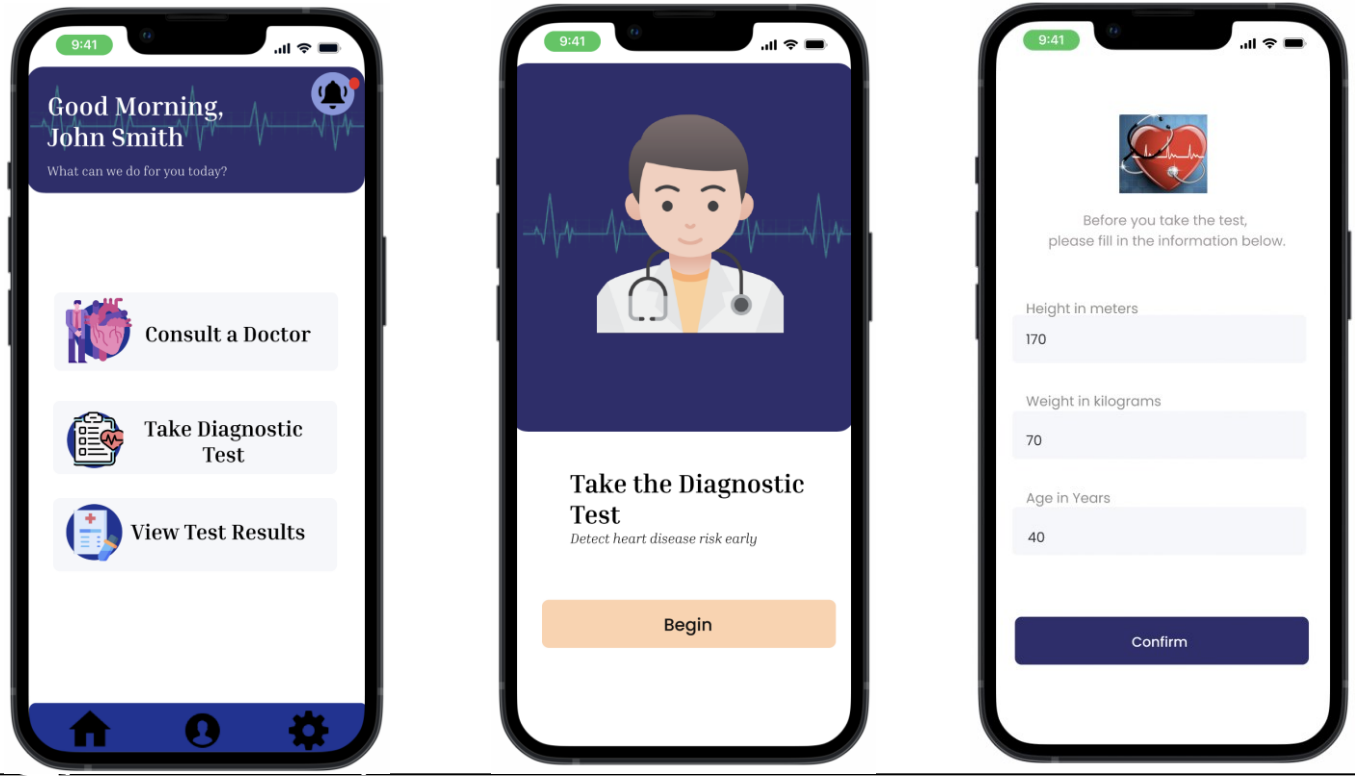
HEART DISEASE RISK DETECTOR APP

First, users create an account



Nagammai Senthil Kumar

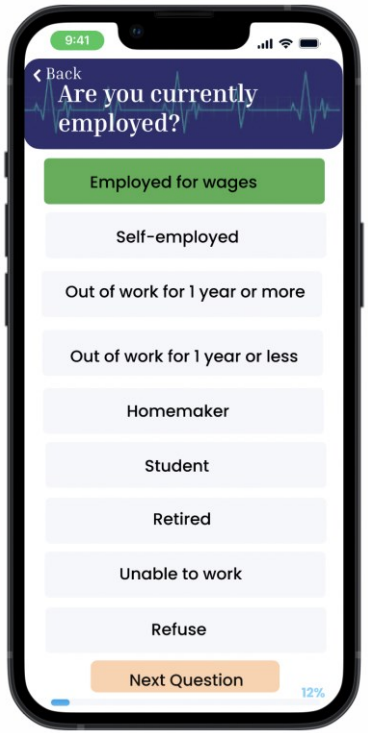
HEART DISEASE RISK DETECTOR APP

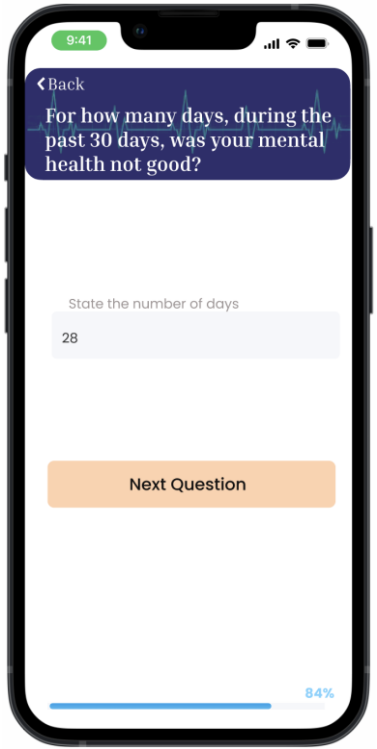


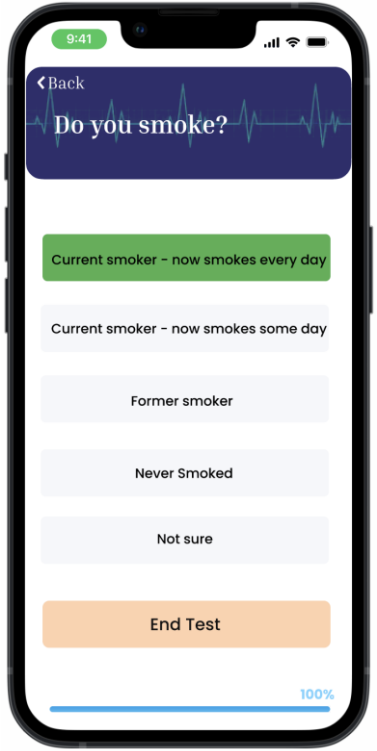
Nagammai Senthil Kumar

HEART DISEASE RISK DETECTOR APP

Users take the given test, providing key non-medical data







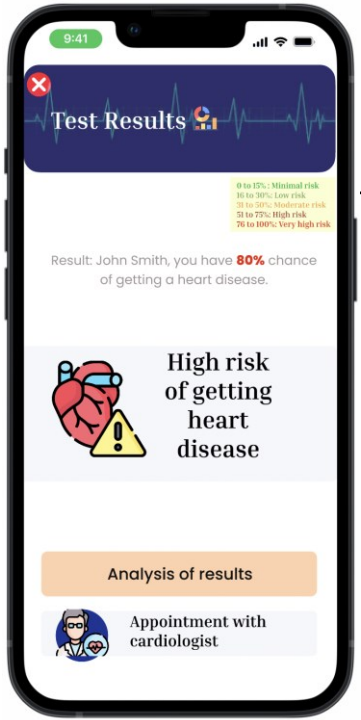
Nagammai Senthil Kumar

HEART DISEASE RISK DETECTOR APP

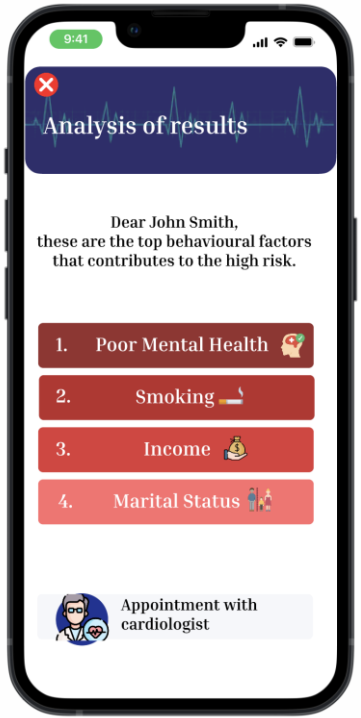
Data collected from the users is run through the CART model

Nagammai Senthil Kumar

HEART DISEASE RISK DETECTOR APP



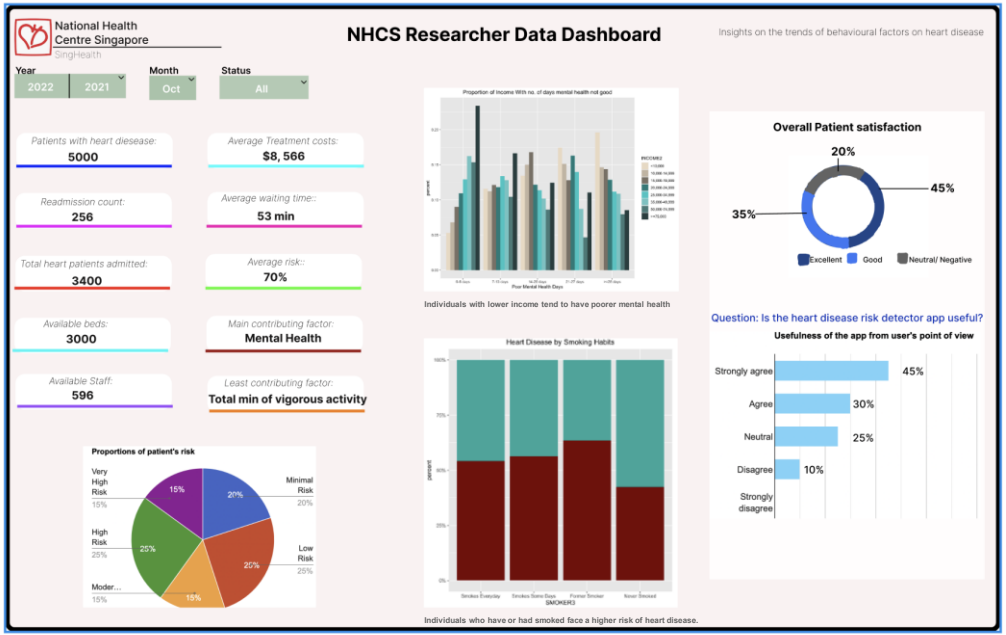
0 to 15% : Minimal risk
16 to 30%: Low risk
31 to 50%: Moderate risk
51 to 75%: High risk
76 to 100%: Very high risk



Results of the test will be displayed based on the CART model and patients should take the necessary actions accordingly.

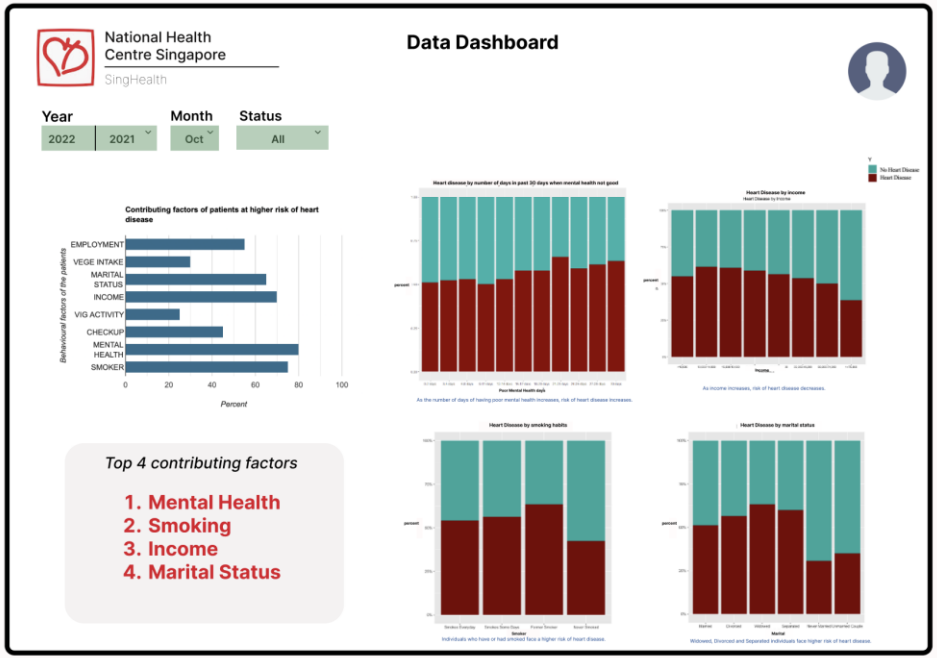
Nagammai Senthil Kumar

NHCS DASHBOARD MOCK-UP



NHCS Researcher's Data Dashboard

APP USER'S DASHBOARD MOCK-UP



User interface

Nagammai Senthil Kumar

VALUE PROPOSITION

Valued benefits over existing models:

1. Efficiency and cost effective
2. Comprehensible and user friendly
3. Alleviates stress on NHCS and the medical professionals

Nagammai Senthil Kumar

VALUE PROPOSITION

1. Efficiency and cost effectiveness

Nagammai Senthil Kumar

VALUE PROPOSITION

2. Comprehensible and user-friendly

Nagammai Senthil Kumar

VALUE PROPOSITION

3. Alleviates stress on NHCS and doctors

LIMITATIONS

1. Success of our app
2. Presence of missing values in the dataset
3. Our app may pose difficulty in terms of usage for elderlies

Nagammai Senthil Kumar

CONCLUSION

In order to make well-informed and timely decisions, NHCS could implement the Heart Disease Risk Detector app which allows users to self-check their risk of heart disease.

THANK YOU!