Kiribati Household Income and Expenditure Survey (HEIS)

The objectives of the survey are to:

* Provide information on income and expenditure distribution within the population
* Provide data for a nutritional analysis of household food consumption

# CHAPTER 1

## Business Understanding

* What is the best way to allocate the limited Kiribati HEIS budget to its residents income and expenditure distribution within the population?

### Objectives;

* Identify cost-saving opportunities
* Assess the financial impact of compliance costs
* Evaluate the financial implications of regulatory changes

# CHAPTER 2

## Analytical approach

Determining analytical type based on business question

Business question is to determine probabilities of an action?

Use a predictive model.

Use descriptive model.

No! /Question is to show

Yes! Relationship.

.

.

Since our objectives are based on learning human behaviour, an appropriate response would be to use clustering associate approaches. A decision tree classification model to identify the combination of decisions that led to costly/cost saving outcomes. Use a decision tree to determine the cost effect of an expenditure.

Was it necessary at the moment?

Label as “Costly”

Type of expenditure

No!

Yes!

# CHAPTER 3

## Data requirements

Select a suitable time frame

2005 population census – 2006 HEIS

# CHAPTER 4

## Data Collection

Identify data sources to know where to find data elements that are needed.

Data source, <https://pacific-data.sprep.org/resource/final-report-kiribati-2006-hies>

Sampling design used a multi-stage selection methodology

# CHAPTER 5

## Data understanding

Statistics on each variable. Mean median mode max min standard deviation

Correlations formed

Histogram to understand distribution

Comparison was done between household and population counts of the 2005 population cencus and the 2006 HEIS

# CHAPTER 6

## Data preparation (time consuming)

Data cleaning

Automate some of the data collection and preparation processes will come in handy

Transforming data to make it easier to work with

Address missing or invalid values, remove duplicates,

Average household income.

Average household expenditure.

# CHAPTER 7

## Modelling

What is the purpose of data modelling?

What the characteristics of this process

Descriptive model – based on an action

Predictive model – tries to yield yes/no outcomes, uses a training set

Build a predictive model,

Question; is an expenditure costly? Yes/no

How could the accuracy of the model be improved to predict the yes outcome?

# CHAPTER 8

## Evaluation

Done during model development and before deployment

Assess quality of model

Does it meet the initial request?

### Diagnostic measure phase

To ensure the model is working as intended

### Statistical significance testing

To ensure data is properly handled and interpreted within the model

# CHAPTER 9

## Deployment

Deploy model to an environment for test drive

Test environment or a limited group of users

Easy to use

Train staff/test environment

# CHAPTER 10

## Feedback

Feedback from users will help refine the model and assess it for performance and impact

Make adjustments as long as solution is required

The more you know, the more you will want to know