**Data Visualization**

**COS30045**

**Rising and falling health concerns within New Zealand**

**Authors and key information:**

* Joel Downie 104540652
* JULIAN GIORGILLI JULIAN GIORGILLI 103618985
* Due 7 June 2024 semester 1

**Website:**

<https://swinburne.instructure.com/courses/56073/assignments/580947>

**Table of content:**

1. Introduction

1.1: Background and Motivation

1.2: Visualisation Purpose

1. Data

2.1: Data Source

2.2: Data Processing

1. Visualization design

**Introduction:**

**1.1 Background and Motivation:** The main focus group we intend for our data to carter for are new Zealand high school students aiming to provide them with general health related trends such as food consumption and body weight statistics within their country compared to others and discuss any important risks involved with the data observed over the past few years. This is important to educate the students on factors leading to increased food consumption/body weight and other substances providing them with the knowledge necessary to understand the risks of an unhealthy lifestyle and techniques on how to keep a more balanced and healthy lifestyle choices.

**1.2 Visualisation Purpose:** The purpose of our visualisations will be to provide a visual display of our data to the students in a far more readable and easy to follow format compared to a spreadsheet. This will help provide a more clear picture to the students on the current health statistics of their country making it easier for them to understand any issues and other key information helping give them the information needed to make more educated decisions for how they lead there life in the future aiming to reduce.

The website will have these important features.

* A comparison between

# 2 – Data

## 2.1 – Data Source

The two datasets regarding Food consumption (share of population consuming vegetables daily) and alcohol consumption, are collected from OECD and feature data ranging from 2012-2022. OECD (Organization for Economic Co-operation and Development) is a global organization that works towards building bigger and better policies, for better lives. The overarching goal of the OECD is to “shape policies that foster prosperity, equality, opportunity and well-being for all.”

The datasets are in tabular form and consist of the following attributes:

### Dataset: Alcohol Consumption

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Description** | **Type** |
| Reference Area | This field contains the country, which the data is from | Categorical (Text) |
| Frequency of observation | This field contains the frequency of observation (in which the data is observed) | Categorical (Text) |
| Measure | This field contains the measure, being Alcohol Consumption | Categorical (Text) |
| Age | This field contains the age of the population in the dataset, all being 15 and over | Categorical (Text) |
| Unit of measure | This fields contain the unit, in which the data is measured | Categorical (Text) |
| TIME\_PERIOD | This field contains the time period or year in which the data represents | Ordinal (Number) |
| OBS\_VALUE | This field contains the observation value, (the value of the data we are looking for) | Continuous (Number) |
| Observation Status | This field contains the status of the observation value. | Categorical (Text) |

### Dataset: Food consumption and consumption

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Description** | **Type** |
| Reference Area | This field contains the country, which the data is from | Categorical (Text) |
| Frequency of observation | This field contains the frequency of observation (in which the data is observed) | Categorical (Text) |
| Measure | This field contains the measure, being food supply and consumption | Categorical (Text) |
| Age | This field contains the age of the population in the dataset | Categorical (Text) |
| Sex | This field contains the sex of the population in the dataset | Categorical (Text) |
| Unit of measure | This fields contain the unit, in which the data is measured | Categorical (Text) |
| Measurement method | this field contains the measurement method of the dataset | Categorical (Text) |
| TIME\_PERIOD | this field contains the year, or time period of the data | Ordinal (Number) |
| OBS value | This field contains the observation value, (the value of the data we are looking for) | Continous (Number) |
| Observation status | This field contains the status of the observation value. | Categorical (Text) |

## 2.2 – Data Processing

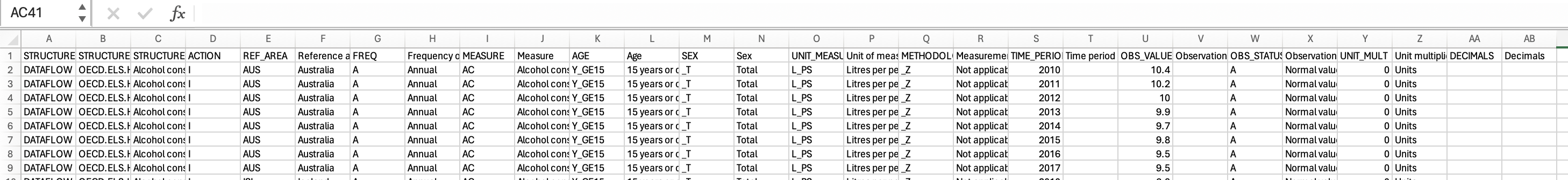
In the analysis process, we utilized the OECD data explorer to filter and extract relevant datasets, tailored to our overarching theme of health. The OECD data explorer allowed me to selectively navigate through extensive data categories and refine data points based on various parameters, such as country, age and sex. It also allowed me to select a time period of 2012-2022 for both datasets.

A screenshot of a computer

Description automatically generated

Following the data acquisition phase, I utilized Excel, a versatile platform for data manipulation, to further refine the datasets. By using Excel's robust data management capabilities, I was able to efficiently remove unnecessary columns that were not relevant to the analysis. This step was crucial as it helped in eliminating redundant information, thereby streamlining the datasets for easier handling and analysis.

**Before:**



After:

A table with text on it

Description automatically generated

The combination of OECD Data Explorer's precise data filtering capabilities and Excel's powerful data cleansing features significantly enhanced the overall quality and usability of the data, making it an invaluable asset for conducting comprehensive and accurate research.

Joels work(ADJUST FORMAT LATER)

**Data:**

2.1: Our data has been sourced from OECD health statistics found here ([https://data-explorer.oecd.org/?fs[0]=Topic%2C0%7CHealth%23HEA%23&pg=0&fc=Topic&bp=true&snb=46](https://data-explorer.oecd.org/?fs%5b0%5d=Topic%2C0%7CHealth%23HEA%23&pg=0&fc=Topic&bp=true&snb=46))

**Share of population who are daily smokers:**

An average has been made as an alternative display if we would like to use a pie chart

A screenshot of a computer

Description automatically generated

**Percentage of population who are obese:**

The red data has been filled in as estimated data to help keep the data consistent when it displays.

A screenshot of a table

Description automatically generated

# 3 – Visualisation Design

## Sketch 1: Line Chart - Alcohol Consumption over time: Norway vs Costa Rica

A graph with lines and a red line

Description automatically generated

* Plotting years on the x-axis between 2011-2021
* Litres per person on the y-axis
* Different colours for Norway and Costa Rica to easily differentiate and compare them.
* This will allow viewers to easily see how consumption has changed over time and compare the trends between the two countries.

## Sketch 2: Bar Chart – Average Alcohol Consumption by country

A graph of bar graph

Description automatically generated with medium confidence

* Calculate average alcohol consumption per country.
* Country on x-axis
* Average consumption by litres per person on y-axis.
* This type of visualization is straightforward and effective for quickly understanding differences in average consumption between the multiple locations.

## Sketch 3: stacked area chart for food supply over time by country

A graph of different colored lines

Description automatically generated

* This chart could show the total food supply over time for each country listed.
* Each country would have its own layer in the chart, which could be color-coded to distinguish between the countries.
* This visualization would be great for observing trends in food supply changes over time and comparing them across different nations.

## Sketch 4: Clustered Bar Chart of Daily Consumption Rates:

A graph of different colored bars

Description automatically generated with medium confidence

* This chart compares the daily consumption rates of fruits and vegetables across different countries, segmented by gender.
* Each bar represents a country, with the bar divided into segments that show the proportion of males and females consuming fruits and vegetables daily.
* Use different colors or patterns to differentiate between fruits and vegetables within each bar. Further, differentiate the segments within these categories by gender using shades or texturing.
* This visualization is excellent for observing the dietary habits across different demographics within each country. It allows for an immediate visual comparison of gender-based consumption patterns between different nations.

Joelswork(ADJUST FORMAT LATER)

**Visualization designs:**

**Percentage of population who are obese:**

**A close-up of a graph

Description automatically generated**

**Share of population who are daily smokers and an average of the Share of population who are daily smokers:**

**A diagram of a pie chart

Description automatically generated**