Data Visualisation Project

Assignment 1 – Report

World Happiness: Analysis of Key Score Attributes

Module code : SPEC9995: 2022-23

Ciaran Finnegan
Student No : d21124026
TU060 Yr. 2
Data Science MSc

28/10/2022

Table of Contents

1 Assignment Overview

1.1 High Level Description – Global Happiness Dataset

This document covers the design, implementation, and observations on all parts of the November 2022 CA(1) for the Data Visualisation module (TU060 – DS – Year 2).

This assignment is based on a starting point with a Kaggle dataset that tracks a 'Happiness Index' for most countries in the world; (https://www.kaggle.com/datasets/unsdsn/world-happiness)[1].

Although the Kaggle dataset contains a number of attributes that explain the contribution of GDP, and other factors, towards what makes a nation happy (their position on the index), the data is not readily understandable.

By combining the original Kaggle dataset with other datasets with key information metrics for countries around the world, it is possible to build up a more meaningful set of visuals.

These graphs will elaborate on which parts of the globe are generally more 'happy', and where the least happy nations can be found. Insights on some the key elements that combine to drive happiness will be displayed in the dashboards produced by this assignment.

The datasets (and key attributes) that have been combined with the source Kaggle dataset are;

- Our World in Data – GDP per capita (2015)[2].
- Fragile States Index – Political stability per nation (2015)[3].
- CAI World Fact Data – Life expectancy at birth (2105)[4].

Section 5 of this document lists the reference locations for all of these datasets.

For reference, the final augmented 'happiness' dataset used for this assignment (***Global_Happiness_Metrics_CA1.xlsx***), along with all the contributing datasets, is stored on GitHub at this location:

https://github.com/JackDaedalus/DataVizCA1_Datasources_2022.git

1.2 Intended Audience

All reporting dashboards were developed and presented in Tableau.

The intended audience is anyone interested in the key factors that influence the quality of live in nations across the globe.

The term 'happiness', as used in this CA report, is really just a shorthand for a measure of how well a country provides for the material and social wellbeing of its citizens.

The dashboards in this report are intended to provide a visual insight into some of the major causes that drive this measure of 'happiness' across the world.

This report is a snapshot into the scores for global happiness in **2015**. Each of the supplementary datasets were filtered to extract only data from that year. The reason for selecting 2015 was because this was the most recent year for which accurate GDP numbers were available in this selection of datasets.

The choice of 2015 also reflects a period after the economic downturn of 2008-2011 but before the disturbance of the Covid-19 pandemic. Therefore, the year 2015 might be assumed to be a slightly better reflection of 'happiness' attributes in the early 21st century, as it somewhat avoids periods of significant disruption.

2 Part 1: Data Exploration

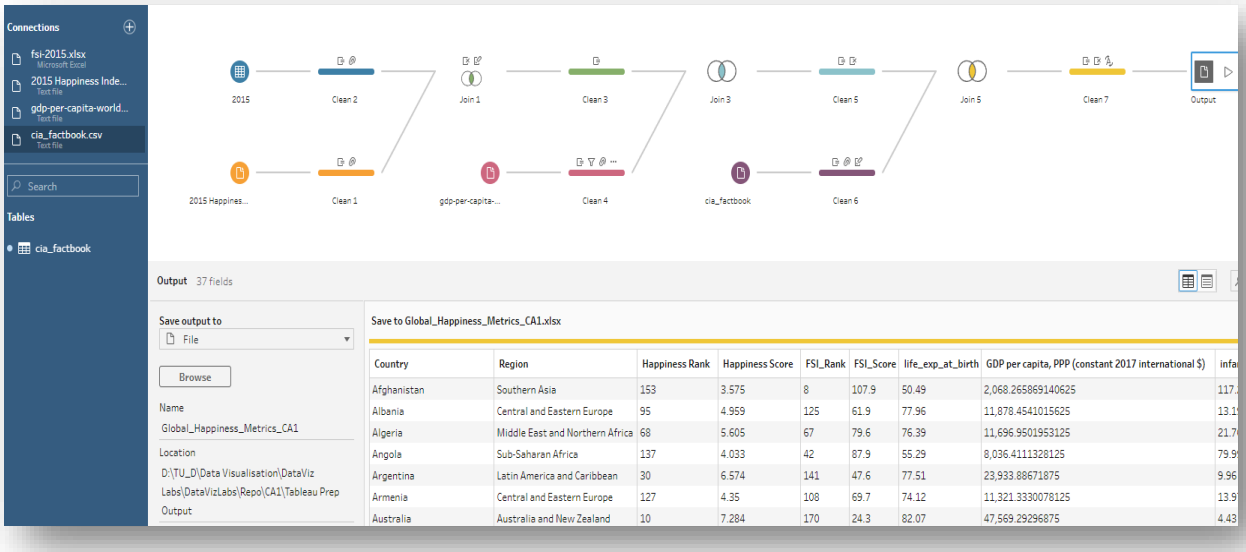
2.1 Dataset Structure – Cleaning and Joining Datasets

The quality of the information contained within the datasets listed in Section 1.1 is generally good, with relatively few records that are missing attribute missing data.

However, in order to streamline the process of creating the final output dataset for visualisation, this assignment used the features provided by the **Tableau Prep Builder** tool.

The joins across the datasets were driven by the name of each country, so a progressive set of steps were required to *trim* and alter this text information to correctly join the data.

The screenshot below shows the process followed by Tableau Data Prep to join all five datasets into a single final dataset.



The layout of the final dataset gathers the required attributes together (other columns contain additional reference material, but the screenshot above displays the primary data points for the visualisations in this assignment).

2.2 Exploration Process

Section 2.1 (above) described the process to augment the original Kaggle Global Happiness dataset.

Although the Kaggle dataset provides measures as to why factors such as GDP, trust in government, and life expectancy contribute to a nation's sense of 'happiness', the additional datasets attributes provide a set of more explicit, and understandable, metrics.

The Exploration dashboard provides a graphical display of these key metrics. The relationships between these data elements and 'happiness' are examined in more depth in the Insight Dashboard (Section 3).

2.3 Data Exploration: Key Results

The primary exploratory data elements extracted from our augmented dataset, and displayed on the first dashboard, are;

- Countries with the Top/Bottom 10 'Happiness' Scores.
- A Geographical display of the relative 'happiness' scores across the countries of the world.
- A relative display of political fragility scores for countries across the world.
- To improve the quality of the analysis, countries have also been assigned to specific identifiable geo-political regions. The exploratory dashboard provides a graphical view of;
  - Where these regions can be found on the globe.
  - An overview of the countries that constitute each region.
  - A view of the average GDP for each region (Tableau has built in drill down features to allow for a lower level interrogation of this data).

The actual number of countries in the world is just over 200. Our exploratory analysis is limited to 158 because not every country has data tracked in the Kaggle dataset. It was necessary to focus on those countries with which we could build a complete set of attributes.

Looking at the graphical spread of data coverage in our graphs, it is reasonable to assume that our dataset sample is large and broad enough to make the global inferences in our Insights Dashboard.

2.4 Tableau Public Dashboard - Location

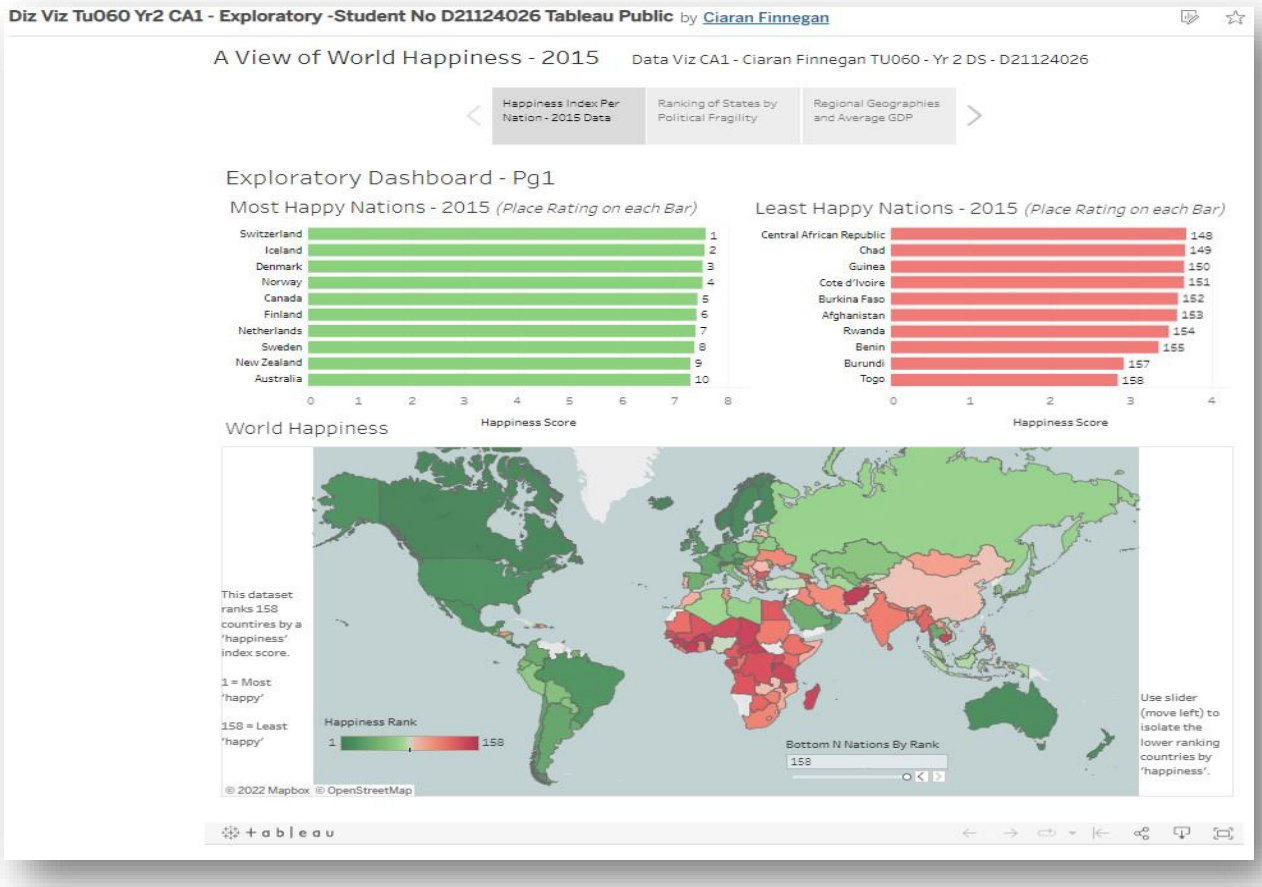
The World Happiness Exploratory Dashboard for this assignment is accessed on the Tableau Public server at this location;

https://public.tableau.com/views/DizVizTu060Yr2CA1-Exploratory-StudentNoD21124026TableauPublic/AViewofWorldHappiness?:language=en-US&:display_count=n&:origin=viz_share_link

2.5 Tableau Public Dashboard – Screenshot Overview

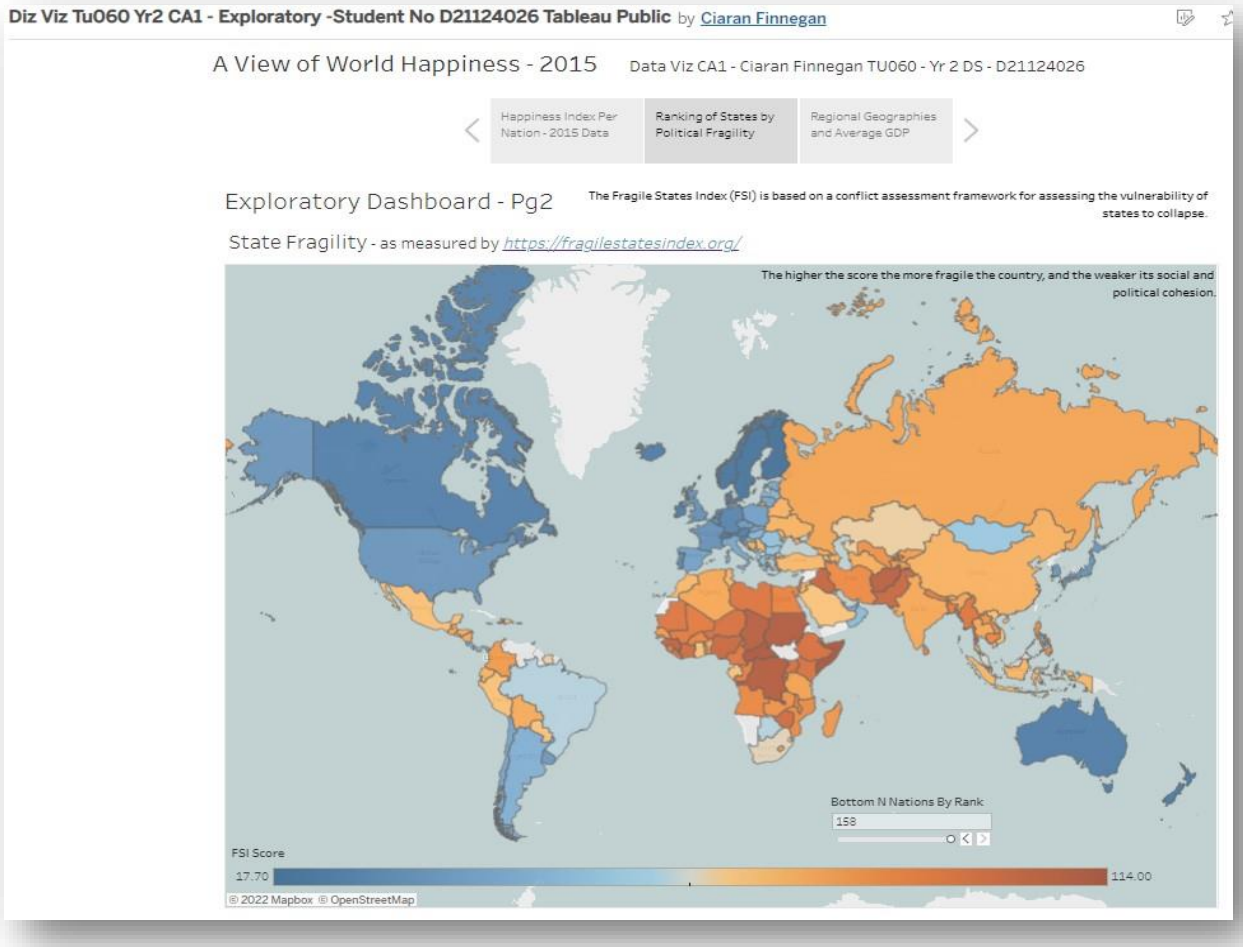
The Exploratory Dashboard for this assignment has been constructed as a Tableau three-part 'story', reflecting the results breakdown as described in Section 2.3.

2.5.1 Exploratory Dashboard – Pg1 *Nations Ranked by Happiness in 2015*



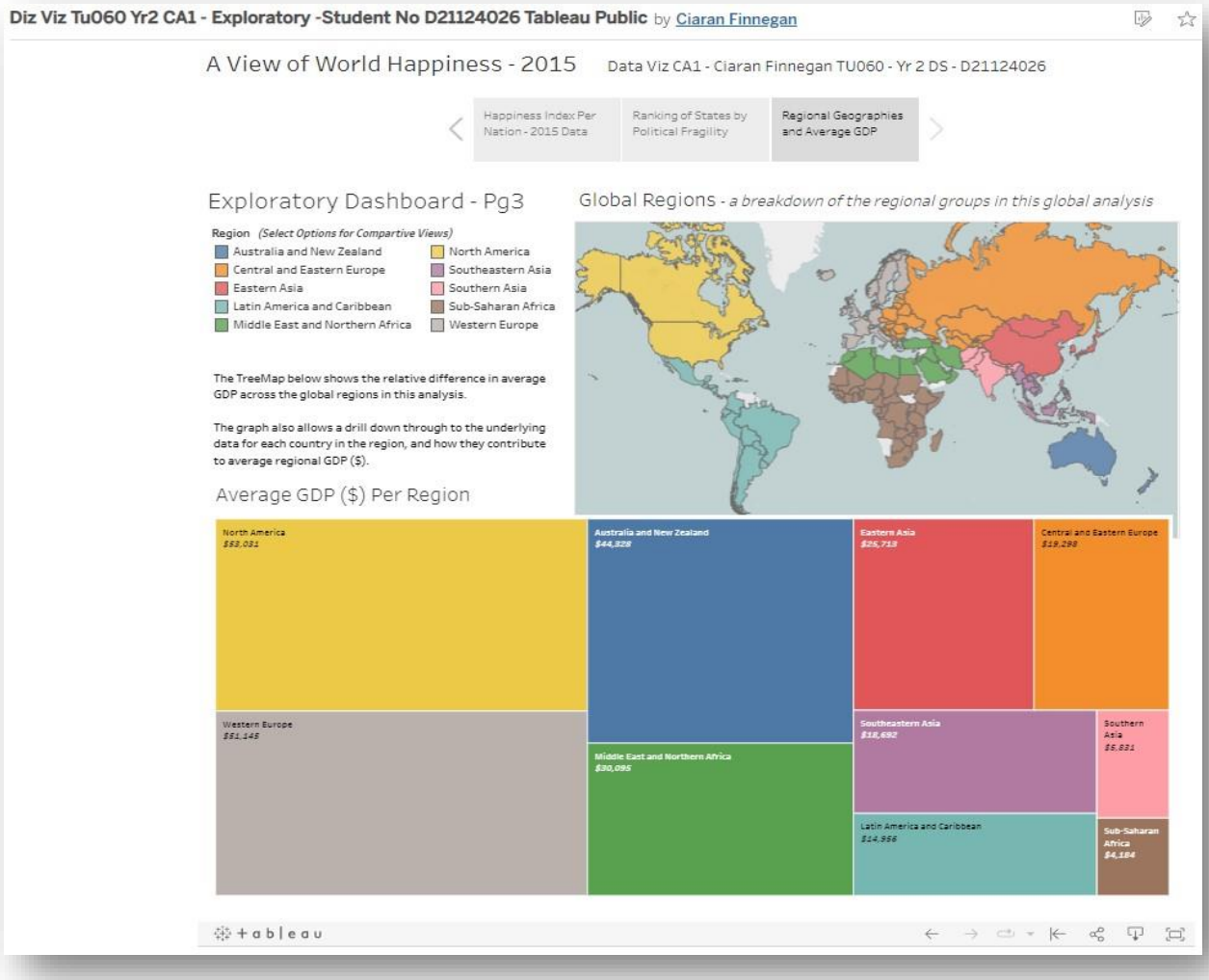
- The user can filter on the bar charts to isolate individual countries in the World Map.
- The Slider control on the map allows the user to display the range of less 'happy' nations.

2.5.2 Exploratory Dashboard – Pg2 *States Ranked by Political Fragility in 2015*



- The Slider control on the map allows the user to display the range of less politically stable nations.

2.5.3 Exploratory Dashboard – Pg3 *Regional Breakdown / Average GDP 2015*



- The user can dynamically filter on geo-political regions.
- Drill down options are provided by Tableau to elaborate on the regional/country data behind the individual cells in the TreeMap.

3 Part 2: Data Insights

3.1 Obtaining Insights from the Data

The augmented Kaggle World Happiness dataset allow us, for this assignment, to infer the relationship that the following national data attributes have towards a countries sense of happiness;

1. Wealth (measured in GDP)
2. Political stability and inclusiveness
3. Life expectancy at birth

It could be argued that the general insights in this assignment dashboard into what makes a nation's citizens 'happy' are relatively obvious.

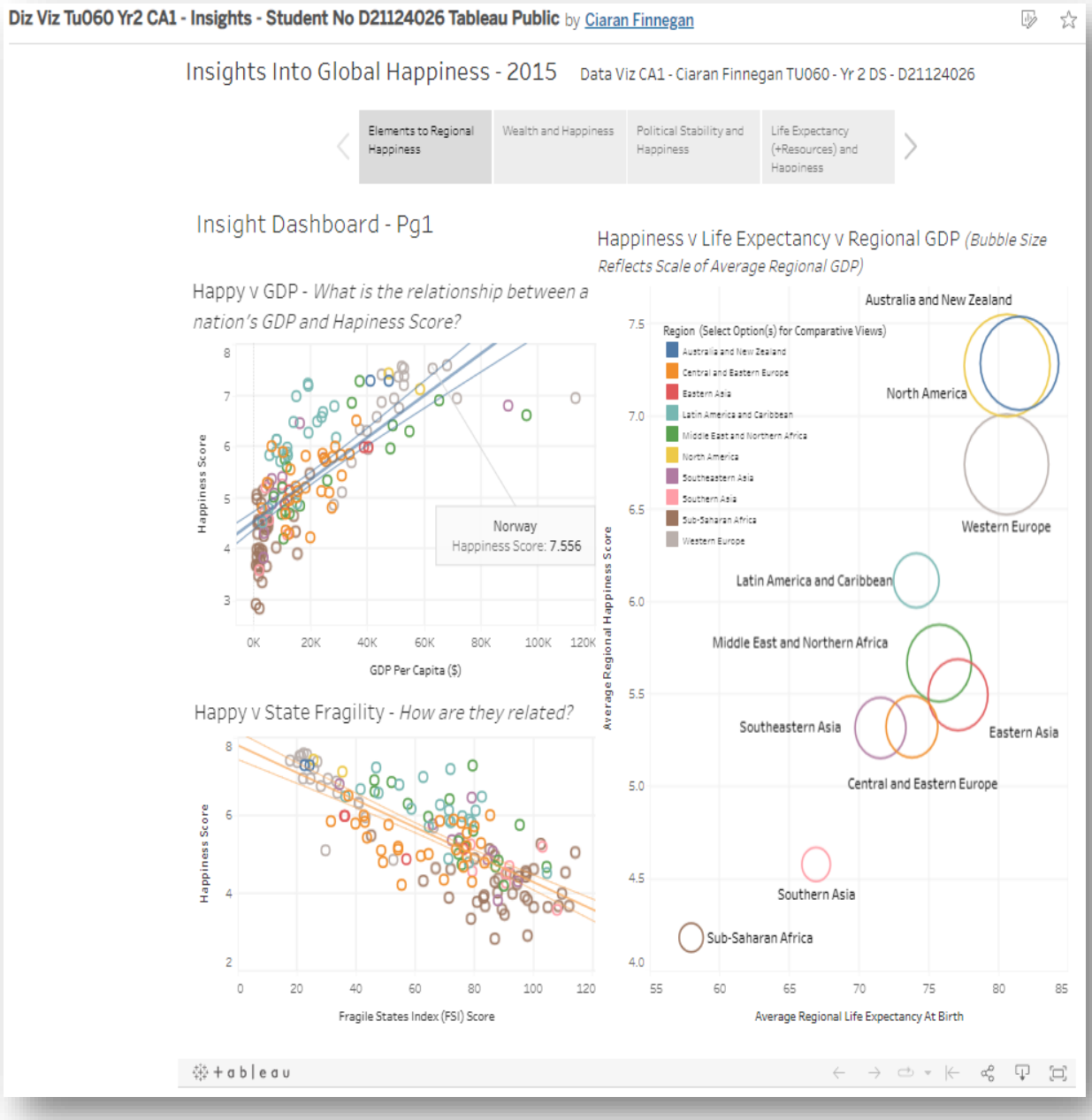
However, the Insights Dashboards calls out some interesting characteristics in the data, along with emphasising the impact of the global wealth gap.

Again, the Insight Dashboard follows a Tableau 'story' format.

The first page is a general dashboard with three graphs that model the attributes listed above against the 2015 Kaggle Happiness score for individual countries.

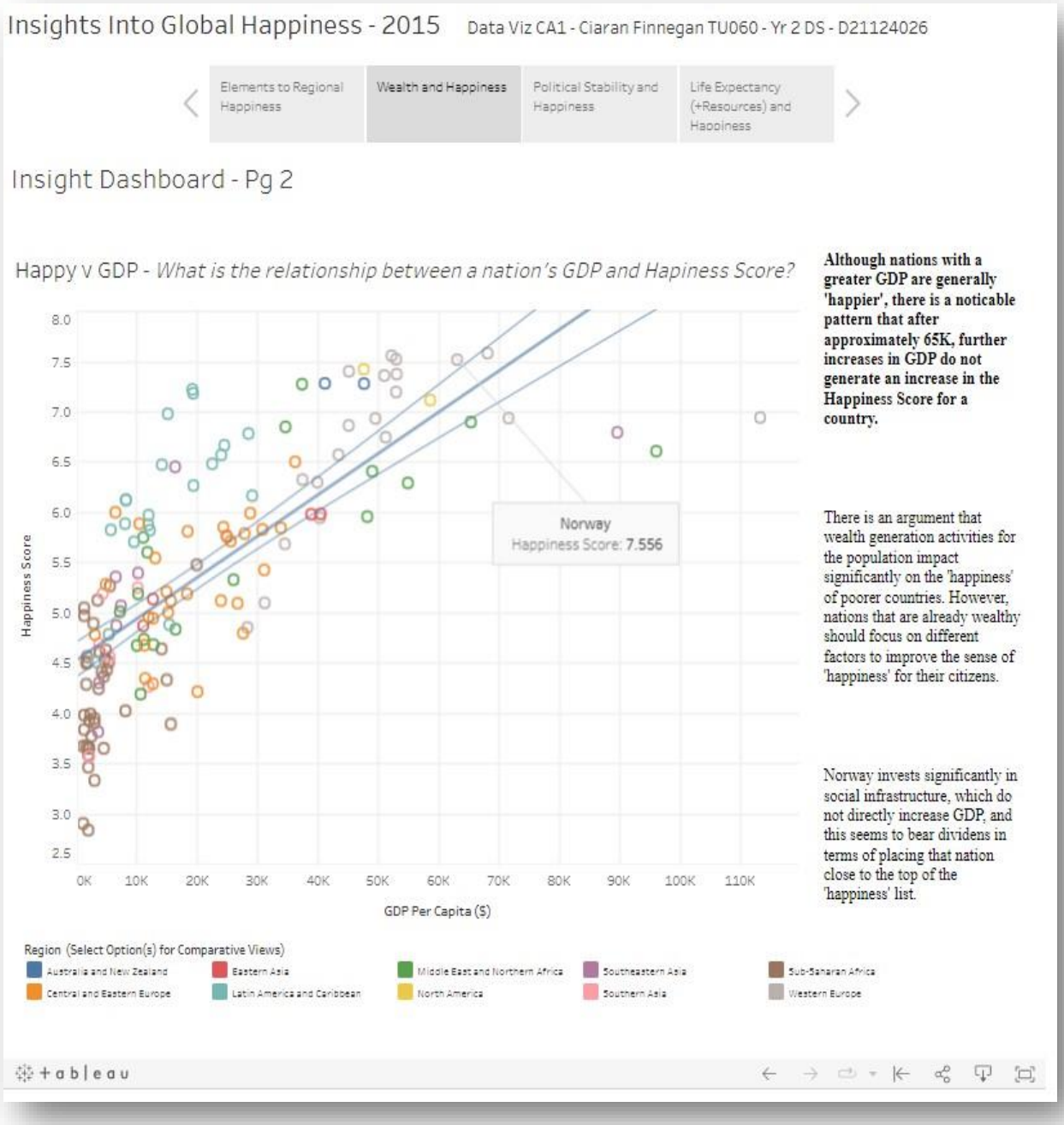
The following three pages elaborate on these relationships in further detail, calling out key points of interest.

3.1.1 Insights Dashboard – Pg1 *Elements of Regional Happiness in 2015*



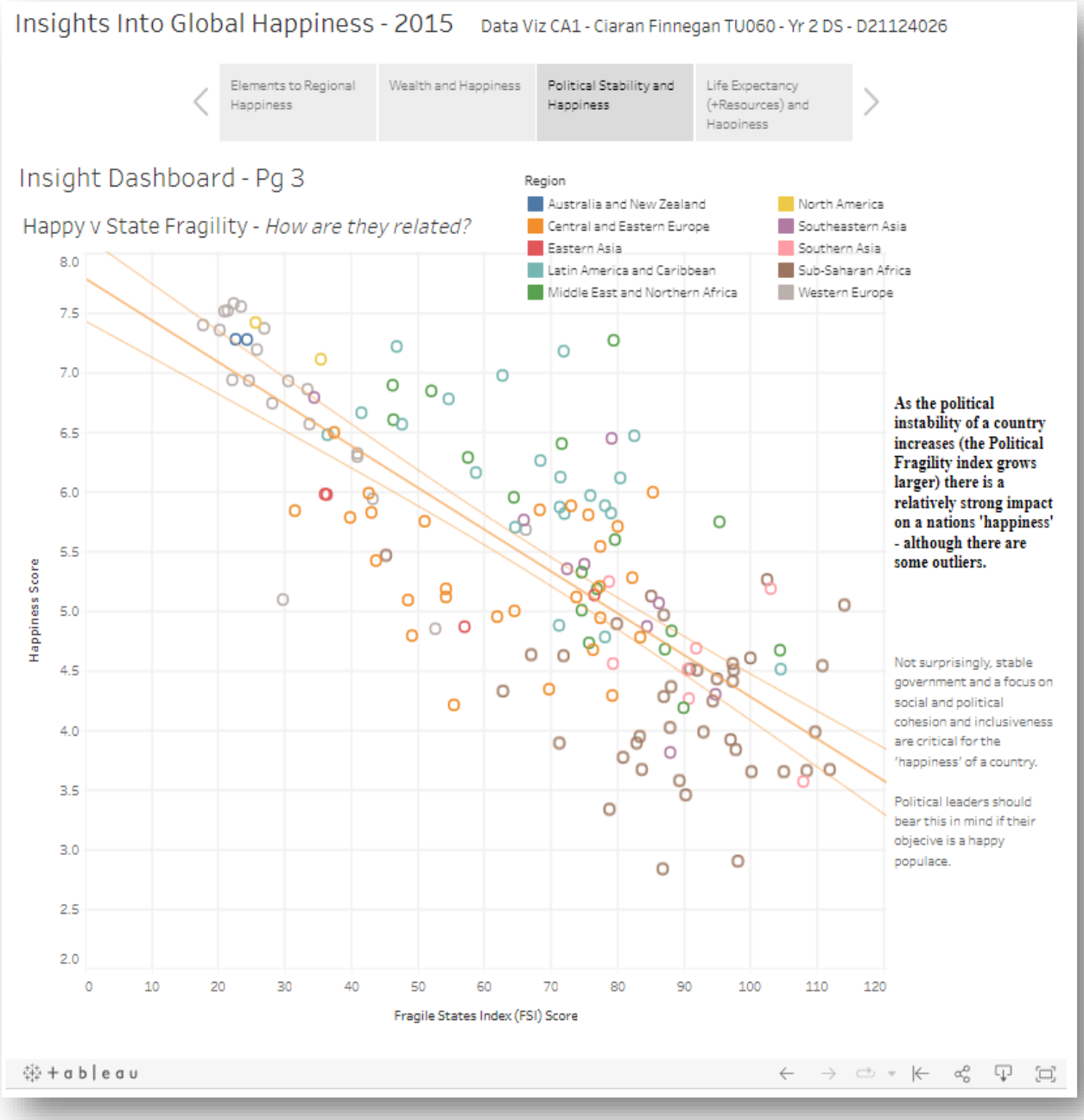
- The user can dynamically filter on geo-political regions across the graphs.

3.1.2 Insights Dashboard – Pg2 *Focus on Wealth and Happiness in 2015*



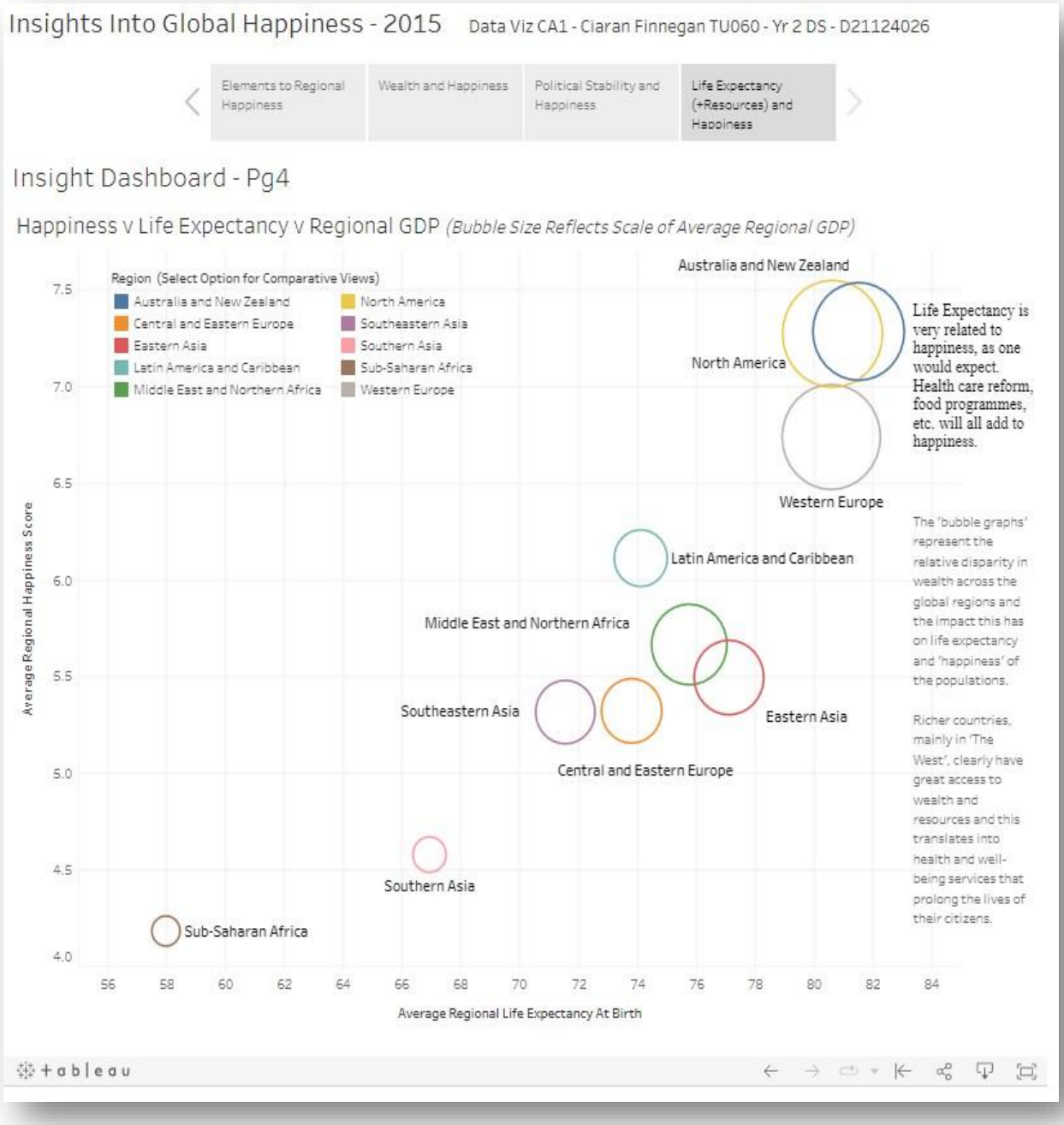
- The user can dynamically filter on geo-political regions to isolate nation groups in the scatter plot graph.

3.1.3 Insights Dashboard – Pg3 *Focus on Stability and Happiness in 2015*



- The user can dynamically filter on geo-political regions to isolate nation groups in the scatter plot graph.

3.1.4 Insights Dashboard – Pg4 *Life Expectancy/GDP and Happiness in 2015*



3.2 Tableau Public Dashboard - Location

The World Happiness Insights Dashboard for this assignment is accessed on the Tableau Public server at this location;

https://public.tableau.com/views/DizVizTu060Yr2CA1-Insights-StudentNoD21124026TableauPublic/InsightsIntoGlobalHappiness?:language=en-US&:display_count=n&:origin=viz_share_link

4 Conclusions

4.1 Interpretations from the World Happiness Visualisations

Although there are many more datasets that could be combined to the original Kaggle World Happiness dataset, it is reasonable to state the following inferences based on the dashboards in this assignment;

1. The gradual increase in happiness is consistent across all political regimes, with certain outliers/exceptions, as GDP increases.

2. Poorer nations benefit most from an increase in wealth generation in terms of increasing happiness. By filtering on regions, the viewer can see that the trend line of increasing happiness is more correlated to increase in wealth for countries in Sub-Saharan Africa or Latin America. (An increase in GDP does not always translate directly to increased income for citizens but it is a good indicator of a nation moving out of poverty).

3. Richer nations also stop feeling more 'happy' after a certain plateau of GDP is reached (65K). After that point wealth generation seems to become ineffective in making a nation happy - although this is only seen with the very small number of nations in this bracket.

4. The negative correlation of political instability against 'happiness' is generally as strong as the inverse positive relation of happiness/wealth. One interesting point (and topical in late 2022) is that although a country like Qatar currently receives a significant amount of negative international criticism about a lack of internal freedom, it provides stable government that translate well into a happiness score. (This is best viewed by filtering on the '*Middle East and Northern Africa*' region in the Insight Dashboard),

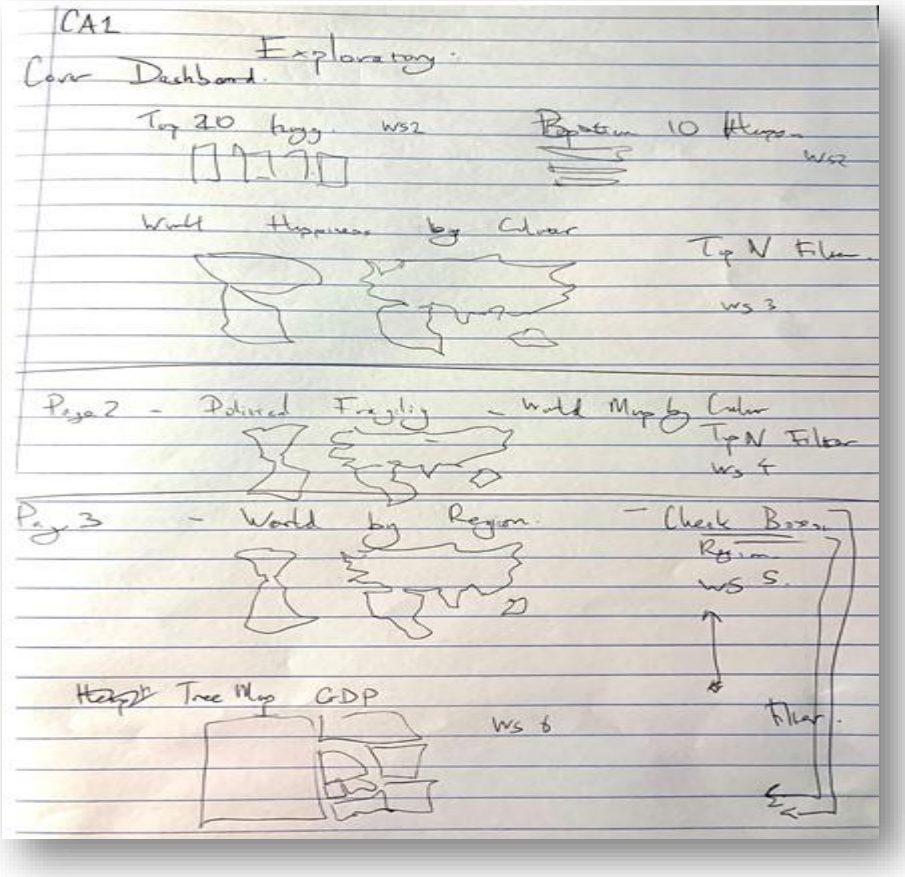
5. Life Expectancy is very related to happiness, as one would expect. Health care reform, food programmes, etc. will all add to longevity and, hence, happiness.

4.2 Iterations in the World Happiness Data Visualisations

The dashboards for this assignment were all worked out on paper in advance, so the actual time spent in building the Tableau dashboards did not require significant iteration except for;

- Format options to ensure data labels appeared on all GDP 'bubbles' on the Insight Dashboard.
- Resizing and moving legends to better fit the dashboard layout.
- Resizing World Maps on dashboards to increase display area.

The image below is an example of the paper design for the Exploratory Dashboard.



The details may be slightly obscured, but the original design proposed;

- Vertical bar charts for the 'Most Happy Nations'. This was replaced with horizontal bars because the display of country names was not truncated.
- A 'Top ***N***' filter option. This was replaced with a 'Bottom ***N***' slider, which gave a more interesting dynamic view of declining happiness across nation states.

5 References

1. Kaggle. (2021). *World Happiness Report*. Kaggle.com. Retrieved 2021, from https://www.kaggle.com/unsdsn/world-happiness The World Happiness Report is a landmark survey of the state of global happiness. The happiness scores and rankings use data from the Gallup World Poll. The scores are based on answers to the main life evaluation question asked in the poll.

2. Oxford Martin School. (n.d.). *GDP per capita*. Our World in Data. Retrieved October 27, 2022, from https://ourworldindata.org/grapher/gdp-per-capita-worldbank

3. The Fund For Peace. (n.d.). *Fragile States Index - The Fund For Peace*. Fragile States Index. Retrieved October 27, 2022, from https://fragilestatesindex.org/

4. OpenIntro. (2021). *CIA Factbook Details on Countries*. Openintro.org. Retrieved 2021, from https://www.openintro.org/data/index.php?data=cia_factbook Country-level statistics from the US Central Intelligence Agency (CIA).