

**WORLD HAPPINESS REPORT (Data Analysis and Web Development)**

**INTERNSHIP REPORT**

**Quarter IV (Year 1)**

***Submitted by***

**Nivedhidha I E0220003**

***In partial fulfilment for the award of the degree of***

**BACHELOR OF TECHNOLOGY**

**in**

**COMPUTER SCIENCE AND ENGINEERING**

**(Cyber Security & Internet of Things)**

**Sri Ramachandra Engineering and Technology,**

**Sri Ramachandra Institute of Higher Education and Research,**

**Porur, Chennai -600116**

**JULY 2021**



**BONAFIDE CERTIFICATE**

Certified that this project report **“WORLD HAPPINESS REPORT (Data Analysis and Web Development)”** is the bonafide work of **Nivedhidha I Reg No. E0220003** who carried out the internship work under my supervision.

**Signature of Faculty Mentor Signature of Vice-Principal**

|  |  |
| --- | --- |
| **Prof. Shiva Shankar G.**  Lecturer  Sri Ramachandra Engineering and Technology  Porur, Chennai-600116 | **Prof. M. Prema**  Vice-Principal  Sri Ramachandra Engineering and Technology  Porur, Chennai-600116 |

**Evaluation Date:**

**Table of Contents**

**Title** **Page**

1. Problem Statement 5
   1. Problem Statement 5
   2. Solution 5
2. Domain Introduction 5
   1. Data Analysis 5
   2. Data Visualization 6
   3. Web Development 6
3. Objective 6
4. Survey 7
   1. About World Happiness Report 7
   2. About The Dataset 7
5. Workflow 8
   1. Data Analysis 8
   2. Web Development 9
   3. Website Layout 10
6. Technology Used 11
   1. Data Analytics 11
   2. Web Development 11
7. Implementation 13
   1. Data Visualization - Tableau 13
   2. Data Analysis and Machine Learning - Python 20
   3. Website Development 20
8. Output 20
   1. Tableau 20
   2. Python (ML Model) 38
   3. Website 46
9. Sample Code 50
10. Timeline 50
11. Conclusion 51
12. Target Audience 51
13. Future Scope 51
    1. Questions Which Can Be Addressed Later 51
    2. Improvements To Be Made In Website 52
14. References 52

**ACKNOWLEDGEMENT**

I express my sincere gratitude to our Chancellor, Vice-Chancellor, our Provost **Dr. V. Raju**, and our Vice-Principal **Prof. M. Prema** for their support and for providing the required facilities for carrying out this study.

I wish to thank my faculty mentor, **Prof. Shiva Shankar G., Prof. Ashok Kumar P.,** and my internship mentor **Mr. Senthil Jayavel**, Department of Computer Science and Engineering, Sri Ramachandra Engineering and Technology for extending help and encouragement throughout the project. Without their continuous guidance and persistent help, this project would not have been a success for me.

I am grateful to the Department of Computer Science and Engineering, Sri Ramachandra Engineering and Technology, our beloved parents and friends for extending the support, who helped us to overcome obstacles in the study.

1. **PROBLEM STATEMENT**
   1. **Problem**

* A country’s happiness has a vital role in the country’s growth
* It is dependent on its citizens happiness
* Nowadays, happiness has been decreasing in many people’s lives due to many reasons, most recent one being covid and quarantine
  1. **Solution**
* Addressed the important questions and done an analysis to give an insight into how the happiness of a country can be improved
* Made a website to make the analysis available for a wider audience

1. **DOMAIN INTRODUCTION**

Over time, happiness has been reducing around us, due to various factors mainly being stress. We wanted to know about the world situation regarding this and suggest the major factors in which countries could focus on, to make them a happy one. Hence, we decided to work with this dataset.

1. **Data Analytics:**

Data analytics is the process of examining data sets to analyse, find trends and conclude the insights that it contains. Increasingly data analytics is used with the aid of specialized systems and software.

It is widely used in commercial industries to enable organizations to make more informed business decisions. It is also used by scientists and researchers to verify or disprove scientific models, theories, and hypotheses.

1. **Data Visualization:**

Data visualization is the graphical representation of information and data by using visual elements like charts, graphs and maps. Data visualization tools provide a more approachable way to see and understand trends, outliers and patterns found in the data.

In the world of Big Data, data visualization tools are essential to analyse massive amounts of information and make data-driven decisions.

## **Web Development:**

Web development refers to the building, creating, and maintaining of websites. It includes aspects such as web design, web hosting, web programming, and database management.

1. **OBJECTIVE**

* Learning all the basic concepts required for Data Analysis and Frontend Development
* Understanding the chosen Dataset by doing a survey about it
* Deciding what questions to address and create graphs and visualization for Analysis
* Create Dashboards and Storyboards for better understanding of the viewers
* Designing a UI design and making the website (Embedding all the visualizations into website)
* Learning AWS and hosting of Website in AWS

1. **SURVEY**

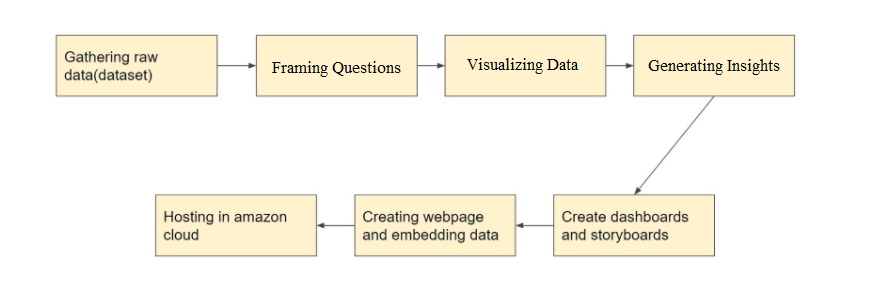
**About World Happiness Report**

* The Index reflects a new worldwide demand for more attention to happiness as a criterion for government policy.
* It is also an eye-opener for those who think that the pursuit of a healthy GDP number alone is sufficient to mark the happiness of a nation.
* For the making of the World Happiness Report, nationally representative samples of respondents are asked to rate their own current lives on a 0 to 10 scale.
* The data used to rank countries in each report are drawn from the Gallup World Poll, as well as other sources such as the World Values Survey, in some of the reports.
* Experts in fields including economics, psychology, survey analysis, and national statistics, describe how measurements of well-being can be used effectively to assess the progress of nations.

**About The Dataset**

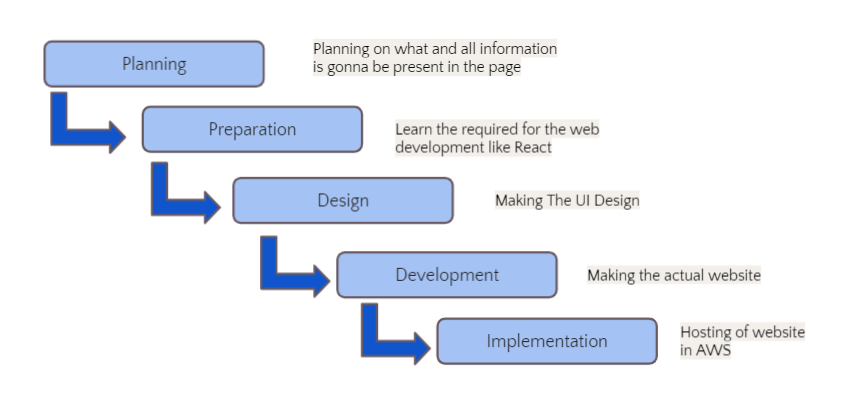
* The dataset provides the happiness rank for nearly 150 countries for the 5 years.
* The rank is calculated based on various parameters like GDP, Life Expectancy and the rank is given by people of the sample for different questions asked.
* The questions were mainly based on a yes or no question given in Gallup World Poll of which a national average was taken.
* The factors were:
  + Social Support
  + Economy (GDP per capita)
  + Freedom
  + Trust Towards Government
  + Health (Life Expectancy)
  + Generosity

1. **WORKFLOW**
   1. **Data Analysis**

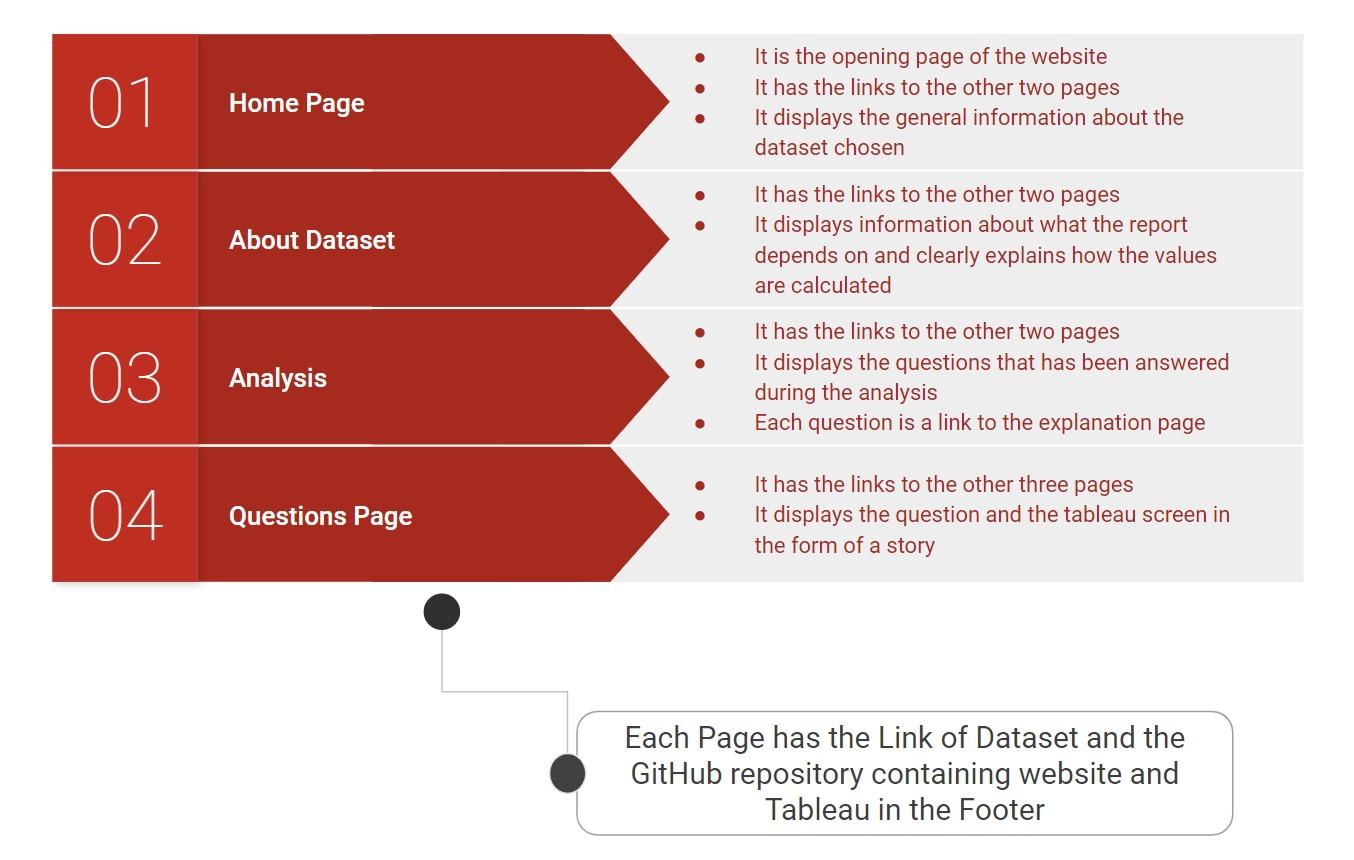


1. Gathering Raw Data - First we should find a proper dataset which is dependable and should do the required cleansing to make it ready for analysing.
2. Framing Questions - Next we should frame questions from the dataset which we want to answer through our analysis.
3. Visualizing Data - Next we should visualize the dataset by creating various graphs.
4. Generating Insights - Next we should generate insights from the created graphs and find out the answers for questions formed before.
5. Creating Dashboards and Storyboards - Next we should create dashboards or storyboards as per requirement for creating a storyline to explain how we got our insight to others.
6. Creating Website and Embedding Data - Next we should create a website using an IDE to make the analysis available for everyone.
7. Hosting in AWS - Finally we should host the website in AWS to make the website available for anyone interested in the analysis.

* 1. **Web Development**



* + 1. Planning - First we should plan on what and all information is going to be present on the website.
    2. Preparation - Then we should start preparing by learning the required IDE like Angular and React JS.
    3. Design - Next we should make the UI design of the website.
    4. Development - Then we should make the website using HTML, CSS, Angular and React.
    5. Implementation - In the end we should host the website so that it becomes available for anyone interested in the movie.
  1. **Website Layout**



* There are totally 4 types of web pages present in the website.
* Each page has the links of GitHub repository and Kaggle dataset.
* Home page contains general information about the world happiness report.
* The About Dataset page has the details about each factor of the dataset.
* Analysis page has the questions which have links to each question of the analysis page.
* There are totally 8 Question pages containing a tableau screen in the form of a story.

1. **TECHNOLOGY USED**

## **Data Analytics**

1. Tableau

Tableau is a visual analytics platform transforming the way we use data to solve problems - empowering people and organizations to make the most of their data. It is one of the fastest evolving data visualization tools. It is very fast to deploy, easy to learn and very intuitive to use for a customer.

### MySQL

MySQL can cost-effectively help you deliver high performance, scalable database applications. It is a relational DBMS based on SQL. It is open-source and free software under the GNU license. It is supported by Oracle Company.

### JUPYTER NOTEBOOK

Jupyter is a free, open-source, interactive web tool known as a computational notebook, which researchers can use to combine software code, computational output, explanatory text and multimedia resources in a single document. It can be used for all sorts of data science tasks including data cleaning and transformation, numerical simulation, exploratory data analysis, data visualization, statistical modelling, machine learning, deep learning, and much more.

## 

## **Web Development**

### ADOBE XD

Adobe XD is a powerful and easy-to-use vector-based experience design platform that gives teams the tools they need to craft the world's best experiences collaboratively.

### VISUAL STUDIO CODE

Visual Studio Code is a source code editor that can be used with a variety of programming languages, including Java, JavaScript, Go, Node.js, Python and C++. It is based on the Electron framework, which is used to develop NodeJS Web applications that run on the Blink layout engine.

### REACT JS

### ReactJS is an open-source JavaScript library that is used for building user interfaces specifically for single-page applications. It's used for handling the view layer for web and mobile apps. React also allows us to create reusable UI components.

### GITHUB

GitHub is a code hosting platform for version control and collaboration. It allows people from multiple locations to work together on projects from anywhere.

#### 

#### AWS (AMAZON WEB SERVICES)

Amazon Web Services (AWS) is a secure cloud services platform, offering computing power, database storage, content delivery and other functionality to help businesses scale and grow. Deliver static and dynamic files quickly around the world using a Content Delivery Network (CDN).

1. **IMPLEMENTATION**

## **Data Visualization - Tableau**

1. What is the overall situation in the world regarding happiness?

Graph 1 - (A map graph which shows the regions and the happiness scores of all the countries) We analysed those darker shaded countries were mostly from the regions North America & ANZ, and Western Europe which implied that they had higher happiness scores in comparison to other regions.

Graph 2 - (A tree map which shows the average score of each factor) Happiness score was mostly affected by Social Support and Economy (GDP per capita), and least affected by Generosity and Trust.

Graph 3 - (A scatter plot which compares the top 2 factors with happiness scores with respect to regions) Economy and GDP per capita equally contributed to the Happiness Score.

Conclusion - In the world, the happiest regions were North America & ANZ, and Western Europe and the highest impacting factors were Social Support and Economy.

1. Which countries are better positioned in each of these aspects?

Graph 1 - (A bar graph of the Top 10 countries with respect to GDP) Qatar has the highest score of 1.7 followed by Luxembourg and Singapore. We can notice that almost all the countries have high happiness scores too.

Graph 2 - (A bar graph of the Top 10 countries with respect to Freedom) Uzbekistan has the highest score but has a very low happiness score. At the same time, other countries of the same graph have high happiness scores too.

Graph 3 - (A bar graph of the Top 10 countries with respect to Generosity) Myanmar has the highest score with respect to Generosity. The top ten countries have both low and high scores with respect to happiness scores.

Graph 4 - (A bar graph of the Top 10 countries with respect to Life Expectancy) Singapore and Hong Kong have the highest scores amongst the other countries. This implies that these countries have very good health infrastructure and higher life expectancy. Almost all the countries have a high happiness score.

Graph 5 - (A bar graph of the Top 10 countries with respect to Social Support) Iceland has the highest score of social support which implies that people in Iceland have enhanced quality of life in adverse life situations due to the support from their family and friends. Almost all the countries have a high happiness score.

Graph 6 - (A bar graph of the Top 10 countries with respect to Government Corruption) Rwanda and Singapore have the highest score. Important point to note is that, though these countries have the highest position, their score is less than 0.5 which is very less when compared to all the other factors. Also the happiness score is variable for the top ten countries.

Conclusion - Happiness score correlates with the three main factors, namely, Economy, Health and Social Support.

1. Finding the reason why happiness rank changed over the years in different regions?

Graph 1 - (A table containing countries which has highly varying happiness ranks and their ranks for each year) We could split the countries into two categories, Improving and Deteriorating based on the way it varies.

Graph 2 - (A bubble chart for both the variations with respect to all the factors) Social Support and Economy had made a major impact in both the categories.

Graph 3 - (A dual-axis line graph to compare both categories based on Social Support) Both the categories are improving in their scores but Improving is higher than the Deteriorating category.

Graph 4 - (A dual-axis line graph to compare both categories based on Economy) Both the categories are stable in their scores but Improving is higher than the Deteriorating category.

Graph 5 - (A dual-axis line graph to compare both categories based on Generosity) Both the categories are deteriorating in their scores, but Deteriorating is higher than the Improving category.

Conclusion - Generosity has not had any proper impact in the change of happiness rank and Social Support has a higher impact when compared to Economy regarding the variations.

1. In what ways, Generosity and Health play a role in impacting the world happiness score?

Graph 1 - (A histogram for Life Expectancy and the Number of Countries for each score) Most of the countries have their Life Expectancy between 0.50 to 0.90. Life Expectancy is high in most of the countries.

Graph 2 - (A histogram for Generosity and the Number of Countries for each score) Most of the countries have their Generosity score between 0.05 to 0.30. Generosity score, on an average, is low in most of the countries as per people's opinion.

Graph 3 - (A box plot to compare both the factors side by side with respect to different regions) The score given by people for Generosity is low when compared to Health. North America & ANZ is at the top in both the factors, whereas Sub-Saharan Africa and Central & East Europe are the bottom most regions in terms of Life Expectancy and Generosity.

Graph 4 - (A linear regression trend line over a scatter plot that shows comparison between Generosity and Life Expectancy) Most of the scatter points lie near the trendline and the trendline has a low inclination angle.

Conclusion - There is a weak positive linear correlation between Generosity and Life expectancy.

1. Does a country's GDP affect the happiness rate?

Graph 1 - (A linear regression trend line over a scatter plot that shows a comparison between Happiness score and GDP) The trendline shows a large inclination and points are closer to the line. There is a strong positive correlation between GDP and happiness score.

Graph 2 - (A tree map with respect to the Happiness Score in regards with each region) North America & ANZ and Western Europe have the highest score. On the other hand, South Asia and Sub-Saharan Africa have the least score in both factors. So then we created a group on the above basis.

Graph 3 - (A scatter plot splits all the countries into 3 categories namely top, bottom and middle with respect to region) Countries in the top group show a strong positive correlation. The economy has played a major role towards the happiness score of these countries. The countries in the bottom group show a weak positive correlation. A low economy is not the only reason for the low happiness scores of these countries.

Graph 4 - (A map graph which shows how the economy varies in top and bottom countries) The top countries don’t show much of a change in their colour code. The bottom countries show high variations in their colour code. The top countries are more stable than the bottom countries.

Conclusion - Even though GDP is one of the main factors which affects the happiness score, the level of impact is different for each region.

1. What is the trend followed by India’s happiness score in the past 5 years?

Graph 1 - (2-line graphs compared - one showing change in happiness rank and the other showing change in happiness score of India) Happiness score of India has been decreasing continuously. Happiness score would give a much clearer picture regarding the changes that have happened in India.

Graph 2 - (A bar chart having scores of each factor of India displayed) Economy and Social Support were the 2 main factors contributing to India’s happiness score.

Graph 3 - (A bar graph shows the change in Trust score year wise) Trust has continuously increased from 2015 to 2018 and had a sudden downfall in the year 2019. In the year 2019, one of the main reasons for the decrease in the happiness score would be the decrease in trust towards the government. Studies say that the Central Government was accused of destroying democratic institutions and processes after the election was conducted.

Graph 4 - (A bar graph shows the change in Freedom score year wise) Freedom has continuously increased from 2015 to 2018 and had a sudden downfall in the year 2019. In the year 2019, one of the main reasons for the decrease in happiness score would also be the People's opinion on the Freedom they have.

Graph 5 - (A bar graph shows the change in Social Support score year wise) Social Support score has decreased in the year 2016 and has had an increase for the next 3 years. In the year 2016, one of the main reasons for the decrease in happiness score would be the Social Support score.

Graph 6 - (A bar graph shows the change in Generosity score year wise) Generosity score has continuously decreased till the year 2018 and had an increase in the year 2019. In the first 4 years, one of the main reasons for the decrease in happiness score would be the Generosity score.

Graph 7 - (A bar graph shows the change in Life Expectancy year wise) Health has continuously improved from 2015 to 2017 and had a sudden downfall in the year 2018 and again raised in the year 2019. In the year 2016, one of the main reasons for the decrease in happiness score would be due to Health. The lower level of Health can be attributed to the low level of public investments in preventive health facilities such as sanitation and waste management, as well as in medical care facilities such as primary health centres and health professionals.

Graph 8 - (A bar graph shows the change in GDP year wise) GDP has continuously increased from 2015 to 2017 and had a sudden downfall in the year 2018 and again raised in the year 2019. In the year 2018, one of the main reasons for the decrease in happiness score would be the GDP value. The lower GDP growth figures are attributed to weak domestic consumption so, the Indian economy reported a continuous decline in GDP growth rate over the year 2018.

Conclusion - Each year a different set of factors has affected the happiness score of India.

1. Which country had more trust in the government with respect to freedom?

Graph 1 - (A histogram for Trust score and the Count of Countries given the score) Most of the countries have given their score between 0.02 to 0.14. Trust score, on an average, is low in most of the countries as per people's opinion.

Graph 2 - (A histogram for Freedom score and the Count of Countries given the score) Most of the countries have given their score between 0.30 to 0.60. Freedom score, on an average, is high in most of the countries as per people's opinion.

Graph 3 - (A dual axis between line and bar graph for Freedom and Trust score) The factors didn’t have any correlation with each other when we see year wise. Also, the sense of Freedom is higher when compared to the trust towards the government. Another point to note is Trust towards the government has negligible change, whereas Freedom is varying highly.

Graph 4 - (Dual axis graph having scatter plots of both the factors region wise) Region of North America & ANZ has the highest score in both the factors, hence has the highest happiness score.

Graph 5 & Conclusion - (Two scatter plots having Trust and Freedom of top ten countries based on Freedom) Denmark has the highest trust towards the government with respect to Freedom.

1. An analysis to find which are the factors causing fewer happiness scores in a few regions and prediction on how it would be if each factor is improved.

Graph 1 - (A bar graph showing the happiness scores of the bottom most countries) Burundi has the lowest happiness score.

Graph 2 - (A pie chart showing the intensity of how each factor affects happiness of the bottom countries) The factors having lowest score are Generosity and Trust.

Conclusion - These countries are supposed to focus more on their Freedom, Trust and Generosity scores to go to better positions on the world scale. Even though the Economy, Health and Social Support are high compared to other factors, it still requires concentration as it is low when compared to other countries.

## **Data Analysis and Machine Learning - Python**

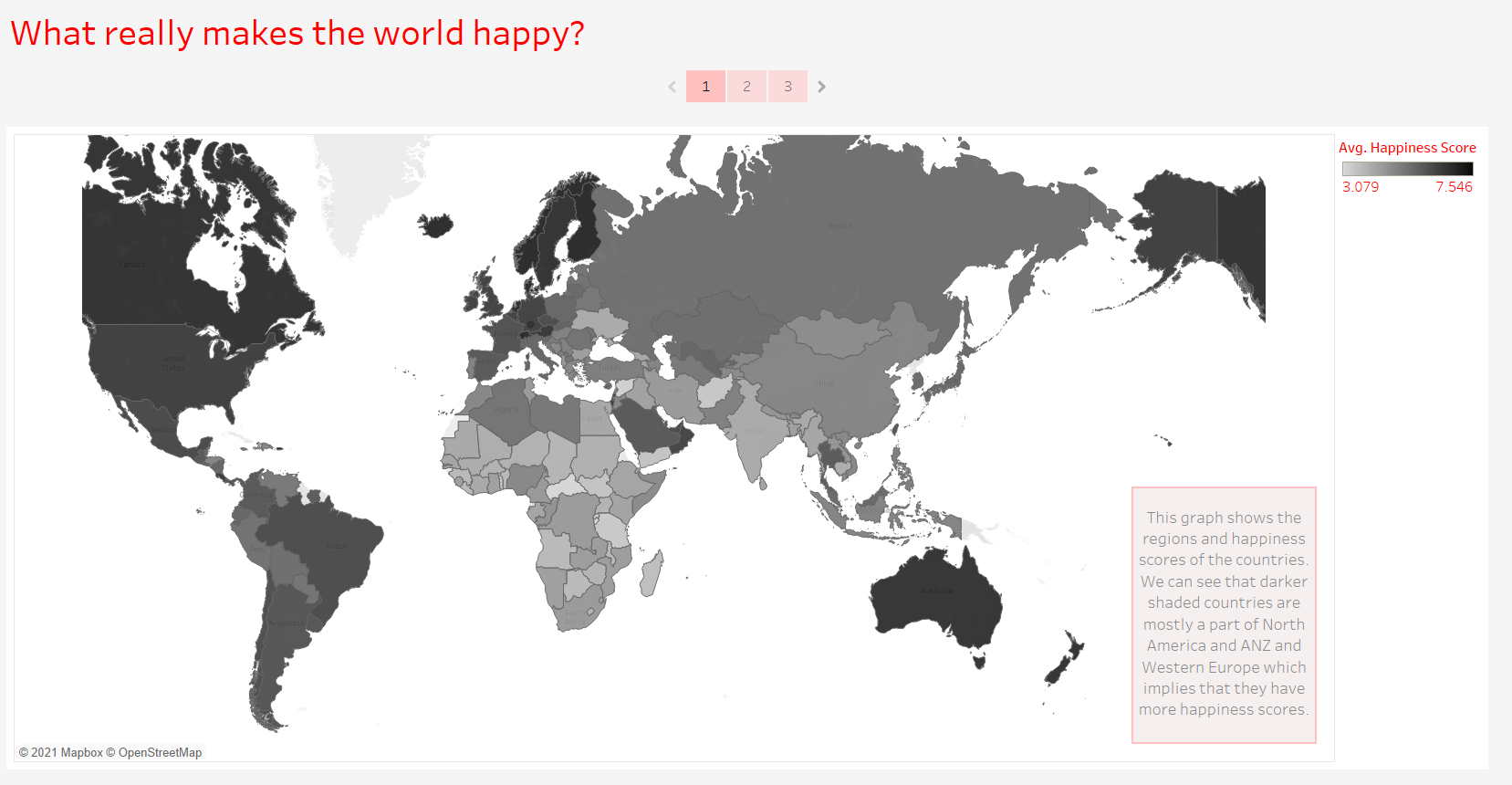
* An imaginary country named Custom Land was created by getting input from users for all the factors and predicting its happiness score.
* To provide a proper visualization, a bar chart is used in which the happiness score of the imaginary country is inserted into its position when kept in descending order and highlighted using a different colour after finding its rank.

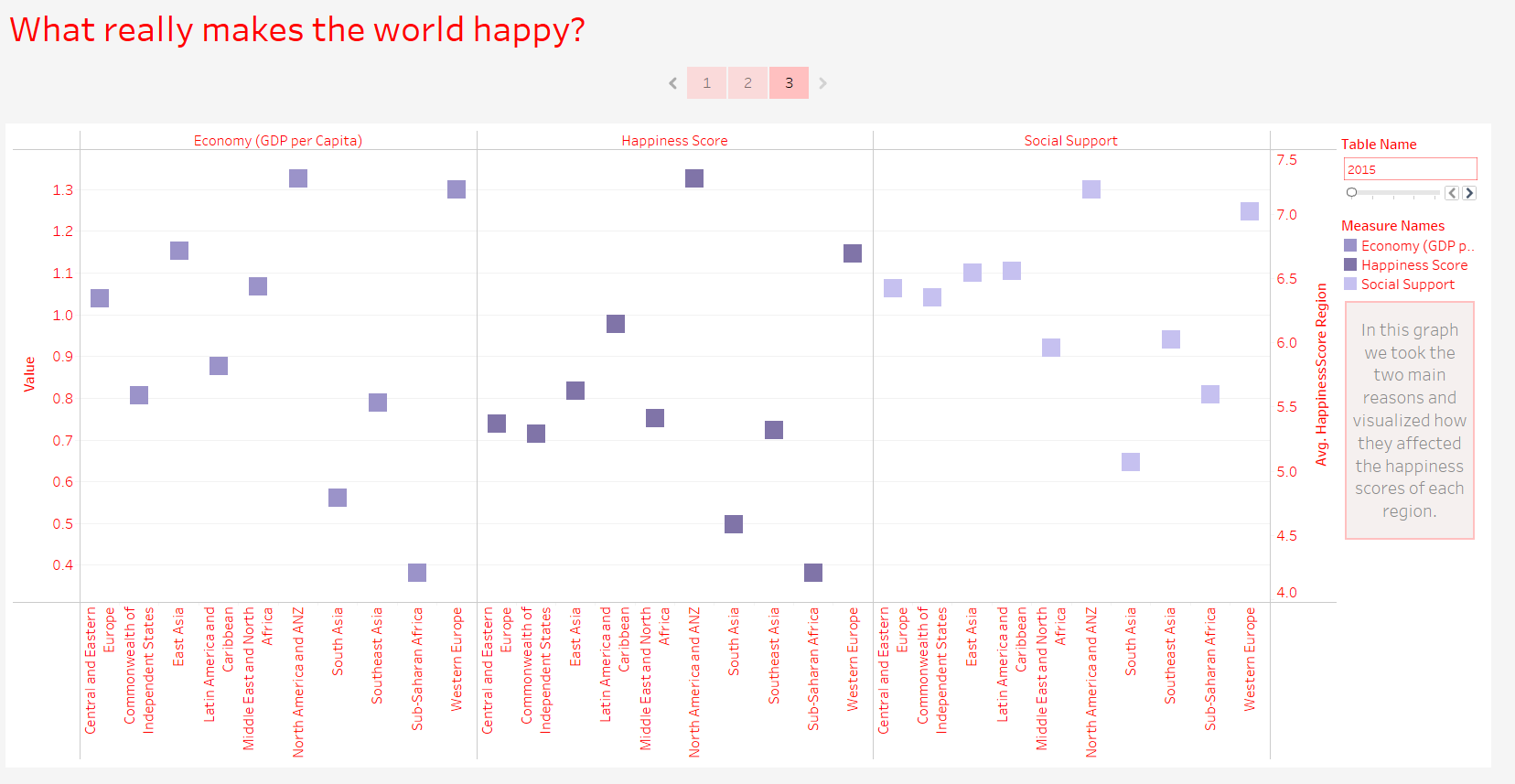
## **Website Development**

* A UI design of the website was made using ADOBE XD
* A static website was created, using ReactJS, in which a description of WHR and its main factors were given along with a page dedicated for showing the questions and all the Storyboards of the visualizations we made are embedded in it

1. **OUTPUT**
   1. **Tableau**

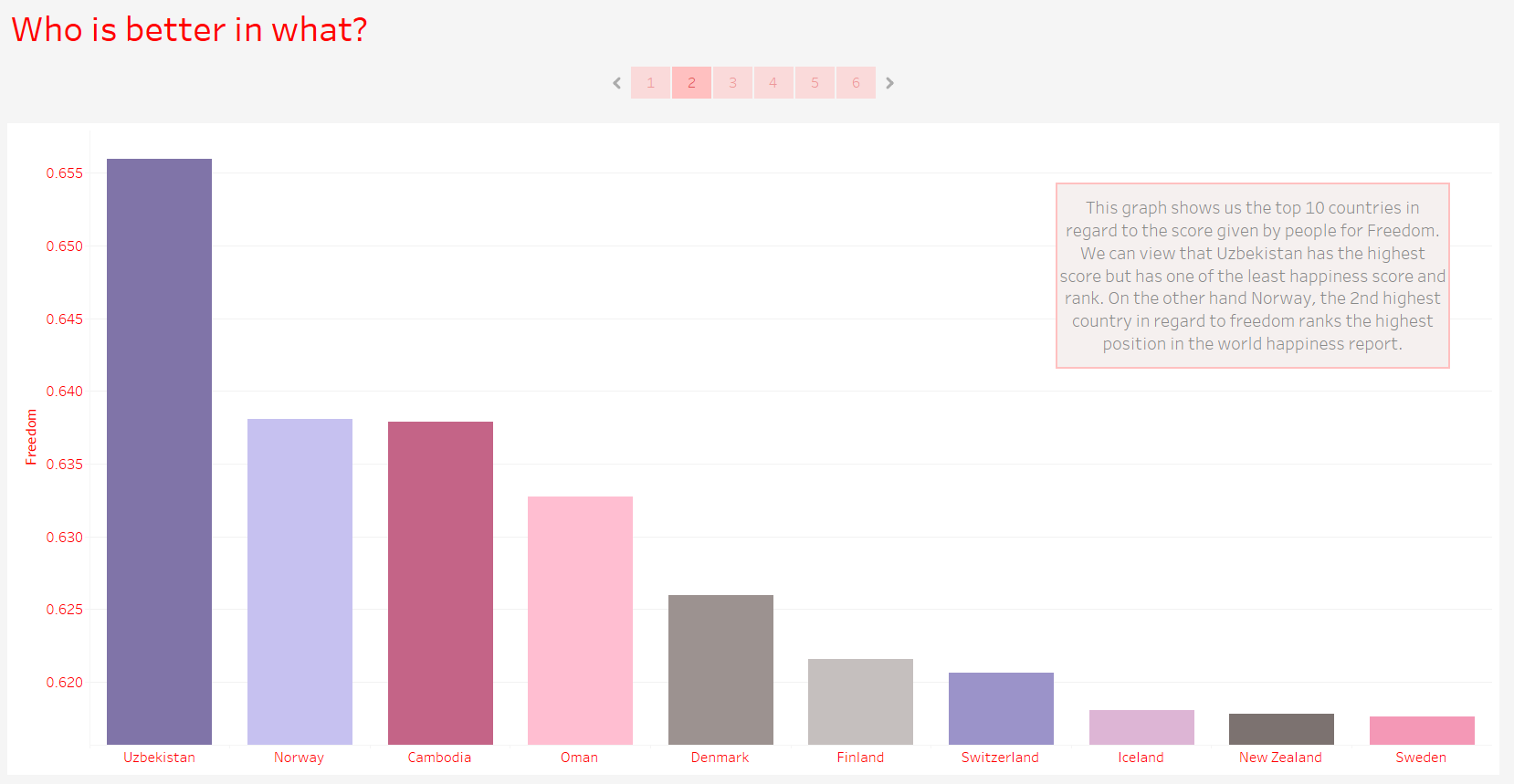
Question 0:

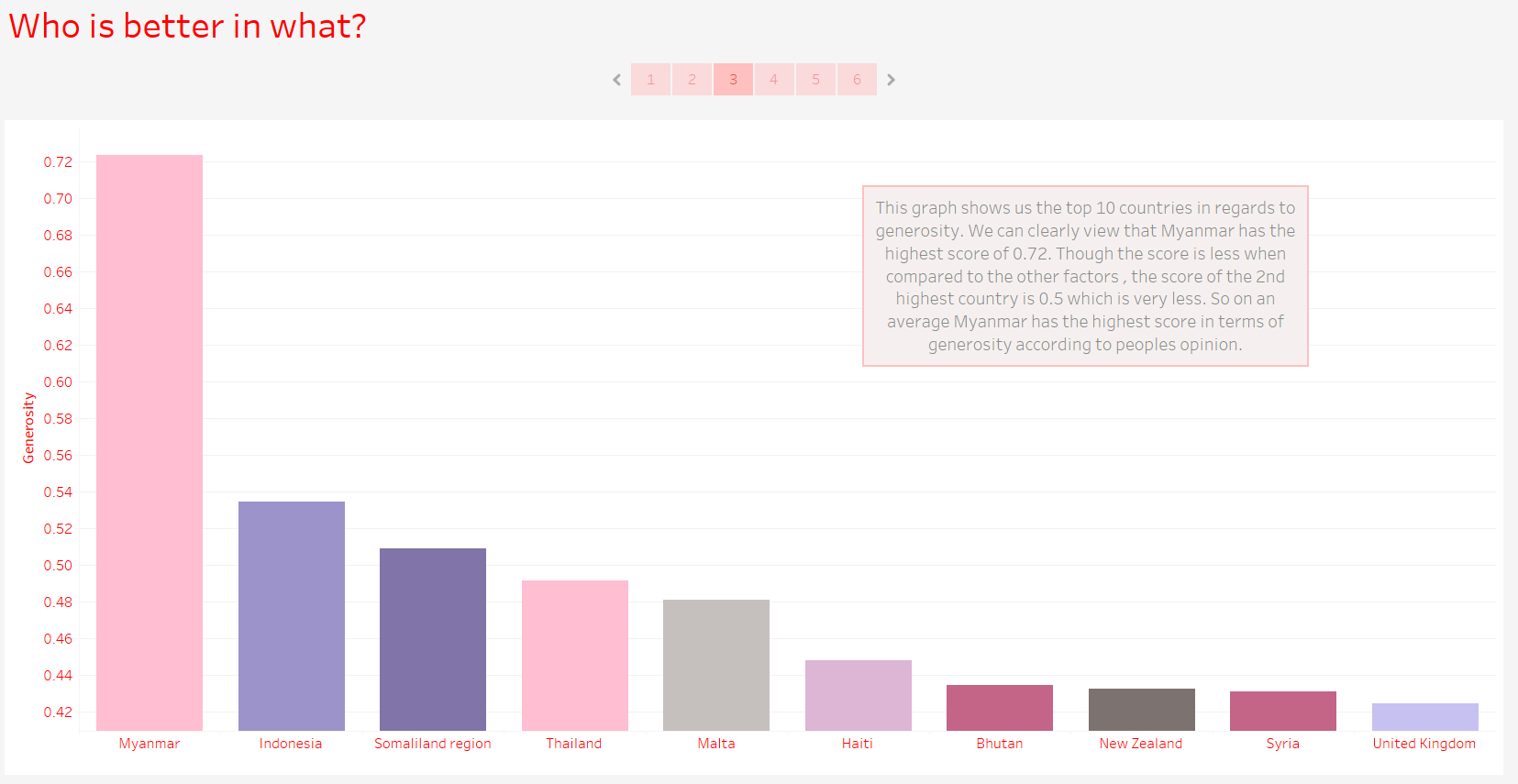
****

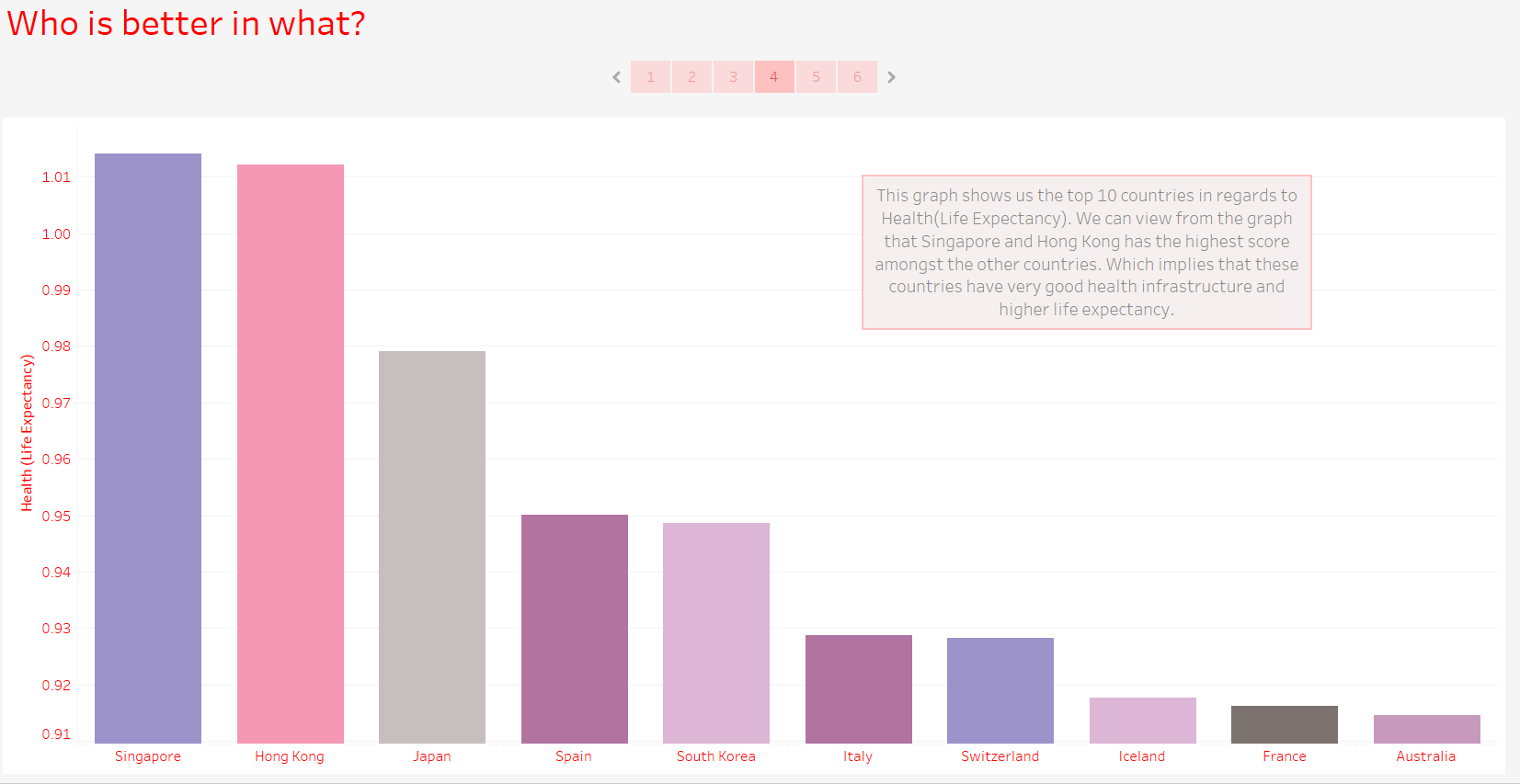
****

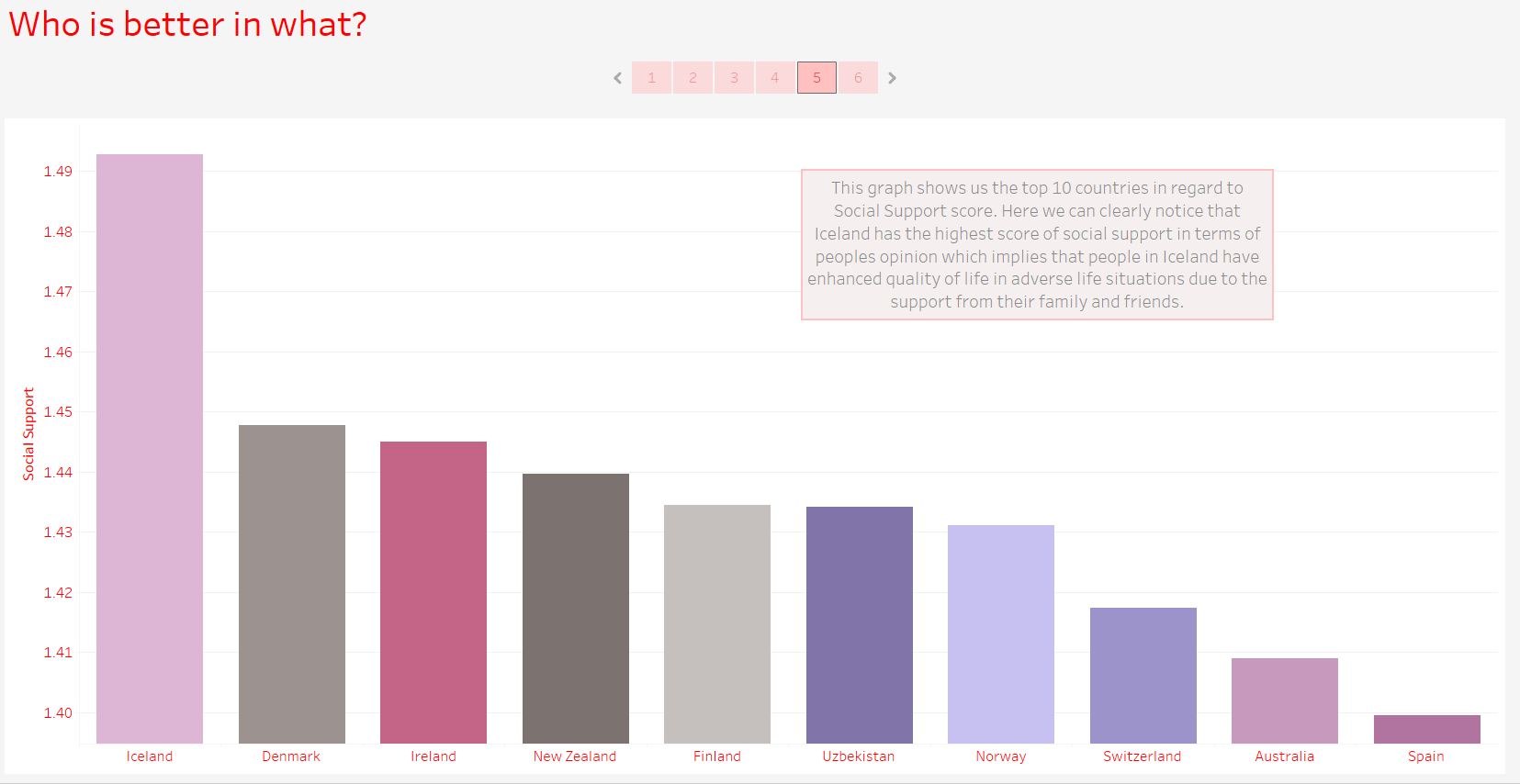
Question 1:

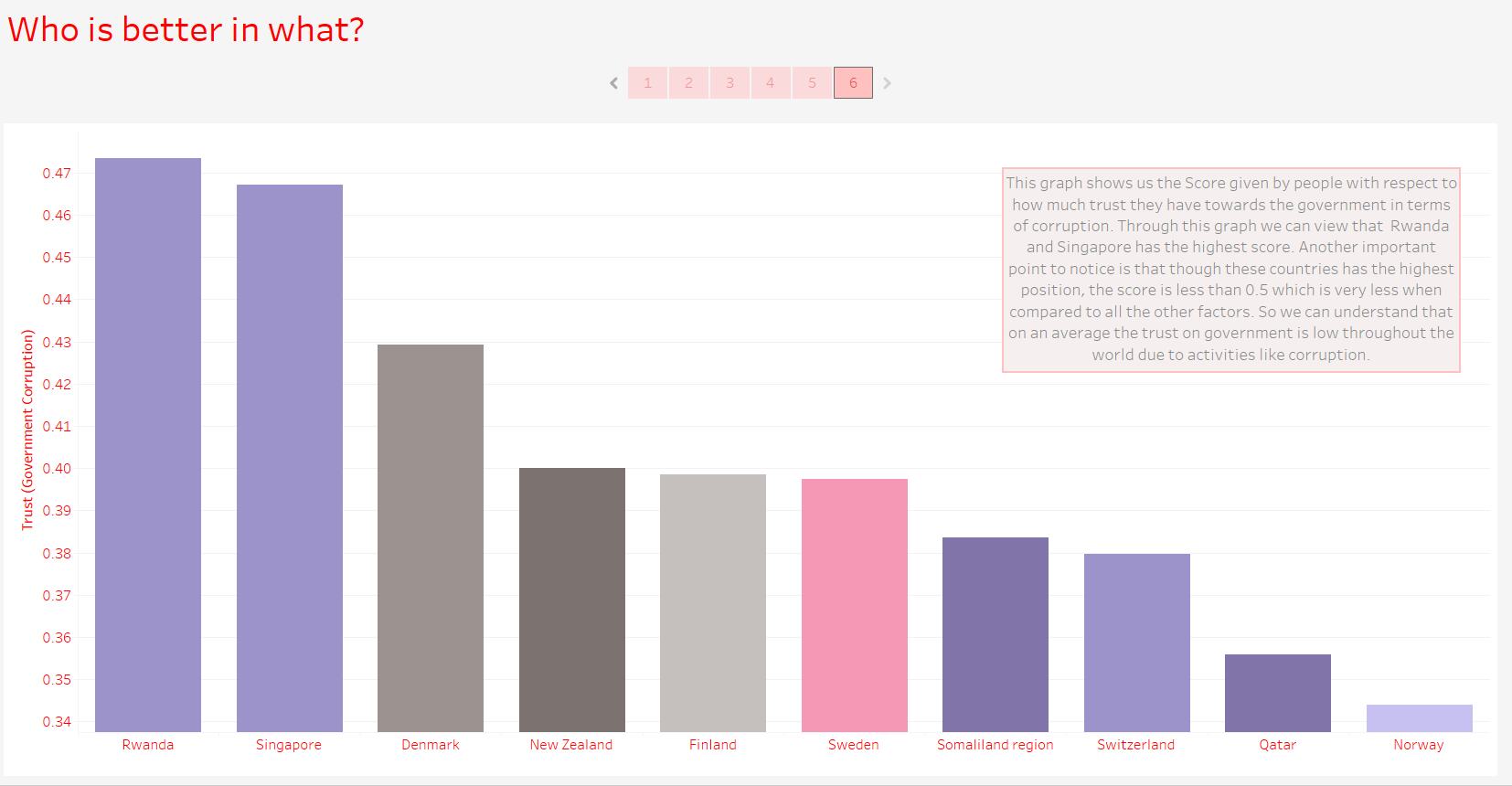




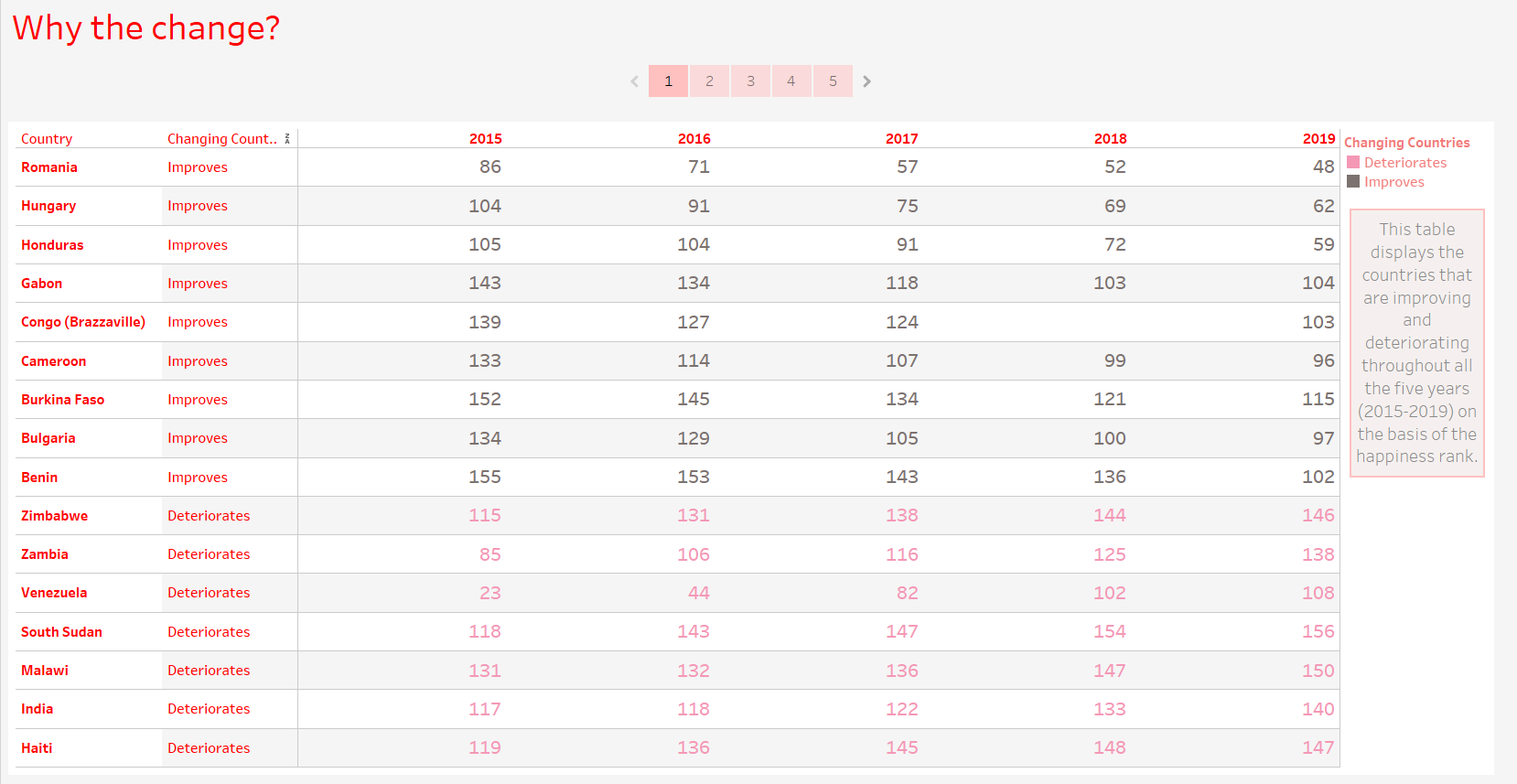


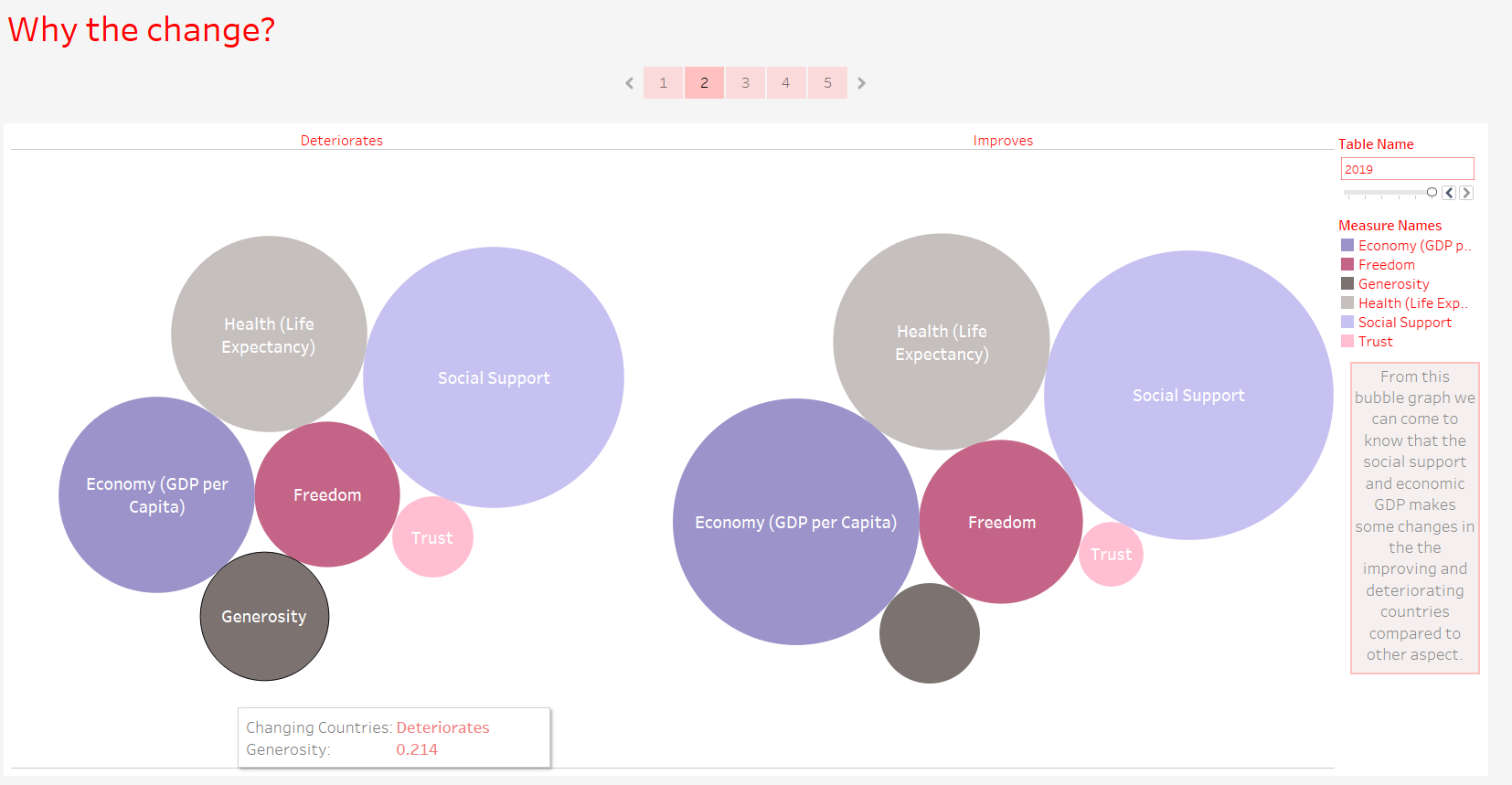


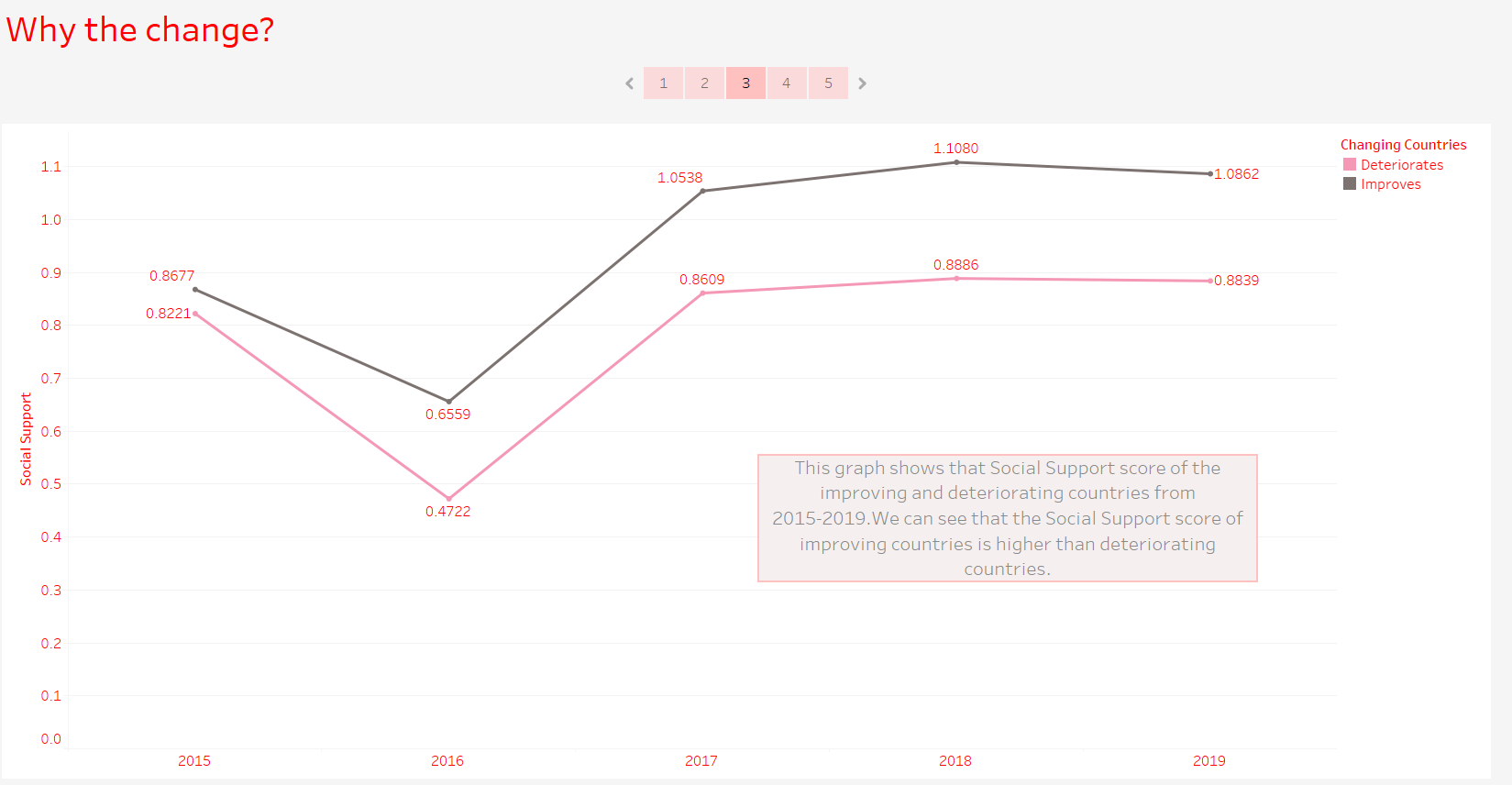


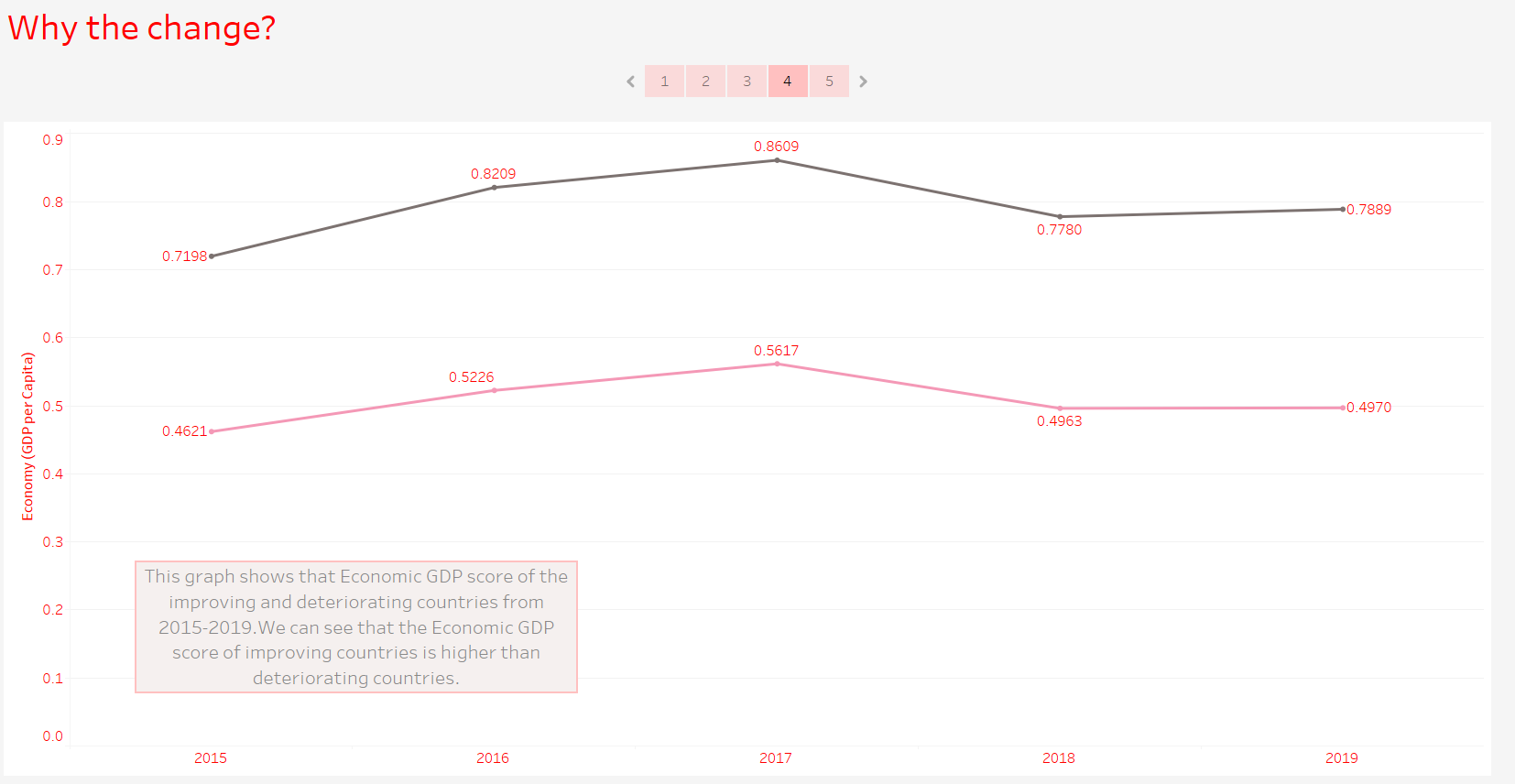


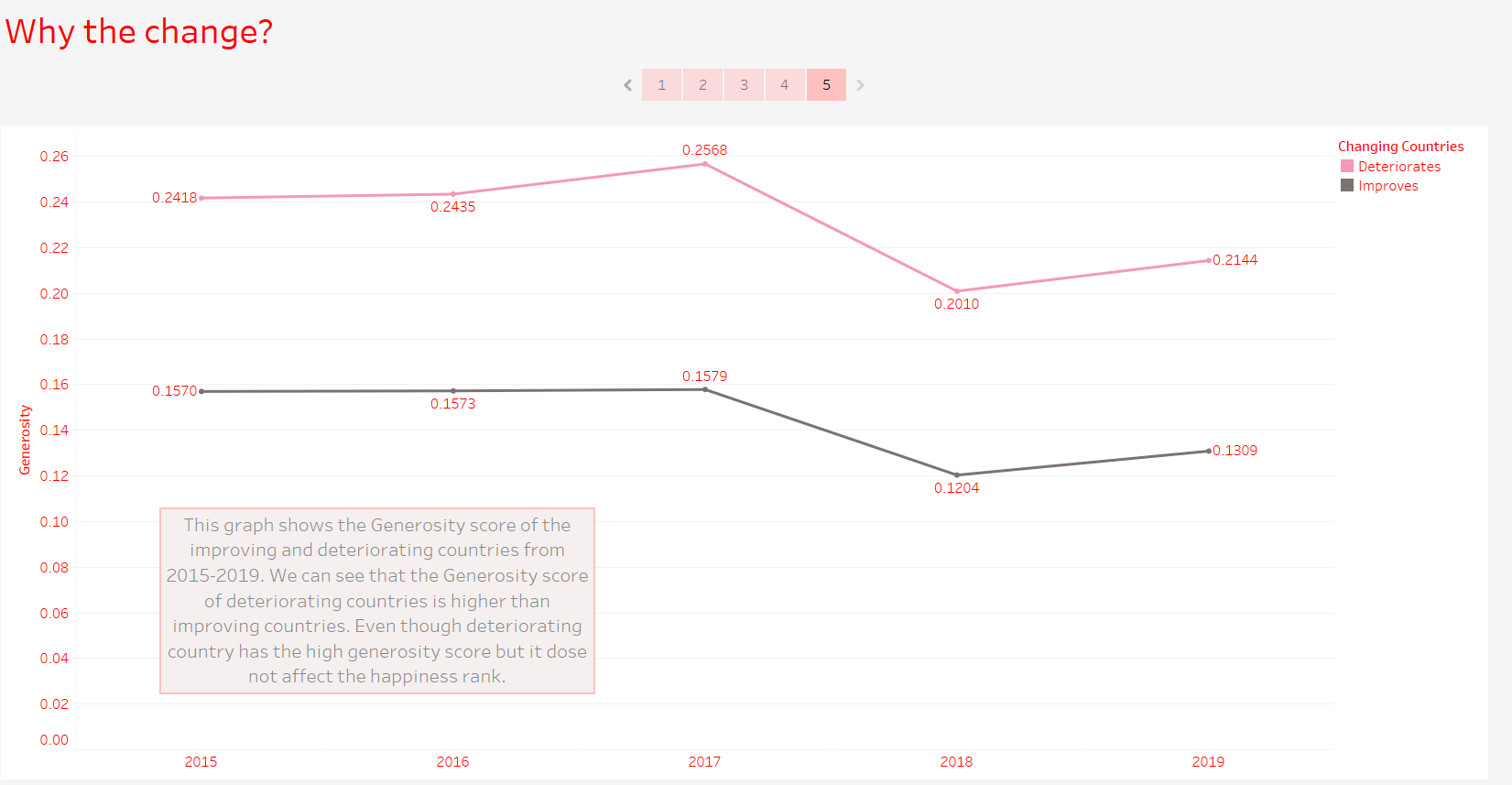
Question 2:



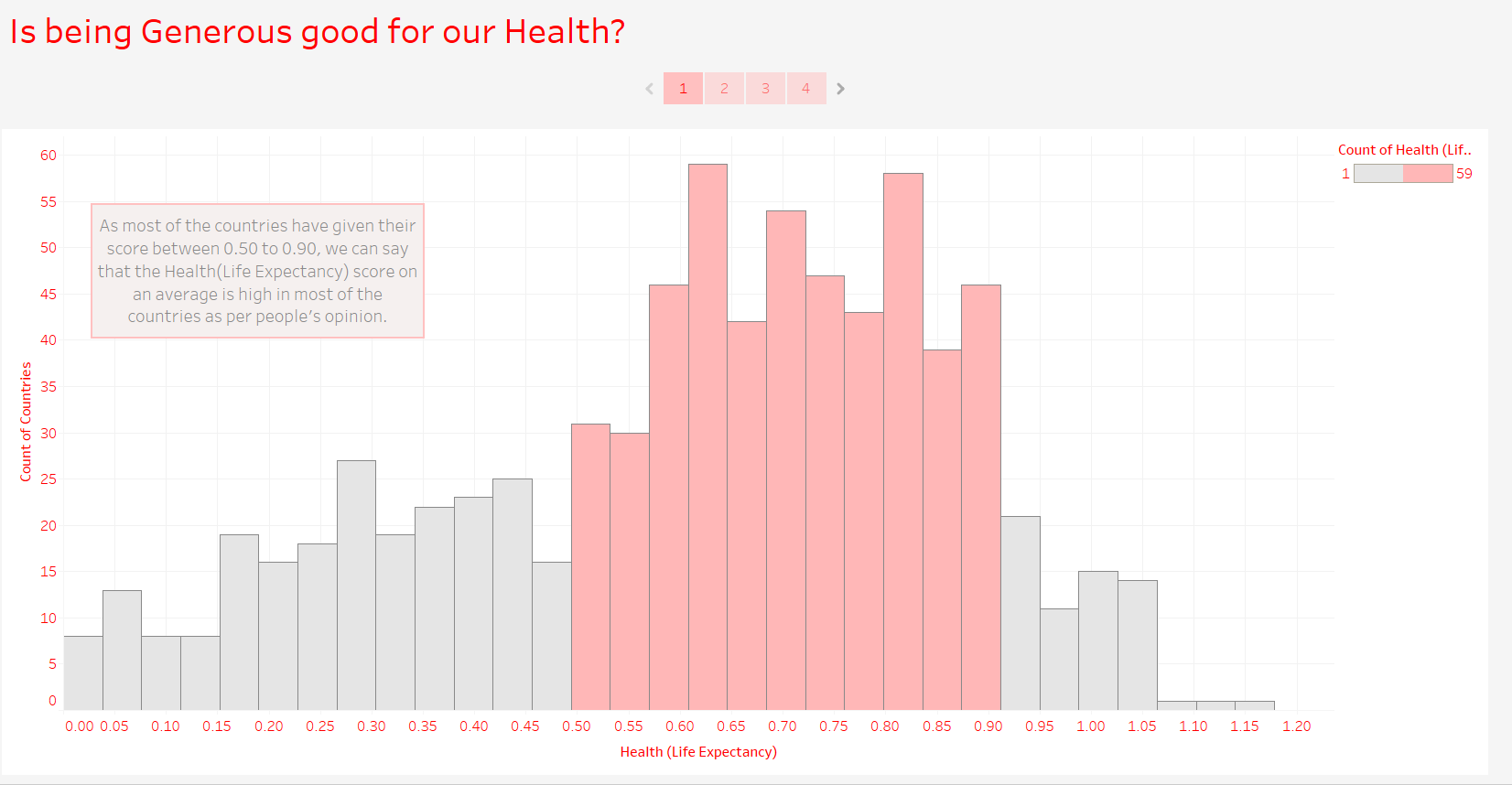


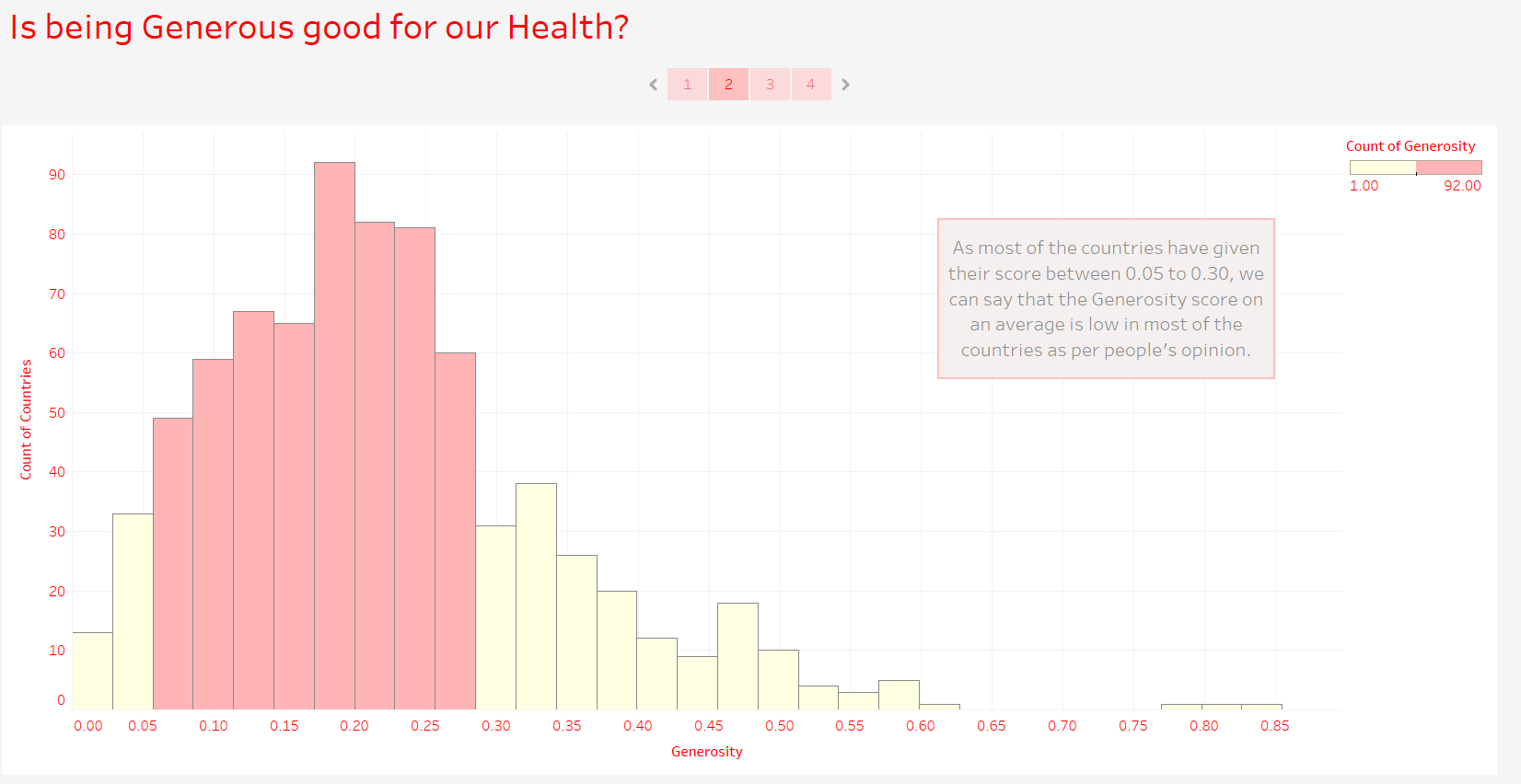


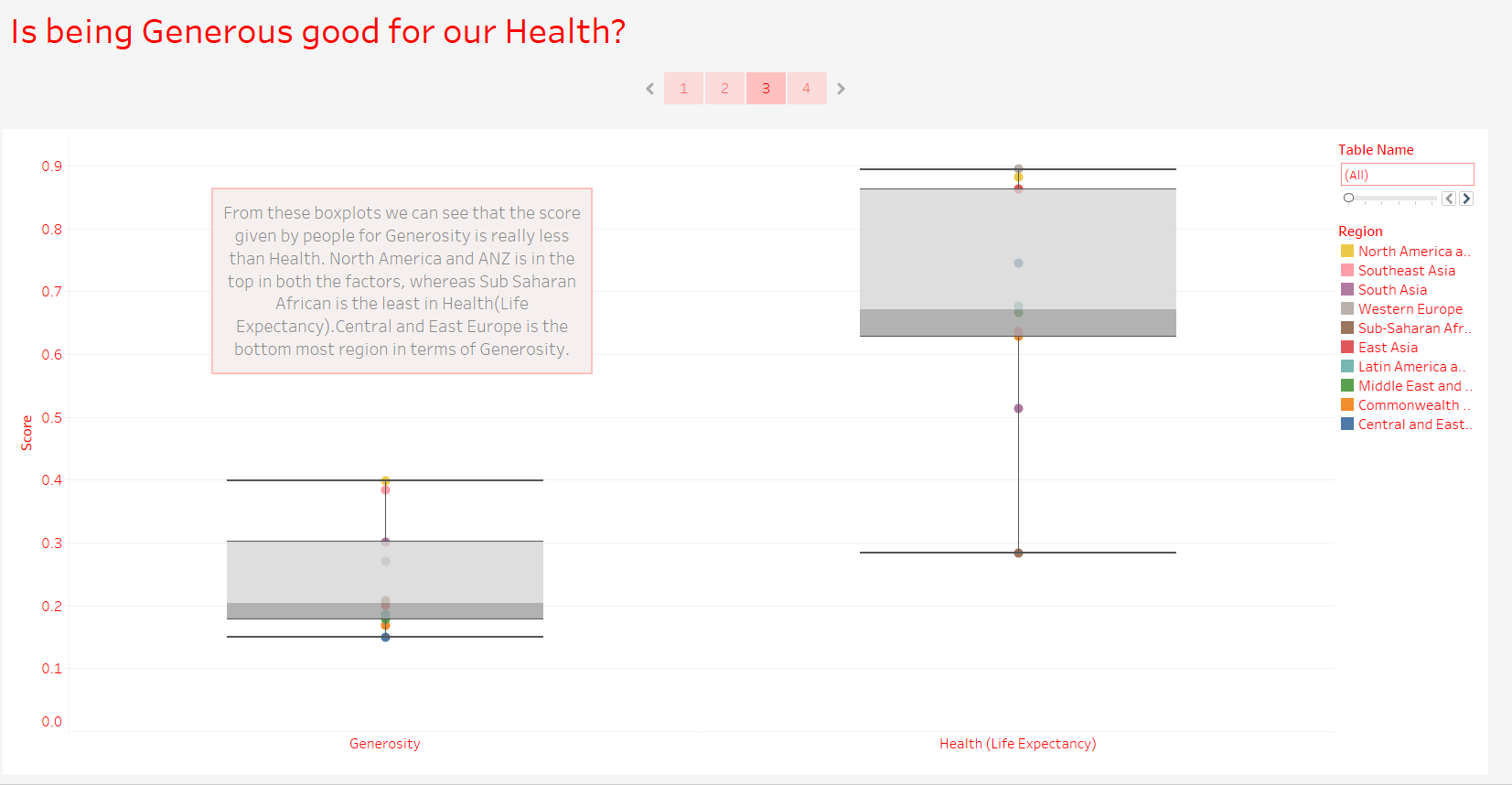


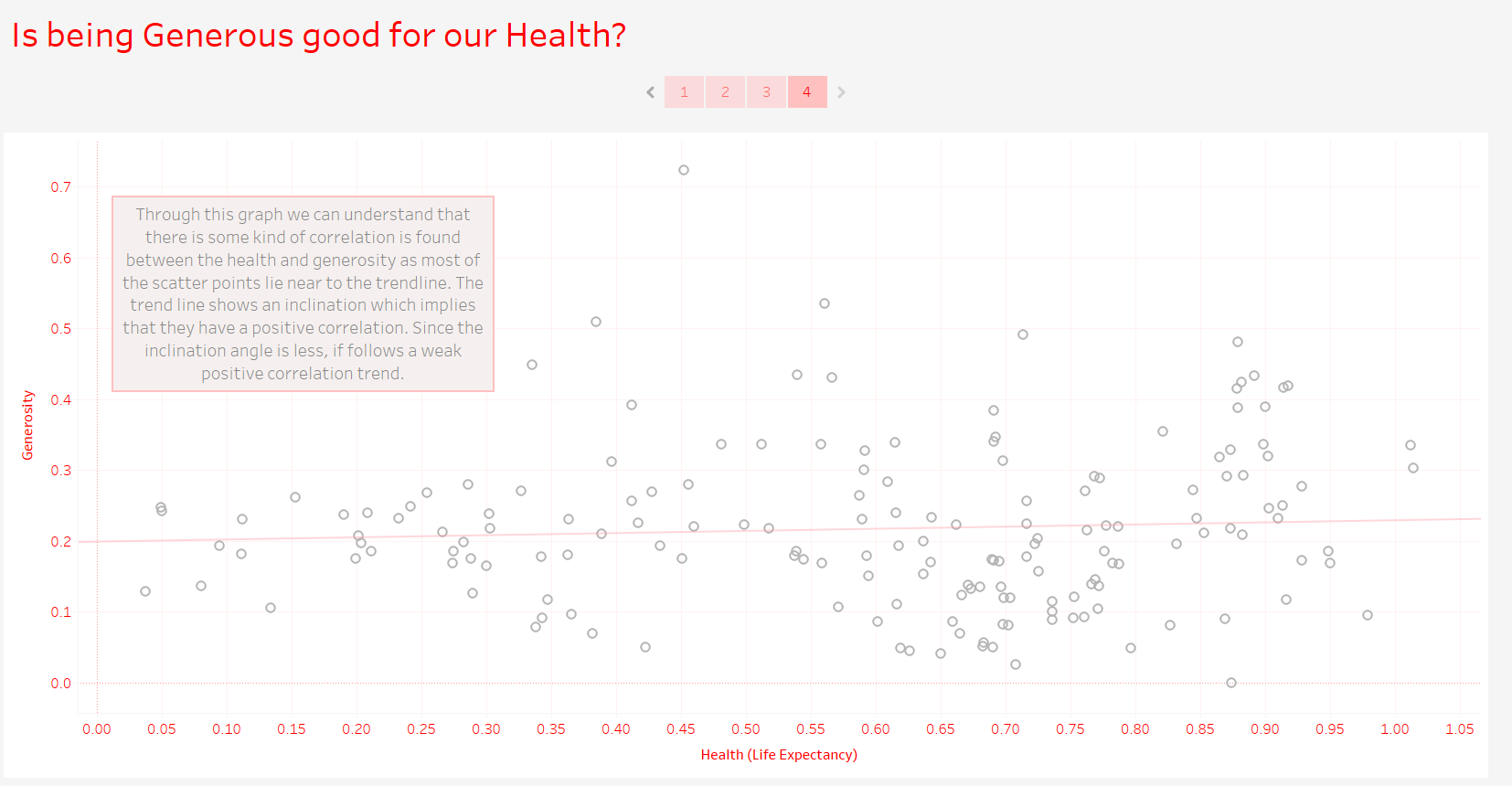


Question 3:



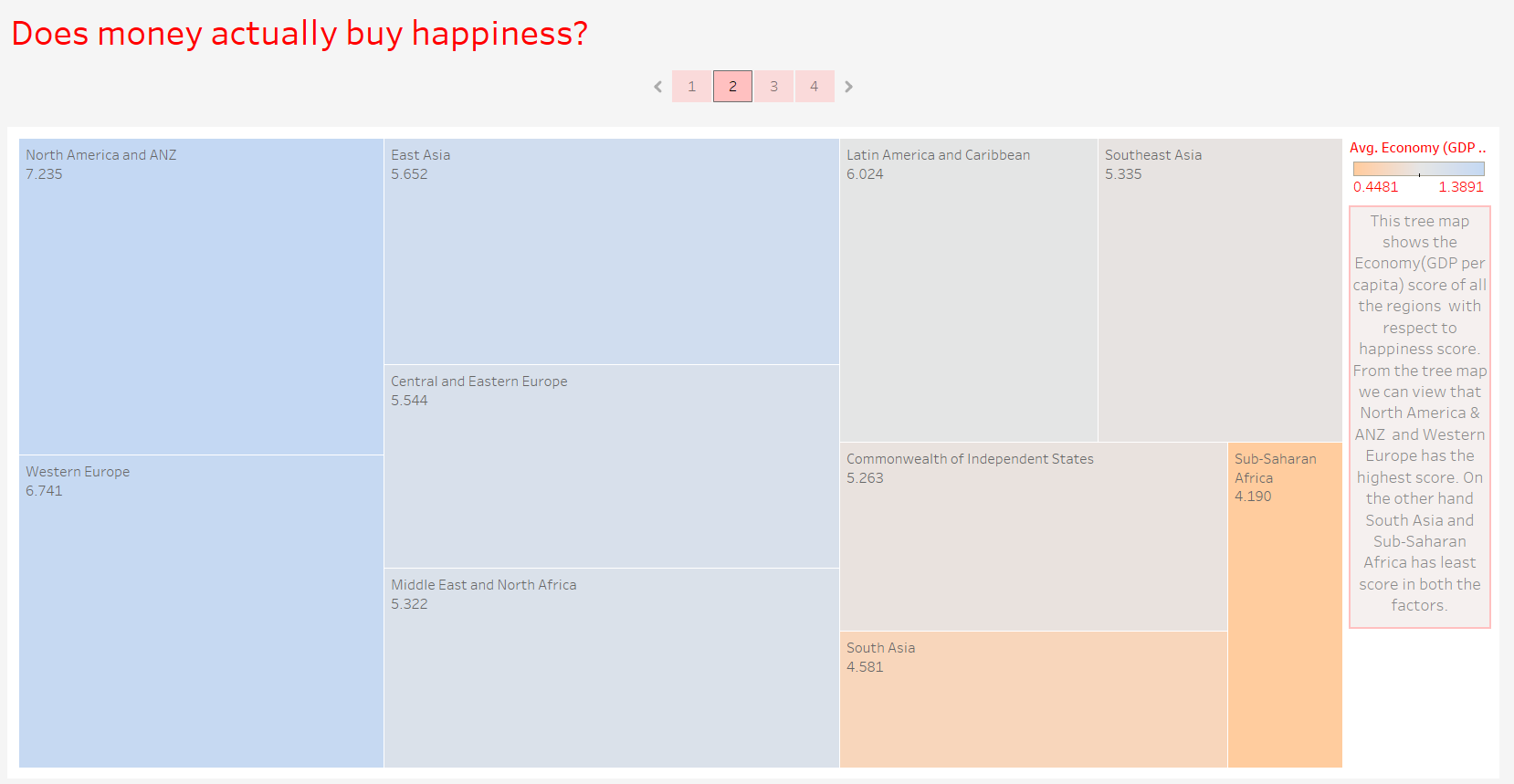


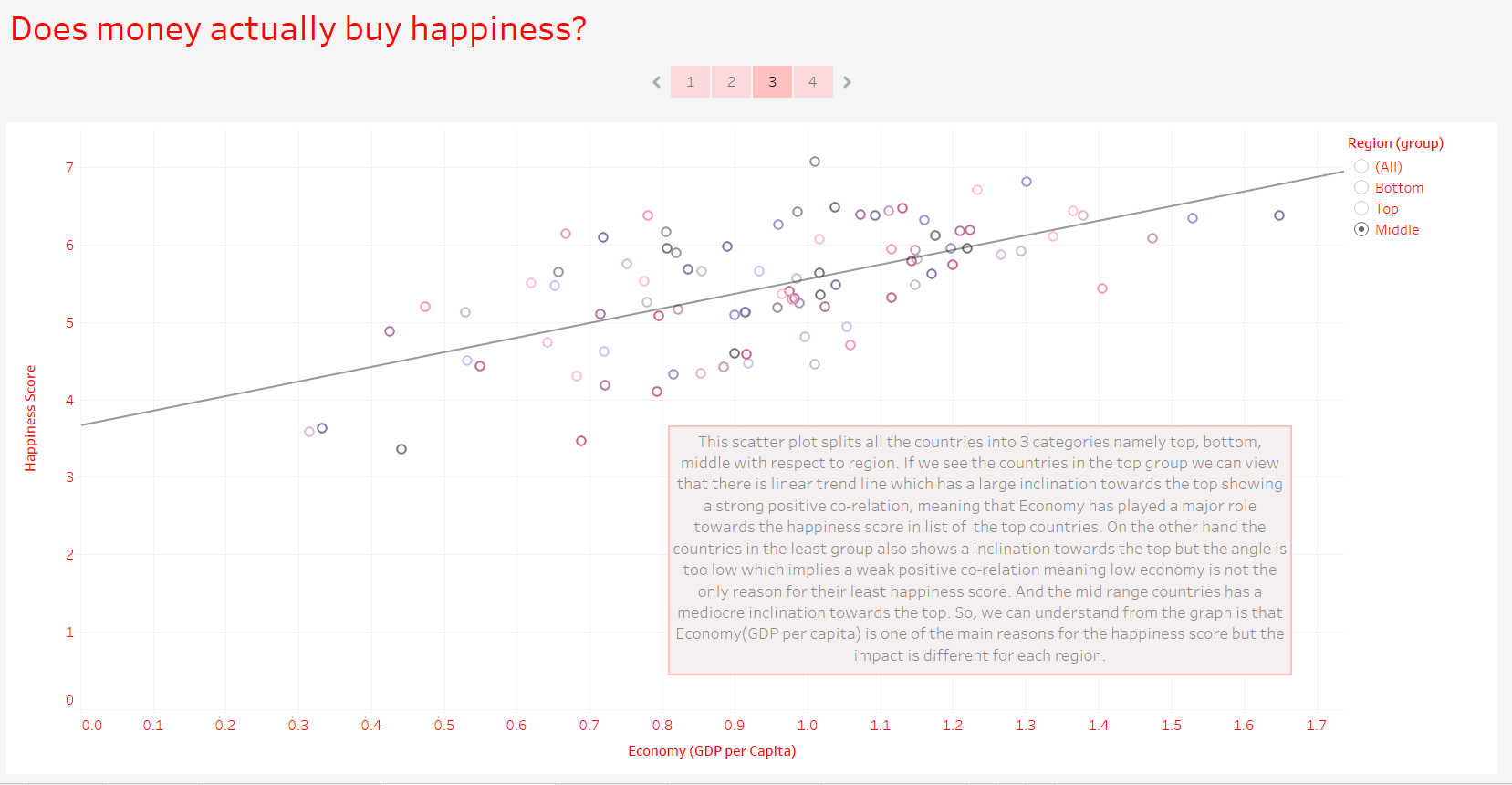


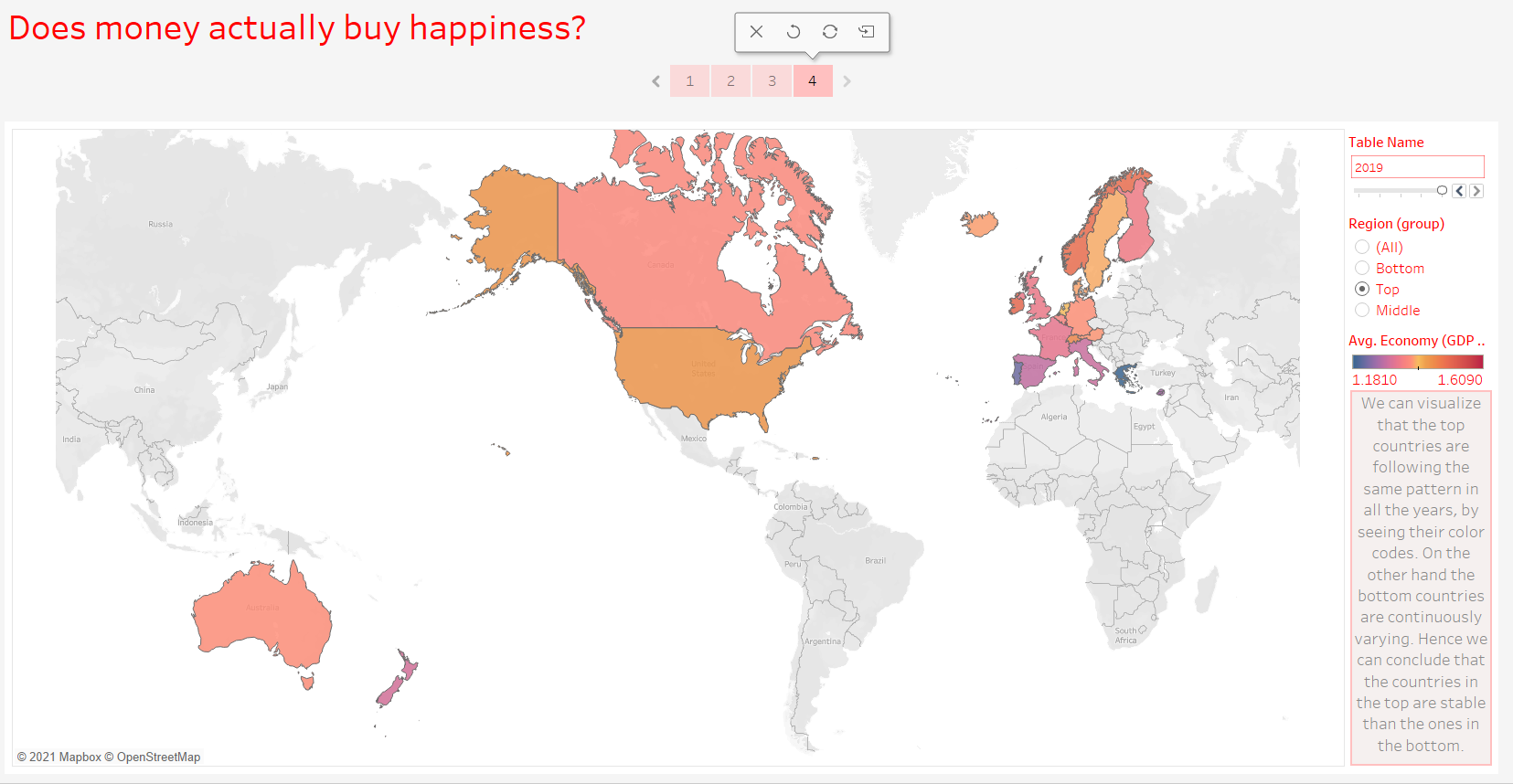


Question 4:







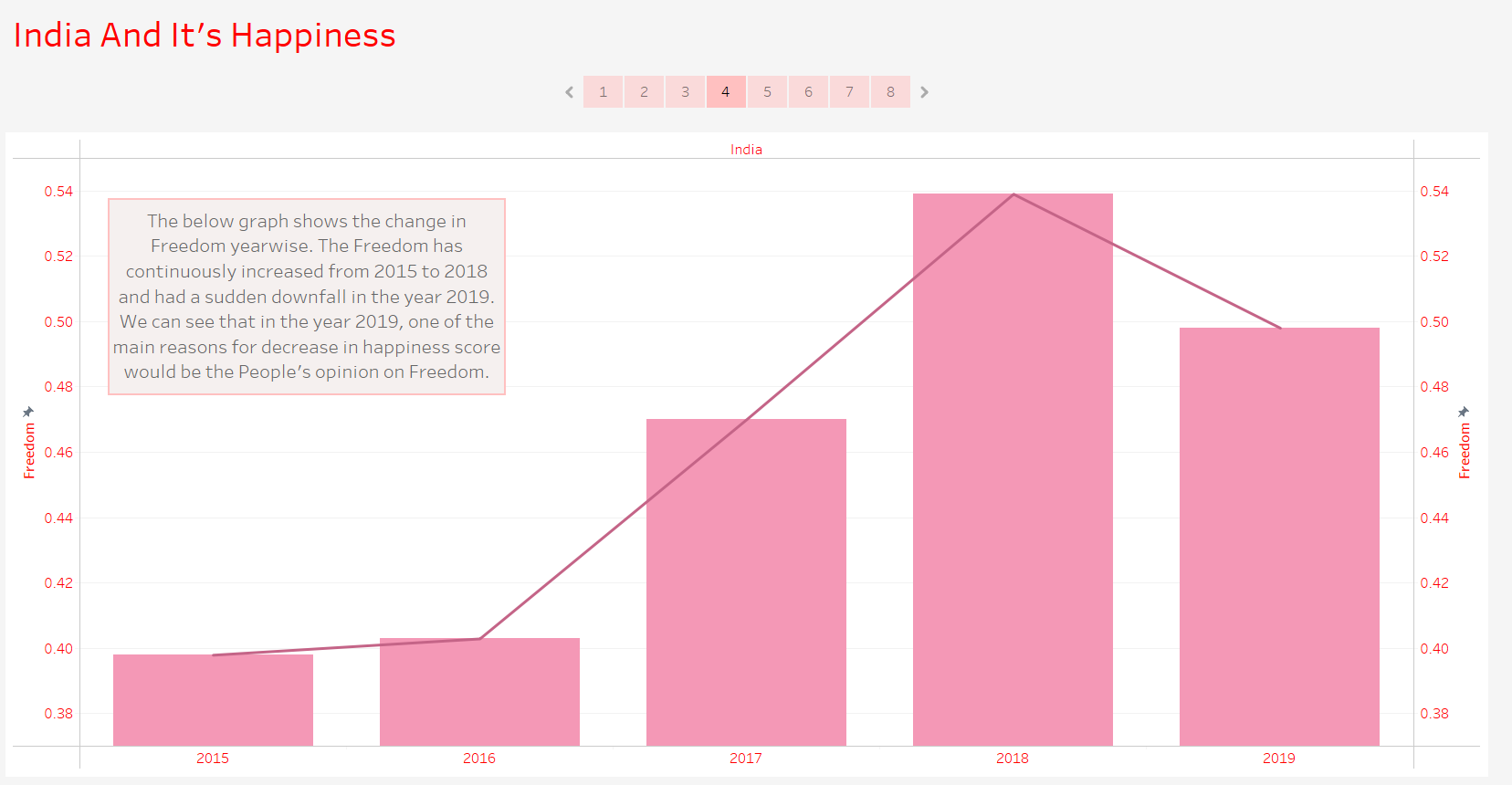


Question 5:

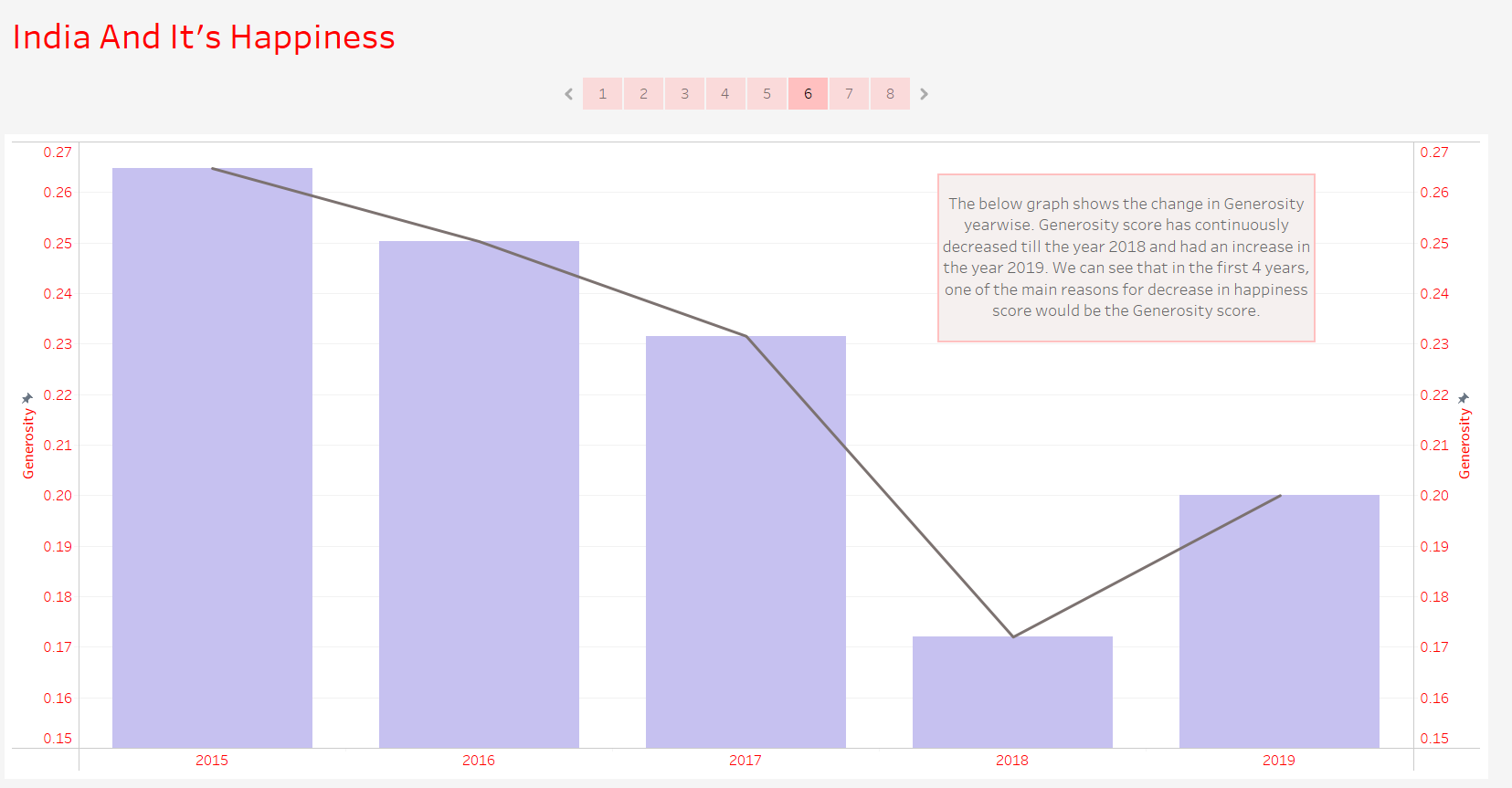


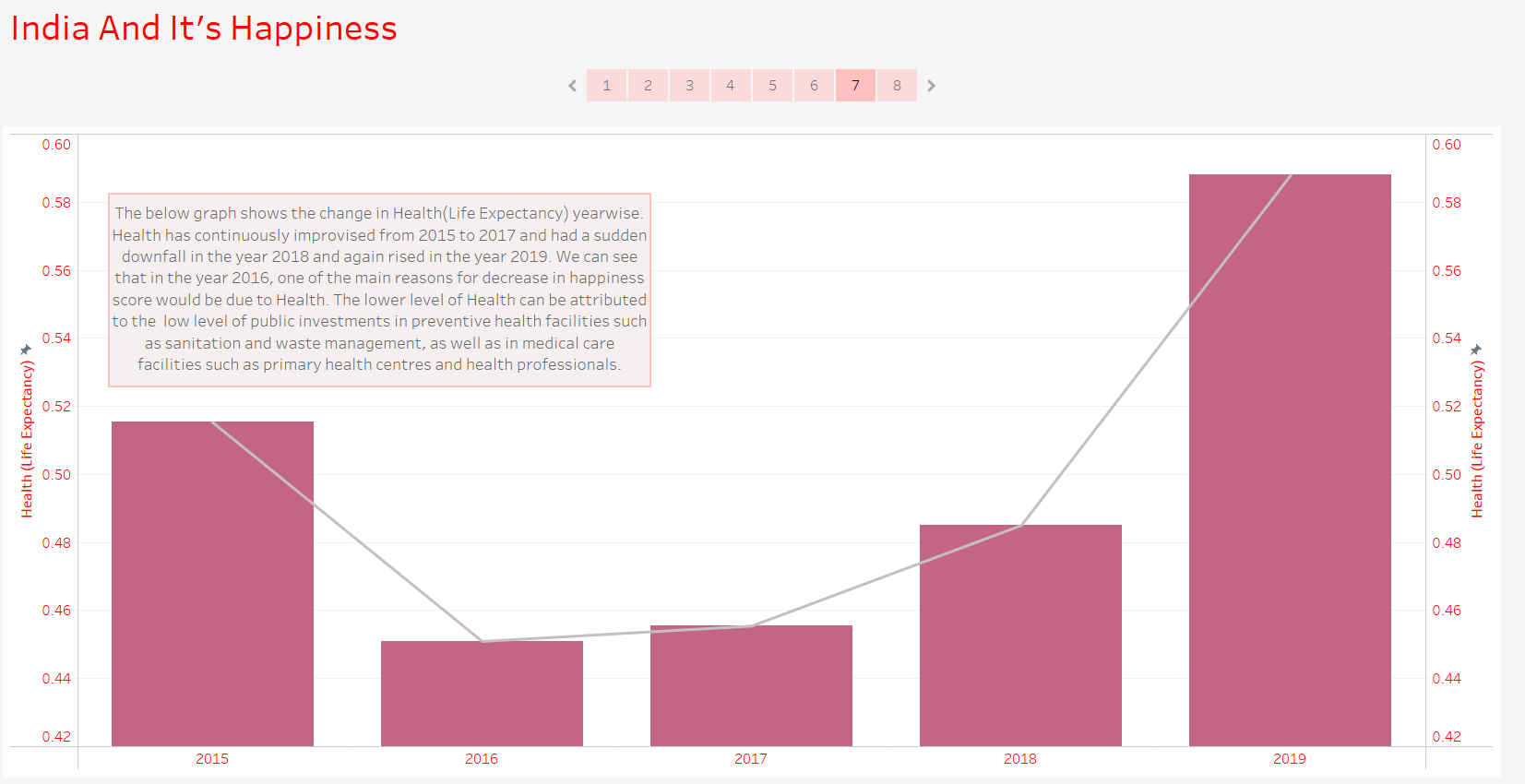


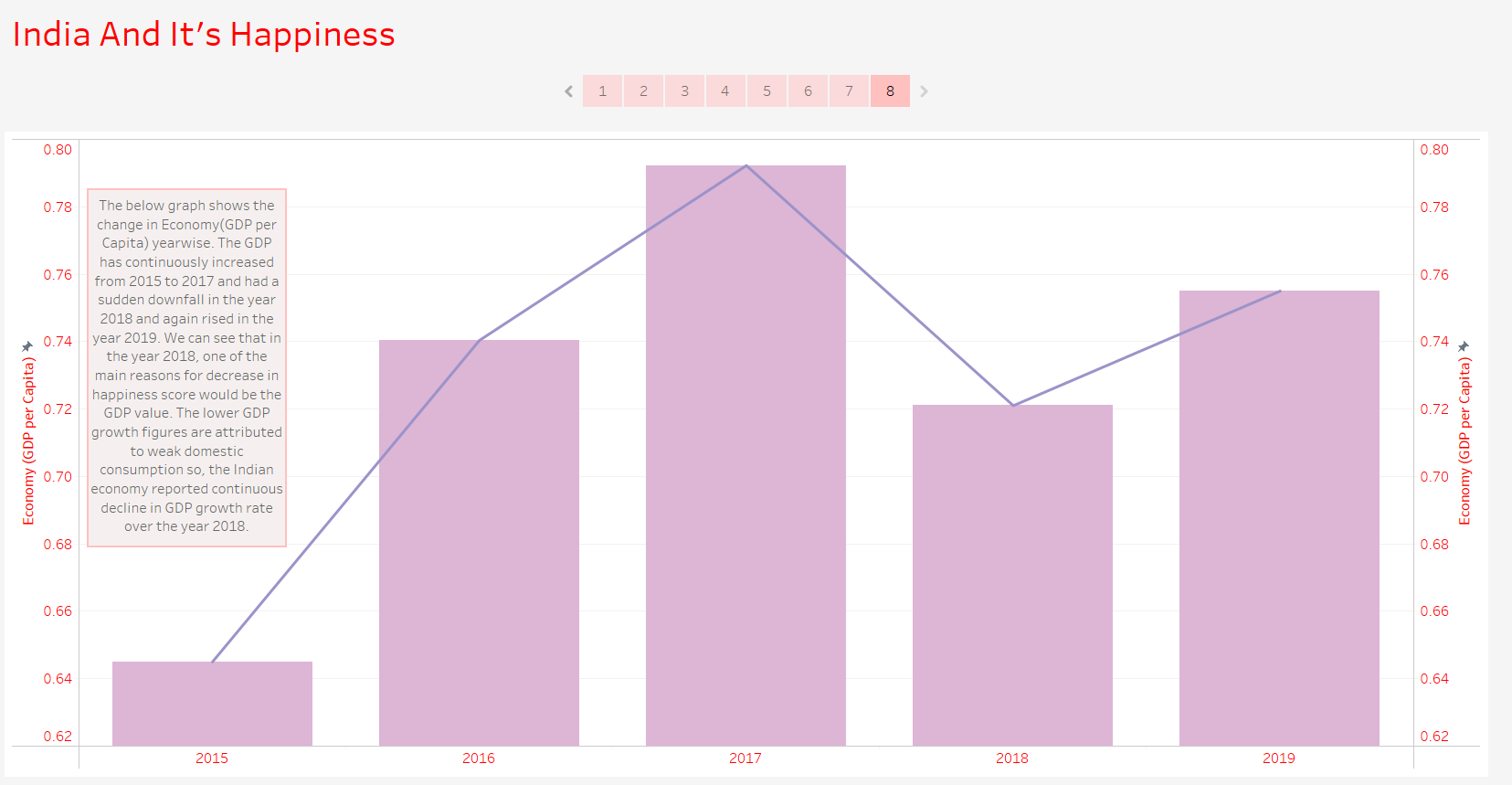




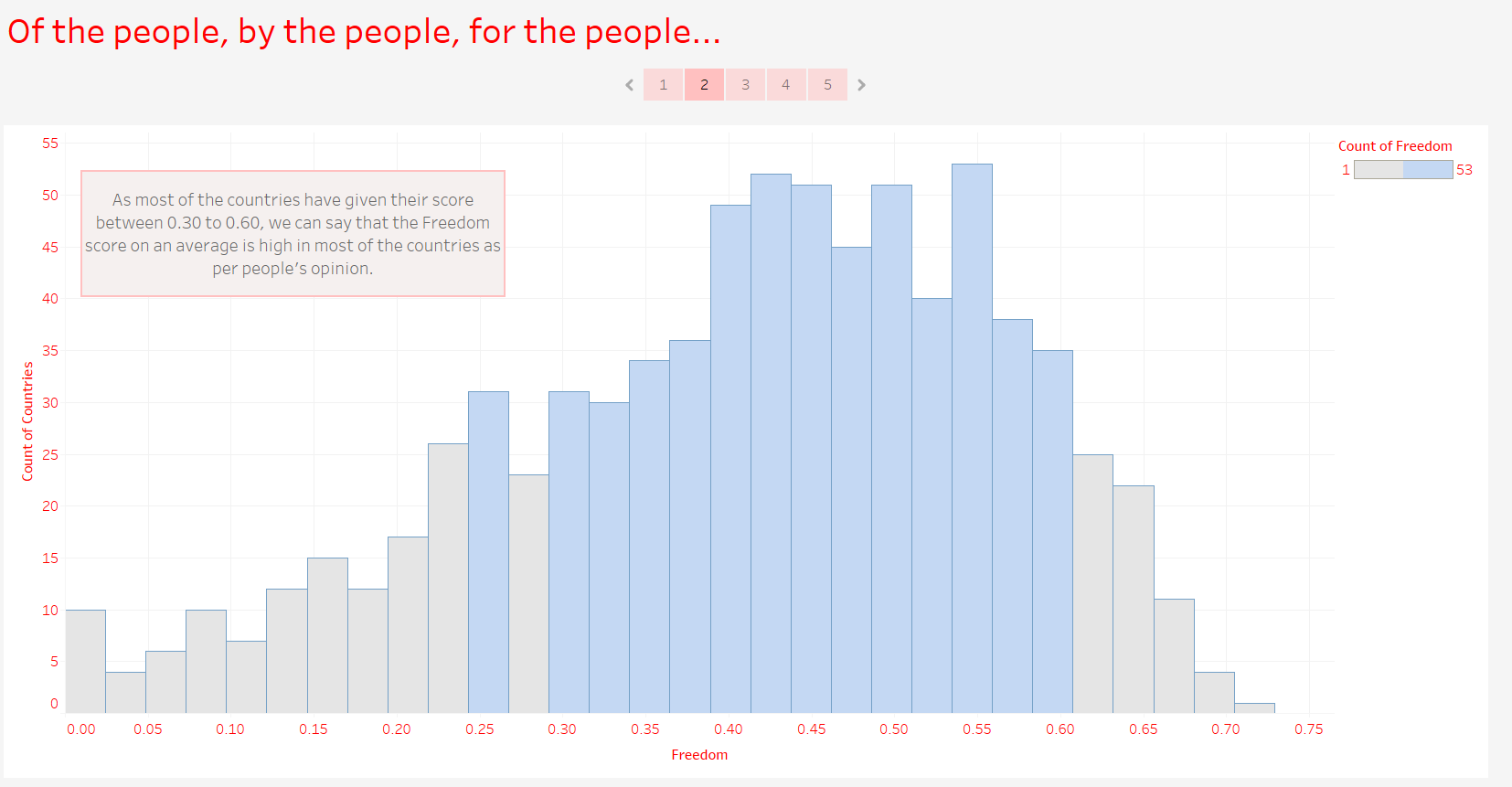




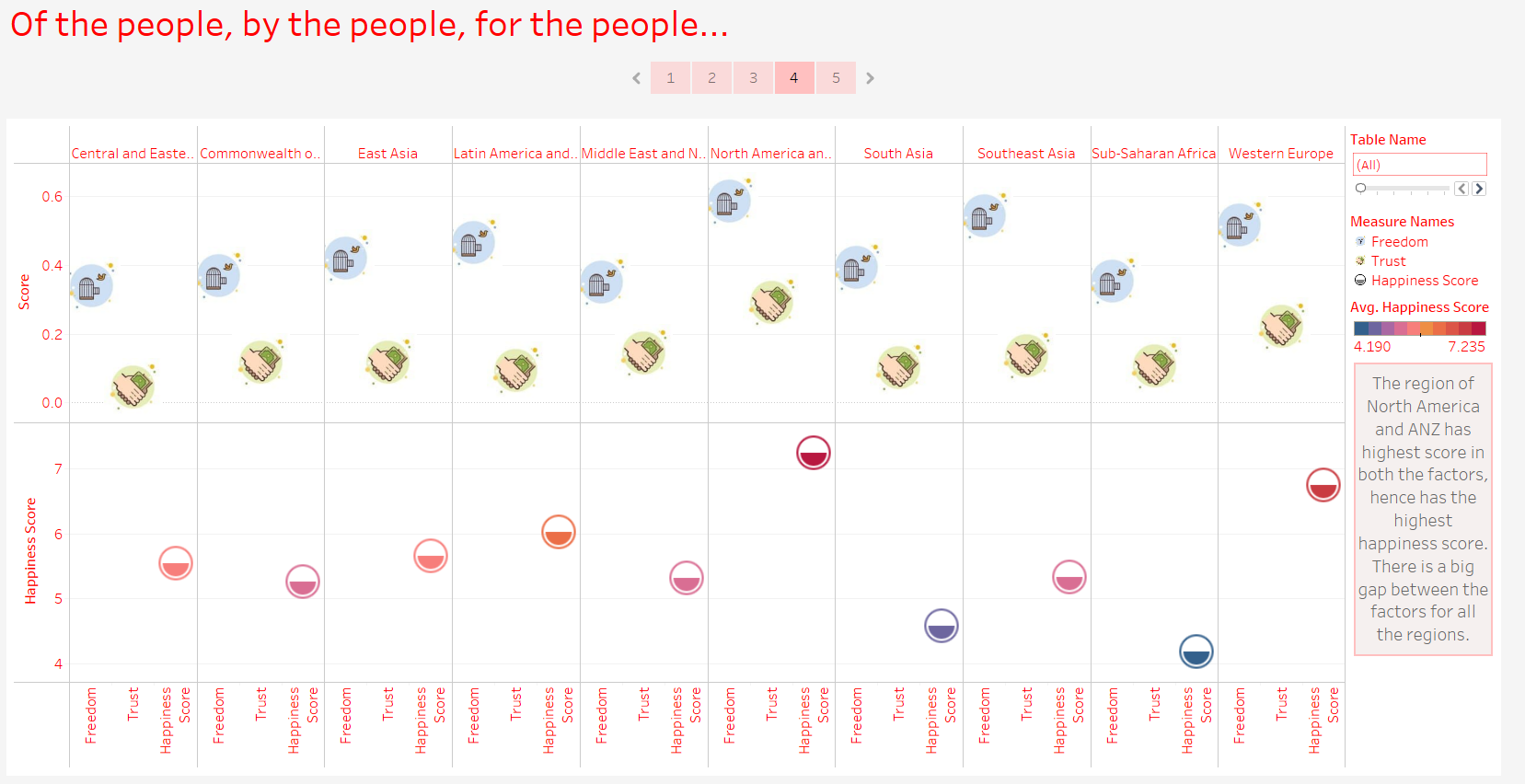


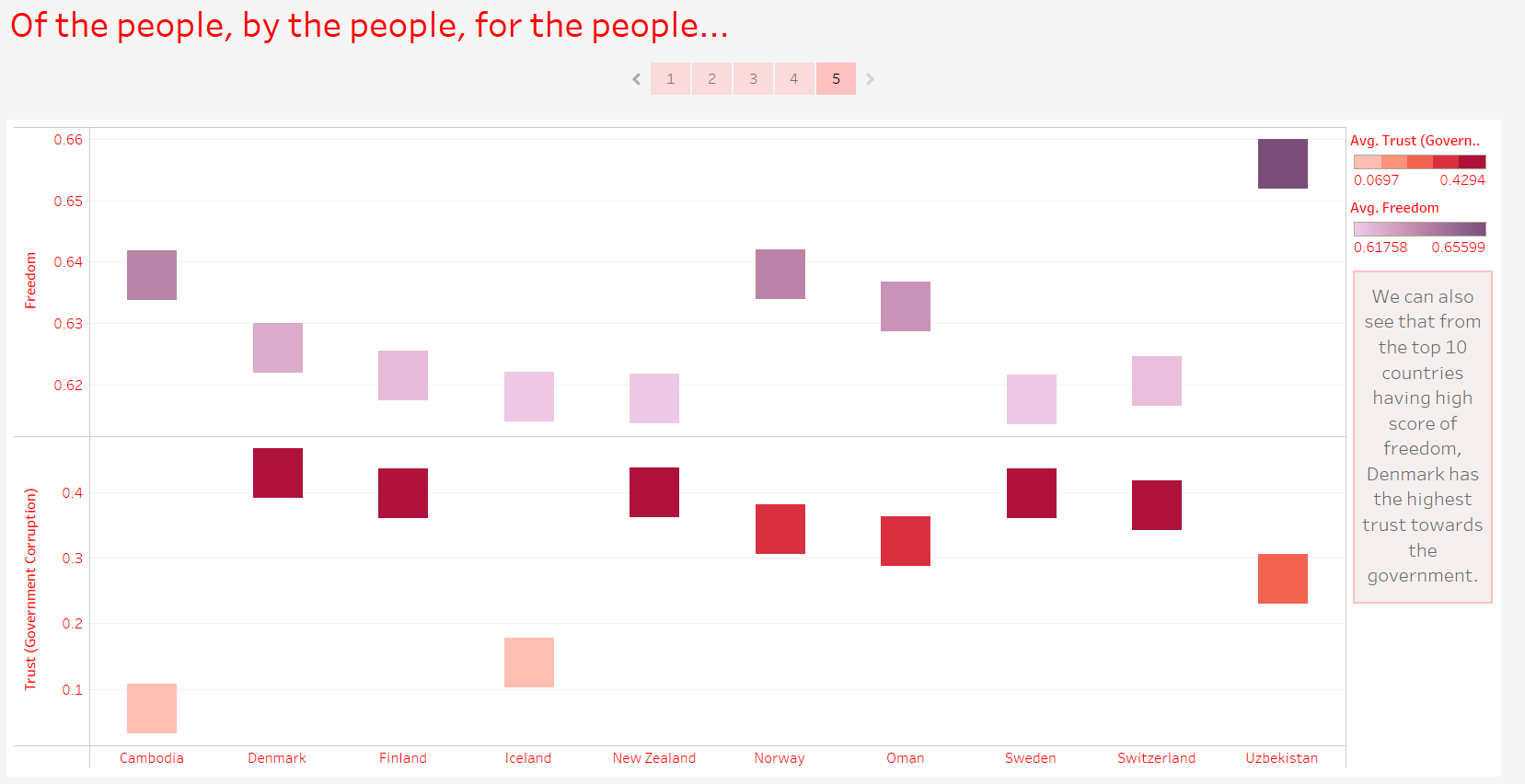


Question 6:

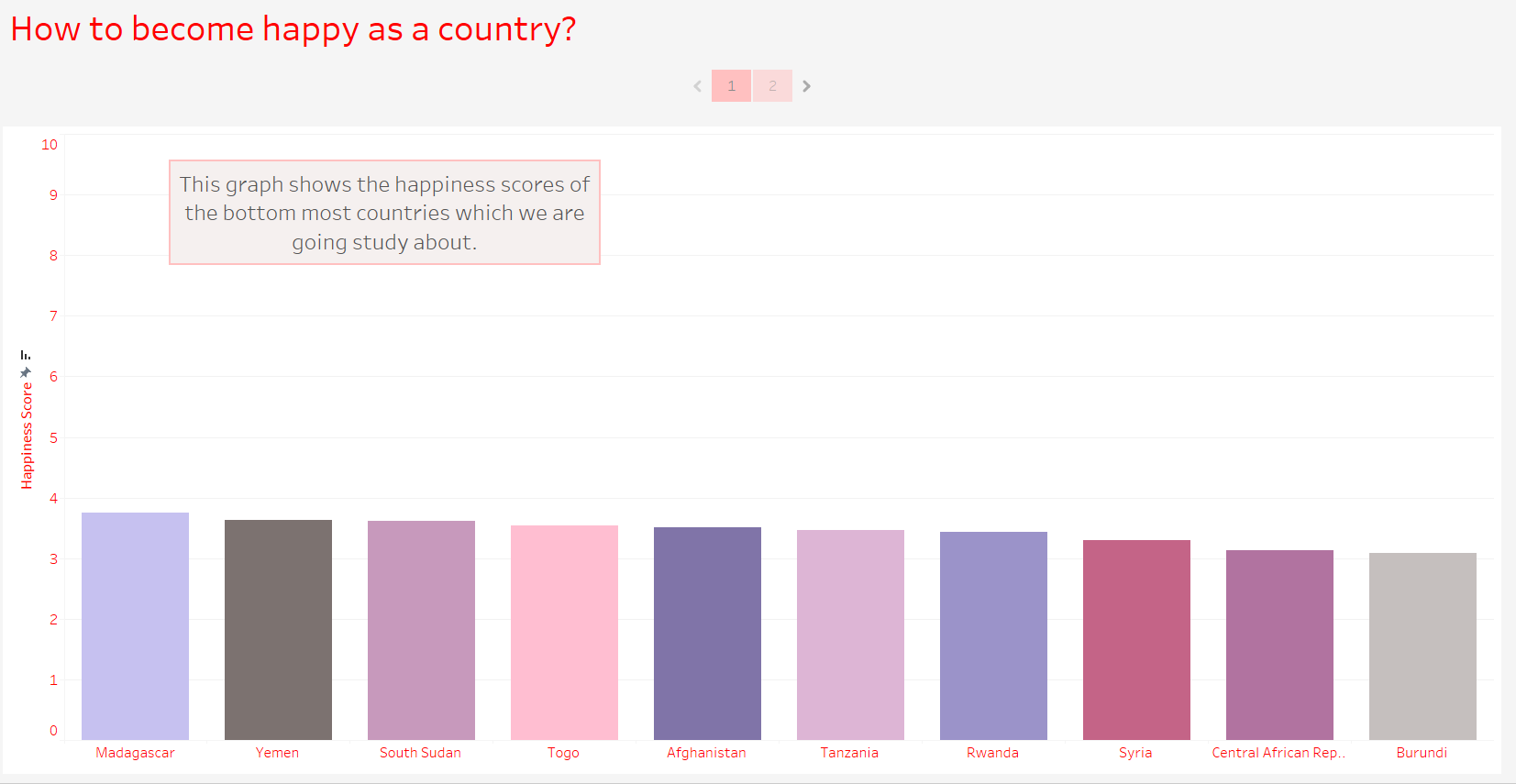


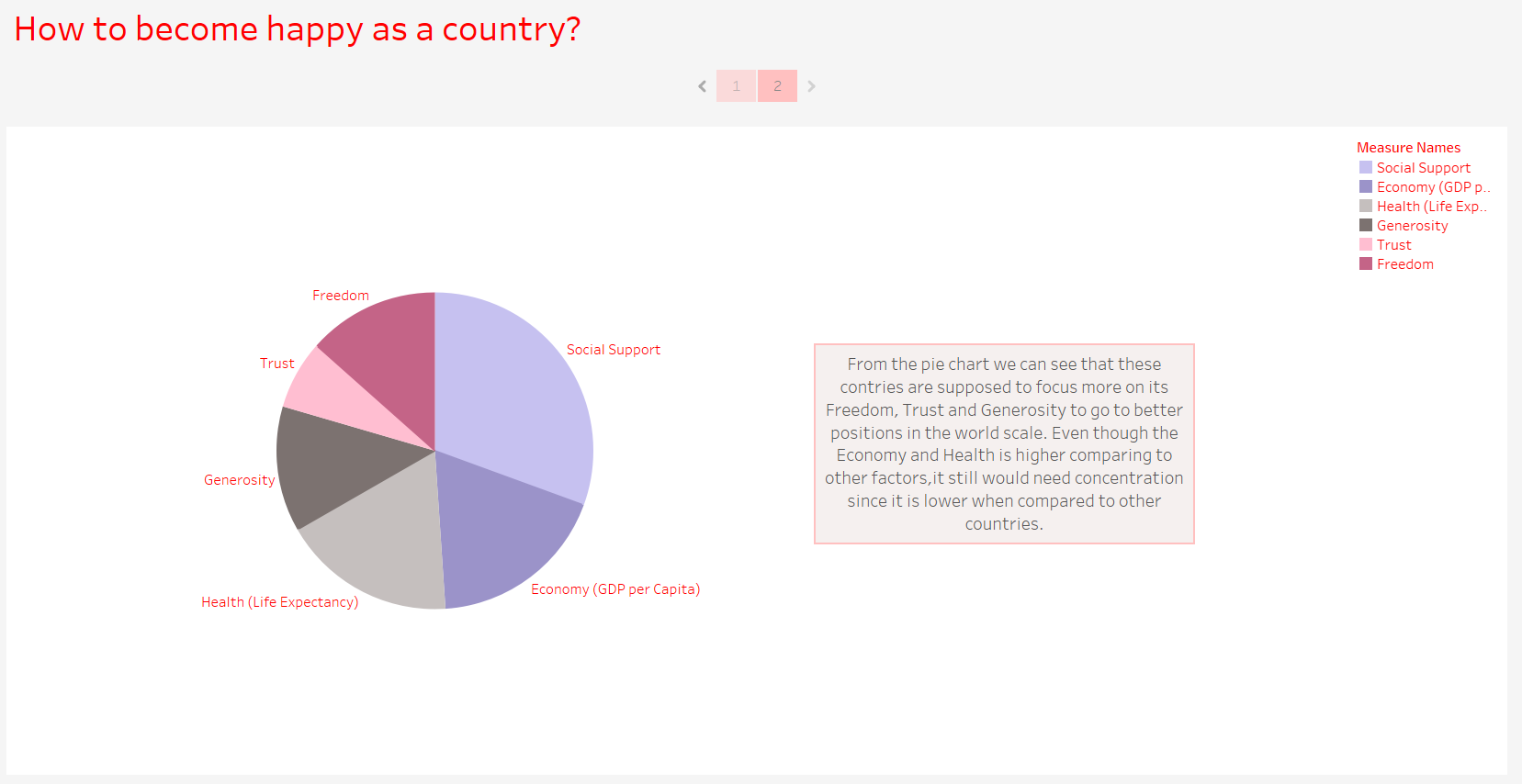






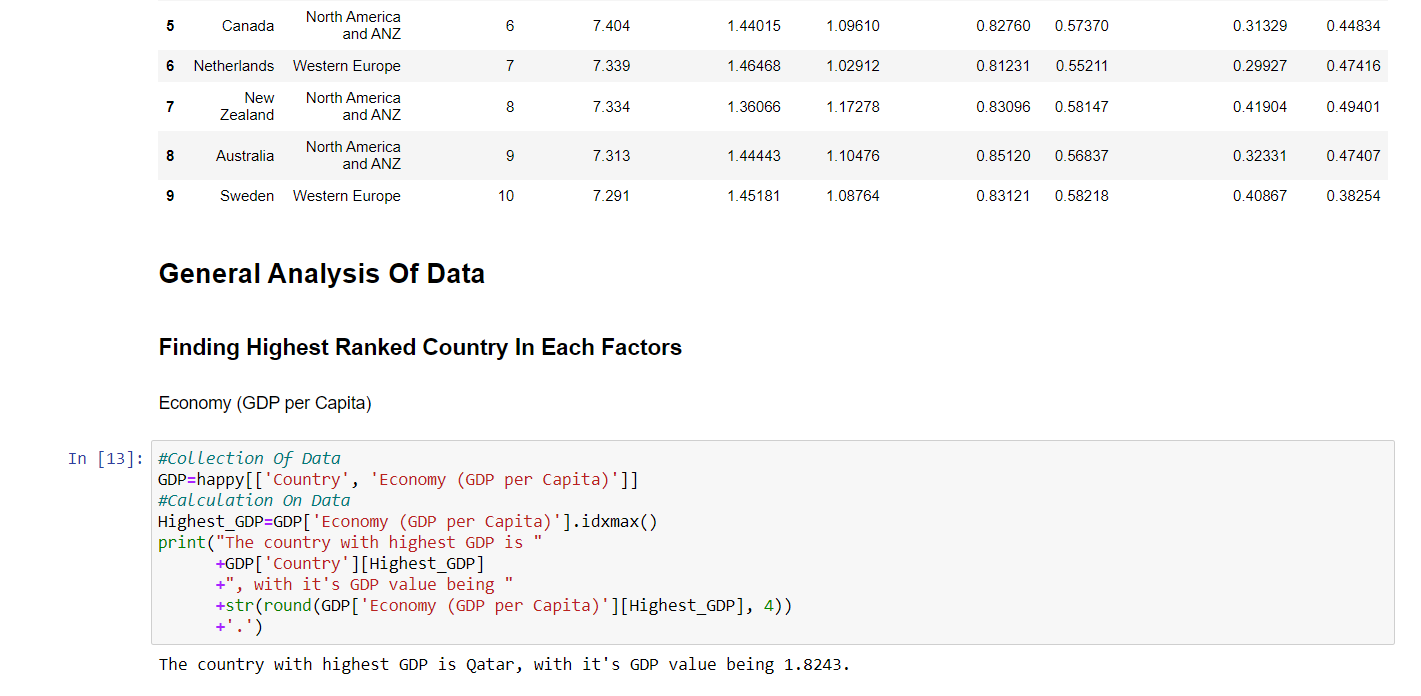
Question 7:



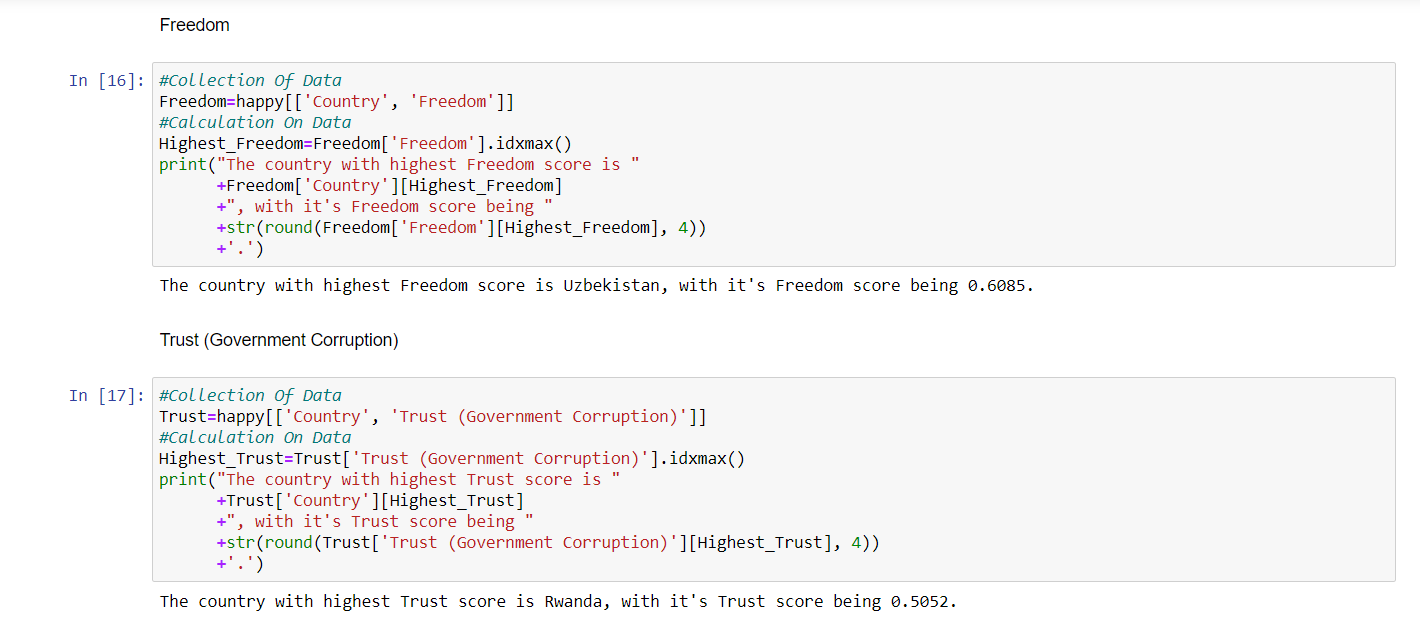


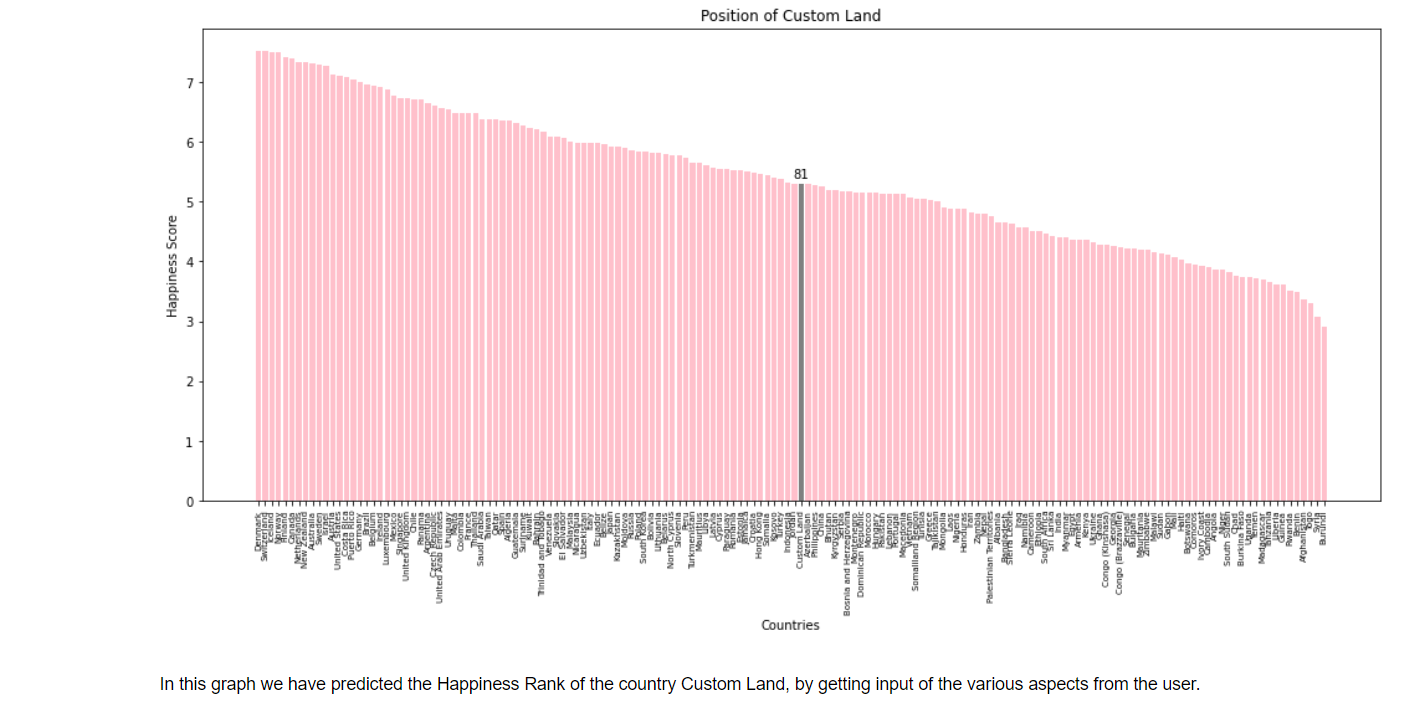
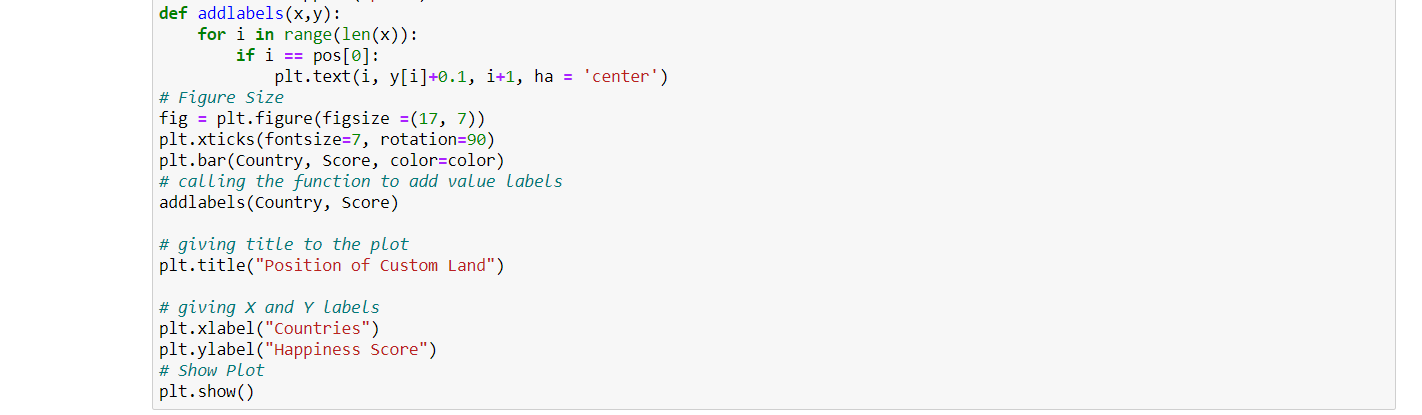
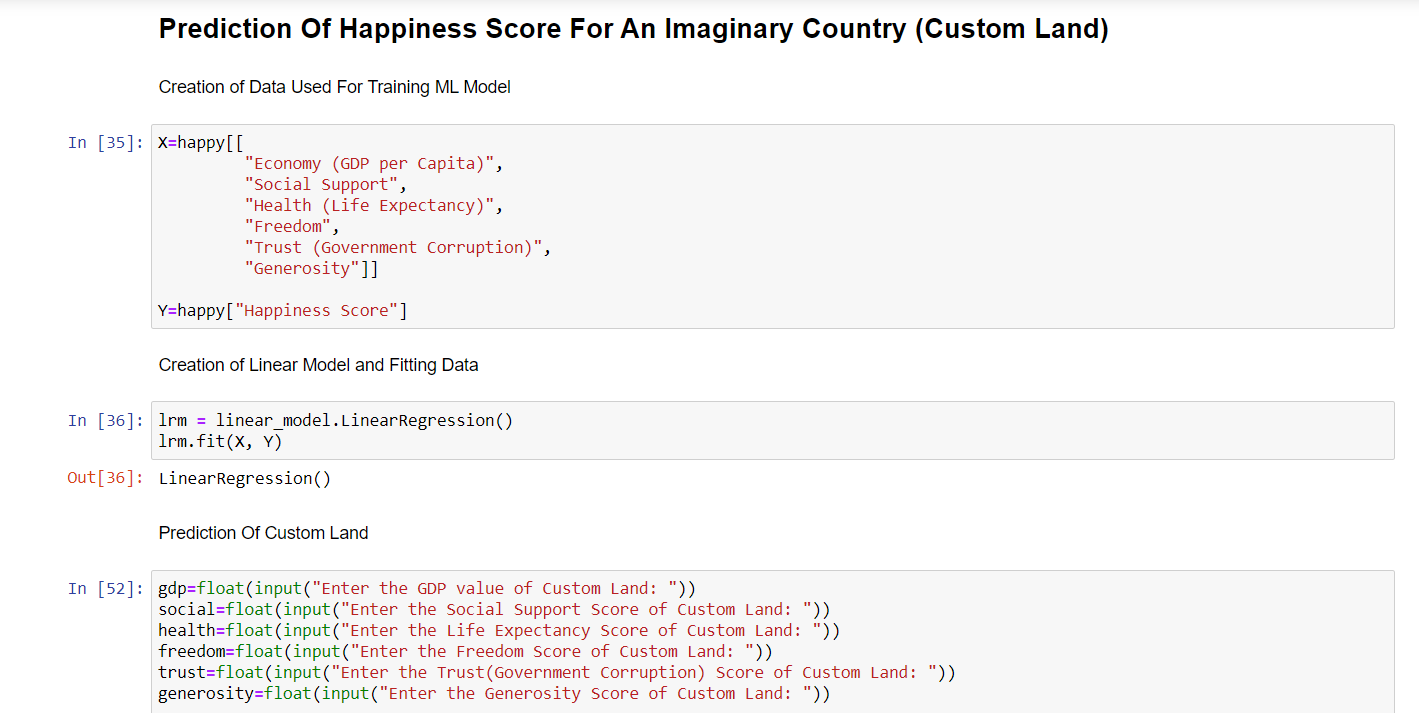
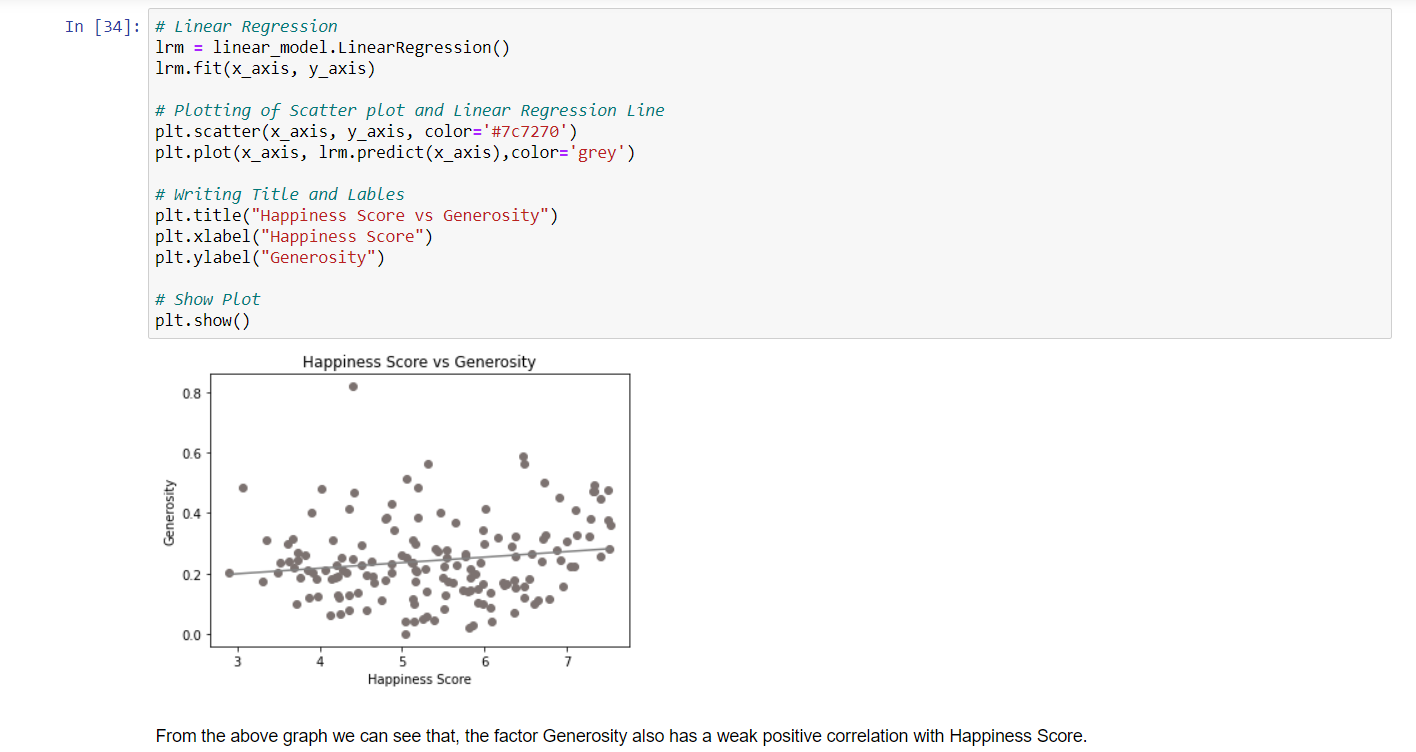
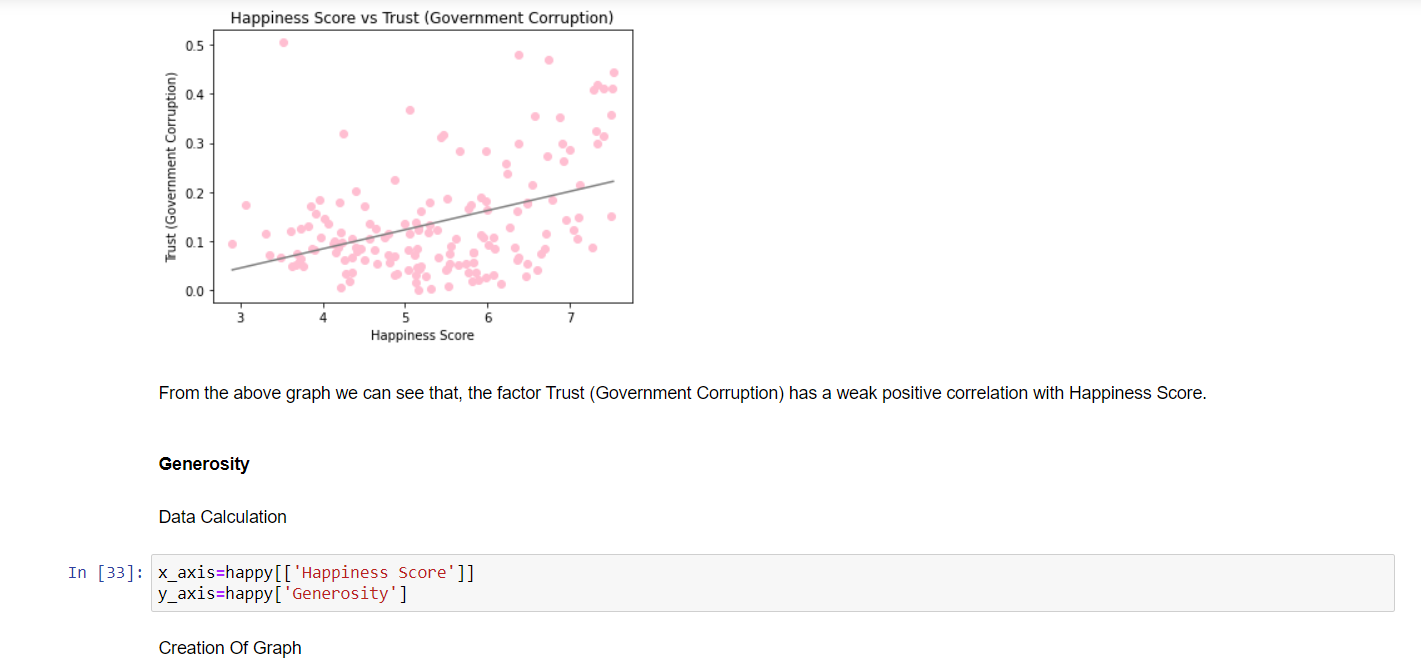
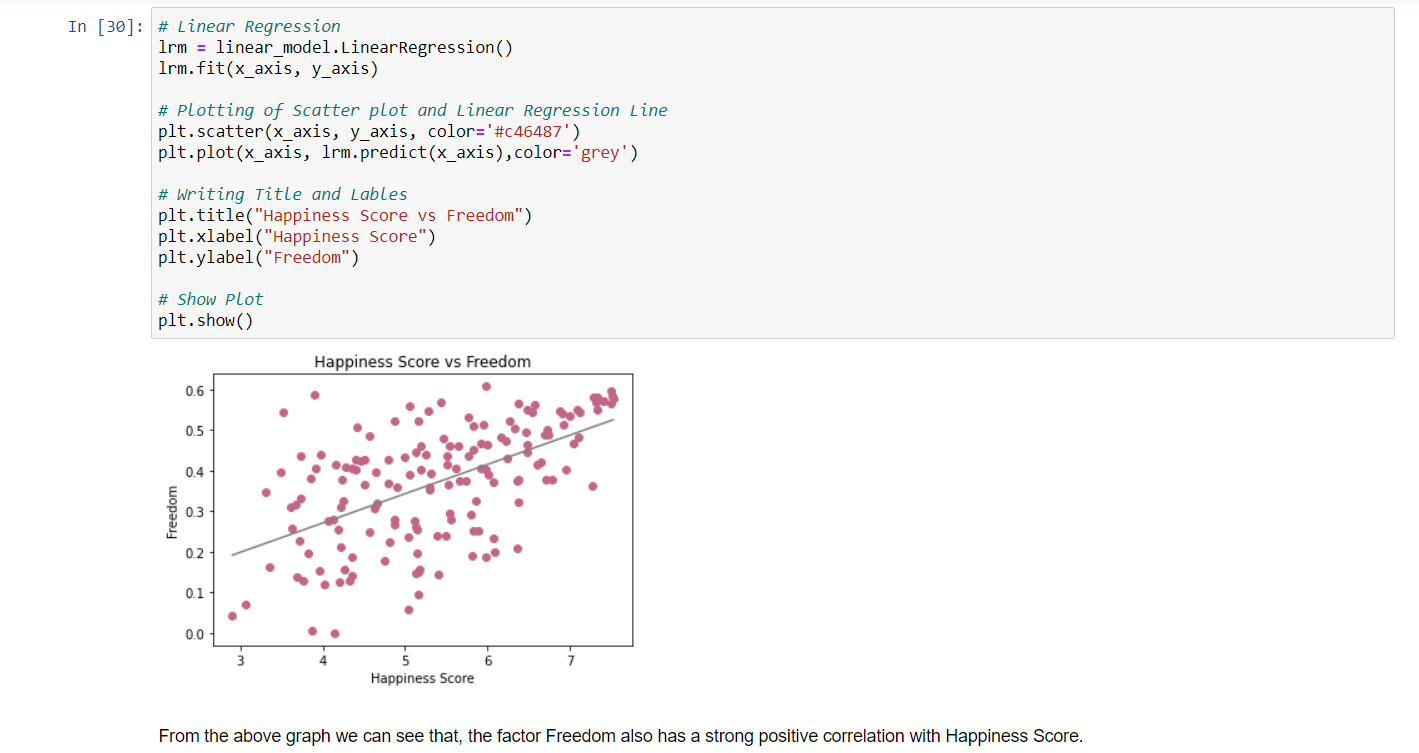
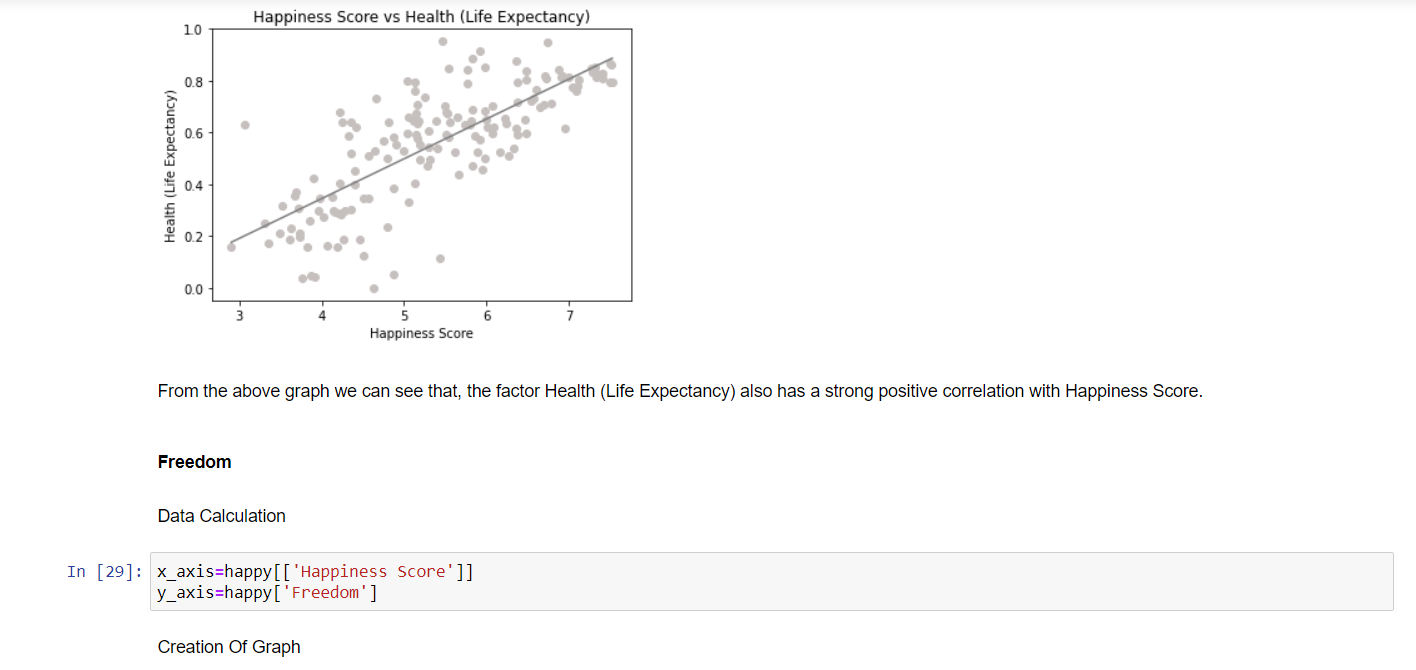
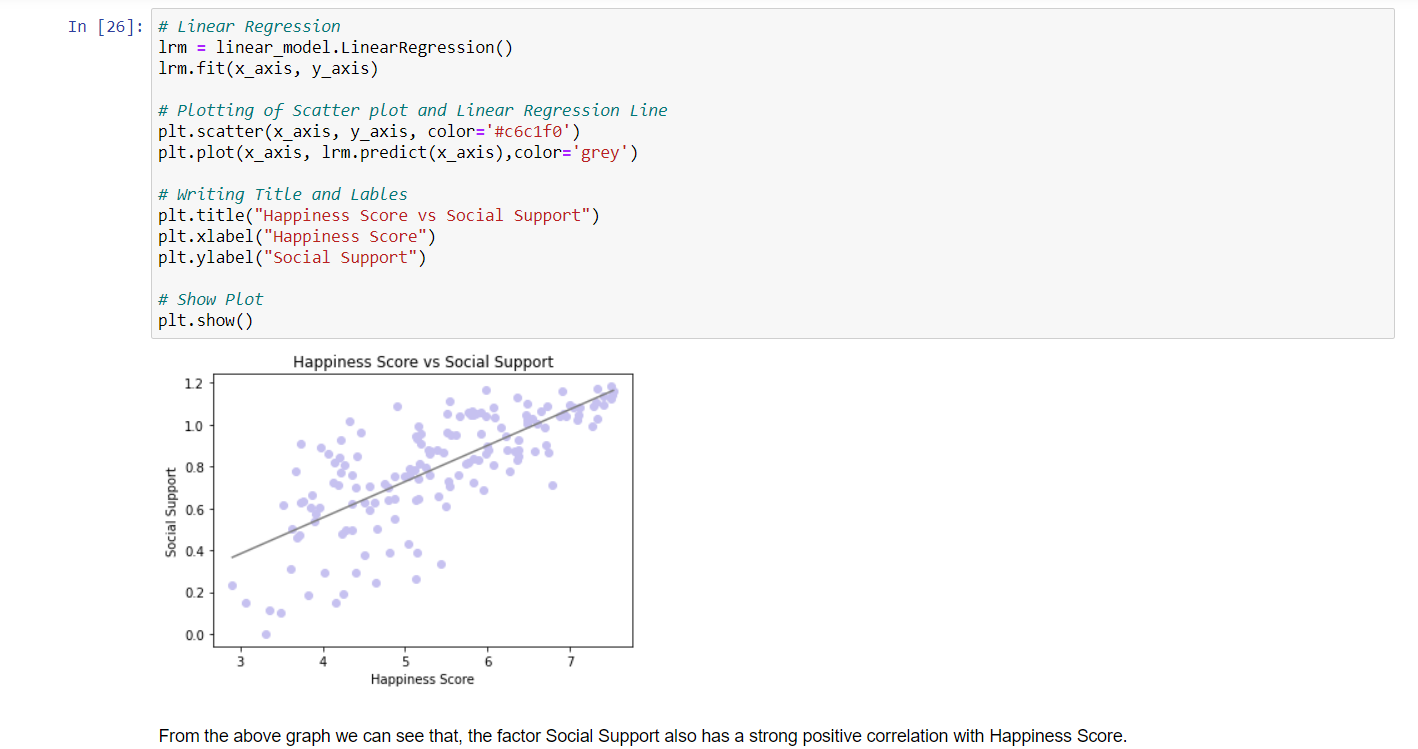
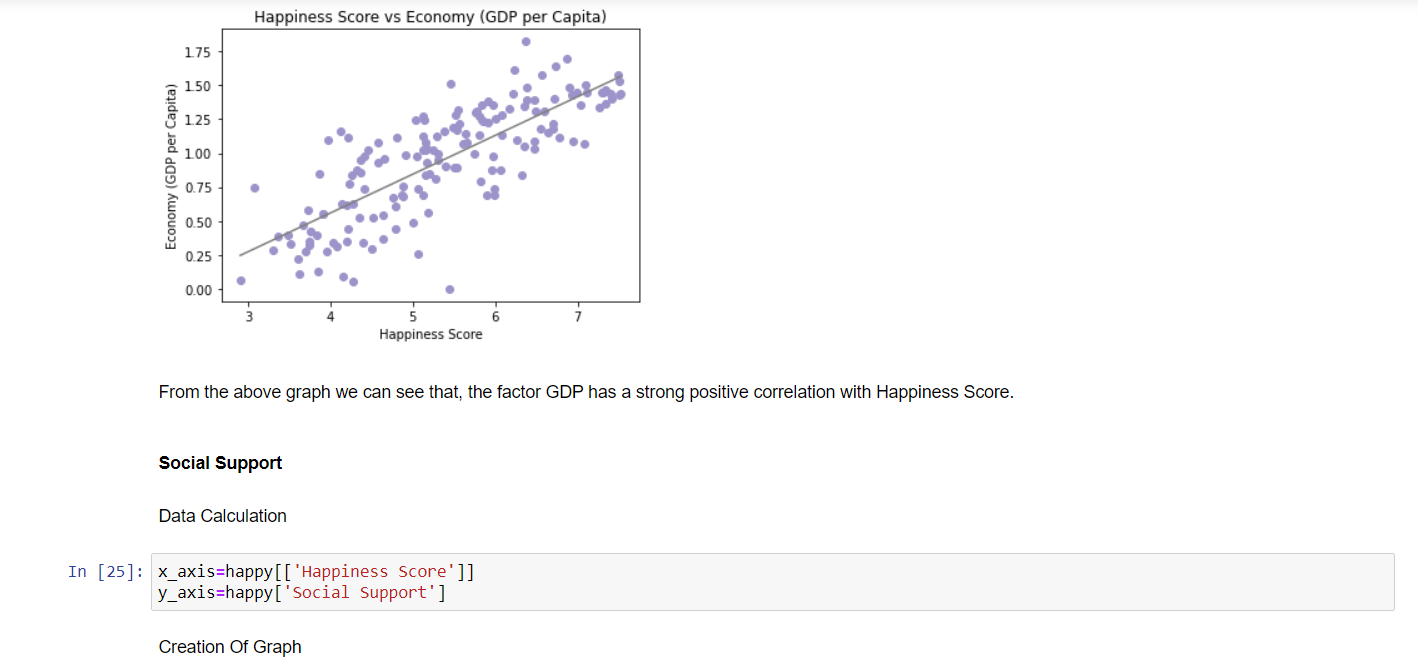
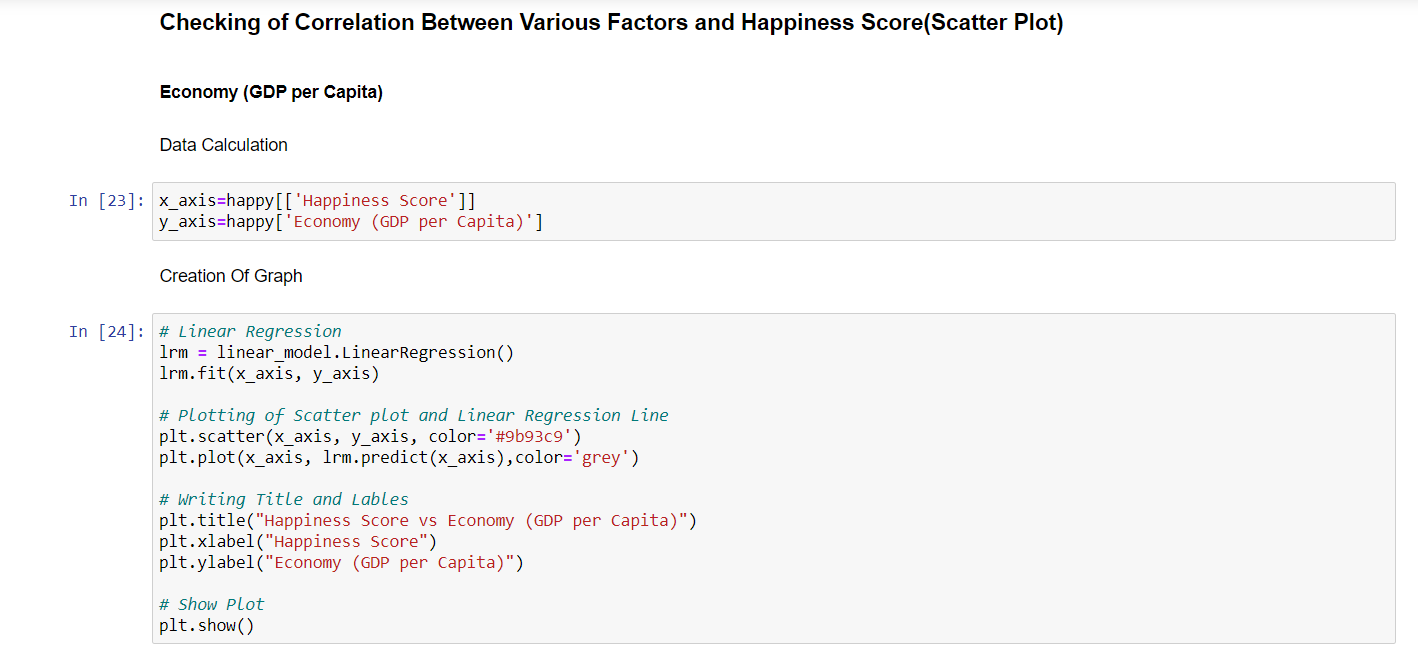
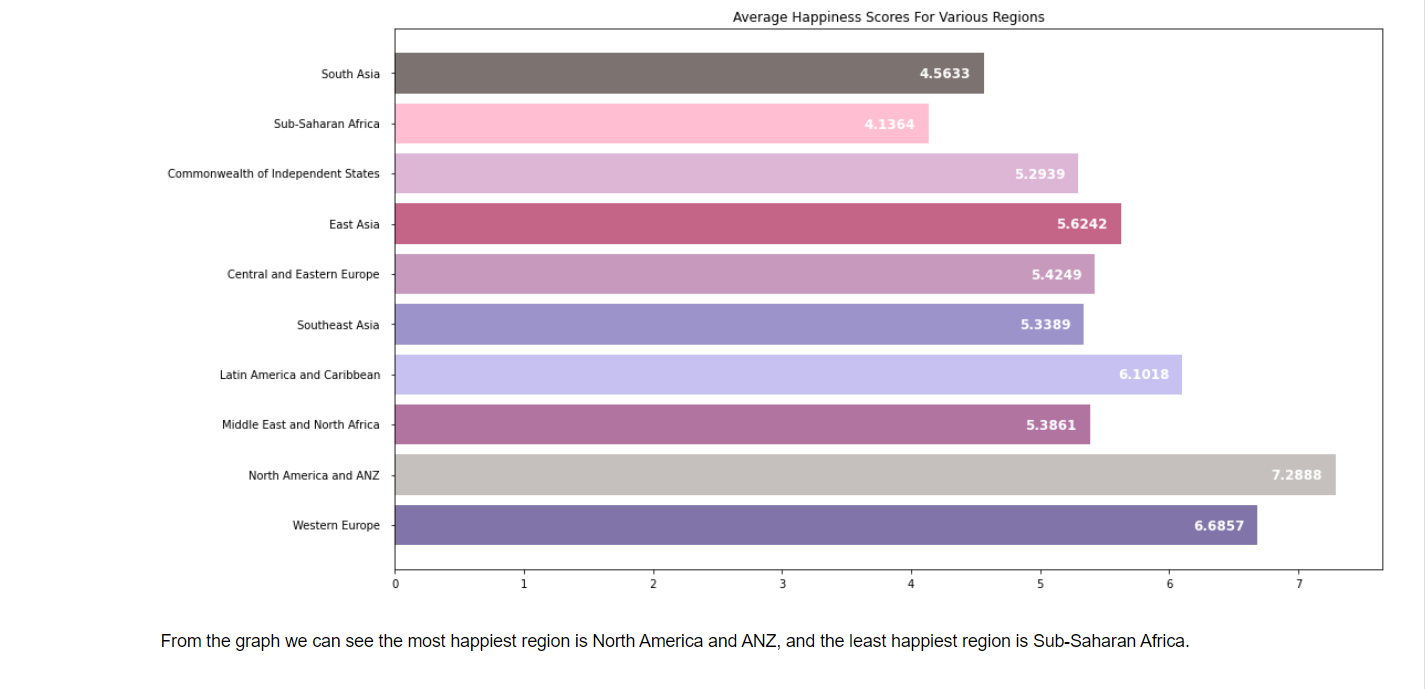
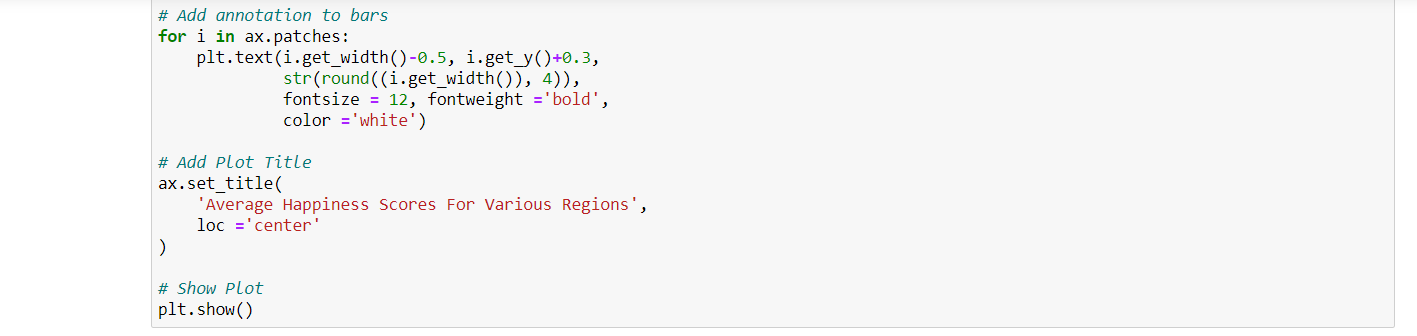
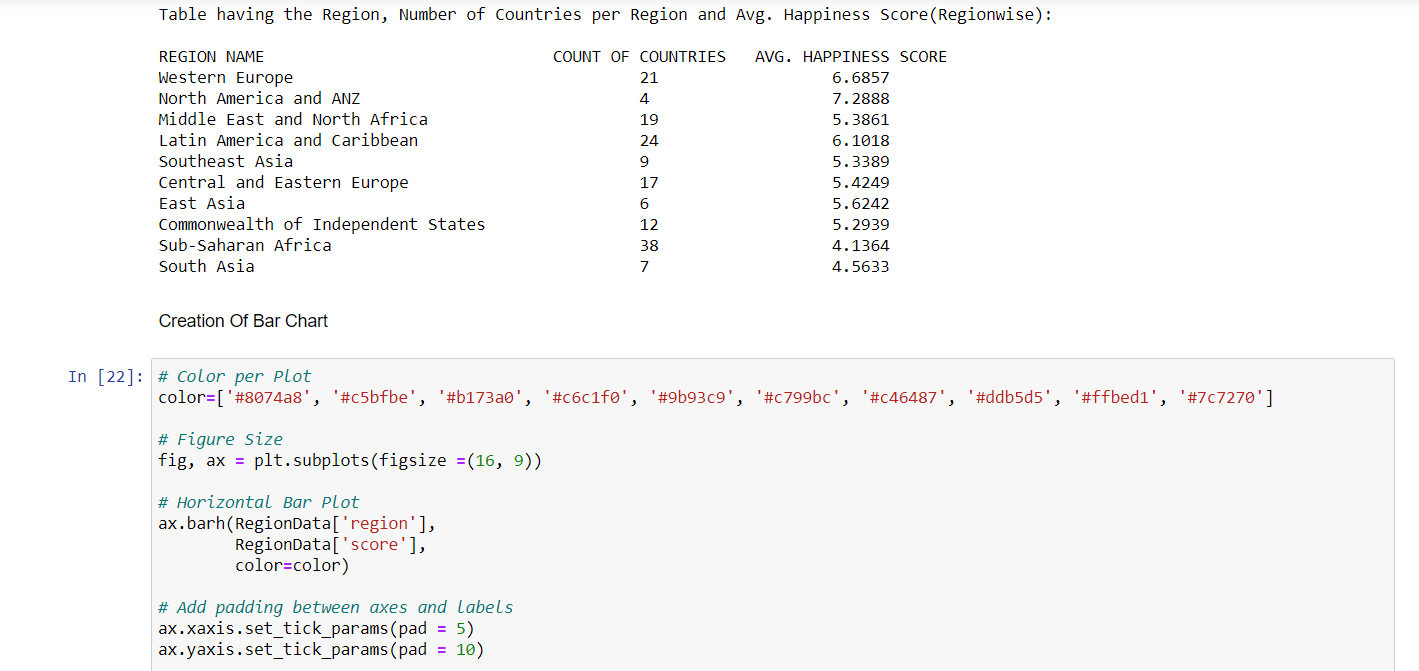
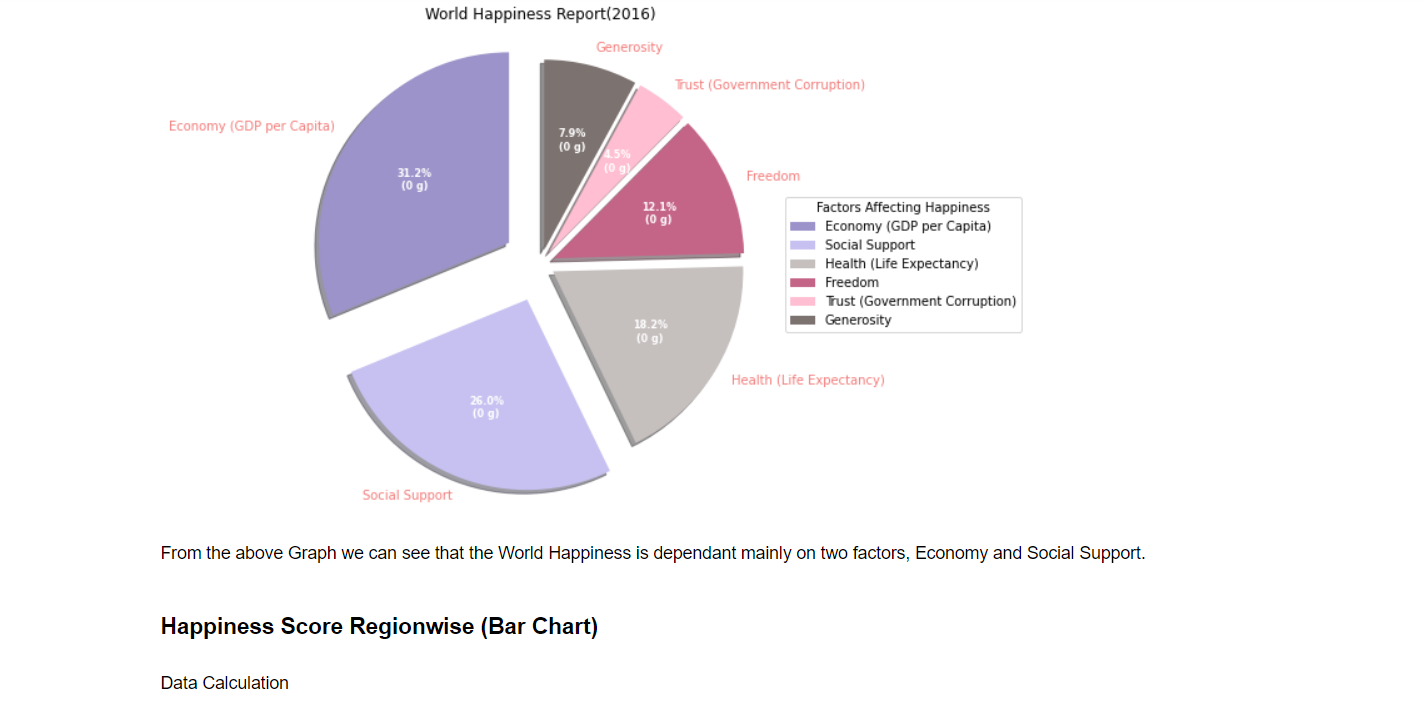
* 1. **Python (ML Model)**

****

****

****

****

****

* 1. **Website**

1. Home Page

****

1. About Dataset Page



1. Questions Page



1. Analysis Page Sample



1. **SAMPLE CODE**

function App () {

return (

<Router>

<div>

<Header/>

<Switch>

<Route exact path='/'><Home/></Route>

<Route path='/about'><About/></Route>

<Route exact path='/analysis'><Question/></Route>

<Route path='/analysis/q0'><Q0/></Route>

<Route path='/analysis/q1'><Q1/></Route>

<Route path='/analysis/q2'><Q2/></Route>

<Route path='/analysis/q3'><Q3/></Route>

<Route path='/analysis/q4'><Q4/></Route>

<Route path='/analysis/q5'><Q5/></Route>

<Route path='/analysis/q6'><Q6/></Route>

<Route path='/analysis/q7'><Q7/></Route>

</Switch>

<Footer/>

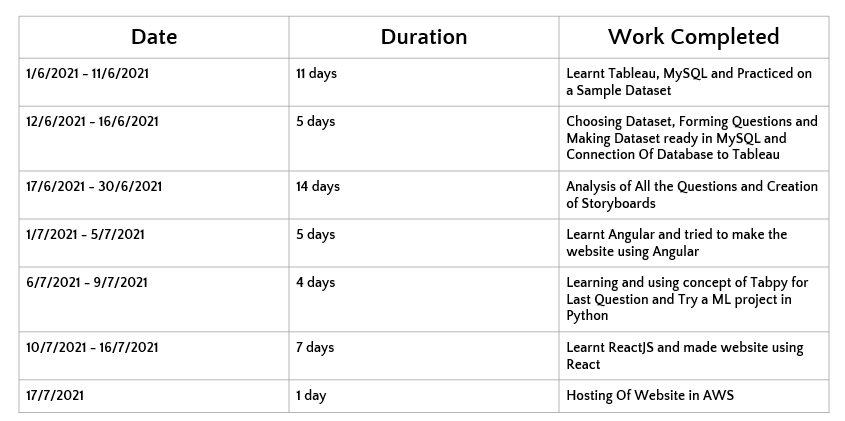
</div>

</Router>

);

}

1. **TIMELINE**

****

1. **CONCLUSION**

The world happiness report depends on the average of scores that people gave, based on their opinions about how different factors affect their happiness. From the overall analysis of the World Happiness Report, we could conclude that the happiness scores vary based on different factors in different regions. On the basis of the analysis that we made, we found that Economy (GDP per capita) and Social Support are the main factors on which happiness score depends on.

1. **TARGET AUDIENCE**

* People who are interested in learning about World Happiness Report
* People who want to know about how the Report is made and based on what factors scores are given
* People who want to know about how different factors influence the happiness of the country
* People who are learning data analysis
* Data analysis enthusiasts

1. **FUTURE SCOPE**

**Questions Which Can Be Addressed Later**

1. What is the percentage of each factor, while taken for calculating the happiness score?
2. What impact does Social Support Score have on the Happiness Rank?
3. Is there any correlation between Social Support and Health?
4. How Covid has brought a major change in people’s Health and did that affect the Happiness Score? (If we get the 2021-year dataset!)

**Improvements To Be Made In Website**

* Make the website Responsive (For making it easy to view in both Tablet and Phone)
* Make the website Dynamic (Integration of Backend)
* Make a survey page which has a form in which users can enter values using which happiness score can be calculated

1. **REFERENCE**

* [**https://www.kaggle.com/mathurinache/world-happiness-report?selec t=2015.csv**](https://www.kaggle.com/mathurinache/world-happiness-report?select=2015.csv) **- for Collecting Dataset**
* [**https://en.wikipedia.org/wiki/World\_Happiness\_Report**](https://en.wikipedia.org/wiki/World_Happiness_Report) **- for Understanding the Dataset**
* [**https://www.youtube.com/watch?v=aHaOIvR00So**](https://www.youtube.com/watch?v=aHaOIvR00So) **- for Learning Tableau**
* [**https://www.geeksforgeeks.org/linear-regression-python-implementation/**](https://www.geeksforgeeks.org/linear-regression-python-implementation/) **- for referring regarding Python ML model**
* [**https://youtu.be/W6NZfCO5SIk**](https://youtu.be/W6NZfCO5SIk) **- for learning JS basics**
* [**https://youtu.be/PFmuCDHHpwk**](https://youtu.be/PFmuCDHHpwk) **- for learning JS OOPS concept**
* [**https://youtu.be/NCwa\_xi0Uuc**](https://youtu.be/NCwa_xi0Uuc%20) **- for learning ES6**
* [**https://m.youtube.com/watch?v=CI2kX2EFhWc**](https://m.youtube.com/watch?v=CI2kX2EFhWc) **- for Learning React JS**
* [**https://www.npmjs.com/package/tableau-react-embed#usage**](https://www.npmjs.com/package/tableau-react-embed#usage) **- for Integration of Tableau into Website**
* [**https://www.datacamp.com/community/tutorials/git-push-pull**](https://www.datacamp.com/community/tutorials/git-push-pull) **- for Using GitHub Repository**
* [**https://youtu.be/lB4DTqMEumY**](https://youtu.be/lB4DTqMEumY) **- for Learning Hosting in AWS**