Visualization of Worldwide Mental Health Disorder

CS5346: Information Visualization

Authors: Wang Ruixiao (A0225545E), Wang Bingluo (A0224877R) and Chen Nan (A0224882Y)

2021

1 Team Member and Contribution

- Wang Binluo (A0224877R): Made the map chart section and wrote the corresponding parts of the report.
- Wang Ruixiao (A0225545E): Made the word cloud and scatter chart; wrote the corresponding parts of the report; and created the poster.
- Chen Nan (A0224882Y): Collected and pre-processed datasets; made the correlation charts and aggregated all created charts into a Tableau story workbook; wrote corresponding parts of the report.

2 Purpose

Mental health refers to cognitive, behavioral, and emotional well being [1]. It's exactly how people think, feel, act, just as consequential as our physical health. In line with the World Health Organization (WHO), mental health includes "subjective well being, perceived self efficacy, autonomy, competence, inter-generational dependence, and self actualization of one's intellectual and emotional potential, among others" [2]. It allows people to address with challenges of life, discover one's own potential, work productively, and contribute to their communities. It 's essential at every stage of human life, from childhood and adolescence through adulthood.

However, mental health disorders are causing more and more harassment to people in recent years and exhibiting a trend of younger and diversified diseases as the data presented. For example, the proportion of young Americans undergoing different types of mental health disorders has risen obviously over the past decade, with no corresponding increase in older adults, according to research published by the American Psychological Association [3]. Unlike physical illnesses, mental disorder is more insidious. Patients with mental illnesses are often ashamed to seek professional medical treatment.

Fortunately, this phenomenon has been greatly improved, and the awareness of mental illness has also been popularized. On the one hand , it is due to the increase of people's awareness of mental diseases. On the other hand, the emergence of database technology and data visualization technology also makes it more intuitive to present the phenomenon of mental illness for further analysis.

The targeted group of our visualizations will be the government's public health department and health professionals, who can use these charts to see the current status of various types of mental disorders in their country as well as the relationships between these disorders and their governance and economic factors. The pubic can also use our visualizations to compare the current status of mental health in their country with the rest of the world. Finally, we summarize the basic information of this visualization framework in Table 1, and the complete view of the visualization framework is in Appendix.

Theme	Visualization of worldwide mental health disorders
Developing tool	Tableau
Link	https://public.tableau.com/views/Final_Project_16186526140260/Story1

Table 1: Basic information of this project.

3 Dataset Description and Queries

3.1 Dataset and data attributes

- **Prevalence of mental health disorders by disorder type [4].** This dataset contains tabular data, which describes the percentage of the total population with a given mental health disorder. It contains five types of mental health disorders: anxiety, depression, schizophrenia, bipolar disorder, eating disorder, drug use disorder, and alcohol use disorder. The data type and attribute type of this dataset is summarized in Table 2.
- **Mental health governance data by country [5].** This dataset includes information about governance policy on mental health for each country, indicating whether a country has stand-alone laws or plans for mental health. It's categorical data, the value is whether True or False.
- **Mental health service availability.**
  - **Human resources data by country [6].** This dataset contains information about the provision of human resources for mental health in each country. It contains the number of psychiatrists, nurses, social workers, and psychologists working in mental health sector per 100,000 population. It's quantitative data.
  - **Facilities data by country [7].** This dataset contains information about the number of facilities dedicated for mental health per 100,000 population in each country, including mental hospitals, mental health units, mental outpatient facilities, mental health day treatment facilities, and community residential facilities. It's quantitative data.
  - **Beds data by country [8].** This dataset stores information about the availability of beds in different kinds of mental health facilities. It's quantitative data.
- **Government expenditures on mental health as a percentage of total government expenditures on health** (%) **[9]**. This dataset records information about the percentage of total government budget on health that spent on public mental health. It's quantitative data.

3.2 Queries

1) What is the situation of different types of mental disorders in each country in terms of the number of disorder cases per 100,000 population? We considered alcohol use disorder, anxiety disorder, bipolar disorder, depressive disorder, drug use disorder, eating disorder and schizophrenia

Attribute	Data Type	Attribute Type	Description
Entity	Item	Categorical	Country name.
Code	Item	Categorical	Country code, not used.
Year	Attribute	Quantitative	Time information.
Prevalence - *	Attribute	Quantitative	Prevalence of *.

Table 2: The data type and attribute type of dataset "Prevalence of mental health disorders by disorder type"

2) In each country, which type of mental disorder has the highest number of cases?

3) What is the relationship between the number of cases with different types of mental disorders in each country?

4) For each types of mental disorder, which countries have the highest number of cases per 100,000 people?

5) What is the relationship between the number of disorder cases and some governance or economic factors, such as whether there is stand-alone mental health legislation, and the government expenditures on mental health as a percentage of total government expenditures on public health?

4 Different Design ideas considered

For the first two queries, since we have geospatial attribute in our dataset, we choose maps to describe the state of mental health disorders in each country, alone with a pie chart to show the distribution of case numbers of each mental disorder type in this country.

We also considered stacked bar chart and chord diagram to describe which country has how many mental health cases.

We use a correlation matrix to show the relationships between different types of mental disorders for the third query. For two elements with high correlation coefficient, we specifically use a scatter chart with a trend line to show their relationship.

As for the fourth query, we pick the form of word cloud because the country name is a text attribute.

In the last visualization, we choose scatter to present the number of mental disorder cases of a country and its governance or economic factors. The points are clustered according to its coordinate.

5 Data Pre-processing

We divide the data-preprocessing process into two parts: one is to pre-process the dataset about the prevalence of mental health disorders [4], and the other is to clean other datasets and aggregate them into one dataset.

Since the dataset about mental health disorder prevalence is well processed by the provider, we do not need to spend much effort on cleaning it. The pre-processing steps are listed as below:

Singapore	SGP	1990	0.2636717	0.7170189	0.373698	3.5981277	0.8708784	3.9267323	0.4783849	8.8792487
Singapore	SGP	1991	0.2639003	0.7178909	0.3805031	3.6019584	0.8716734	3.9562397	0.4797154	8.9204925
Singapore	SGP	1992	0.264229	0.7188251	0.3865267	3.6067072	0.8728936	3.9819791	0.4809943	8.9582671
Singapore	SGP	1993	0.2645916	0.7196904	0.393371	3.6115908	0.8738066	4.0001741	0.4821619	8.9894178
Singapore	SGP	1994	0.2649724	0.7205214	0.3999145	3.6165901	0.8742238	4.0080244	0.4831761	9.0100228
Singapore	SGP	1995	0.2653472	0.721299	0.4069095	3.6214482	0.8752532	4.000114	0.4839777	9.0151179
Singapore	SGP	1996	0.2660152	0.7224551	0.4133402	3.6278979	0.8766296	3.9820016	0.4851665	9.0117101
Singapore	SGP	1997	0.2671375	0.7241899	0.4203563	3.6371525	0.8802125	3.9651944	0.4870134	9.0140307
Singapore	SGP	1998	0.268427	0.7261035	0.4277679	3.6475799	0.8822976	3.9453632	0.4889322	9.0152415
Singapore	SGP	1999	0.2695856	0.7277608	0.4348987	3.6569503	0.885081	3.921154	0.4904383	9.0103494
Singapore	SGP	2000	0.2703364	0.7287551	0.440829	3.6633357	0.8860097	3.8949249	0.4910847	8.9981812
Singapore	SGP	2001	0.2708232	0.729245	0.445971	3.6672969	0.8876013	3.8366383	0.4888176	8.9499744
Singapore	SGP	2002	0.2713442	0.7296727	0.451158	3.6712358	0.8881229	3.7427167	0.4832476	8.8661274
Singapore	SGP	2003	0.2718439	0.7300367	0.4567067	3.6756722	0.8886552	3.6372797	0.4764572	8.7715393
Singapore	SGP	2004	0.2722839	0.7303653	0.4622397	3.6803796	0.8900165	3.5479104	0.4707824	8.6931789
Singapore	SGP	2005	0.2725748	0.7305227	0.4685406	3.6848841	0.8906006	3.4946217	0.4683777	8.6511439
Singapore	SGP	2006	0.2726878	0.730728	0.4749166	3.6901039	0.8920233	3.472043	0.4686989	8.6404793
Singapore	SGP	2007	0.2726599	0.7310867	0.4831692	3.696331	0.8957455	3.4509384	0.4694999	8.6341853
Singapore	SGP	2008	0.2725695	0.7314846	0.491488	3.7033385	0.8996242	3.4348514	0.4703687	8.633732
Singapore	SGP	2009	0.2725132	0.7318304	0.5001035	3.7100131	0.9035143	3.4227507	0.4711898	8.6372109
Singapore	SGP	2010	0.2725708	0.7319918	0.5078859	3.7149777	0.9059223	3.4178248	0.4717833	8.645251
Singapore	SGP	2011	0.2727116	0.7319693	0.5147904	3.7189096	0.9072197	3.4171601	0.4726366	8.655541
Singapore	SGP	2012	0.2728597	0.7318654	0.521609	3.7226305	0.9087179	3.4175034	0.4743613	8.6664679
Singapore	SGP	2013	0.2730132	0.7316932	0.5286816	3.7259056	0.9098704	3.4191858	0.4766697	8.6784793
Singapore	SGP	2014	0.2732077	0.7315064	0.5358159	3.7284902	0.9108076	3.4214806	0.4796079	8.6905007
Singapore	SGP	2015	0.2734453	0.7312933	0.5423945	3.729619	0.9121666	3.4257204	0.4831588	8.7024726
Singapore	SGP	2016	0.2737376	0.7310098	0.5476877	3.7300905	0.9131575	3.4299597	0.4871476	8.7124852
Singapore	SGP	2017	0.2740899	0.7305922	0.5537116	3.7304671	0.9146695	3.4382711	0.4913691	8.7271319

Figure 1: Dataset "Prevalence of mental health disorders by disorder type" after pre-processing.

1. Remove irrelevant entities, such as Central Asia and High-income countries.

2. Aggregate all types of mental health disorders to get a new attribute: mental health disorder as a whole.

The pre-processed data set is shown in Figure 1, which totally contains 195 countries with year ranging from 1990 to 2017. The data type and attribute type of it is the same as that in Table 2. When dealing with the other datasets, which are not processed by the provider as well as the preceding one, we encountered two major problems. The first one is that some datasets record information of different years. Take the one shown in Figure 2 as an example, this dataset is about the existence of stand-alone law or policy for mental health problem where records for different countries are not aligned with each other in year: rows of Armenia contain information about 2014 and 2016 while those of Australia contain information about 2014 and 2015. To solve this problem, we extract the latest data record as the indicator. For the records shown in Figure 2, we take record about 2016 for Armenia, record about 2015 for Australia, and so on. Another issue is that the set of countries in one dataset might not be aligned with each other, that is, different datasets have different numbers of countries, which makes it hard for us to aggregate different datasets together. For instance, the dataset about human resources for mental health disorders contains records of 154 countries while the dataset about government expenditures only contains information of 78 countries. To address this issue, we apply outer join to combine these datasets into a single dataset and ignore the NaN elements produced during the joining process. Part of the aggregated dataset is displayed in Figure 3.

Country	Year	Stand-alone	Year the law	Stand-alone	Publication year of the policy or plan (latest revision)			
Afghanistan	2016	Yes	1987	Yes	2016			
Afghanistan	2014	Yes		Yes				
Albania	2016	Yes	2012	Yes	2013			
Albania	2014	Yes		Yes				
Algeria	2016	Yes	1905	No				
Algeria	2014	Yes		Yes				
Angola	2016	No		No				
Antigua and	2016	Yes	1957	Yes	2013			
Argentina	2016	Yes	2013	Yes	2016			
Argentina	2014	Yes		Yes				
Armenia	2016	Yes	2004	Yes	2014			
Armenia	2014	Yes		Yes				
Australia	2015	Yes		Yes	2009			
Australia	2014	Yes		Yes				

Figure 2: Dataset about the existence of stand-alone law or policy for each country. This dataset stores information of different years.

Country	Mental hosp	Mental healt	Mental healt	Mental healt	Community r	Latest stand	Latest stand	Psychiatrists	Nurses work	Social worke	Psychologist	Extent of im	Beds in men	Beds in com	Beds for me	Total bed	Government
Afghanistan	0.003	0.012	0.006	0	0	Yes	Yes	0.231	0.098	0	0.296	Larger scale	0.296	0	0.296	0.592	
Albania	0.068	0.068	0.41	0	0.445	Yes	Yes	1.471	6.876	1.06	1.231	None	16.762	4.584	4.276	25.622	
Algeria	0.048	0.068	0.048	0	0	Yes	No					Larger scale	10.456	0	1.507	11.963	
Angola	0.011	0	0	0	0.014	No	No	0.057	0.66	0.022	0.179		0.585	0.337	0.187	1.109	
Antigua and	1.001	0	0	0	0	Yes	Yes	1.001	7.005	4.003	0		136.105	0	0	136.105	8
Argentina	0.937	1.071	1.72	0	0.152	Yes	Yes	21.705	0	0	222.572		40.566	0	2.718	43.284	
Armenia	0	0.069	1.371	0.034	0	Yes	Yes	3.84	11.245	0.274	0.788	Limited	0	0	1.714	1.714	19.5
Australia	0.071	0.601	0	1.105	0.752	Yes	Yes	13.525	90.582	0	103.036	Larger scale	7.214	10.383	21.761	39.358	7.8
Austria	0.092	0.265	3.457	2.616	4.448	No	No					Larger scale	26.087	59.664	19.485	105.236	
Azerbaijan	0.114	0.021	0.156	0.094	0.031	Yes	Yes	3.452	6.717	0.114	1.165	Larger scale	31.193	2.911	0.416	34.52	3.5
Bahamas	0.259	0.259	0.517	0.259	0	Yes	Yes						97.457	0	3.102	100.559	
Bahrain	0.146	0	1.239	0.729	0.146	No	Yes	5.467	27.918	1.458	1.239	Larger scale	19.317	1.093	0	20.41	3.2
Bangladesh	0.001	0.035	0.055	0	0.045	No	No	0.13	0.873	0	0.124	None	0.434	2.236	0.315	2.985	0.5

Figure 3: Part of the aggregated dataset.

6 Result of each component and its visual encoding

6.1 Banner

At the beginning of the visualization, we "hang" a banner showing how serious the mental health problem we are facing, as shown in Figure 4. This chart displays three figures where the center one shows the number of people suffering from mental health disorders in 2017, the left one shows the share of global population with the disorder in 2017, and the right one shows the rate of increase from 2007 to 2017 [10].

Figure 4: Banner showing the seriousness of mental health issues around the world.

6.2 Map

There are two maps in the this section. The first map reflects the distribution of mental health disorders in the world, as shown in Figure5. There are two dimensions: country and mental disorder type, and we used two filters to screen them. For example, when year 2000 is selected, the distribution of

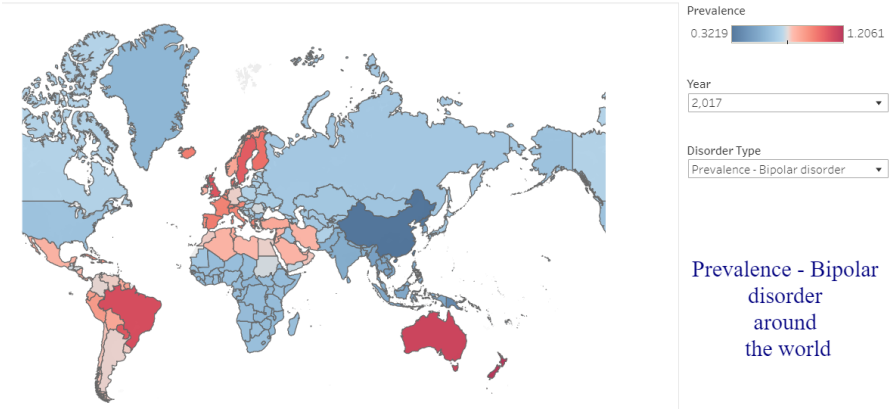

Figure 5: Map showing the distribution of mental health disorders in the world.

the proportion of mental illness patients to their own country in 2000 will be displayed. When a cause of mental health is selected, like anxiety disorders, the proportion of mental illness patients in different countries caused by that cause will emerge. The depth of the color on the map represents the number of patients. In other words, the darker the color, the more the number of patients in this country, and vice versa.

The second one is a combination of pie chart and geographical map, as shown in the Figure6. The pie chart represents the proportion of different causes of disorders in a country when the date is determined. And the situation of each country is represented by a pie chart. As we can see, the higher the proportion of diseases, the relatively larger the pie chart is.

6.3 Correlation Map

This chart is used to explore the correlations between different mental health disorders, as shown in Figure 7. We take advantage of this chart to answer the fourth query. We calculated the Pearson's correlation coefficient between two disorders as an indicator of the linear relationship between them, and the computation of this correlation coefficient will be discussed in Chapter 7.

In this chart, the marks for different correlation coefficient value are areas, and the channels for different marks are horizontal as well as vertical position to convey identity information, and color to encode the magnitude information of the correlation. Note that here we use blue color to represent negative correlation coefficients and red color to represent positive ones, while using the saturation of color to encode the absolute value of coefficients.

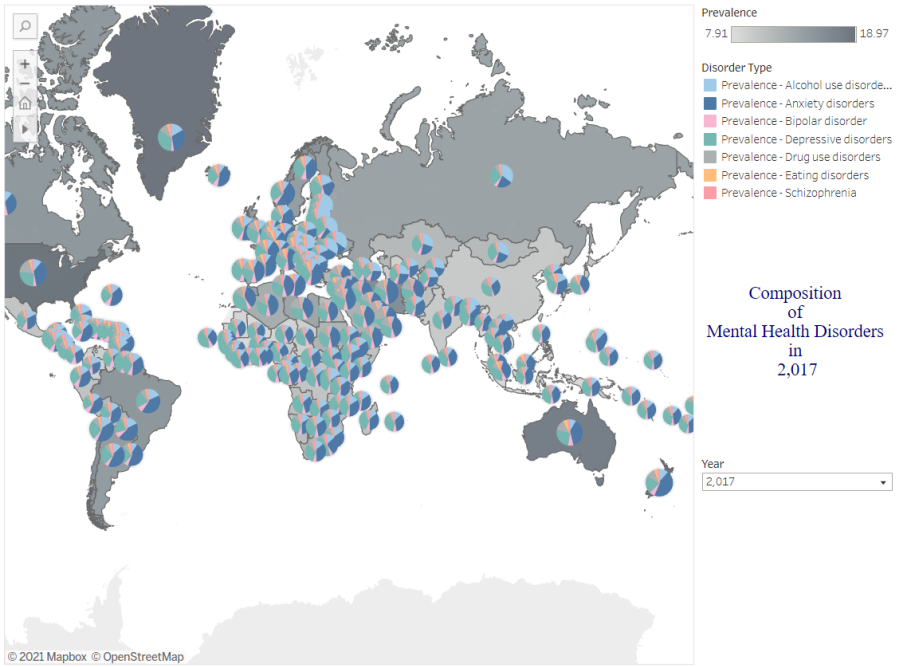

Figure 6: The pie chart map showing the proportion of different causes of disorders in a certain country.

In this chart, users have three ways to interact with the component. Users could use the country selector to select the country of interest to further analyze. Also, users could use the year filter to filter out uninterested year range and focus on a certain year range. Besides, users could click on the cell in the chart, which will trigger the visualization of the correlation in the line chart below, which will be discussed in the next section.

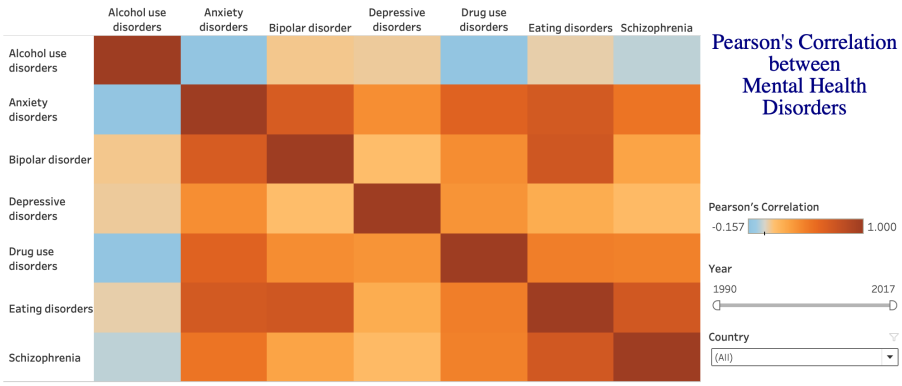

Figure 7: Correlation map for different mental health disorders.

6.4 Scatter Plot to Visualize Correlation

When users click a cell in the Correlation Charts, the line chart below it will visualize the correlation in the plane, which is shown in Figure 8. It provides us an intuitive understanding of how these two mental health problems interact with each other.

In this chart, the marks are points where each point encodes a record for one year. The channels for different marks are horizontal positions to encode the magnitude information of the prevalence of one mental health disorder and vertical positions to encode the magnitude information of the prevalence of the other disorder. Besides, we also calculate a linear function to approximate the relationship between the two selected disorders, which is displayed as a line in the chart.

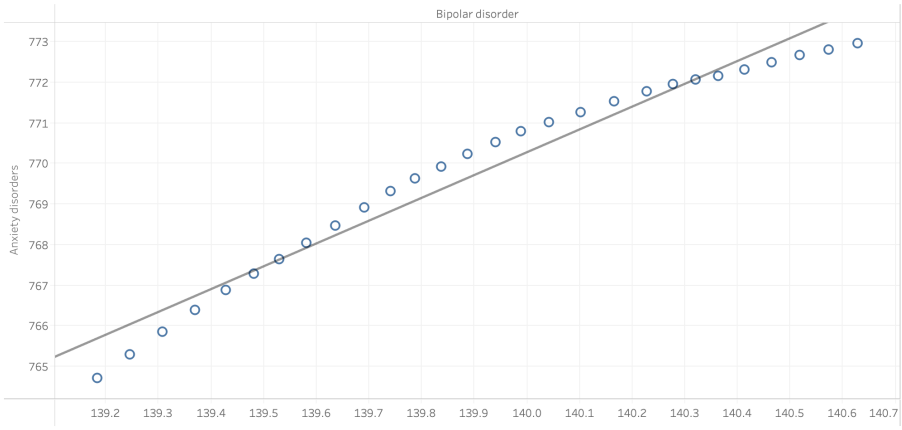

Figure 8: Scatter plot visualizing the correlation relationship between anxiety disorder and bipolar disorder.

6.5 Word Cloud

In this visualization, we use word cloud to reflect the number of cases with different mental disorder types in each country per 100,000 population, as shown in Figure 9.

The color and the size of each country name represent the number of cases. The filters on the left can be used to select the mental disorder type and view country information under different ranges of number of cases.

6.6 Scatter Chart with Clusters

This chart is related to the fifth query above. As in Figure 10, it shows the severity of mental disorders in each countries and its economic and public service factors.

The x-axis stands for the number of cases with different types of mental disorder, and the y-axis stands for the economic and public service factors

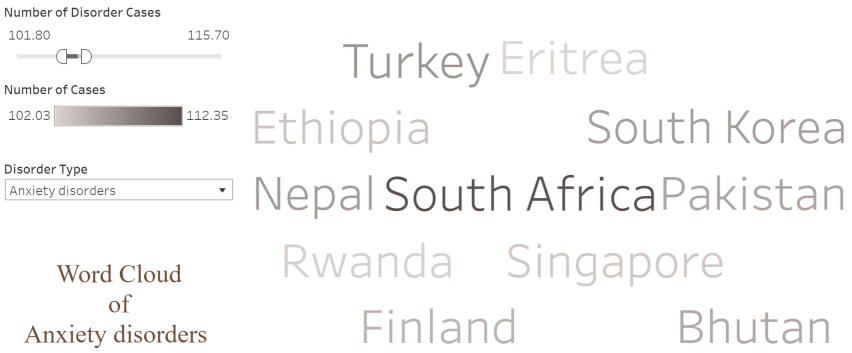

Figure 9: word cloud for different mental disorder types

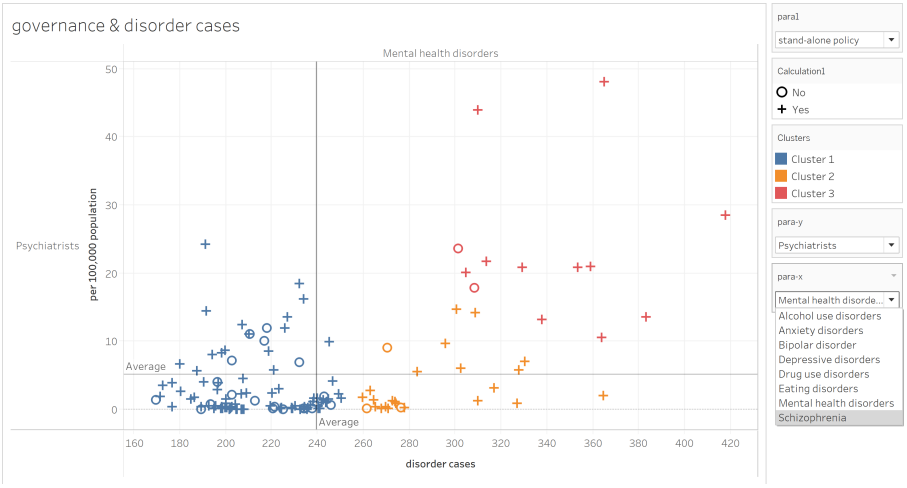

Figure 10: mental health with governance factors

including the government expenditures on mental health as a percentage of total government expenditures on health, and the human resources and facilities data of each country. The parameters on the right can be used to select the attributes represented by x and y axes.

Each scatter point in the chart represents a country and the shape of the point indicates whether this country has stand-alone laws or policies for mental health (cross means true and circle means false). The points are automatically clustered into three groups according to their coordinates, and showed in

different colors.

7 Step-wise Visualization Process

7.1 Creation of map

For the first map, we made two parameters to select the date dimension and the mental disorder types. When a year and a mental disorder type is selected, the map will show the distribution of that mental illness in different countries around the world during that year. The parameters of year and mental disorder type are shown in Figure 11 and the calculation fields are shown in Figure 12. The calculation on year attribute is a binary data, and we filter it to show only values with "true" calculation result on the map.

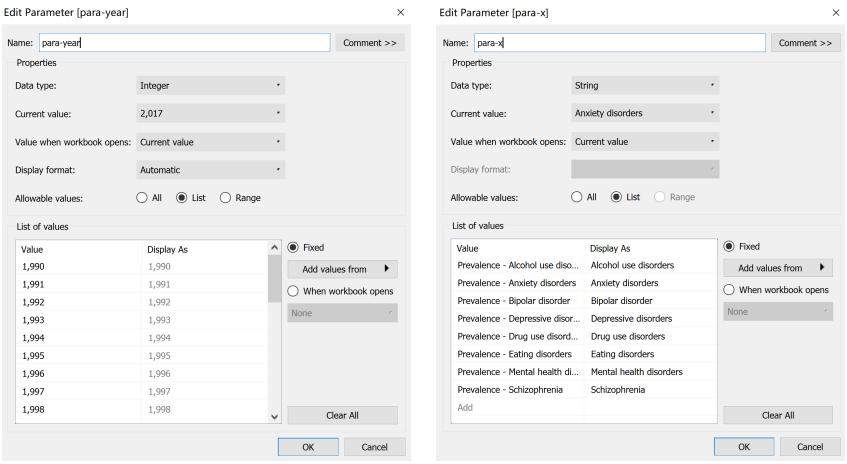

Figure 11: Parameters

```
CASE [para1]
WHEN "Alcohol used disorder" THEN AVG([Prevalence - Alcohol use disorders])
WHEN "Anxiety" THEN AVG([Prevalence - Anxiety disorders])
WHEN "Bipolar" THEN AVG([Prevalence - Bipolar disorder])
WHEN "Depressive" THEN AVG([Prevalence - Depressive disorders])
WHEN "Drug" THEN AVG([Prevalence - Drug use disorders])
WHEN "Eating" THEN AVG([Prevalence - Eating disorders])
WHEN "Mental" THEN AVG([Prevalence - Mental health disorders])
WHEN "Schizophrenia" THEN AVG([Prevalence - Schizophrenia])
WHEN "sum" THEN SUM([Prevalence - Alcohol use disorders]+[Prevalence - Anxiety disorders]
+[Prevalence - Bipolar disorder]+[Prevalence - Depressive disorders]+[Prevalence - Drug use disorders]
+[Prevalence - Eating disorders]+[Prevalence - Mental health disorders]+[Prevalence - Schizophrenia])
ELSE 0
END
```

SUM(**expression**)

```
([Year] == [para-year])
```

Figure 12: The calculation fields

7.2 Creation of correlation charts

The computation of the Pearson's Correlation Coefficient r is shown in Equation1.

$$r = \frac{\sum(x_i - \hat{x})(y_i - \hat{y})}{\sqrt{\sum(x_i - \hat{x})^2 \sum(y_i - \hat{y})^2}} \tag{1}$$

where x_i and y_i are the values of variable x and y, respectively, and $\hat{x}$ and $\hat{y}$ are the mean values of variable x and y, respectively. In tableau, there is a built-in function provided for us to calculate this correlation coefficient between different columns, as shown in Figure 13.

Figure 13: The calculation of the Pearson's correlation coefficient in Tableau.

7.3 Creation of word cloud

For the word cloud, we use the same parameters and calculation fields as shown in Figure 11 and 12. Meanwhile, we use a customized filter on the range of number of cases, as Figure 14 shows.

7.4 Creation of clustered scatter chart

For this visualization, we also use the parameters and calculation fields mentioned in section 7.1, and add a new parameter for y-axis, which is shown in Figure 15.

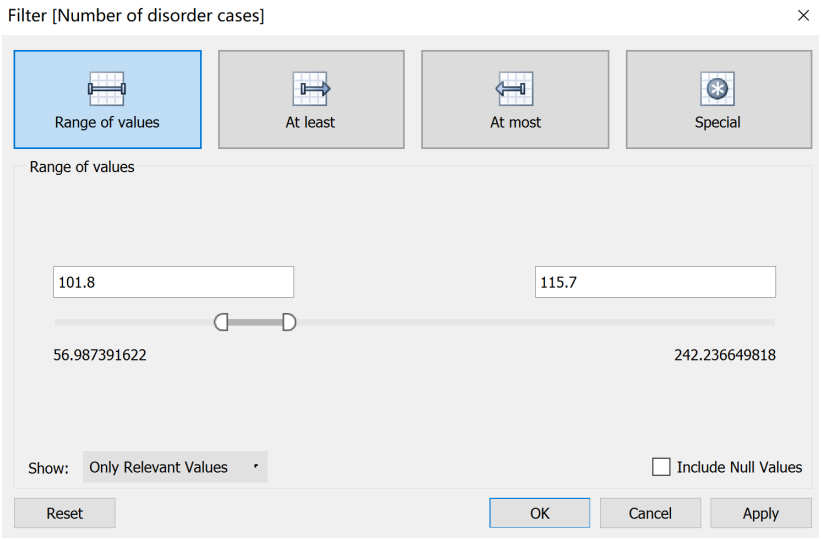

Figure 14: Filter

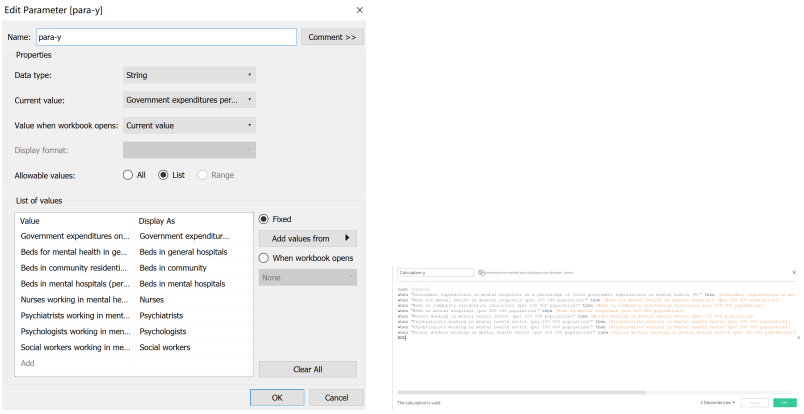

Figure 15: y-axis parameter

Besides, we use cluster function in Tableau to automatically cluster the points into three groups according to their coordinates.

8 Charts Justification

8.1 Map

Why maps? When we got the data, we found that these countries have given complete map coding, and the data of almost all countries are perfect. Therefore, we want to know the situation of people suffering from mental illness in different countries.

Obviously, if we take the country as one of the coordinate axes, bar charts, pie charts, broken line charts and so on are obviously not suitable. After all, there are more than 200 countries. If they are all displayed on the coordinate axes, the pictures will become dense, and we can't simply and directly understand the information. This is related to human beings can only grasp 3-5 focus information at a time. In addition, most people can only remember countries like the United States, China and Russia. Unless they are geographers, not so many people can remember where more than 200 countries are. Therefore, the map gives the audience an intuitive dimension to understand the location of different countries.

So, we use two filter to screen our dimensions, and the lightness of color in map to present the prevalence among different countries in the dimension of time and causes.

Bibliography

[1] What is mental health? https://www.medicalnewstoday.com/articles/154543.

[2] The world health report 2001 - mental health: New understanding, new hope. https://www.who.int/whr/2001/en/.

[3] Jean M Twenge, A Bell Cooper, Thomas E Joiner, Mary E Duffy, and Sarah G Binau. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of abnormal psychology*, 128(3):185, 2019.

[4] Prevalence by mental and substance use disorder. https://ourworldindata.org/grapher/prevalence-by-mental-and-substance-use-disorder?tab=table&country= OWID_WRL.

[5] Mental health governance data by country. https://apps.who.int/gho/data/node.main.MHPOLFIN?lang=en.

[6] Human resources data by country. https://apps.who.int/gho/data/node.main.MHHR?lang=en.

[7] Facilities data by country. https://apps.who.int/gho/data/node.main.MHFAC?lang=en.

[8] Beds data by country. https://apps.who.int/gho/data/node.main.MHBEDS?lang=en.

[9] https://www.who.int/data/gho/data/indicators/indicator-details/GHO/government-expenditures-on-mental-health-as-a-percentage-of-total-government-expenditures-on-health-(-).

[10] Hannah Ritchie and Max Roser. Mental health. *Our World in Data*, 2018. https://ourworldindata.org/mental-health.

Appendix

Complete view of the visualization

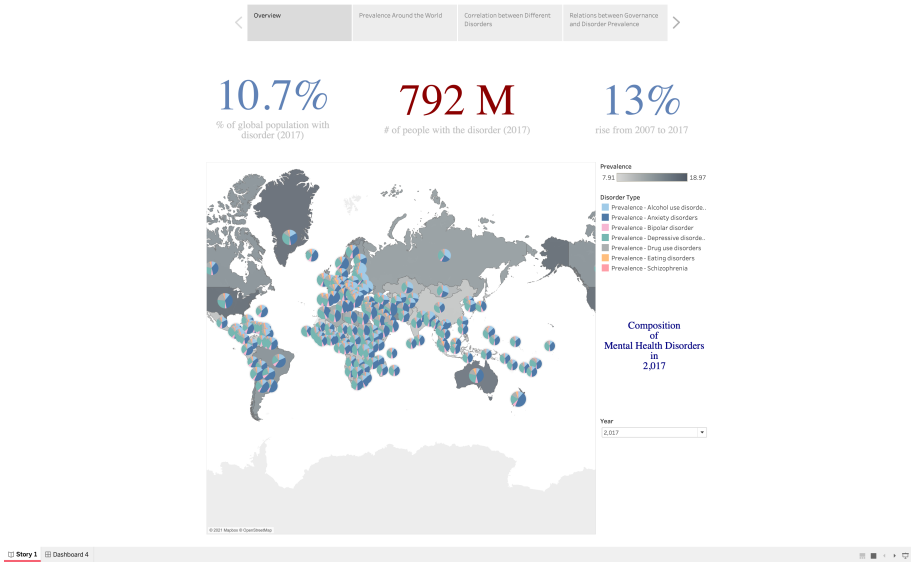

Figure 16: Story 1.

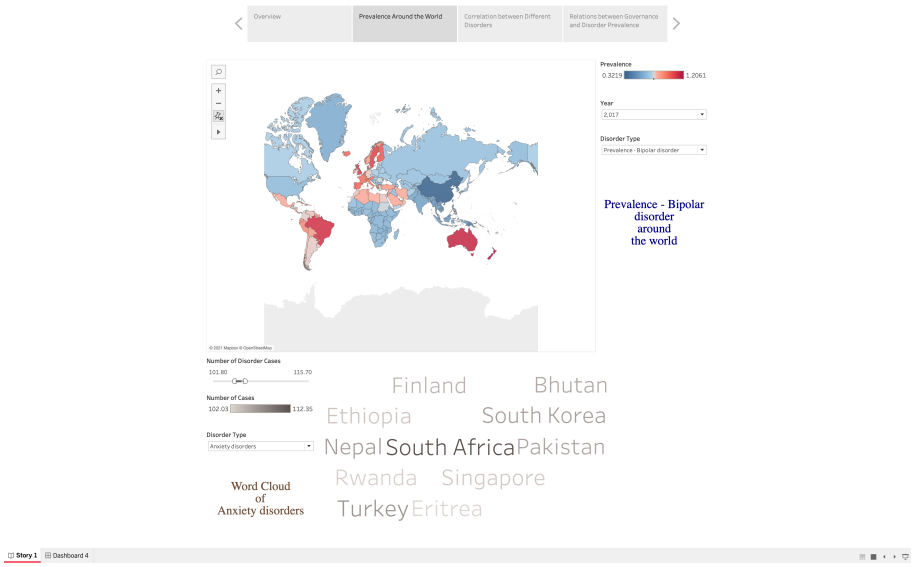

Figure 17: Story 2.

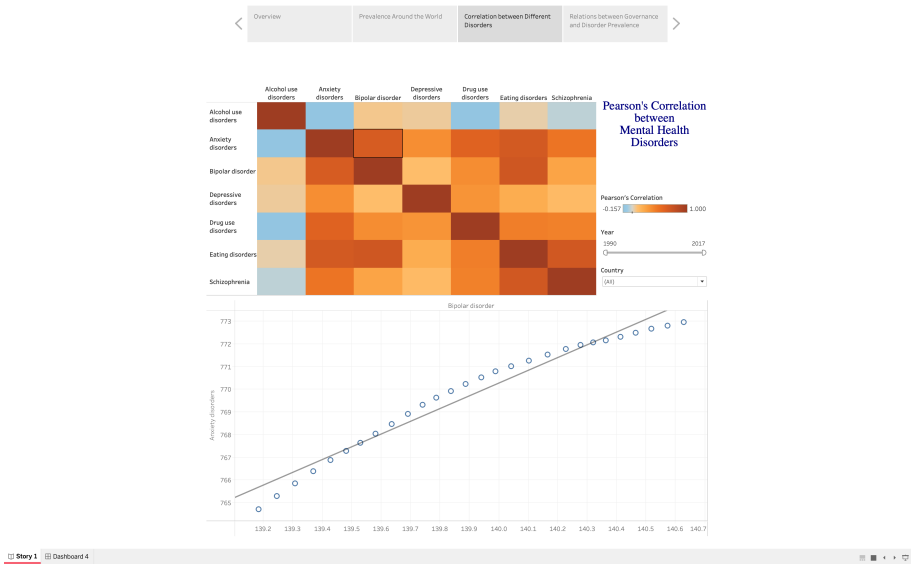

Figure 18: Story 3.

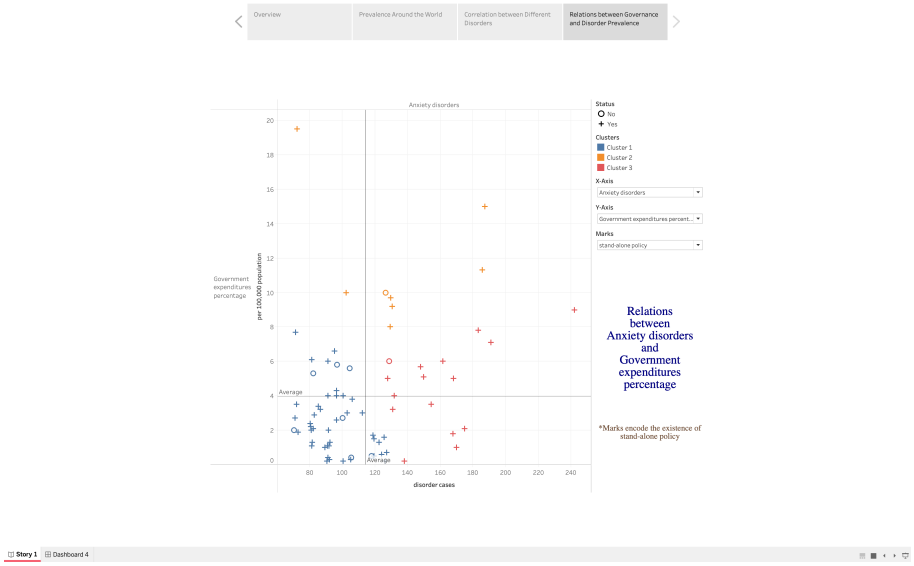

Figure 19: Story 4.