Weather and its Effects on Mental Illness

Ms. Sinatra

MDM4UE

Apr. 9 2021

Kristine Oh & Shirley Lin

Table of Contents

Introduction

Hypothesis: The weather of a country has an impact on the mental health of its citizens. Specifically, as the sunlight duration increases, the number of mental illness cases decreases; as the rainfall increases, the number of mental illness cases increases.

We compared 3 different sets of data, worldwide, in the UK, and in the US. The worldwide data was comparing different countries, the UK data compared different months, and the US data compared both different months and different states.

The worldwide data we used was the prevalence of mental health issues, in percent of the population, in each country in 2014[1]. We then sorted these into 4 different climates, based on the country's mean latitude. Countries within 15° of the equator were considered Tropical, 15°-30° were Dry, 30°-45° were Temperate and above 45° were Continental. This allowed us to plot comparative box plots for each climate type. We then plotted graphs of the prevalence of mental disorders against both the countries' average daily sunlight[2,3] and average yearly precipitation[4], and found the correlation between them.

The UK data was the number of people in contact with mental health services across 1-month periods, from January 2018 to February 2020[5,6]. This data included the number of adults, children, those with autism, new referrals, and those subjected to the Mental Health Act (without agreement, suggests urgent cases), in addition to the total numbers. This data was then plotted against the average daily sunlight hours[7] and total monthly rainfall[8] for the same time period. In particular, the graphs of adult contacts against sunlight and rainfall stood out.

The US data was the percentage of people reporting signs of either depression or anxiety, surveyed on a weekly basis from April 2020 to March 2021[9]. The data for each US state was then plotted against the average annual sunlight[10] and rainfall[11] for each state.

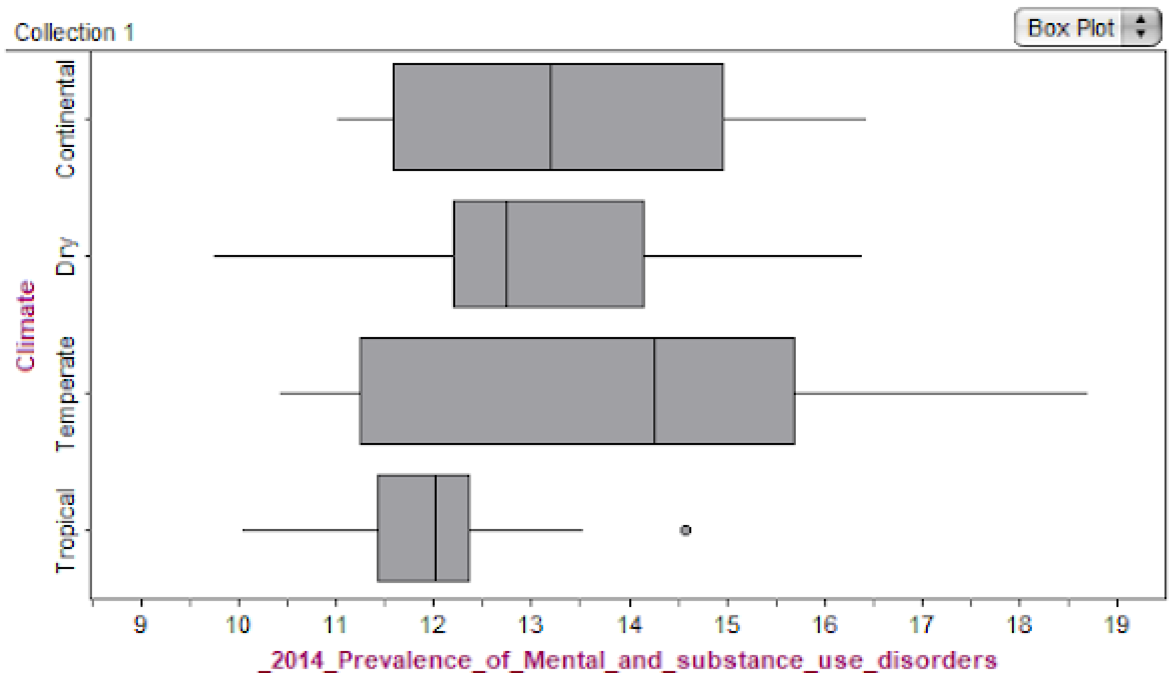
---

[1] Global mental health: five key insights which emerge from the data
[2] Wikipedia's List of cities by sunshine duration
[3] If the data table includes multiple cities of a country, we take their average as a country's yearly sunshine duration.
[4] Average annual precipitation
[5] Mental Health Services Monthly Statistics
[6] This data was collected from the Key Facts section of the report
[7] UK daylight hours by month 2020
[8] Average rainfall UK 2020
[9] Mental Health - Household Pulse Survey - COVID-19
[10] Sunniest States 2021
[11] US States Ranked by Average Precipitation

The depression/anxiety data was then further sorted by month for each state. This was done by taking the average of the weekly percentages for each month[12]. The state monthly data was then plotted against average monthly sunlight and rainfall[13] for the state.
Lastly, an average of the monthly data for all states was plotted, taking the averages of the depression/anxiety data, the sunlight data and the rainfall data. In this way, we could compare both differences in weather in a single location, as well as different locations, using the same symptom data.

[12] If the week crosses two months, it counts towards the earlier month. For example, May 28 - June 2 counts towards the percentage for May
[13] Climate United States - Normals and averages

One-Variable Analysis



	Climate				Row Summary
	Continental	Dry	Temperate	Tropical	
S1 = mean ( )	13.364286	12.94	13.653056	11.958605	12.905556
S2 = count ( )	28	28	36	43	135
S3 = min ( )	11	9.74	10.42	10.02	9.74
S4 = Q1 ( )	11.58	12.125	11.22	11.42	11.43
S5 = median ( )	13.195	12.74	13.925	12.03	12.3
S6 = Q3 ( )	14.985	13.895	15.645	12.39	14.37
S7 = max ( )	16.43	16.38	18.69	14.58	18.69
S8 = stdDev ( )	1.8826644	1.5176444	2.5726005	0.86096176	1.9029802
S9 = iqr ( )	3.405	1.77	4.425	0.97	2.94

Tropical countries have the lowest IQR and median, showing a small spread and low prevalence. There is an outlier present, Guyana, in South America. Guyana's high prevalence is likely from other factors that have a stronger effect on mental health than weather, such as a high percentage of people living in poverty, combined with a lack of healthcare professionals. On top of that, there is a very limited access to treatment facilities, and seeking help is stigmatised and one may even be persecuted.[14]

Tropical climates often have high amounts of sun and high temperatures. If they are in a monsoon area, they have a wet and a dry season, otherwise, they have moderately high rainfall year-round.

[14] 7 Facts About Mental Health in Guyana

This supports our hypothesis, as Tropical countries have both a high amount of sunlight and a low prevalence of mental disorders, though Tropical climates also have high rainfall.

Conversely, Temperate countries have the highest IQR and median, and thus the largest spread and prevalence rates. The data is skewed to the left, with no outliers. The highest prevalence rates are found in temperate climates, with a maximum of 18.69%, in New Zealand.
Temperate countries have 4 seasons, with moderate rainfall and mild summers and winters. Due to the seasons, they also have a moderately high-temperature range. They have less annual sunlight than Dry and Tropical countries, but higher than Continental, due to their latitude. Additionally, they have moderate rainfall, higher than Continental and Dry, but lower than Tropical.
They can be thought of as average in terms of weather, which might explain why the spread is so wide; since their weather is mild, other factors, such as government support, may have a larger role in mental wellbeing than the weather.

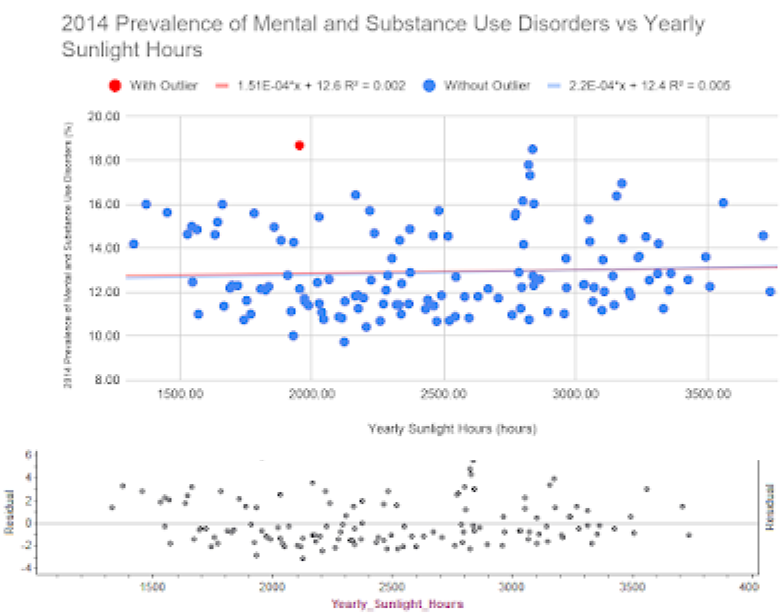
Dry countries have the second-lowest IQR and median, after Tropical. The data is skewed to the right, with no outliers. The country with the lowest prevalence of mental disorders has a Dry climate, being Vietnam.
Dry climates have high sunlight and very low rainfall. The temperature varies greatly, being very hot in the day, and very cold at night. They do not experience seasons to the extent of Temperate and Continental climates, but there are noticeable temperature changes.
Once again, the high sunlight together with the low prevalence of mental disorders supports our hypothesis. However, the fact that rainfall is lower than Tropical yet the median prevalence of mental disorders is higher, suggests that the relationship between rainfall and mental disorders may be opposite that of our hypothesis.
The low IQR of both Dry and Tropical climates suggests that high amounts of sunlight, a trait shared by both climates, affect mental health more than other factors, thus leading to smaller spreads. This is the opposite of Temperate and Continental climates, which have high IQRs, which we speculate is due to other factors being stronger than weather.

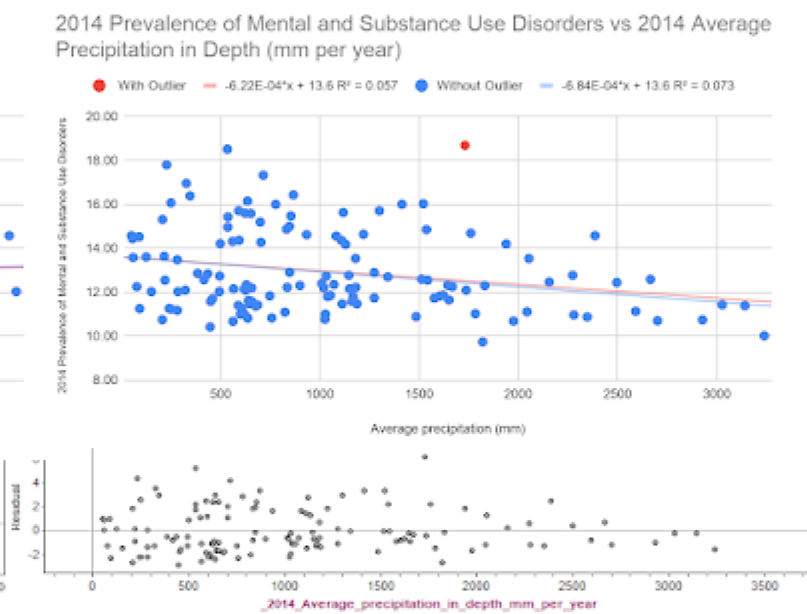
Continental countries have the second highest IQR and median, after Temperate. The prevalence data is fairly symmetrical, with a slight skew to the right and no outliers. While they have a higher IQR, their range is smaller than that of Dry countries.
Continental climates are very similar to Temperate climates, though with harsher winters and warmer summers, leading to a larger temperature range. Additionally, they have the lowest annual sunlight, due to being at higher latitudes.
As these countries have harsher weather than Temperate climates, it is possible that the weather could have more effect on mental health. The low rainfall with lower prevalence rates than

Temperate, as well as the low sunlight with higher prevalence rates than Tropical and Dry, support our hypothesis.

Two-Variable Analysis

World Data:





These graphs show the percentage prevalence of mental and substance use disorders against both yearly average sunlight hours, and yearly average precipitation. Each point is a different country.

Sunlight:
The slope is positive, indicating as the yearly sun hours increases, the prevalence increases, opposing our hypothesis. The model predicts that for every extra hour of yearly sun, the prevalence increases by 0.000151%, as shown by the slope value of 0.000151% per yearly sun hour. The model also predicts that, for a country with 0 hours of annual sunlight, there would be a 12.6% prevalence rate, though there is no such country.
The R^2 is 0.5%, with the r at 0.07, which implies little association between prevalence of mental illness and yearly sunlight hours.
The residual plot has no clear pattern, showing that the linear model is appropriate.

Rainfall:
The negative slope indicates that as the average annual precipitation increases, the prevalence of mental health and substance use disorders decreases, which goes against our hypothesis. The value of the slope is -0.000622% per mm of yearly average precipitation, showing the model

predicts that the prevalence of mental disorder decreases by 0.000622% with every 1mm increase in precipitation.
The R^2 is 5.7%, which shows that 5.7% of the variation in prevalence can be accounted for by the variation in precipitation using this model. Although precipitation has a stronger association with prevalence than sunlight hours, the effect is still small.
The residual plot is scattered, which shows that the linear regression is appropriate.

Outlier:
New Zealand can be considered as a high residual point in both graphs. It is slightly influential. Looking at prevalence vs sunlight hours, the R^2 increases from 0.2% to 0.5% when removing New Zealand. As for prevalence vs precipitation, the R^2 increases from 5.7% to 7.3% after it is removed. The absolute value of the slope also increases from -0.000622 to -0.000684, indicating that the model predicts that the prevalence of mental disorders decreases by 0.000684% with every 1mm increase in precipitation.
Despite being a developed country, New Zealand has high rates of teenage pregnancy, as well as poverty. The culture in New Zealand also relates mental issues with weakness, preventing people from seeking help and recovering.[15] These factors may play a part in New Zealand's high prevalence, thus making it an outlier. This is also supported by the country with the second highest prevalence, Australia, with a similar culture and beliefs.

Both graphs:

Correlation Coefficient:

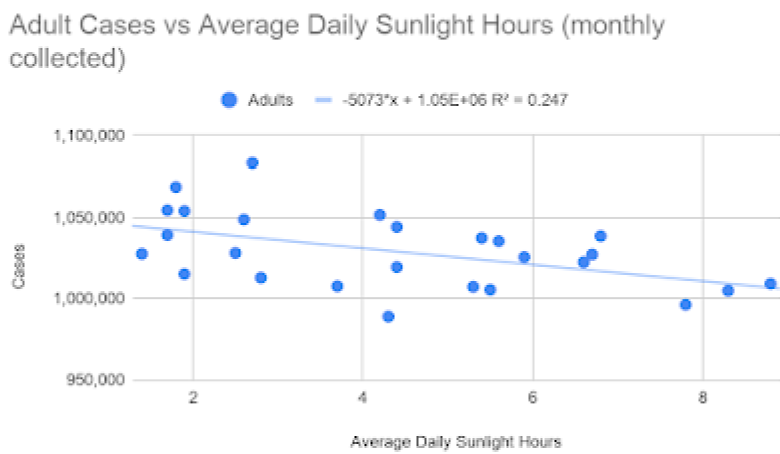
	Sunlight duration	Precipitation
With outlier	0.045	0.071
Without outlier	0.239	0.270

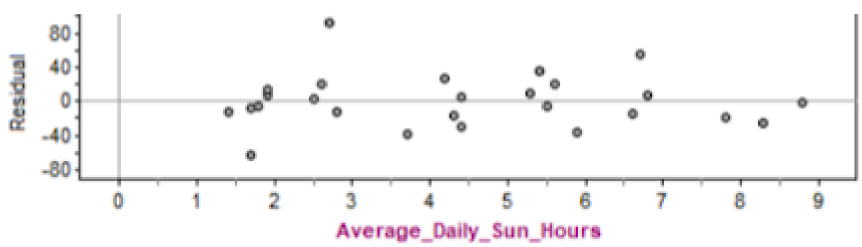
Overall, both models' correlation coefficients are rather low, which suggests that weather has little effect on the prevalence of mental health disorders between different countries.
Social stigma might be one reason that some countries have low percentages of prevalence, since reporting mental illness is voluntary, and people may be dissuaded to do so.
Weather may be a factor that increases chances of mental illness, but it is not as decisive as other factors such as a country's culture, policy, economic status or living environment.

[15] What's behind New Zealand's shocking youth suicide rate?

The worldwide data thus shows little association, and so we would like to continue our investigation by narrowing down the data to a single country. In doing so, we can control the lurking variables.

UK Data:

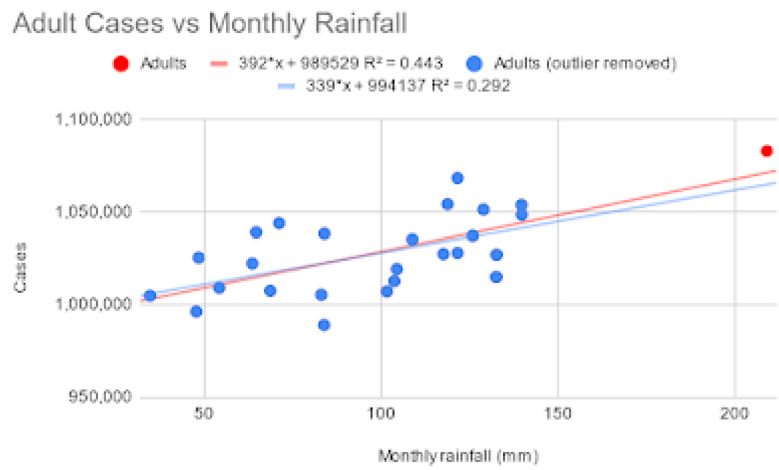


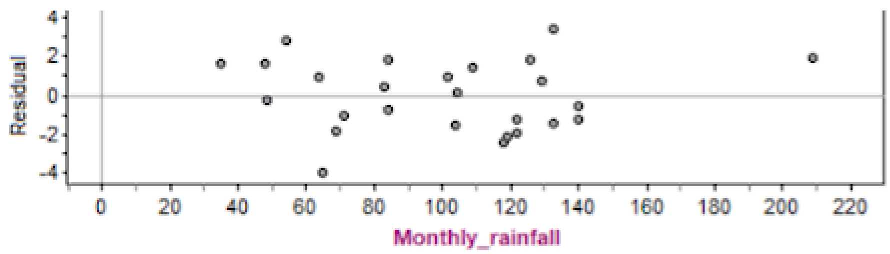


This graph shows the total number of adults in contact with UK Mental Health Services compared to average daily sunlight hours. Each point on the graph is a different month, from Jan 2018 to Feb 2020.

The slope of the graph is negative, showing that as sun hours increase, adult cases decrease, which supports our hypothesis. The slope is -5073 cases per sunlight hour, suggesting that for an average 1 hour increase in sun, the number of adults in contact with mental health services decreases by 5073. The intercept of the graph is about 1,050,000 cases, suggesting that if there were an average of 0 hours of daily sunlight, there would be 1,050,000 cases. However, it is unreasonable to expect to see a month with no sun in the UK.

The residual plot is scattered, with no clear pattern, showing that the linear model is appropriate. The R^2 is 24.7%, with the correlation being -0.497, suggesting a moderately strong association. The R^2 and r being higher than that of the worldwide graph suggests that the effect of sunlight on mental health is weaker than that of other geographical or societal factors. While the R^2 is somewhat weak, it is substantial enough to suggest that weather has an effect on mental health. The model suggests that about 24.7% of the variation in adult cases is due to differences in average daily sunlight.





This graph shows adult cases in the UK against monthly rainfall. Each point on the graph is a different month, from Jan 2018 to Feb 2020.

There is an outlier in Feb 2020, with unusually high rainfall. This point has high leverage, with a moderate residual, and is thus influential. Without that point, the R^2 decreases from 44.3% to 29.2%. This new R^2 is similar to the UK sunlight graph's, allowing us to draw the same conclusions. The slightly higher R^2 suggests that rainfall has a slightly larger effect than sunlight on mental health, similar to what the worldwide graphs show. Removing the outlier also decreases the correlation coefficient from 0.666 to 0.540, still suggesting a moderately strong association.

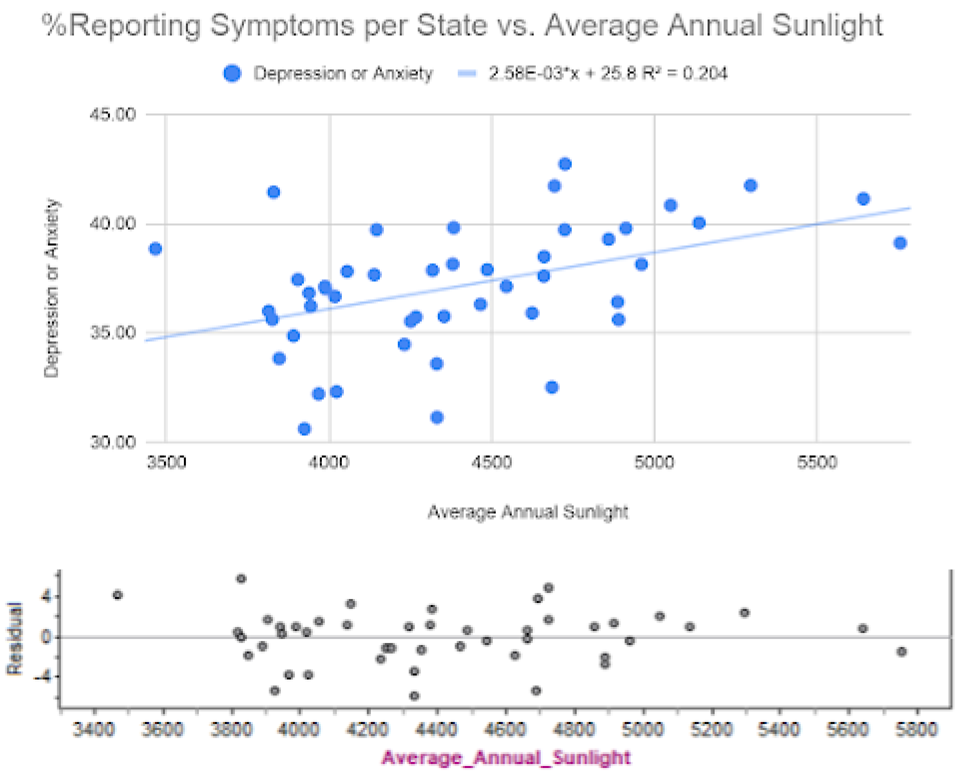
The slope of this graph is positive, showing that as monthly rainfall increases, cases increase, supporting our hypothesis. The value of the slope, 339 cases per mm, shows that this model predicts that for every 1mm increase in rainfall, cases increase by 339. The intercept of the graph is 994,137 cases, showing that the model predicts 994,137 cases in a month with no rainfall.

The residual of this graph is scattered, like the others are, showing appropriateness.

Overall, both models support our hypothesis: as the sunlight duration increases, the number of mental health disorder cases decreases; as the rainfall increases, the number of mental health disorder cases increases. Even after removing the outlier, the models suggest that the associations between mental illness and sunlight, as well as with rainfall, are relatively strong, with rainfall being slightly stronger.

Though the UK data supports our hypothesis, in order to ensure accuracy, and to investigate whether the association is universal, we would like to examine the data of another country.

US Data:

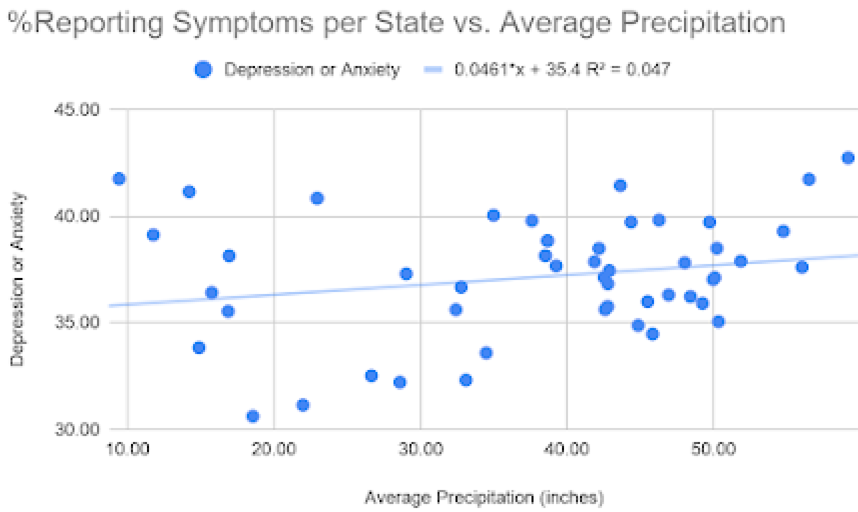


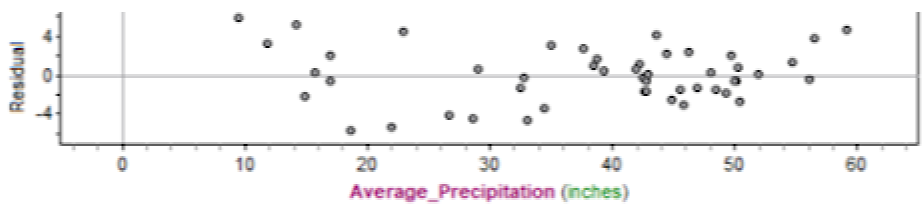
This graph shows the percentage of people in the US reporting symptoms of depression or anxiety during the April 2020 to March 2021 period, against the average annual sunlight. Each point is a different state.

The R^2 shows that 20.4% of the variation in reported depression/anxiety is accounted for by the variation in average annual sunlight hours in this model. The correlation coefficient of 0.452 shows that the association between depression/anxiety and sunlight is moderately strong.

The positive slope in the model predicts that the depression or anxiety rate increases by 0.258% for every 1 hour increase in average annual sunlight hours, which opposes our hypothesis. The intercept of 25.8% shows that the model predicts that 25.8% of the population will report depression/anxiety symptoms for a state receiving no annual sunlight, though to have a state like this is unreasonable.

The linear regression is appropriate as seen from the scattered residual plot.





This graph shows the percentage of people in the US reporting depression or anxiety symptoms during the April 2020 to March 2021, against the average precipitation in inches. Each point represents each state.

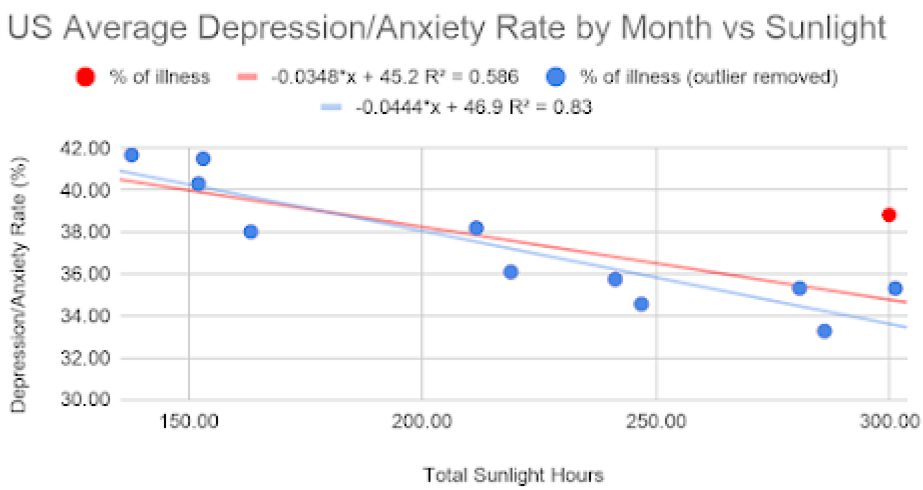
The R^2 shows that 4.7% of the variation in reported symptoms of depression/anxiety is accounted for by the variation in average annual precipitation in this model. The correlation coefficient of 0.213 shows that the association between depression/anxiety and precipitation is not as strong as the association between depression/anxiety and sunlight hours

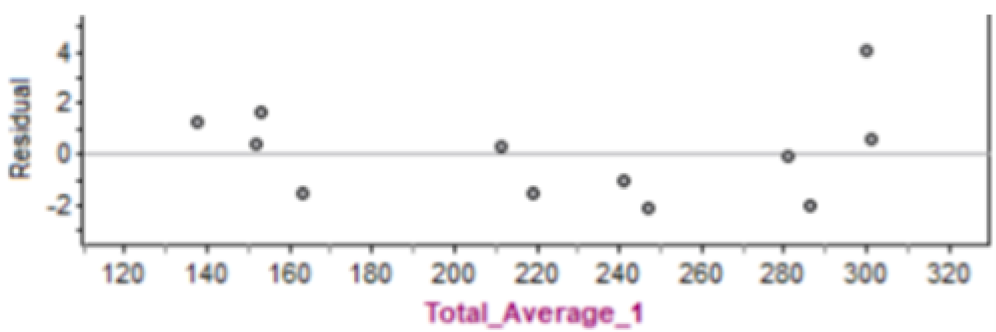
The positive slope in the model predicts that the depression or anxiety rate increases by 0.0461% for every 1 inch increase in average precipitation hours, supporting our hypothesis. The linear regression is appropriate as seen from the scattered residual plot.

Overall, the percentage of people in the US reporting symptoms vs precipitation graph does not show a strong correlation. On the other hand, the percentage of people in the US reporting symptoms vs sunlight graph shows a moderately strong association. This is unlike the worldwide and UK graphs, where the rainfall graphs showed higher correlations than sunlight. However, the model suggests that as average sunlight hours increase, the percentage of depression and anxiety increases, opposing our hypothesis. The reason might be that the states located in the northeast have a better economic status thus a lower depression/anxiety rate, but as they are located in the northern part of the US, the sunlight duration is lower.

Although there are fewer differences in living conditions in the states in the US than the countries worldwide, it is still inevitable to compare states with different policies, economic status or beliefs. As we cannot effectively control these variables, such comparison might not provide us with accurate information to examine the relationship between weather and mental health illness.

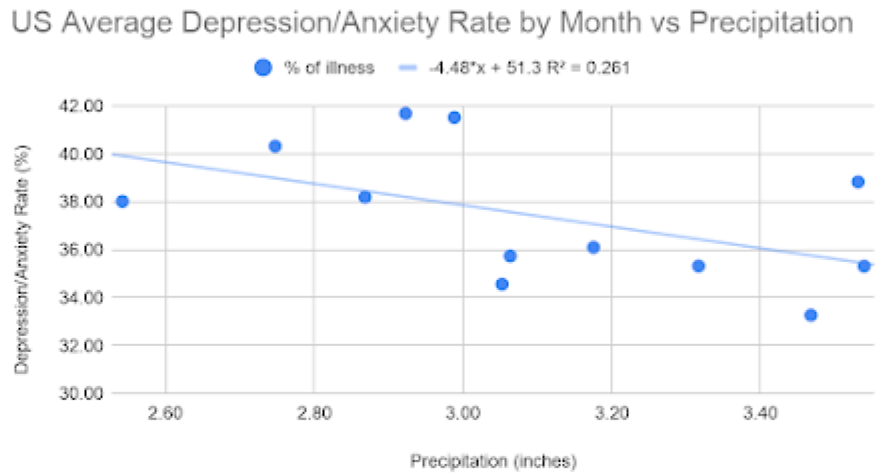
Thus, we would now like to examine the association in a single location across different months, similar to what we did with the UK data.

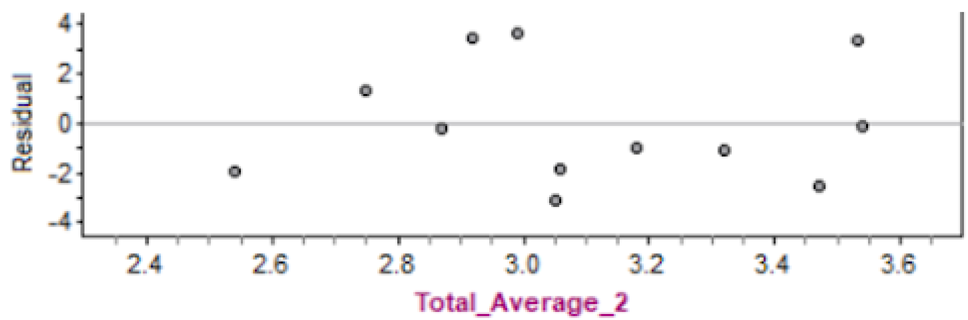




This graph shows the percentage of people in the US reporting depression or anxiety symptoms against total monthly sunlight. Each point is a different month from April 2020 to March 2021. The R^2 of 58.6% shows that this model suggests that 58.6% of the variation in depression/anxiety symptoms is accounted for by the variation of total sunlight hours. The correlation coefficient is -0.766, indicating a strong relationship between the depression/anxiety rate and sunlight hours. The negative slope of -0.0348% per sunlight hour shows that the model predicts that the depression/anxiety rate decreases by 0.0348% for every 1 hour increase of total monthly sunlight.

July 2020 may be considered a high leverage, high residual point with 38.83% of reported symptoms and 300.00 hours of monthly sunlight duration. July 2020 was supposed to be the month for summer travelling; however, due to a peak of COVID-19 cases, many people's vacation got cancelled. It might be a reason why the percentage of symptoms was higher than expected. With the outlier removed, the R^2 increases to 83% and the association strengthens with a correlation coefficient of -0.911. The sign of the slope doesn't change but the absolute value increases, thereby predicting that the depression/anxiety rate decreases by 0.0444% for every 1 hour increase of total monthly sunlight.





This graph shows the percentage of people in the US reporting depression or anxiety symptoms against total precipitation for that month. Each point is a different month from April 2020 to March 2021.

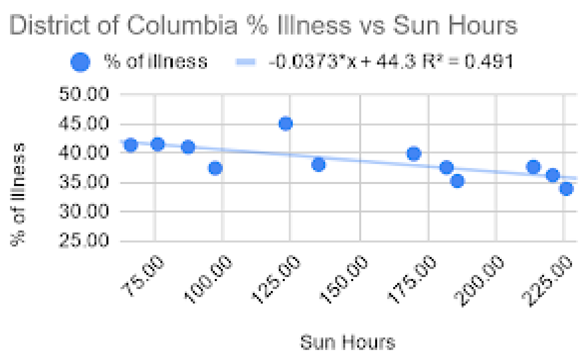
The residual plot looks somewhat like a sin graph, showing a slight curved pattern, therefore, a linear model might not be appropriate. As the graph has concavities both upwards and downwards, re-expressing data might not be useful. In addition, more data need to be collected for the graph to be representative as there are only 12 data points for now.

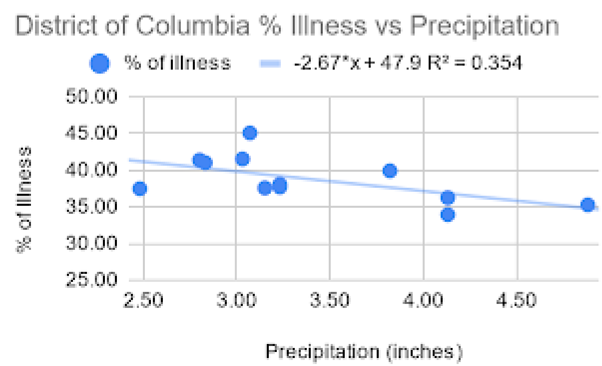
However, assuming the linear model is accurate, this graph shows that for every inch of monthly precipitation, the Depression/Anxiety rate decreases by 4.48%, as shown by the negative slope of -4.48% per inch of precipitation. This goes against our hypothesis. The R^2 and R are 26.1% and -0.511 respectively, suggesting that 26.1% of variation in Depression/Anxiety symptoms are due to differences in monthly precipitation, with a moderate correlation.

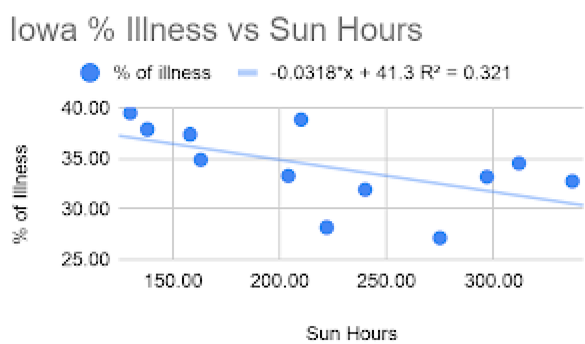
Nonetheless, we should keep in mind that we are unsure of the linear model's appropriateness.

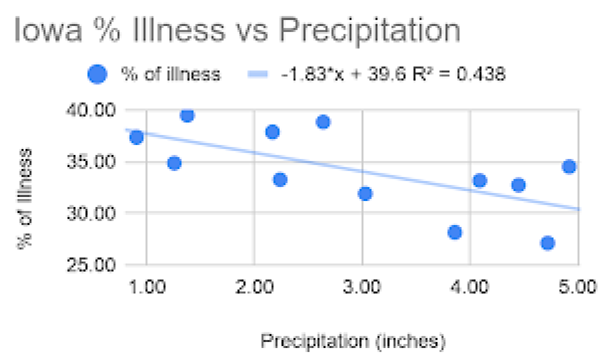
Since we don't have many data points in either graph, we also made several scatter plots for each state, showing the percentage of people reporting symptoms against either total sunlight hours or rainfall for that state. Each point is a different month from April 2020 to March 2021, and each graph looks at one state only. For all of the state graphs, see [this spreadsheet.]

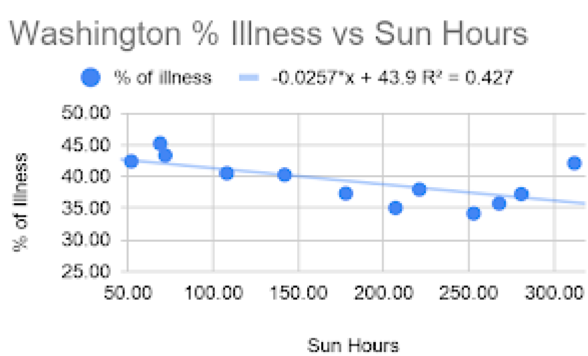
Examples:

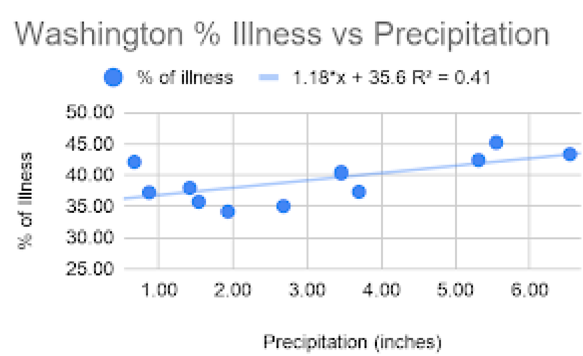


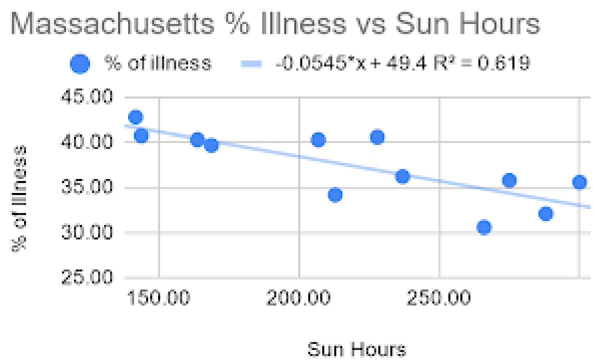


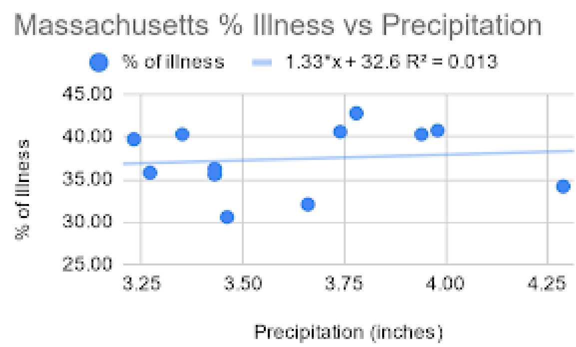












Almost all of the sunlight graphs show a moderately strong negative relationship between the average monthly sunlight hours and the depression/anxiety rate. The graphs suggest that as sunlight hours increase, the rate of depression/anxiety decreases, thus supporting our hypothesis.

The signs of slopes of the rainfall graphs are inconsistent, and therefore, no clear association is shown between the amount of precipitation and the percentage of reported depression/anxiety symptoms, opposing our hypothesis.
A possible reason for this could be the duration or type of the rainfall. Many short drizzles could add up to one big storm precipitation-wise, but would have a different effect on people. Likewise, a long, light storm might have the same amount of rain as a short thunderstorm, but with a different effect on people. Hence, a possible reason for the varying relationships shown in the graphs.

Possible Sources of Error

- Worldwide data
  - Climate types assigned are not the actual climate, but estimated based on latitude
    - Actual weather may not match the assigned climate type
- UK Data
  - Total number of Adult cases include cases of Autism/Learning Disabilities, on top of other, more temporary mental health issues
    - May slightly skew data, some numbers may be higher than they should be, and autism cases may be different month to month
    - As we do not know how many adult cases are due to autism/learning disabilities, we cannot correct for it
      - Only have knowledge of total autism cases, not which ones are adults
      - Some adults may have both autism and another mental health issue, causing overlap
- US Data
  - Data used is post-Covid
    - New factors, such as Covid case numbers, may affect mental health more greatly than pre-Covid lurking factors
    - Medical facilities were overworked, preventing people from getting help
    - Thus, difficult to compare US data, R and R^2 with UK data

Conclusions

- Mental health illness has an association with weather, in terms of rainfall and sunlight duration
    - Only obvious when comparing it monthly in a single location
    - May be due to lurking variables, such as differences in poverty rates in different locations
    - More accuracy if we had more data so we we could better control for lurking variables
    - More accuracy if provided with monthly data from a single location across a longer time period

- Different regions may have different relationships between sunlight/rainfall and mental health
    - Culture regarding the weather may be different
        - Eg: May celebrate rain for crops, or may dislike being stuck inside when raining
    - May be more sensitive to one factor than another country
        - Eg: If sunlight is consistently high, an increase would not affect much, but if it's consistently low, even a small increase may mean a lot
    - Evident in how the slopes for the UK and US graphs had different signs, as well as the different states within the US
        - All had comparable correlation coefficients
    - Hypothesis may be correct for some countries but wrong for others
        - Worldwide data shows weak associations
            - Different countries having different relationships could be another reason for weak associations (on top of possible lurking variables)
        - UK data supports hypothesis
            - Sunlight graphs showed negative slopes, as sunlight increases, mental illness reports decrease
            - Rainfall graphs showed positive slopes, as rainfall increases, mental illness reports increase
        - Some US data supports hypothesis
            - Most states' sunlight graphs support hypothesis
            - Overall monthly sunlight graph strongly supports hypothesis
            - States' rainfall graphs vary greatly
                - No clear association found between rainfall and prevalence of mental health disorders