

The Sunny Side Up: An Interdisciplinary Overview of Sunlight's Effects on our Health

Joseph Toth*

Abstract

Do levels of daily sunlight exposure have a measurable effect on peoples' health? I used to live in Michigan where it is cloudy seemingly every day. This is drastically different from Texas, where the sun is almost always shining. It is well known that an increase in sunlight exposure can drastically affect one's mood; when looking at cloudy areas versus areas full of sunshine, the difference in overall happiness and rates of depression can be quite dramatic. Another factor to consider is that spending more time in direct sunlight can help with vitamin D synthesis. Low vitamin D levels can lead to all sorts of health problems, demonstrating that having direct sun exposure has the potential to be very influential in one's health. To investigate whether sunlight can have a measurable impact on peoples' health and life expectancies, I collected data on average life expectancies and reported happiness levels for all 50 U.S. States, as well as over 70 cities within the United States. These cities were grouped together using K-means clustering to control for confounding variables, and then analyzed. This allowed for an accurate representation of whether sunlight really does affect one's health or not. Data came from the United States Census Bureau, the CDC, and other reliable sources. This research highlights how certain climate areas can affect our health, happiness, and overall life expectancy.

Keywords

Sunlight — Insolation — Life Expectancy — Happiness — Vitamin D — Weather — Health

*Corresponding author: josephtoth@my.unt.edu — Honors College, University of North Texas

Contents

Introduction	1
1 Review of Previous Literature	2
1.1 Happiness and Life Expectancy	2
1.2 Weather Effects on One's Life	2
1.3 Vitamin D Effects on Health	3
2 Methodology	3
2.1 Comparing Variables by State	3
Data Collection • Analysis of Data	
2.2 Analyzing Individual Cities	5
Choice of Cities • Data Collection • Analysis of Data	
3 Results	6
3.1 Results from State Data	6
3.2 Results from City Data	7
4 Discussion	8
4.1 Conclusions	8
4.2 Looking Forward	9
Limitations of This Study • Suggestions for Future Research	
About the Author	10
Acknowledgments	10
References	10

Introduction

In this study I will examine the relationship between the amount of sunlight one receives with their overall health and life expectancy. I will also investigate the connection between sunlight and happiness; happiness and life expectancy have long been known to be positively correlated, so I hypothesize that the relationship between sunlight and happiness will have similar outcomes to comparing sunlight with life expectancy. I spent my entire childhood living in a small town in Michigan where it was always cold – 85°F (29.4°C) is the warmest that it gets on average in peak summer – and clouds were a near constant. There were seldom any sunny days, and on the rare occasion of sunshine, some clouds were still present. I later moved to Texas and immediately noticed a drastic climate difference. In terms of sunlight, Texas is the opposite of Michigan. Over time I noticed that I was generally a more optimistic person after moving to Texas. This could be due to outside factors, but I believe that the change in weather also had a significant role in my attitude change. Therefore, I aim to research whether this could be true by examining the relationships between sunlight, happiness, vitamin D levels, life expectancy and overall health. This research is meaningful in a variety of ways. If my hypothesis is correct, then my research could illustrate that something as simple as the climate factors from where you live can have an impact on your health.

1. Review of Previous Literature

1.1 Happiness and Life Expectancy

They say that the happier you are, the healthier you are. Does being happy really increase your longevity? Bjørnskov [1] studied the trends of happiness and life expectancy in Europe. He collected data on life expectancy at birth from World Bank World Development Indicators, and compared this with data from the Eurobarometer, which measures life satisfaction. With just these two variables, he created a simple baseline to compare his findings when including other variables into his regressions. Bjørnskov's baseline found, with a correlation coefficient of 0.52, that for the world population, as satisfaction with life as a whole increases, so does life expectancy. He then created further regressions that utilized aspects such as economic development by looking at GDP per capita, health expenditures and average education levels. It was found that some variables, such as overall government consumption, did not have much significance in happiness levels, but economic success and health expenditures have a rather high impact on general happiness of a country's citizens. A fault in the study however, was that for some of the more advanced regressions, there was data missing for some countries. Bjørnskov used baseline data for 85 countries, but only 76 of those had full data for all the regressions.

In moderation, stress can be a positive factor in our lives, however, excess stress can lead to serious health problems. If left unchecked, extreme amounts of stress can actually lead to a premature death. How much stress is 'too much' varies from person to person. Individuals handle stress differently, but what is most measurable is the difference in how men and women cope with excessive amounts of stress. In most populations, women tend to handle stress better, and also tend to live longer than men. However, when accounting for life happiness, this life expectancy gap begins to change. As happiness decreases, the life expectancy gap begins to grow [2]. This reaffirms the notion that women tend to handle stress better than men. Long-term stress leads to health problems such as depression, but women are typically able to handle these issues more effectively. When it comes to long-term stress, it can be said that "women get sick, and men die" [3].

If happiness can positively affect one's health, how can happiness generally be increased? When people are happier, they generally tend to possess more confidence. Increased confidence can lead to better life opportunities such as promotions in one's career, or being perceived as more attractive to potential lovers. It is more common for someone to report being very happy if they are married, have a well-paying job, or have a higher education degree [4]. This may imply a positive feedback loop; the happier someone is, the better they perform in life, and the better they perform in life, the happier they are. The reverse may also be true, as those who are unhappy with life actually tend to have higher mortality rates. This positive relationship between life happiness and performance may have a limit however. According to a study

performed by Daniel Kahneman and Angus Deaton, higher income levels will generally increase one's happiness. However, once a salary of \$75,000 is reached, happiness plateaus, and further income does not have a statistical effect on life happiness [5].

1.2 Weather Effects on One's Life

For some, winter can be a time of cuddling up by the fireplace, spending time with loved ones, and drinking hot cocoa. For others, winter means getting sick with the flu and shivering from the cruel cold. Winter can be harsh, but can it affect your health? Tina Ho and Andrew Noymer [6] conducted a study to measure if winter weather can actually cause a drop in life expectancy. They divided the year into 'pseudo-seasons', summer and winter, and then calculated life expectancy within those portions of the year. By looking at data for mortality rates over 55 years, it was concluded that life expectancy rises by one year during the summer. The study was conducted only in the United States however, which could show bias, as data may be different in other countries.

Simply being in a good mood or bad mood can affect how you perceive your life as a whole. If you are feeling happy at the moment, you are more likely to perceive your life in a positive manner, and vice versa. People typically are happier and go out more on days with nice weather, so can it be assumed that someone living in an area with year-round nice weather will be happier? Kämpfer conducted a study in Germany where he asked subjects questions pertaining to their life happiness on different days depending on if the weather was good or bad [7]. His study found that on sunny days there was a 6% increase in happiness when compared to days with bad weather. Some believe that sunlight can increase serotonin in the brain, therefore stimulating happy emotions. While Kämpfer found that people tend to be happier on sunny days, this study will try to extend this research to see if sunlight can actually influence life expectancy overall along with one's life satisfaction.

While it is known that sunlight and the weather can affect one's daily mood, can these factors tie into overall happiness? Assuming this is true, does that also mean that good weather and sunshine can help fight depression? The sun rises in the East and sets in the West. As the sun rises in the morning, it floods rooms with an Eastern window with sunlight, but in the evening, the effect on West facing rooms is not as strong. Utilizing this knowledge, a study in Milan, Italy, examined the average length of a hospital stay of bipolar depression patients [8]. Patients in East and West rooms were given identical treatment, and had the same attending physicians; the only difference was whether their windows faced to the East or West. By the end of the study, it was found that patients in East rooms were released from hospital care 3.67 days earlier than those in West rooms. This study was a retrospective study however, which still needs to be confirmed by a prospective study. If this study can be replicated and confirmed, this could

be proof that increased sunlight exposure can help to fight off depression and related illnesses.

1.3 Vitamin D Effects on Health

It is well known that vitamin D plays an important role in our general well-being. While vitamin D can be supplemented through our diets, for most of us, it mainly comes from being outside in the sunlight. In a study on the correlation between mortality rates and vitamin D levels [9], it was determined that by increasing the uptake of vitamin D, one's life expectancy can rise by two years or more. Having low levels of vitamin D can actually increase risk of certain types of cancers, heart disease, and several other diseases. These have been identified as 'vitamin D sensitive diseases'. On average, low levels of vitamin D can increase an area's mortality rate by 20%, showing that vitamin D is vital to our health and well-being.

As we age, our bodies begin to break down and become more susceptible to diseases. As a result, our skin loses its ability to produce vitamin D from sun exposure. Due to this, elderly citizens must supplement their diet with double or triple the amount of vitamin D that their younger counterparts require [10]. Some people however, are reluctant to take supplements for fear of 'vitamin overdose'. While it is possible to overdose on vitamin D, it takes a substantial amount to actually cause damage. An individual can receive as many as five times the recommended amount of vitamin D just by sitting out in the sun, yet there are no recorded instances of vitamin D intoxication from simply sitting outside. Vitamin D supplements however, can sometimes lead to severe problems if misused. If supplements are needed, it is recommended that consumption is limited to 10,000 IU per day or less [11]. However, some vitamin D supplements can come in the form of an injection containing up to 600,000 IU. This excess vitamin D tends to linger in the body for weeks as it is unusable, thus leading to various health concerns such as anorexia, hypertension, vomiting, and more. Despite the fact that vitamin D intoxication is possible, many people tend to have the reverse: vitamin D deficiency.

There are many factors that contribute to the vitamin D deficiency epidemic. As time progresses, our bodies lose their natural ability to generate vitamin D; despite this, many young people are also suffering from vitamin D deficiency. This may be partly due to our evolving technology and better understanding of the dangers of UV rays from the sun. For example, think back to a bright summer day in your childhood. Just as you're about to run out the door to play outside with your friends, you hear your mother yelling, "Don't forget your sunscreen!" While sunscreen helps to prevent sunburns, it has a negative effect as well. When a sunscreen with SPF 30 or higher is applied to the skin, 95% to 99% of the skin's ability to produce vitamin D is lost [12]. Continuous application of sunscreen can lead to insufficient vitamin D production, despite the sun beating down on you relentlessly. Those with darker skin are also at a disadvantage; while their skin is

adapted to not burn in heavy sunlight, their increased melanin pigment leads to a side effect of being up to 90% less efficient in absorption of vitamin D. Many experts recommend that moderate sunlight exposure is actually beneficial [11]. The word 'moderate' is a vague term however; this simply means being out in the sun, but not long enough to encounter a sunburn. By being out in the sun for a moderate time period, it is possible for the majority of people to naturally produce all the vitamin D they need. Winter months tend to mean less available sunlight, but recent research has shown that humans may actually be able to store some naturally produced vitamin D in the liver and fat to make up for lost sunlight time in the winter.

2. Methodology

Hypothesis Before the collection and analysis of data was complete, a hypothesis was formulated based on previous literature. Results from related studies seemed to have a common trend; therefore, it was suspected that there may indeed be a positive relationship between sunlight and its effect on happiness and life expectancy. While measuring this relationship directly in the field can be hard, costly, and time consuming, analysis of already existing data can still yield significant results. While there may be a positive relationship between sunlight and health, it is too difficult to predict the extent of this relationship. If this hypothesis is correct, further regressions may show similar relationships between sunlight and other health factors.

In order to fully understand the relationships between sunlight, happiness and life expectancy, two main methods of analysis were employed. First, it was necessary to read about how each variable is measured and determine the best way to obtain data. Next, to get a general overview of the trend between the three variables, data grouped by states in the United States were collected. This allowed for a correlation of initial data, and also gave a path to create several maps for a visual representation of each variable. Lastly, after analyzing data by state, it was then necessary to gather and analyze data for individual cities in the United States. Grouping cities by demographics and other factors granted control over confounding variables, allowing for a more accurate representation of the variables of interest.

2.1 Comparing Variables by State

In December 2018, initial data was collected for happiness, life expectancy and average daily sunlight. Each of the three data sets were based by state¹; this allowed for a quick overview of the relationships between the variables. This was not intended to be a very accurate method though, as some states like Texas are quite large, while others like Rhode Island are very small. Instead, the purpose of collecting data

¹Alaska and Hawaii were not included in this study. This was to avoid having outliers in the data, as these states are thousands of miles from the Continental United States.

grouped by state was to get an initial overview of the relationships between the variables and also create color coded maps of the United States to visualize these trends.

2.1.1 Data Collection

Happiness Happiness is not an easy construct to measure. There is no international “unit” to measure happiness in the same way that we measure height and weight. Sometimes, people lie or exaggerate about how happy they are. Some people care so much about their outward appearance, that if anonymity is not guaranteed, they may not admit that they are unhappy with their lives. Other times, people may have the exact same lifestyle and situations, but may report different levels of happiness. The question, “How happy are you?” is a very open-ended and subjective question. It is possible to ask two different people how happy they are and get completely different answers, yet these people may actually feel the same way inside. Therefore, researchers have come up with many methods to measure happiness.

The data for happiness were collected from a Wallet Hub study titled *Happiest States in America*[13]. Their researchers gave each state a “Happiness Score” – out of a maximum score of 100 points – based on a large variety of 31 relevant factors between the following three categories: Emotional and Physical Well-Being, Work Environment, and Community. Each factor was also given a corresponding weight, as some health factors are known to have larger effects on mental well-being; for example, numbers of hours worked were weighted triple that of the physical health index. Data for the Wallet Hub study were collected from trusted government resources such as the United States Census Bureau, Bureau of Labor Statistics, the CDC and more. The Wallet Hub study presented a very thorough and well thought out methodology for measuring happiness, and therefore the data presented was deemed useful for further studies. For the purposes of this research, the final happiness score for each state was collected and compared to the other variables of interest. The average happiness score across the entire United States was noted to be 52.0.

Life Expectancy Data for life expectancy in years were very easy to collect. Life expectancy across large areas is easy to measure, and as such, data for this is widely available. The data for this project was sourced from Wikipedia on their “List of U.S. States and Territories by Life Expectancy” page[14]. In the reference list for the page, one can find a link to the 2011 United States Life Tables issued by the CDC[15], allowing for easy verification of the data presented on the Wikipedia page. From this data, the average life expectancy within the United States was 78.63 years.

Insolation The data for average sunlight were sourced from the CDC WONDER online database[16]. CDC WONDER is a massive public health information system that is published by the Centers for Disease Control and Prevention (CDC) and is freely available for anyone to access. The data sourced

were recorded as daily insolation (sunlight) averages in kJ/m^2 . These numbers can easily be converted to kWh/m^2 ($1 \text{ kWh} = 3600\text{kJ}$). In layman’s terms, one kWh of sunlight is the same as one hour of peak sunlight without any cloud cover. According to the data, the mean sunlight level for the continental United States is 4.41543 kWh/m^2 .

2.1.2 Analysis of Data

After collecting everything into one spreadsheet, the data were uploaded into R-Studio to visualize the connections between the variables. To start, the name of each variable was changed to ease the coding in R-Studio: Happiness Levels were now HAP, Life Expectancy was now LE, and Sunlight Levels were now SUN. Utilizing a Pearson Correlation, it was possible to create a linear representation of the relationship between amounts of sunlight and its effects on happiness and life expectancies. The Pearson Correlation was also useful to get an idea of the magnitude of these relationships. The results from this correlation may not be completely accurate however; as stated earlier, the size of states can vary immensely, so this data may not be the most accurate method of representation. However, the data gave a way to create several maps of the United States (Figures 1a, 1b, and 1c), which allows for a visual of the variables across the country.

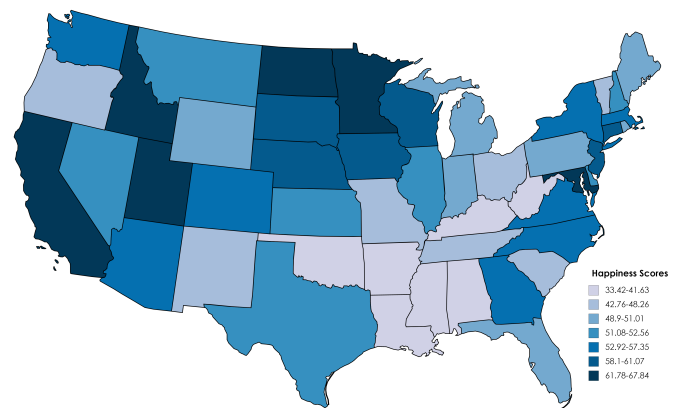


Figure 1a. Happiness Scores by U.S. State

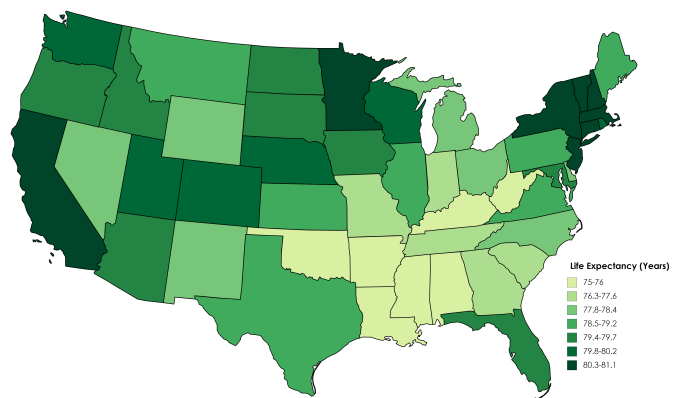


Figure 1b. Life Expectancy By U.S. State

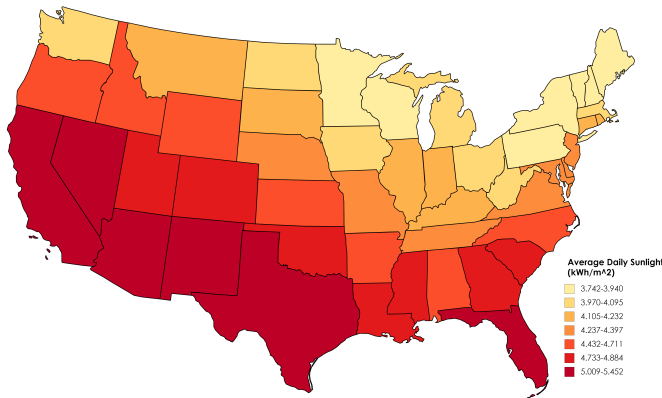


Figure 1c. Sunlight Levels by U.S. State

2.2 Analyzing Individual Cities

After examining initial state data, it was clear that it was necessary to dig deeper. Several cities in the United States were matched together by similarity in terms of demographics. After grouping cities together, data for the variables of interest corresponding to each city were then collected for analysis. By limiting the areas to individual cities and grouping them by demographics, this study was able to control for many confounding variables, thus leading to stronger results.

2.2.1 Choice of Cities

Information from the U.S. Census[17] was examined for over 70 cities across the United States. Specific demographics used for city matching were income distributions, education levels, and poverty levels. Income was measured as the average income for each city, and education was measured as percentage of the city population that had a Bachelor's degree or higher. Poverty levels for each city were determined by the percentage of the city population living under the poverty threshold². After collecting the necessary data from the U.S. Census, K-Means Clustering was used to group these cities into five groups based on their similarities to each other. The following groups of cities were determined to be suitable for analysis:

1. San Francisco, CA; San Jose, CA
2. Austin, TX; Bakersfield, CA; Boston, MA; Charlotte, NC; Chicago, IL; Denver, CO; Fort Worth, TX; Las Vegas, NV; Los Angeles, CA; Madison, WI; Manchester, NH; Minneapolis, MN; Modesto, CA; Nashville, TN; New York, NY; Omaha, NE; Phoenix, AZ; Port St. Lucie, FL; Portland, OR; Raleigh, NC; Sacramento, CA; Salt Lake City, UT
3. Baton Rouge, LA; Birmingham, AL; Brownsville, TX; Cincinnati, OH; Cleveland, OH; Detroit, MI; Knoxville,

²Detailed information on how the U.S. Census determines poverty thresholds and poverty rates can be found on the *How the Census Bureau Measures Poverty* page[18]

TN; Memphis, TN; Miami, FL; Milwaukee, WI; Mobile, AL; New Orleans, LA; Newark, NJ; Philadelphia, PA; Providence, RI; St. Louis, MO; Tucson, AZ

4. Albuquerque, NM; Atlanta, GA; Baltimore, MD; Bridgeport, CT; Columbia, SC; Columbus, OH; Dallas, TX; Des Moines, IA; El Paso, TX; Fayetteville, NC; Fresno, CA; Grand Rapids, MI; Greensboro, NC; Houston, TX; Indianapolis, IN; Jacksonville, FL; Kansas City, MO; Little Rock, AR; Oklahoma City, OK; Orlando, FL; Pittsburgh, PA; Portland, ME; Richmond, VA; San Antonio, TX; Spokane, WA; Tampa, FL; Tulsa, OK
5. San Diego, CA; Seattle, WA; Virginia Beach, VA

To get a visual of where each of these cities were located and how spread out each cluster was, a map was created with each city color coded to its respective cluster (Figure 2). Cluster 1 was plotted in red, Cluster 2 in blue, Cluster 3 in green, Cluster 4 in orange, and Cluster 5 in purple.

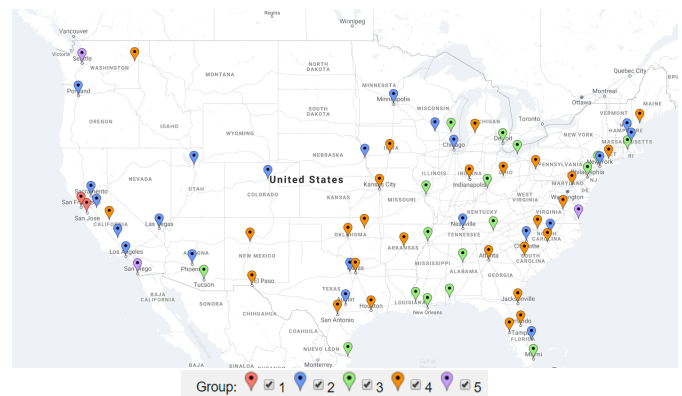


Figure 2. City Cluster Map

This map is also accessible at: <http://bit.ly/SunMap>

Each group of cities was selected for a reason. Most of the cities within each group are in varying climate regions of the United States. Some cities in each group may have similar climates, but they are in separate parts of the country, which may still yield useful results. The next step was to collect data for each city on the three variables of interest.

2.2.2 Data Collection

Happiness The data for happiness came from the same Wallet Hub source that was used to analyze the United States as a whole. The Wallet Hub researchers did an additional study[19] to determine the happiest cities in the United States. Each of these cities were evaluated and ranked using the same methodology as was used for the *Happiest States in America* study. This makes it very easy to compare happiness levels of an individual city to the state it resides in.

Life Expectancy Research into life expectancy for individual cities was a little complicated. It is important to note that measures for life expectancy are constantly changing. Additionally, life expectancy can change quite rapidly when

looking at smaller and smaller areas. According to a study by Virginia Commonwealth University’s Center on Society and Health, life expectancy can differ by as much as 20 years in neighborhoods only five miles apart[20]. These large differences in life expectancies between neighborhoods are due to many factors, including safety conditions, environmental factors, education and income, and other social factors.

Due to the complexities of measuring life expectancy at levels as small as a single city, finding accurate data took a bit of work. Finally, I came across the Health Inequality Project, which uses big data to measure life expectancy by income across the nation[21]. Within their website was data that had life expectancies for the top 100 most populated commuting zones in the United States. Luckily, this directly corresponded with the data for each of this study’s relevant cities.

Insolation Insolation data for each city were collected from Solar Energy Local[22]. Solar Energy Local is a comprehensive source of solar data, statistics, resources, and other information for areas across the United States. One of the main purposes of this site is to measure solar energy levels across the nation to help people determine whether buying solar panels would be a smart idea or not. The information provided is also useful in this research however, as the website also provides solar radiation data measured in kWh/m².

2.2.3 Analysis of Data

After grouping cities by clustering variables and collecting data for all three experimental variables, it was time to compare the relationships between these and see if there is evidence of sunlight playing a role in overall health. First, the averages of each variable were taken for all five clusters in addition to the population as a whole and put into two tables. One table represented the Clustering Variables: Income, Education, and Poverty (**Table 1**), and the other represented the Testing Variables: Happiness, Life Expectancy, and Sunlight (**Table 2**). Then, the percentage above or below the population mean was calculated for each cluster’s variables. These numbers were then separated into two columns set side by side, one for the clustering variables, and one for the test variables (**Table 3**).

	Income	Education	Poverty
Cluster 1	96463.5	0.4855	0.1085
Cluster 2	56789.82	0.370409	0.174
Cluster 3	36599.00	0.262294	0.277529
Cluster 4	47074.11	0.328148	0.203333
Cluster 5	73866.67	0.469667	0.116667
Entire Data-set	50099.87	0.335887	0.205676

Table 1. Clustering Variable Averages

	Happiness	Life Expectancy	Sunlight
Cluster 1	72.86	81.25	5.44
Cluster 2	59.019	79.6318	5.1618
Cluster 3	46.613	79.1471	4.8418
Cluster 4	53.850	78.9185	5.0452
Cluster 5	64.347	79.8333	4.7833
Entire Data-set	54.697	79.2986	5.0327

Table 2. Test Variable Averages

3. Results

3.1 Results from State Data

Using the initial data sorted by state, the correlations between each of the three variables were examined through a Pearson correlation method (**Table 4**). As shown in previous studies, happiness and life expectancy have a very strong correlation; the state data that was used in this study showed this correlation has a magnitude of 0.817, a statistically strong relationship. Interestingly though, average daily sunlight has a slight negative correlation to both happiness and life expectancy: -0.139 and -0.234 respectively. This study will further investigate whether this will hold true for the data collected by cities; the city data analysis controls for several confounding variables, while the state data analysis does not. Therefore, there may be a significant change in the correlations between the variables for each section of this study. Moreover, while the correlation between sunlight and happiness is definitely significant ($p < .01$), the correlations between sunlight with happiness and life expectancy are not so significant. Sunlight correlating with happiness yielded a p-value equal to 0.3446, while sunlight correlating with life expectancy yielded a p-value equal to 0.1094. This further shows how confounding variables may be skewing the results. A large confounding factor that could be imposing a negative correlation for sunlight with happiness and life expectancy may be money; residents of southern states such as Alabama and Mississippi, which tend to get more sunlight than northern states, also tend to have less money than residents of other states, thus leading to a smaller happiness score.

A quick plot (**Figure 3a**) was made using the three variables against each other through the ‘pairs’ command in R-Studio. This helped to visualize the relationship between each variable. In the plot, happiness and life expectancy have a very clear positive relationship. When looking at sunlight against the other two variables though, the relationship is not so clear from a visual standpoint alone, partly because the data has confounding factors that are not accounted for. While there is an overall negative correlation when looking at sunlight against happiness and life expectancy, the data shown in the initial graphs is quite sporadic and full of outliers. **Figures 3b, 3c, and 3d** show expanded versions of the plots seen in **Figure 3a** with the addition of a regression line.

	HAP	LE	SUN
HAP	1.0000000	0.8170564	-0.1394099
LE	0.8170564	1.0000000	-0.2340071
SUN	-0.1394099	-0.2340071	1.0000000

Clustering Variables

Cluster 1:

Above average income by 92.54%

Above average education by 44.54%

Below average poverty by 47.2%

Cluster 2:

Above average income by 13.35%

Above average education by 10.28%

Below average poverty by 15.4%

Cluster 3:

Below average income by 26.95%

Below average education by 21.91%

Above average poverty by 34.94%

Cluster 4:

Below average income by 6.04%

Below average education by 2.3%

Below average poverty by 1.14%

Cluster 5:

Above average income by 47.44%

Above average education by 39.83%

Below average poverty by 43.28%

Test Variables

Cluster 1:

Above average happiness by 33.21%

Above average life expectancy by 2.46%

Above average sunlight by 8.09%

Cluster 2:

Above average happiness by 7.9%

Above average life expectancy by 0.42%

Above average sunlight by 2.57%

Cluster 3:

Below average happiness by 14.78%

Below average life expectancy by 0.19%

Below average sunlight by 3.79%

Cluster 4:

Below average happiness by 1.55%

Below average life expectancy by 0.48%

Above average sunlight by 0.25%

Cluster 5:

Above average happiness by 17.64%

Above average life expectancy by 0.67%

Below average sunlight by 4.95%

P-values:

SUN vs HAP: 0.3446

SUN vs LE: 0.1094

HAP vs LE: $p < 0.01$

Table 4. Preliminary Correlations

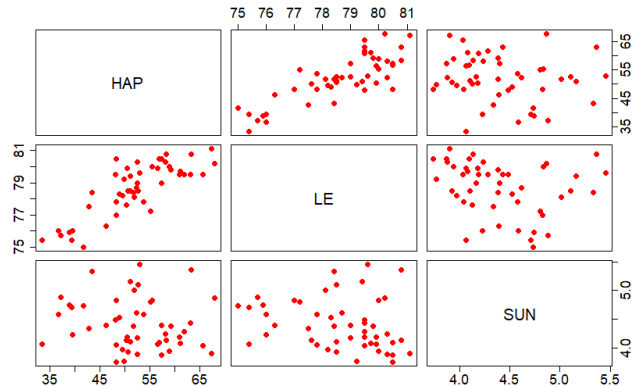


Figure 3a. Preliminary Correlation Graphing

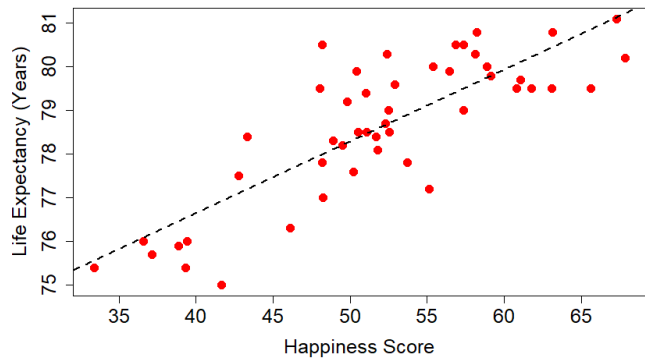


Figure 3b. Happiness vs Life Expectancy

Table 3. Percentages Above and Below Population Means

3.2 Results from City Data

When looking at **Table 3**, one can see some clear trends. These trends are easier to view in a condensed form of the data shown in the table:

- Cluster 1 has above average happiness and sunlight. Life expectancy is also above average.
- Cluster 2 has above average happiness and sunlight. Life expectancy is about average.
- Cluster 3 has below average happiness and sunlight. Life expectancy is about average.

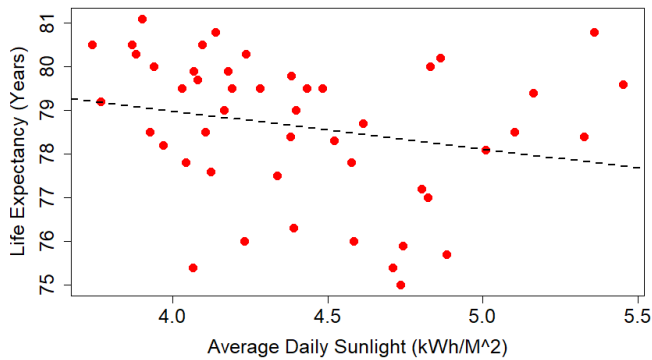


Figure 3c. Sunlight vs Life Expectancy

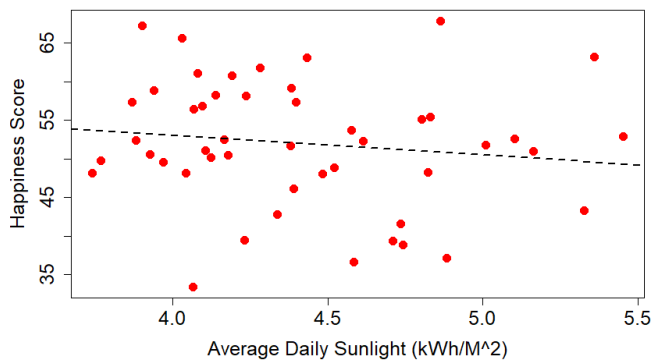


Figure 3d. Sunlight vs Happiness

- Cluster 4 has below average happiness and about average sunlight. Life expectancy is about average again.
- Cluster 5 has below average sunlight, but above average happiness. Life expectancy is still about average.

With the exception of Cluster 1, life expectancy is about average all around. Whether sunlight is high or low, life expectancy stays within 1% of the average. These results suggest that sunlight may not have a direct correlation with life expectancy. Interestingly though, happiness and sunlight seem to have a correlation. When sunlight is above average, happiness is also above average. Conversely, when sunlight is below average, happiness is also below average. Just as Cluster 1 did not follow the life expectancy trend, Cluster 5 does not follow the trend with happiness. However, this may be due to Clusters 1 and 5 containing some of the more extreme case cities; Clusters 1 and 5 both have exceptionally high income and education rates, tied in with particularly low poverty rates (see Table 3).

A simple regression of the sunlight and happiness averages from Table 3, using sunlight as the independent variable and happiness as the dependent variable, yields the following equation: $y = 2.1698x + 7.5463$. What this means, is that for

every percentage point increase above the average in sunlight, one can expect happiness to increase by 2.1698 percentage points. Additionally, if sunlight is exactly at average, one can expect happiness to be 7.5463% above average. As seen in Figure 4a however, there is a large p-value of 0.26229 and a low R^2 value of 0.38725. These values are being caused by the outlier clusters, Cluster 1 and Cluster 5. When removing these from the equation, a much different model is produced: $y = 3.5335x - 1.6675$. This model, shown in Figure 4b, has high statistical significance. The R^2 is nearly 1.0, and $p < .05$.

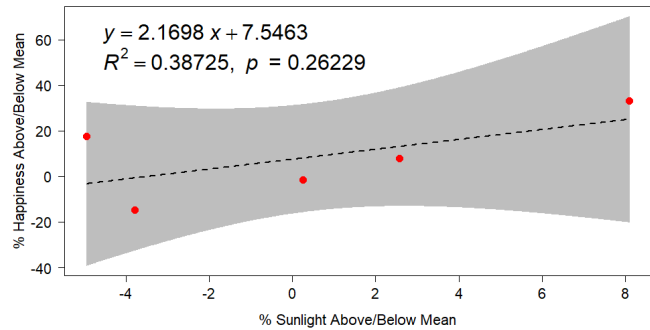


Figure 4a. Regression Model Using Outlier Clusters

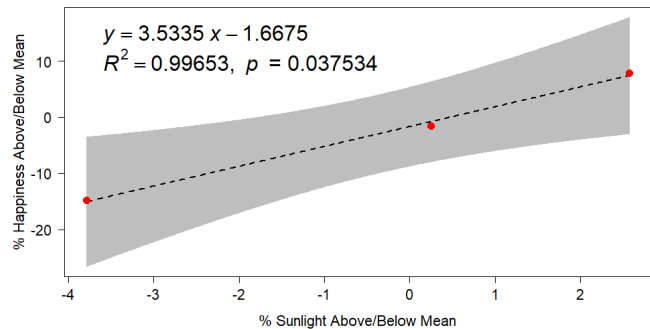


Figure 4b. Regression Model Without the Outlier Clusters

4. Discussion

4.1 Conclusions

While much of the by-state analysis results can be dismissed due to high p-values, it is important to take away some key notes. First, as noted by previous research articles, happiness and life expectancy are clearly shown to have a strong positive relationship. Simply put, happier people tend to be healthier and live longer. The slightly negative correlations between sunlight and happiness/life expectancy shown in the by state analysis can be discarded due to high p-values. As shown in the correlation graphs (Figures 3c and 3d), the data is overly sporadic and inconsistent. This is due to several confounding factors – some states are large, while others are

small; populations of many states are clustered in cities, meaning larger states could have skewed data; variables such as income, poverty, and education rates are not controlled for; etc. While some states may be similar in many aspects, no two states are quite "the same", especially when trying to analyze data in an accurate and statistical manner.

Since the by-state analysis was largely inconclusive, the by-city analysis with K-means clustering was absolutely necessary. This allowed for several confounding variables to be controlled for, thus leading to more significant results. Before even running a regression, two trends were easily seen in **Table 3**. First, while happiness and sunlight had some variation, the life expectancy variable did not change much between all five clusters. With the exception of Cluster 1, which was deemed to be an outlier cluster, life expectancy for each cluster was within 1% of the mean of all the cities examined. This leads to the conclusion that the level of sunlight one receives does not have a direct effect on their life expectancy. Interestingly though, the data from **Table 3** does show a positive trend between sunlight and happiness. While the model from **Figure 4a** must be discarded due to its high p-value and low R, the model from **Figure 4b** has high statistical significance. This shows with a great confidence that sunlight and happiness have a positive correlation.

Using the statistically significant model of $y = 3.5335x - 1.6675$, some inferences can be drawn about the United States. To summarize the model, this shows that for every 1% increase above the mean in sunlight, happiness is expected to rise another 3.5335% above the mean. The negative y-intercept of 1.6675 means that in an area with sunlight exactly at average, happiness is expected to be slightly below the mean for the country. Recalling that the average sunlight is 4.41543 kWh for United States as a whole, along with 52 as the average happiness score, this model shows the following: to have happiness in an area be above average, sunlight should be more than 0.47% above average, or above 4.4362 kWh of average sunlight. This is not far from the mean though; since sunlight is only a factor in happiness and not the sole cause, this number does not hold much meaning. It should also be considered, that areas with below average sunlight could indeed have higher than average happiness scores. As discussed in the literature review, happiness can be affected by various factors such as education, income, stress, and more. Therefore, a city with low sunlight, but an abnormally large average income, could very well have above average reported happiness.

The final results of this paper show the following relationships with sunlight: sunlight and happiness are positively correlated, while sunlight and life expectancy are not directly correlated. However, while sunlight does not directly affect life expectancy, happiness has been proven many times to have a strong relationship with life expectancy. Therefore, if sunlight has a direct correlation with happiness, should it not at least have an indirect correlation with life expectancy?

The relationship between sunlight and life expectancy should not be ignored. Instead, this relationship should be further analyzed; it may be possible that this relationship was not quantifiable due to other factors. Moreover, increased happiness has been shown in other studies to have positive correlations with various health factors. Therefore, sunlight having a positive correlation with happiness should in turn mean that increased sunlight can positively affect other health factors as well, even if these are indirect relationships. One example of a large health factor that is already known to be affected by sunlight is vitamin D synthesis. Overall, this study shows that sunlight certainly has a direct positive relationship with happiness, and possible indirect relationships with life expectancy and other health factors. Despite this, there are many other variables that play a factor in our health, so sunlight alone cannot be used to measure one's health.

4.2 Looking Forward

4.2.1 Limitations of This Study

Due to the limitations of data that was both readily available and reliable, this study was restricted to analyzing just the United States. Therefore, while the conclusions are significant for the United States, results may be different if looking at other parts of the world, even if just looking at close-by countries such as Canada or Mexico. Additionally, while the literature review had some great findings on vitamin D effects, it was not possible to find existing data. Financial limitations restricted cross-country travel and lab-work, so this study was also unable to collect new data for vitamin D levels; therefore, this variable had to be left out of the data analysis. While it would be interesting to be able to use this variable, it would be quite costly to collect all the data, and would also be difficult to measure as well. Vitamin D levels are typically measured by drawing blood, and due to HIPAA restrictions, doctors cannot simply release that data. Collection of vitamin D data would therefore require volunteers.

In April of 2019, I met a gentleman named Gilbert Hu at Collin College in Plano when presenting my preliminary findings. He gave some highly useful advice. First, he pointed out that the way happiness was measured by the WalletHub researchers was an overly subjective method. While for the purposes of this project, conclusions can be drawn using this data, there is a far more accurate method to measure happiness. It is important to note that serotonin and dopamine are two chemicals in the brain that have a large role in the feeling that we describe as "happiness". Additionally, as stated earlier, vitamin D plays a large role in our overall health, and has a strong correlation with the amount of sunlight that we receive. Therefore, using a large sample of volunteers, levels of serotonin, dopamine and vitamin D could be measured, giving a more measurable view of happiness. Combining lab data with a subjective survey could also give a more quantifiable method of measuring happiness, even if it is still not perfect. A large enough sample of volunteers would also eliminate the issue of volunteer bias.

4.2.2 Suggestions for Future Research

Future researchers and studies should analyze other areas of the world to see if the results are similar to those in the United States. Additionally, it would be interesting to see what the results would be for a similar study that looked at the world as a whole. After completing studies for other regions of the world, it may be possible to simply combine all the data for an analysis of the world as a whole. Looking at other regions of the world may show different results for the relationship between sunlight and life expectancy as well; although no quantifiable direct relationship was shown, it is still possible that an indirect relationship between the two variables exists. Relationships between sunlight and other health factors could also be researched, along with relationships between climate factors other than sunlight and corresponding health factors.

Utilizing the advice from Mr. Hu, a second study could be performed as a continuation of this study. Obtaining a more quantifiable measure of happiness would yield more significant results when comparing with sunlight, life expectancy, and other health factors. The data used in this study for sunlight and life expectancy could easily be used in the second study, if completed without much layover time. Climate change usually takes many years to take place, and life expectancy in the United States has plateaued in recent years. There are many other ways to quantify happiness by measuring chemical receptors in the brain. Various lab methods could be used and combined to get a highly accurate measure of happiness, which could then be compared to sunlight along with other health variables.

About the Author

Joseph Toth Joseph is an Honors College student at the University of North Texas. He is pursuing a Bachelor of Arts in Biology, with a minor in Chemistry, and expects to graduate in the spring semester of 2021. Upon graduating from UNT, Joseph plans to attend medical school to become a physician. This research article was the culmination of Joseph's Honors College education, being his Capstone Thesis project. After publishing, Joseph also plans to dive into epidemiology research guided by his former mentor, Dr. Miles.

Acknowledgments

Thomas Miles, PhD Dr. Miles is a Lecturer at the University of North Texas (UNT). He primarily teaches undergraduate research courses for the Honors College. After taking his Introduction to Research course during my Freshman year at UNT, I discovered my passion for research and began working on my thesis proposal. Dr. Miles and I quickly built a friendly relationship, and when I was ready to move onto the Honors Capstone Thesis course, I asked him to be my mentor. Dr. Miles has assisted with this project enormously by writing code for R-Studio, and helping me to understand the statistics behind all the data.

Sean Ryan Sean is an Academic Counselor for the Honors College at UNT. Throughout my time at UNT, Sean has provided me with invaluable advice for varying topics. The most beneficial advice he has given me though, has been for this research project. Sean helped with proposing much of the methodology used in the paper; he was the one who suggested to group areas together by demographics to minimize confounding variables, and he was also the one to suggest using a computer program to do so. Sean's suggestions have greatly influenced the strength of the arguments made in the paper.

Bena Toth I would like to thank my mother for her assistance with editing and revising the final drafts of the paper. She meticulously combed through the entire document to make sure there were no grammar or spelling issues, and that everything flowed naturally. She also aided with understanding some of the statistics behind the by-city analysis.

Elizabeth Landeros I am eternally grateful for the support given to me by my fiancée, Elizabeth. Elizabeth has been an enormous help with this project. Every time I hit a dead end in my writing, she motivated me to get back on track. She also aided with brainstorming synonyms to improve the language used in the paper. Moreover, Elizabeth was the one who came up with the catchy title for this paper, "The Sunny Side Up".

Stephen Cofer I would like to thank my cousin Stephen for his assistance with editing the final stages of this paper. As an attorney, he often reads through complicated documents and legalese. Stephen's skills in reading carefully over documents proved to be quite useful, as he was able to catch some easy-to-miss typographical errors.

References

- [1] Bjørnskov C. Healthy and happy in Europe? On the association between happiness and life expectancy over time. *Soc Sci Med* 2008;66(8):1750-1759. <https://doi.org/10.1016/j.socscimed.2008.01.031>
- [2] Kageyama J. Happiness and Sex Difference in Life Expectancy. *Journal of Happiness Studies* 2012;13(5):947-967. <https://doi.org/10.1007/s10902-011-9301-7>
- [3] Nathanson CA. Sex, illness, and medical care: A review of data, theory, and method. *Social Science & Medicine* (1967) 1977;11(1):13-25. [https://doi.org/10.1016/0037-7856\(77\)90141-X](https://doi.org/10.1016/0037-7856(77)90141-X)
- [4] Lawrence EM, Rogers RG, Wadsworth T. Happiness and longevity in the United States. *Social Science & Medicine* 2015;145:115-119. <https://doi.org/10.1016/j.socscimed.2015.09.020>
- [5] Kahneman D, Deaton A. High income improves evaluation of life but not emotional well-being. *Proc Natl Acad Sci U S A* 2010;107(38):16489-16493. <https://doi.org/10.1073/pnas.1011492107>
- [6] Ho T, Noymer A. Summertime, and the livin' is easy: Winter and summer pseudoseasonal life expectancy in

- the United States. *Demographic Research* 2017;37:45. 10.4054/DemRes.2017.37.45
- [7] Kämpfer S, Mutz M. On the Sunny Side of Life: Sunshine Effects on Life Satisfaction. *Soc Indicators Res* 2011 2013;110(2):579-595. <https://doi.org/10.1007/s11205-011-9945-z>
- [8] Benedetti F, Colombo C, Barbini B, Campori E, Smeraldi E. Morning sunlight reduces length of hospitalization in bipolar depression. *Journal of Affective Disorders* 2001;62(3):221-223. [https://doi.org/10.1016/S0165-0327\(00\)00149-X](https://doi.org/10.1016/S0165-0327(00)00149-X)
- [9] Grant WB. An estimate of the global reduction in mortality rates through doubling vitamin D levels. *Eur J Clin Nutr* 2011;65(9):1016-1026. <https://doi.org/10.1038/ejcn.2011.68>
- [10] Heaney RP. Vitamin D in Health and Disease. *Clinical Journal of the American Society of Nephrology* 2008;3(5):1535-1541. <https://doi.org/10.2215/CJN.01160308>
- [11] Razzaque MS. Sunlight exposure: Do health benefits outweigh harm? *J Steroid Biochem Mol Biol* 2016;175:44-48. <https://doi.org/10.1016/j.jsbmb.2016.09.004>
- [12] Hossein-nezhad A, Holick MF. Vitamin D for health: a global perspective. *Mayo Clin Proc* 2013;88(7):720-755. <https://doi.org/10.1016/j.mayocp.2013.05.011>
- [13] McCann A. Happiest States in America. *WalletHub*. 2018. Available from: <https://wallethub.com/edu/happiest-states/6959/>
- [14] Wikipedia contributors. (2018, December 13). List of U.S. states and territories by life expectancy. In *Wikipedia, The Free Encyclopedia*. Retrieved from: https://en.wikipedia.org/w/index.php?title=List_of_U.S._states_and_territories_by_life_expectancy&oldid=873513284
- [15] Centers for Disease Control and Prevention. (2011). *United States Life Tables, 2011*. Retrieved from: https://www.cdc.gov/nchs/products/life_tables.htm
- [16] North America Land Data Assimilation System (NLDAS) Daily Sunlight (Insolation) years 1979-2011 on CDC WONDER Online Database, released 2012.
- [17] <https://www.census.gov/>
Date accessed: April 7, 2019.
- [18] US Census Bureau. (2018, August 16). How the Census Bureau Measures Poverty. Retrieved from <https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html>
- [19] McCann A. 2018's Happiest Cities in America. *WalletHub*. 2018. Available from: <https://wallethub.com/edu/happiest-places-to-live/32619/>
- [20] Chapman D, Kelley L, Woolf SH. Mapping Life Expectancy. Virginia Commonwealth University Center on Society and Health; 2015-2016. Available from: <https://societyhealth.vcu.edu/work/the-projects/mapping-life-expectancy.html>
- [21] Chetty R, Stepner M, Abraham S, et al. The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*. 2016;315(16):1750-1766. doi:10.1001/jama.2016.4226
- [22] <https://solarenergylocal.com/>
Date accessed: April 8, 2019.