More than a Sunburn: An Analysis of Demographical and Geographical Characteristics on Melanoma Incidences

Team: 13057

March 3, 2023

Contents

1 Executive Summary

Ranking in the top five most ubiquitous cancers, melanoma of the skin affects millions of individuals. While this statistic is alarming, it overlooks many details that are crucial to understanding the full story. Who is more at risk? Where is it most risky? When is it most risky? Seeking to uncover the full story, we developed a model to characterize risks based on several factors: age, ethnicity, biological sex, ultraviolet radiation (UVR), level of urbanization, and elevation. Characterizing these factors allows us to analyze groups, regions, and industries most vulnerable and ultimately provide recommendations to abate the number of melanoma cases.

Research reveals that age[16] [39], skin phenotype[20] [27], and biological sex[24] have noticeable influences on the risk of contracting Melanoma. With this in mind, we examined the US's demographic distribution and its trends. We utilized multiple regression techniques like the Gompertz formula and first-order regression. Employing those regressions, geospatial data, and mapping analysis enabled us to identify risky regions on a demographic basis. Based on this analysis, we found that the northern-midwest regions were most at risk.

Furthermore, it was determined that defined city and geographical characteristics like UV index, level of urbanization, and elevation play a substantial role in the frequency of melanoma cases. For UV index, we examined monthly and periodicity patterns for each county to determine peak, minimum, and average UV intensity. We discovered that spring and summer months as well as southern states are generally more perilous than their northern counterparts.

By reason of the dissimilar urbanization qualities (density of buildings, parks, and roads, etc.) amongst counties, one of our hardest barriers to overcome was a method to accurately quantify shade or "time absent from UV exposure." To achieve this, we took advantage of MATLAB's image processing capabilities to identify "darker spots," or shade, and used regression techniques to correlate population density and shade. We ascertain that the midwest United States, which is the least populous, is most vulnerable. Moreover, through these evaluations, we pinpoint 10 AM to 4 PM as the hours of greatest UV exposure and intensity.

Aggregating each factor together by use of ratios enabled us to generate melanoma case estimates: 52 thousand new cases in one year and over half a million cases by the end of ten. In addition to the model's numerical results, we generated various choropleth maps of the Contiguous United States which revealed that many western states (like Colorado and Utah) hold risky profiles. Ultimately, these results addressed the aforementioned who, where, and when questions, and laid the foundation for our recommendations.

Because melanoma is a nationwide obstacle, we tailored our recommendations to larger-scale industries in lieu of smaller scaled groups. First and foremost, we recommend that the northern mid-west farming industry instigate sun protective measures for their employees due to their high-risk profile, which is heavily influenced by minimal shade coverage. Secondly, we advocate for the government to take initiative in sun awareness campaigns to educate individuals to take prevention and early detection measures. Finally, as our model demonstrates, as people get older their risk of contracting melanoma increases exponentially. Thus, we push for the current government's Medicare program to cover more individuals in terms of screening, and for campaigns to spread awareness about the consequences of immoderate sun exposure, and early diagnosis—key steps to mitigate melanoma's irreversible damage.

2 Introduction & Background

2.1 Overview

At a rate of nearly 100,000 cases annually and 20 deaths daily, malignant melanoma is one of the most common types of cancer in the United States[31], ranking in the top five most frequent cancers by occurrence[3]. In fact, current estimates show that one in five Americans will develop skin cancer in their lifetime. In fact, in the past decade, the number of new invasive melanoma cases diagnosed annually has increased by 27%[31]. This growth can be shown in the graph below[4]. Additionally, the average annual cost per melanoma patient grew over twofold: from $\$2,320$ in 2002-2006 to $\$4,780$ from 2006 - 2011[14].

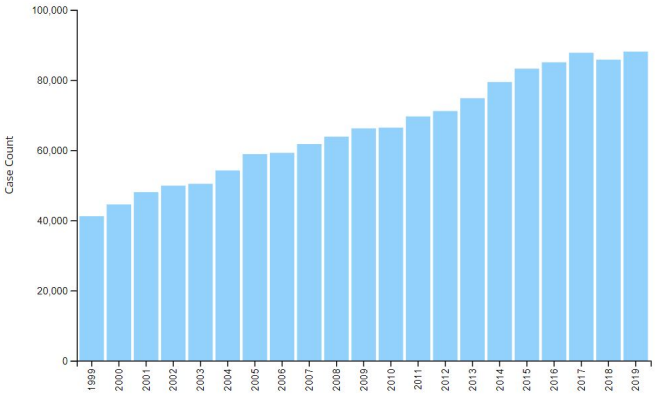

Figure 0: US Melanoma cases over time

Melanoma has the ability to affect almost anyone; however, some groups are inherently at greater risk than others. For example, data shows that the incidence of melanoma among Non-Hispanic Whites is nearly 30 times higher than that of Non-Hispanic Black or Asian/Pacific Islanders. Furthermore, one in 27 men is estimated to be diagnosed with melanoma in their lifetime compared to one in 40 women. While melanoma is one of the most common cancers in people younger than 30 years old, the average age of diagnosis is still 65[18]. Certain characteristics and hereditary traits make some groups evidently more vulnerable to melanoma than others.

The US experiences many melanoma cases every year, and one of the leading contributors to the confirmed number of cases is attributable to excessive and immoderate levels of sun exposure, or more specifically, ultraviolet (UV) radiation, a form of electromagnetic radiation emitted from the sun and other artificial sources. According to the US Department of Health and Human Services, UV exposure stimulates melanocytes to produce melanin, often resulting in a tan or sunburn, and causing changes to skin cells and the DNA within the skin cells. This change can ultimately lead to melanoma. The extent to which this UV exposure increases a person's risk of developing melanoma also depends on other characteristics and hereditary traits like age, ethnicity, and gender as mentioned earlier. These factors also help determine the amount of melanin in someone's skin[36].

2.2 Problem Statement

In our model, we sought to determine which US groups of people and regions are demographically and geographically most vulnerable to contracting melanoma, and forecast the number of melanoma cases in the coming years. Our model explicitly focuses on the frequency aspect (the number of

cases) and does not consider the costs and economic impacts of melanoma (severity). Through these insights, we recommend steps that wide-reaching industries and the government can take to combat the growing omnipresent melanoma.

3 Data Methodology

We drew data from several main sources. Some of which detail location-specific characteristics and others demographic information. These sources were advantageous to separating potential outcomes and identifying frequency.

3.1 Datasets

Our research shows that several primary characteristics and hereditary traits such as age [16] [39], ethnicity [20] [27], and gender put some at greater risk [24]. Thus, we used the Census Bureau's intercensal population data for 3,143 US counties categorized by race, age, and gender from 2010-2015. Although this data source wasn't completely contemporary, it still proved to have very comprehensive data helping us understand historical trends and forecast future trends in the population. As a federally provided database also commonly used for other research purposes [40], we are confident that this source is reliable.

We then synthesized this Census Bureau's population data with corresponding risk estimates provided by the National Cancer Institute's SEER (Surveillance, Epidemiology, and Ends Result Program) [22]. The risk numbers reflected a given person's calculated risk of being diagnosed with melanoma across all ages, ethnicities, and sexes which helped us separate potential outcomes in our model based on different identities, as well as define the frequency of contracting melanoma. However, the risks calculated by the SEER were only based on the SEER 22 areas, states where SEER collected its data from. Despite this, we reasoned that these 22 areas already represented demographically diverse locations and thus we extended its applications to the rest of the US population in our model. SEER is a reliable and authoritative source for US cancer statistics and publishes comprehensive information in an effort to reduce the cancer burden on the US population.

Because one of the leading contributors to contracting melanoma is excessive sun exposure, we obtained data on UV index from the Centers for Disease Control and Prevention's National Environmental Public Health Tracking Network (CDC) which provided exhaustive data on over 3,000 US County's UV indexes throughout all months from 2005-2015 [25]. While we were unable to find more recent data, this was still a very reliable source helping us understand historical patterns, separate potential outcomes, and predict future trends.

In addition to excessive sun exposure, exposure to artificial sources of UV radiation like tanning lamps and beds account for a substantial fraction of US melanoma cases. However, we were unable to find data that demonstrated in some way the effect tanning beds had on the risk of contracting melanoma, and nor were we able to find data on the frequency of tanning bed usage in the US. As a result, we made an assumption to consider this (Assumption 1). More generally, although we found the strength of UV radiation in over 3,000 counties, we also wanted to find data to understand how much time people were exposed to this radiation. We thought that factoring occupations would be

a helpful way to understand levels of sun exposure, however, we were unable to find reliable data comparing occupations to their level of sun exposure in order to build a comprehensive model.

3.2 Shadows Analysis

Because the core of melanoma cases is a result of excessive sun exposure, we also wanted to investigate shadows produced by infrastructure that block sun rays. However, because we were unable to find substantial data on shadows across the entire US, we used Google Earth Pro [13] to simulate shadows in several select areas like Manhattan, New York, and Cambridge, Massachusetts at different hours of the day, and different parts of the year.

Unable to gather images for all counties, we used population density as a proxy for the shadow coverage for the rest of the counties. We used the trend of the amount of shadow measured in the seven cities we looked at and applied that to the entire US. We then gathered each county's population density data from US Census [15], an evidently reliable source. This allowed us to relate population density with shadow percent and predict the frequency of melanoma in different areas of the country. Google Earth is used universally in daily scenarios and is generally very reliable for developing quick and accurate geographical understandings of our surroundings. Google Earth and Google Maps helped us separate potential outcomes as we obtained data for several cities that had distinct shadow data.

3.3 ArcGIS

Because our model is highly geographically oriented, we used ArcGIS's mapping capabilities to help us easily visualize, analyze, and understand relationships and patterns in our data [1]. ArcGIS's mapping capabilities helped us map trends in for all of our model's factors across over 3,000 US Counties. We combined our county data with the USA County Boundary data provided by Esri, ArcGIS's parent company, to help us create choropleth maps. These maps visually represented the frequency of each factor depending on the data collected from each region in the US through a range of light and dark colors.

4 Mathematics Methodology

4.1 Definitions

1. **Urban:**
   is defined as a community with a population density of greater than 1,500 people per square mile.
2. **Rural:**
   is defined as a community with a population density of fewer than 500 people per square mile [9].
3. **Suburban:**
   is defined as a community with a population density between 500 and 1,500 people per square mile.
4. **CONUS:**
   is an abbreviation for *Contiguous United States* which refers to the lower 48 states in North America (including the District of Columbia) [7].

5. **Choropleth Map:**
   is a type of map used to visualize geographic patterns across different defined boundaries by a palette of colors[5].
6. **Level of Urbanization:**
   is our self defined term that describes how urbanized an area is based on the population and infrastructure density.

4.2 Assumptions and Justifications

Assumption 1: Ultraviolet Radiation is the predominant cause of malignant melanoma.

Justification: While some sources of artificial UV radiation (tanning booths, mercury vapor lighting, certain halogen, fluorescent, and incandescent lights, and some types of lasers)[37] are present, they are difficult to quantify.

Assumption 2: Average state elevation is an accurate representation of its county's elevation.

Justification: Due to the lack of elevation data for each county, we are unable to provide extreme precision on the elevation's effect on the county's risk. Therefore, the average elevation of its state will do.

Assumption 3: An individual's genetic or hereditary influence on large-scale melanoma incidents is insignificant.

Justification: Though it was proven that individuals with a family history of melanoma have a higher probability of getting it[11] [23] [30], it is not ubiquitous enough on a nationwide scale. In addition, similar to the reasoning in Assumption 1, it may be hard to quantify and deal with a highly probabilistic nature.

Assumption 4: UV Exposure is strictly proportional to melanoma risk.

Justification: It is widely known that the sun provides Vitamin D which offers a breadth of healthy benefits[43], but if overly exposed, it can be damaging. Because the healthy dosage of sun exposure varies heavily between individuals[29] and thus tremendously hard to measure, we assume a strict positive correlation. In other words, the more UV exposure, the more risk of getting melanoma.

Assumption 5: Our model is limited to the Contiguous United States (Definition 4).

Justification: During our data exploration, we found that some datasets included non-CONUS regions (Alaska, Hawaii, Puerto Rico, etc.) and some that did not. For ease, we solely focus on CONUS for our model.

Assumption 6: Weather and climatic factors are negated.

Justification: Different weather, namely cloudy, rainy, and sunny days can affect a region's UV exposure. Likewise, climatic factors, such as a shrinking stratosphere can affect exposure calculations[38]. Due to the chaotic and long-term nature of these factors, we do not include them in our model.

4.3 Model Overview

Our model combines the six factors: Age, Ethnicity, Biological Sex, UV Radiation, Level of Urbanization, and Elevation. Integrating multiple regression techniques, image analysis, ratio multiplication, and ArcGIS with our provided data sets, we can visualize geographical risk patterns and estimate the multitude of cases across counties. Throughout the model, we consider demographical trends; they are briefly explained in the Mathematics Methodology section, but scrutinized further in the Risk Analysis section. The following diagram (Figure 1) outlines that process.

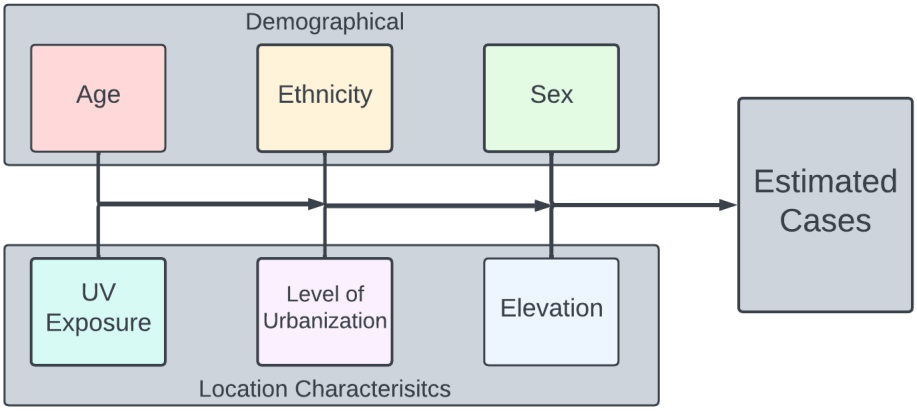

Figure 1: Model Overview Diagram

Table 1 below illustrates the key variables that will be used in this model.

Variables	Description	Units or Values
P_E	Population percentage of specified ethnicity	%
P_S	Population percent of specified sex	%
P_A	Population percentage of specified age group	%
$R_{S,w}$	Overall risk by Sex	unitless
$R_{E,w}$	Overall Ethnicity Risk	unitless
$R_{A,w}$	Overall Age Risk	unitless
R_{UV}	Risk from UV index	unitless
R_{sh}	Risk from shade	unitless
A	Altitude	feet
S	Average Shade Percentage	%
D	Population density	people per square mile

Table 1: Variables

4.4 Age

Aging is one of the leading causes of melanoma with the ages 85+ being many hundred times more susceptible than adolescents[21] [31]. Through this section, we aim to grasp the "riskiness" of a county based on age group populations. By doing this, we can establish which US counties are like to have more melanoma incidents based on age.

Using data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program[21], we modeled the correlation between Age and Risk of contracting melanoma. Their data provided the risk of being diagnosed with melanoma skin cancer solely based on age with the samples from 2017 to 2019.

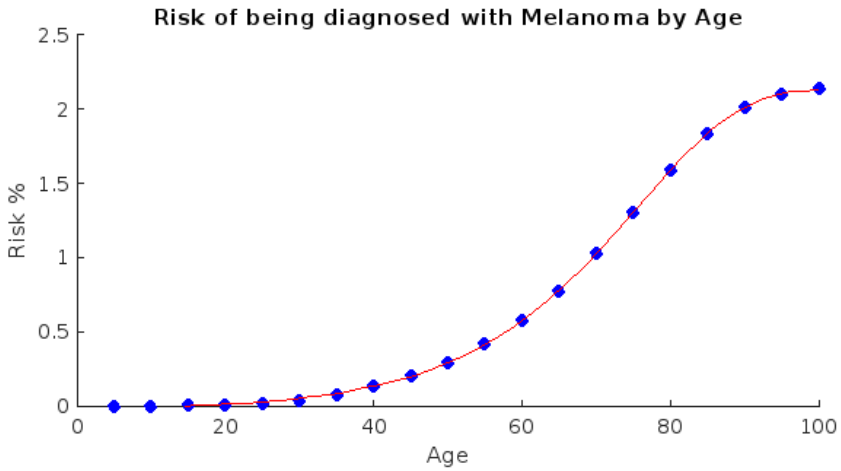

Figure 2: Lifetime Risk of Melanoma Diagnosis

Because the data resembles a logistical growth, we regressed the data using the Gompertz Sigmoidal Curve, a function developed by mathematician Benjamin Gompertz. The regression curve is shown above, and its equation is below.

$$R_{age}(T) = 2.1406 - 2.1675e^{-e^{0.07199(T-75.6635)}} \tag{1}$$

where R_{age} is the risk of being diagnosed with melanoma as a function of T, a specific age. This fit was deemed highly accurate with an R^2 value of approximately 0.99. Note that this function is only valid for ages up to 100, which is reasonable as less than one percent of the US population lives past 100[21].

Performing integral averaging on Equation 1 (represented in Equation 2), we found the average risk estimates for each of the age groups listed in table 2.

$$R_A(T_0, T_f) = \frac{1}{T_f - T_0} \int_{T_0}^{T_f} R_{age}(T)dT \tag{2}$$

where T_0 is the lower bound of the age group and T_f is the upper bound. Note that for the age group 85+ we defined the upper bound to be 100.

Risk	
Age Groups	Risk (%)
0-4	≈ 0
5-9	≈ 0
10-14	≈ 0
15-19	0.0045
20-24	0.0182
25-29	0.0375
30-34	0.0649
35-39	0.1034
40-44	0.1574
45-49	0.2323
50-54	0.3347
55-59	0.4719
60-64	0.6505
65-69	1.0623
70-74	1.1358
75-79	1.4195
80-84	1.6918
85+	7.6408

Table 2: Age Group's risk ratings

Combining age group risks (table 2) with the age group breakdowns in each US county, we can determine which counties are more "risky" than others. Using the US Census's Intercensal Datasets with the provided age groups' population estimates for each county, we used a weighted average equation (Equation 3) to synthesize the overall risk for each county.

$$R_{A,w} = \frac{\sum_{k=1}^{18}(R_A)_k \cdot (P_A)_k}{\sum_{k=1}^{18}(R_A)_k} \tag{3}$$

where R_A is an age group in (table 2), and P_A is the corresponding age group population percentage. Knowing age breakdowns may adapt over time, we conducted first-order regression on each county's age group time series data (2010-2015) to determine 2020 and future breakdowns. The resulting 2020 geographical distribution can be visualized in the ensuing choropleth map (Figure 3). These trends will be further discussed in our Risk Analysis Section (Section 5.5).

4.5 Ethnicity

The risk of contracting melanoma is also influenced by one's ethnicity. In fact, melanoma is over 25 times more common in fair-skinned individuals than those with darker skin phenotypes. The lifetime risk for Non-Hispanic Whites is 2.6 % compared to 0.1 % for Non-Hispanic Blacks and 0.6% for Hispanics[18]. Thus, it's important to consider race when considering who and where is most vulnerable.

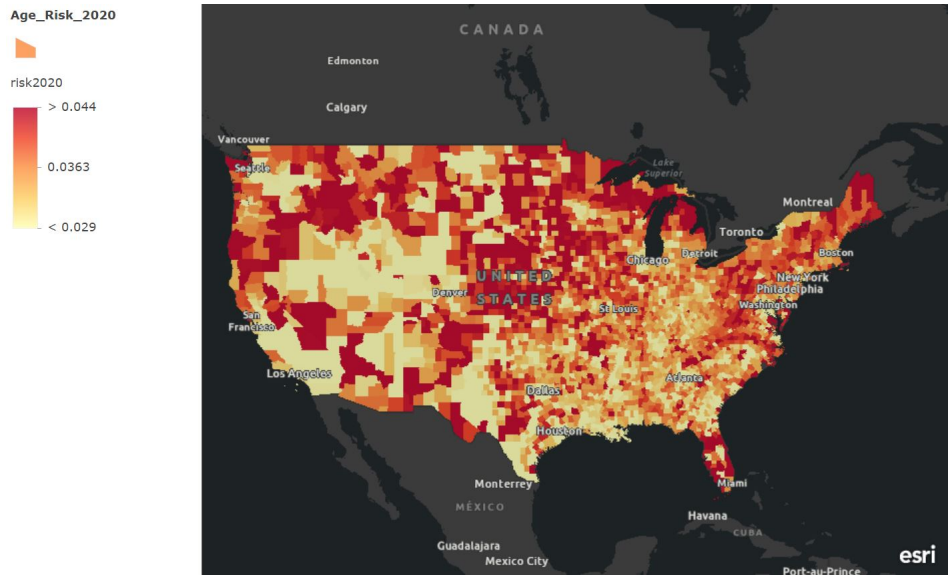

Figure 3: Choropleth Map of Age Risk in 2020

Race / Ethnicity	Risk (%)
Hispanic (any race)	0.5047
Non-Hispanic American Indian / Alaska Native	0.6871
Non-Hispanic Asian / Pacific Islander	0.1495
Non-Hispanic Black	0.0904
Non-Hispanic White	2.9478

Table 3: Risk rating by Race / Ethnicity

Amalgamating the US Census's Intercensal Datasets for each ethnicity's (Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Asian / Pacific Islanders, Non-Hispanic American Indian / Alaska Native, and Hispanic) population estimates and NIH's SEER (Surveillance, Epidemiology, and End Results program) risks of melanoma diagnosis[21] (presented in table 3), we can enumerate a county's average risk as shown in Equation 4 below.

$$R_{E,w} = \frac{\sum_{k=1}^{5}(R_E)_k \cdot (P_E)_k}{\sum_{k=1}^{5}(R_E)_k} \tag{4}$$

where R_E is an ethnicity group's risk and P_E is the corresponding population percentage for $k = \{1, 2, ...5\}$, which corresponds with the 5 different ethnic groups as shown in table 3.
We then performed linear regression using the data from 2010-2015 to forecast each US county's risk for 2020 and 2025 (further analysis will be done in Section 5.5). Using ArcGIS's mapping capabilities, we mapped the risk distribution across CONUS (Definition 4) for 2020 as shown in the choropleth map below (Figure 4).

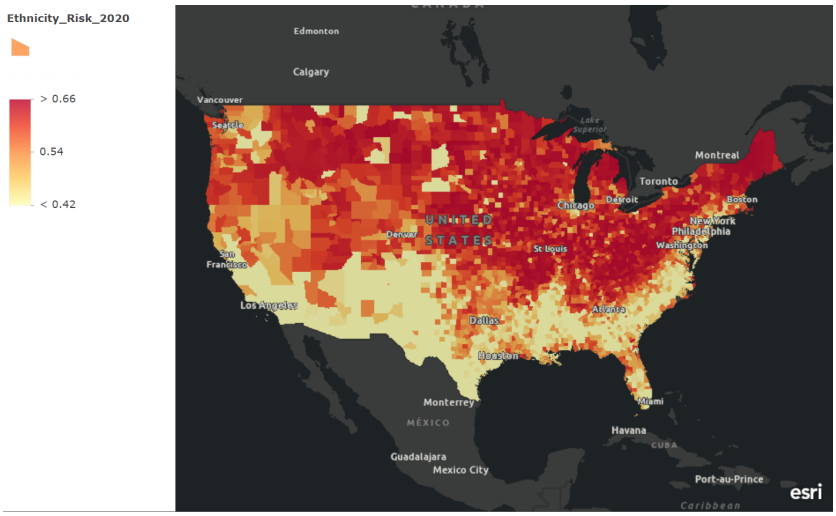

Figure 4: Choropleth Map of Ethnicity Risk in 2020

4.6 Biological Sex

Biological Sex also plays a crucial role in susceptibility to contracting melanoma. In fact, 7,990 people are estimated to die from melanoma in 2023; 2,570 of them are women, and 5,420 are men[18].

Sex	Risk (%)
Female	1.7066
Male	2.6078

Table 4: Risk rating by Sex

Repeating a similar process as age and ethnicity for sex, we can determine the risks across all CONUS counties. Using the risks provided by NIH's SEER program as shown in table 4, we calculated the risk for each county throughout 2010-2015 using the equation for the weighted average formula below:

$$R_{S,w} = \frac{R_m \cdot P_m + R_f \cdot P_f}{R_m + R_f} \tag{5}$$

where R_m and R_f reflect SEER's risks for males and females respectively, and P_m and P_f represent their respective population percentages. Using the calculated risks for all counties throughout 2010-2015, we performed linear regression to forecast risk values for 2020 and 2025 (further analysis will be conducted in Section 5.5).

Finally, we mapped the results for 2020 using ArcGIS as illustrated below in Figure 5.

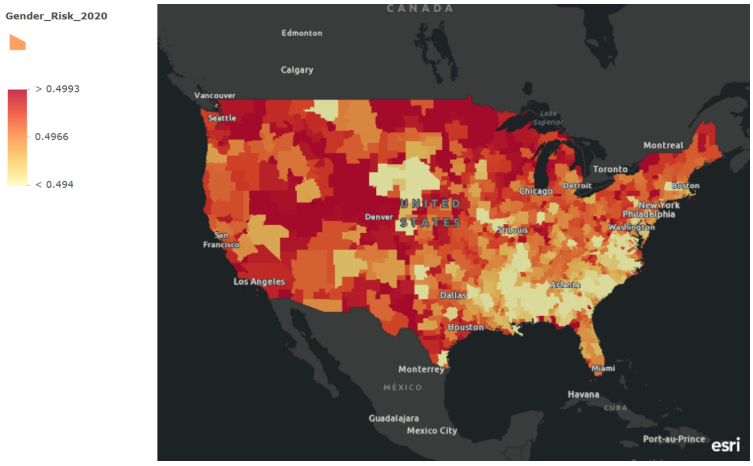

Figure 5: Choropleth Map of Sex Risk in 2020

4.7 UV Radiation

Due to the immense land area of the US, UV Radiation, and thus melanoma risk, varies heavily between different regions. In this section, we aim to quantify UV exposure risk on a UV index scale[41].

To begin, we manipulated data from the CDC[25]. Specifically, the monthly average UV irradiance in $\frac{mW}{m^2}$ at noon from 2005 to 2015, for each of the CONUS's 3000+ counties. Because the data's units were in $\frac{mW}{m^2}$ we first converted it to the UV index scale (1-11+)[28], a more universally interpretable metric. To calculate the UV index, we divided the UV irradiance $\frac{mW}{m^2}$ by the constant $\frac{mW}{m^2}$, defined by the UV index calculation[42].

Subsequently, we garnered monthly time series data for each county and performed sinusoidal regression with a sum of two sines which we discerned to have more favorable R^2 values. Even though climate change and other evolving atmospheric patterns impact UV intensity[38], the sinusoidal property is assumed to persist in future years by assumption 6. The corresponding results for one of the 3000+ CONUS counties (Manhattan, New York) are exhibited in Figure 6 and Equation 6:

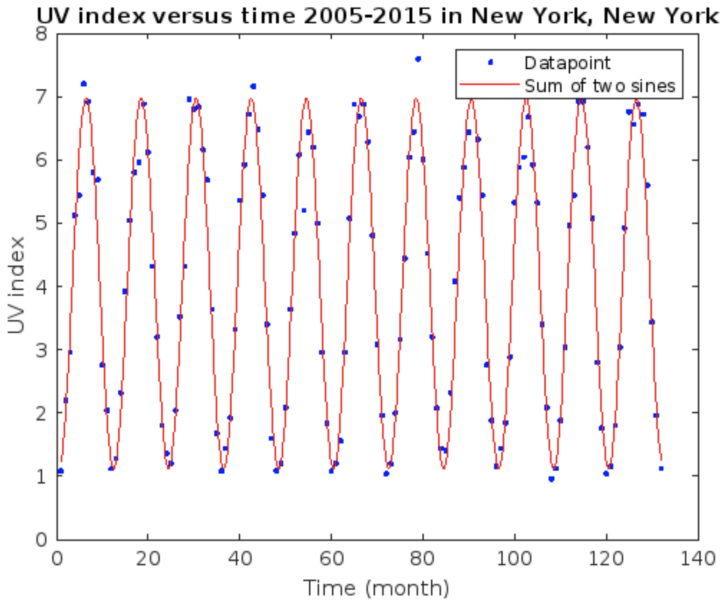

Figure 6: UV Index from 2005-15 for New York County

$$R_{UV}(m) = 4.1250 \cdot \sin(0.0001m + 1.7656) + 2.9232 \cdot \sin(0.5232t - 1.8126) \tag{6}$$

where R_{UV} is the UV risk variable and m is time in months with $m = 1$ equating to January 1st, 2005, $m = 13$ as January 1st, 2006, and so forth. Bear in mind that the range of the y-axis is 1-11+ which parallels the universal UV index scale.

Disregarding the periodicity attribute for now (which will be revisited in Section 5.4.2), we calculate the long-term UV risk using the integral average formula, which outputs a scalar value. Applying this evaluation to each CONUS county, we acquire the following map (Figure 7):

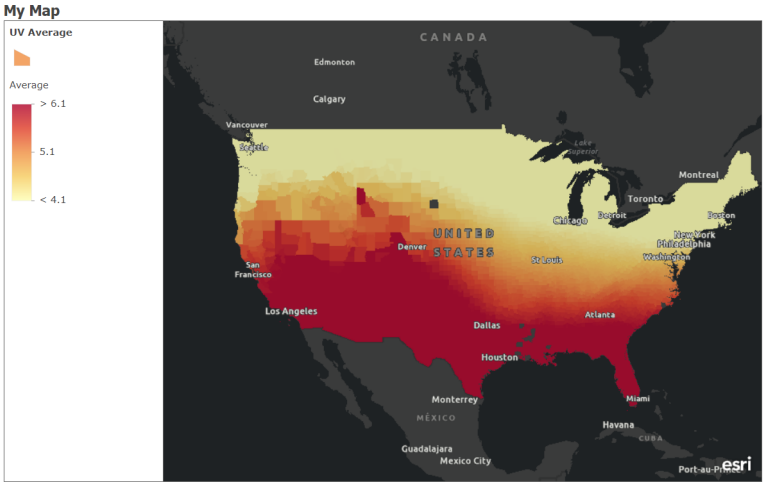

Figure 7: Choropleth map of average UV indexes of CONUS

4.8 Level of Urbanization

4.8.1 Overview

The next location characteristic we model is the level of urbanization (Definition 6). Intuitively speaking, residents of an urban city (high level of urbanization) tend to be less exposed to UV radiation than their rural counterparts (low level of urbanization). To quantify this relationship, we use population density as a proxy for level of urbanization, and thus the amount of time absent from UV exposure. In this section, we will be using Manhattan, New York, as the paradigm.

As a consequence of the lack of "shadow data," we developed an innovative way to go around this barrier: conducting image analysis with MATLAB. The following figure (Figure 8) outlines that process.

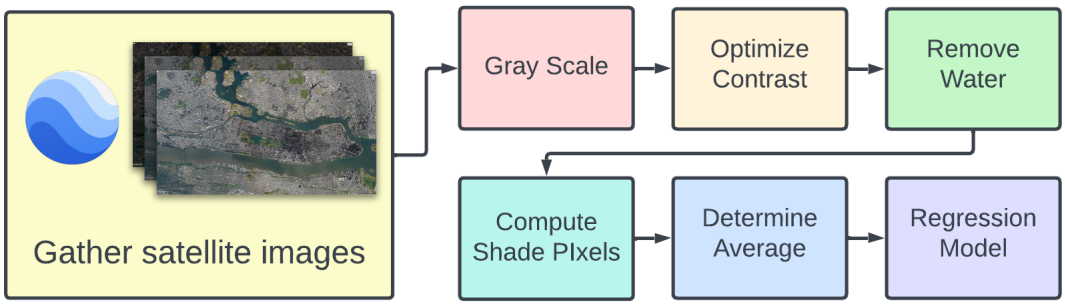

Figure 8: Process for computing and modeling the average shadow percentages for satellite images

4.8.2 Image processing

First and foremost, we extracted hundreds of 3840 x 2160 pixel satellite images from Google Earth pro[13]. Specifically, we collected hourly images for 12 months in a mix of urban, suburban, and rural regions: Manhattan, New York; Boise, Idaho; Cambridge, Massachusetts; Inglewood, California; Chicago, Illinois; Bandera, Texas; and Adair, Oklahoma.

With these images, we utilized MATLAB's image analysis capabilities. To begin, we gray-scaled the 3-dimensional RGB arrays into a 2-dimensional array with pixel values between 0 and 1. However, simply gray scaling does not provide proficient information due to inconsistent contrasts. To solve this, we optimized the contrast to distinguish false shade pixels (alike in color) from "true" shade pixels. Manhattan's example of this is shown below (Figure 9):

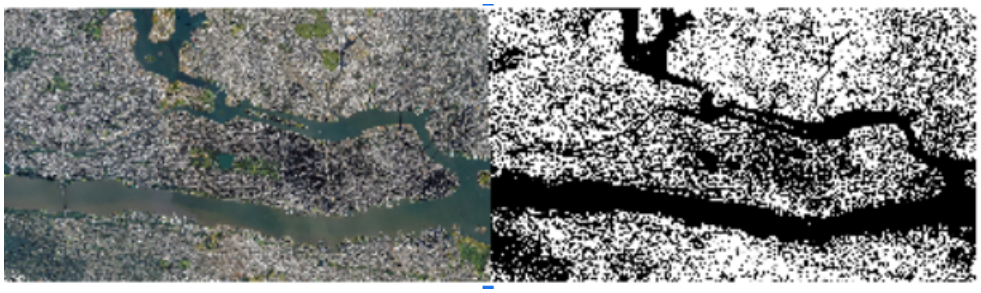

Figure 9: Manhattan satellite image (left) and its grayscaled and contrast-optimized image (right) at 2PM in January

We summed up the number of black pixels but deducted the black pixels created by the bodies of water (in Manhattan's case, the Hudson river) from the calculation. With this number, we calculated the percentage of shade using the following equation (Equation 7),

$$\text{Shade Percent} = \frac{\text{Black Pixels} - \text{Water Pixels}}{8294400} \times 100 \tag{7}$$

where 8294400 is the number of pixels in a 4K resolution (3840 x 2160) image. Rerunning this process for each image, we obtain the following plot:

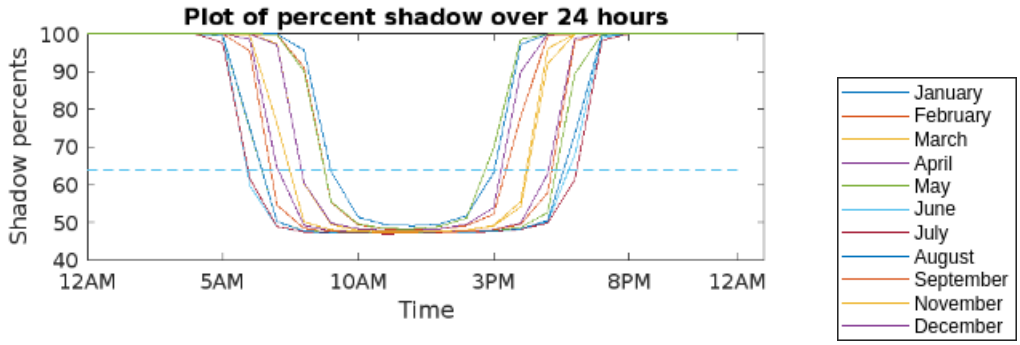

Figure 10: Shadow percent of Manhattan every month

where the horizontal dashed blue line represents the average shade percentage during daylight hours. This average will be used in the regression model in the next section.

4.8.3 Regression

As stated in 4.8.1, we established population density as a proxy for shade. Therefore, to demonstrate a correlation, we reran the procedure described in Section 4.8.2 for Inglewood, California; Boise, Idaho; Cambridge, Massachusetts; Bandera, Texas; Adair, Oklahoma; and Chicago, Illinois. The results are displayed below (Table 5):

Places	Density(people per square mile)	Shade percentage
Manhattan, NY	69,464	63.8
Inglewood, CA	11,885	42.9
Cambridge, Massachusetts	18,520	52.3
Boise, Idaho	2804	40.5
Bandera, Texas	778.55	26.7
Adair, Oklahoma	143	24.31
Chicago (The Loop district)	26700	50.7

Table 5: Sample cities

When testing multiple regression types, we found that a power regression was the strongest R^2 value, illustrated below.

$$S = aD^b + c \tag{8}$$

where $a = 5.085$, $b = 0.214$, and $c = 8.770$, and an R^2 value of 0.9612. D represents a community's population density and S is shade percent.

To test the accuracy of our regression model, we found the shade percentage of Chicago Loop (population density 26770), a neighborhood in Chicago, and compared it to the regression equation (Equation 8). The experimentally found average shade percentage was 50.7%, a mere 3.1% deviation from the predicted value (53.8%), depicted in Figure 11 below:

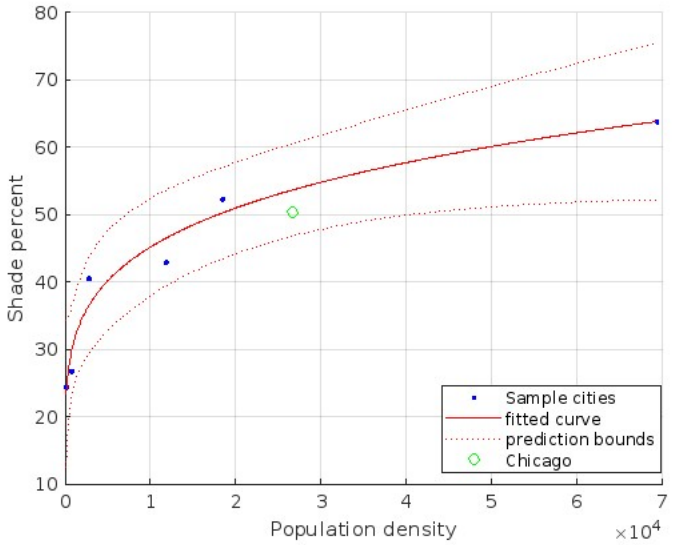

Figure 11: Graph of population density versus shadow percentage with Chicago point

Now that we were confident that there was a correlation between population density and shade percentage, we applied our equation to each US County's population density. Then we mapped the resulting shade percentages (Figure 12).

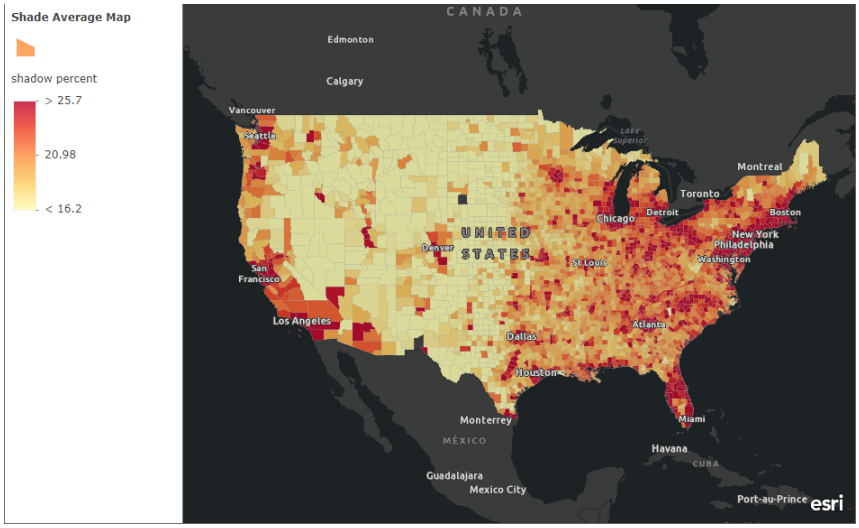

Figure 12: Map of shadow percentages and population density across CONUS

The map shows there is a correlation between densely populated areas and higher shade percentages. This means that rural areas experience much higher levels of sun exposure than urbanized areas. What we take away is that western rural areas, like Nevada and North Dakota, are many times more susceptible to UVR than suburban and urban regions.

Since shade percentage and melanoma risk are inversely correlated, we use the following equation (Equation 9) to get the shade risk rating R_{sh} which will be used later in Section 4.10.

$$R_{sh} = \frac{1}{S} \tag{9}$$

4.9 Elevation

Using state elevation profiles[12] and Assumption 2 regarding county elevation data, we established elevation profiles for each county as illustrated in the following bar plot (Figure 13). The altitude for a state is stored in the variable R_{Al} and R_{US} is the average US elevation.

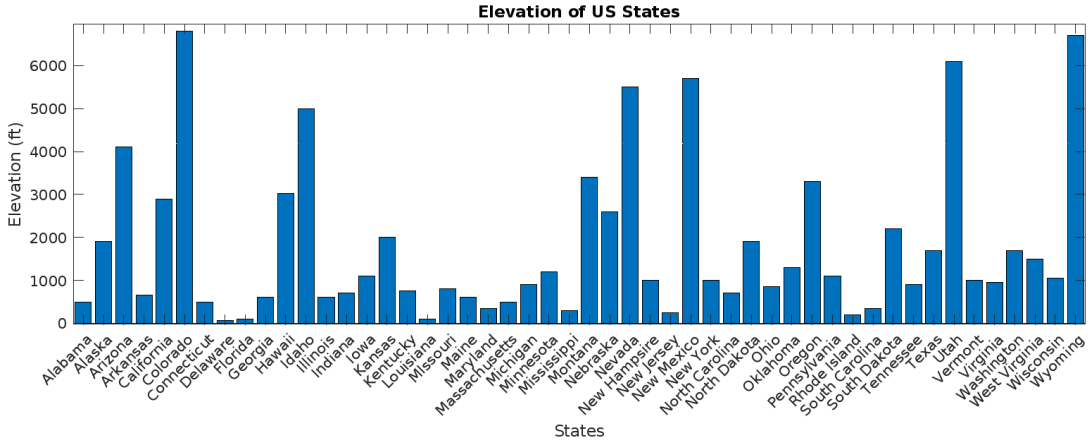

Figure 13: Mean Elevation of US States

4.10 Results

Since each of the six factors has a unique scale, we aggregate them by taking advantage of county-to-nation ratios calculated by Equation 10.

$$I = I_{US} \cdot \prod_{i=1}^{6} \frac{(R_C)_k}{(R_{US})_k} \tag{10}$$

Or expanded,

$$I = I_{US} \left[\frac{(R_{A,w})_C}{(R_{A,w})_{US}} \cdot \frac{(R_{E,w})_C}{(R_{E,w})_{US}} \cdot \frac{(R_{S,w})_C}{(R_{S,w})_{UC}} \cdot \frac{(R_{UV})_C}{(R_{UV})_{US}} \cdot \frac{(R_{sh})_C}{(R_{sh})_{US}} \cdot \frac{A_C}{A_{US}} \right] \tag{11}$$

where I is a county's estimated annual number of new melanoma cases per 100,000 population, I_{US} is the annual number of new melanoma cases per 100,000 population throughout the entire country [2], R_C, and R_{US} is the risk rating for the county and US respectively. R_{US} is calculated by using the same formula but US-specific inputs.

4.11 Strengths and Weakness

4.11.1 Strengths

An apparent strength of our model is its ability to distinguish the risks of rural, suburban, and urban environments expressed in Section 4.8. Because our model takes raw satellite imagery and analyzes it using MATLAB's image processing, we retain maximum "rawness" that could otherwise be altered (not necessarily in a bad way) by a third party's behind-the-scenes data procedures. Being able to circulate through hundreds of images, optimize the contrast, and eliminate non-walkable areas (e.g. bodies of water), proved accurate by a high R^2 value of 0.9612 and Chicago's loop district diminutive 3.3% offset from the predicted value. As shown in Figure 10 from December to January, the amount of shade is almost always 50%, matching the number from New York Times's publication [19].

As entailed by our model, we found that Utah and Colorado have some of the highest melanoma cases in the country, which aligns with the other findings [6] [33]. Our model estimates in 2020 that one county in Utah has 37.26 new cases per 100,000 which are relatively near the recorded 43.4 in Utah [3].

4.11.2 Weaknesses

Due to the lack of recent county demographics as described in Data Methodology (Section 3.1), we had to use the 2010-2015 intercensal county estimates [40] for predictions and risk insofar as possible. As a result, risk ratings for Age ($R_{A,w}$), Ethnicity ($R_{E,w}$), and Sex ($R_{S,w}$) distributions may not accurately reflect measured results.

Assumptions 1-4 can explain some limitations and disparities in our model. For instance, we were unable to account for the effect of artificial tanning beds. While not as potent as solar UVR, tanning beds are widespread across states. These states have varying laws, thus our model may not fully reflect a state's risk. Likewise, for elevation (Assumption 2), mountainous regions like Colorado which have high average elevation may not accurately resemble residential areas since the majority do not live in the mountains.

5 Risk Analysis

5.1 Risk Overview

As mentioned in the executive summary, the reach of this model is to provide a mathematical instrument to uncover the full story behind the risks of melanoma of the skin. This section will digest the model parameters with the objective to examine the frequency, geographical patterns, trends, and driving factors.

5.2 Risk Characterization

In this section, we aim to determine the influence each factor brings, and chiefly, what the driving factors are. To grasp the magnitudes of each factor we conduct the following sensitivity analysis by varying each factor by an interval of ±15%. The outcomes of the changes are shown in table 6. Note that elevation and UV are directly proportional (1:1 ratio), so it is not included in the table.

Variable	+15%	-15%
Ages 65+	+1.785%	-1.785%
Ages 40-60	-0.278%	+0.2788%
Ages 0-40	-1.411%	+1.411%
Male (P_m)	+2.707%	-2.707%
Female (P_f)	-2.707%	+2.707%
Non-Hispanic White	+3.436%	-4.278%
Non-Hispanic Black	-3.246%	+3.221%
Non-Hispanic American Indian/Alaska Native	-0.057%	+0.053%
Non-Hispanic Asian/Pacific Islander	-0.173%	+0.161%
Hispanic	-0.369%	0.346%
Population Density (Rural)	+2.085%	-2.347%
Population Density (Suburban)	+2.231%	-2.512%
Population Density from (Urban)	+2.535%	-2.854%

Table 6: Sensitivity analysis

To get a thorough understanding of the impact, we repeated the calculations on each county and found small variability. All factors considered, we found that the estimated number of melanoma cases does not hugely change the result, but still some factors are more influential than others. If we seclude age for this analysis, we find that an increase or decrease in the elderly population is 6 times more influential than that of the age group 40-60. Conducting a similar analysis, we deduce a larger percentage of Non-Hispanic Whites drives cases up, while an increase in Non-White population drives the number down.

If we hone into the male and female numbers, we notice that they are the second most influential factors. It is important to realize that while this is significant, numbers from Equation 5 reveal

that the population percentage disparity between male and female never exceed 1%.

Since UV and elevation have a 1:1 ratio, a 15% increase would result in a 15% increase, we conclude that the UV and elevation factors are the primary driving factors to melanoma risk. Colorado's case of high elevation exemplifies this and aligns with other findings [34].

5.3 Frequency and expected value

Though we don't consider severity (costs) as stated in our problem statement (Section 2.2), this section will examine the frequency and expected value of this model in terms of the number of melanoma cases. To elucidate the prevalence of melanoma cases, we can visualize the distribution below (Figure 14).

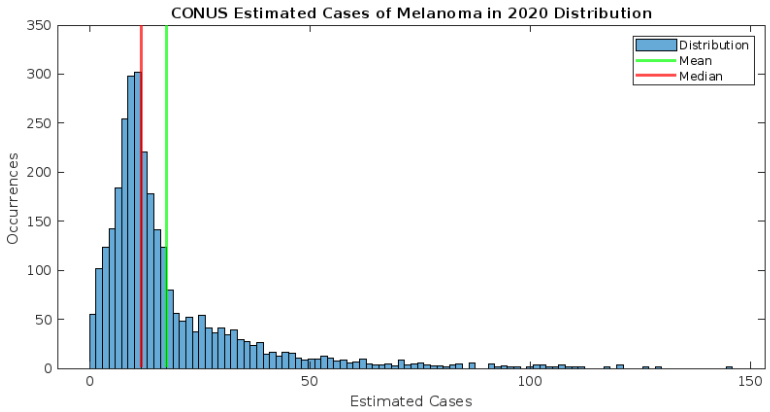

Figure 14: Estimated Cases of Melanoma in CONUS

Mean	Median	Skew
17.4897	11.8615	1.0009

Table 7: Mean, median, and skew

As manifested in Figure 14 and Table 7, the majority of counties tend to have a case count around the mean and median. The significantly larger cases can be attributed with high ratios in all factors. Furthermore, we can visualize this in the below Figure (Figure 16).

1 Year	5 Years	10 Years
52451	262255	524510

Table 8: Estimated melanoma cases in CONUS for 1, 5, and 10 years

Drawing from the aforementioned equation (Equation 11) and distribution (Figure 14), we find that there are 52-, 262-, and 524-thousand new cases after 1, 5, and 10 years respectively (Figure 8). Though the number may feel minuscule in the context of the US's 300+ million population, the accumulation of cases and costs for treatment can rapidly tower.

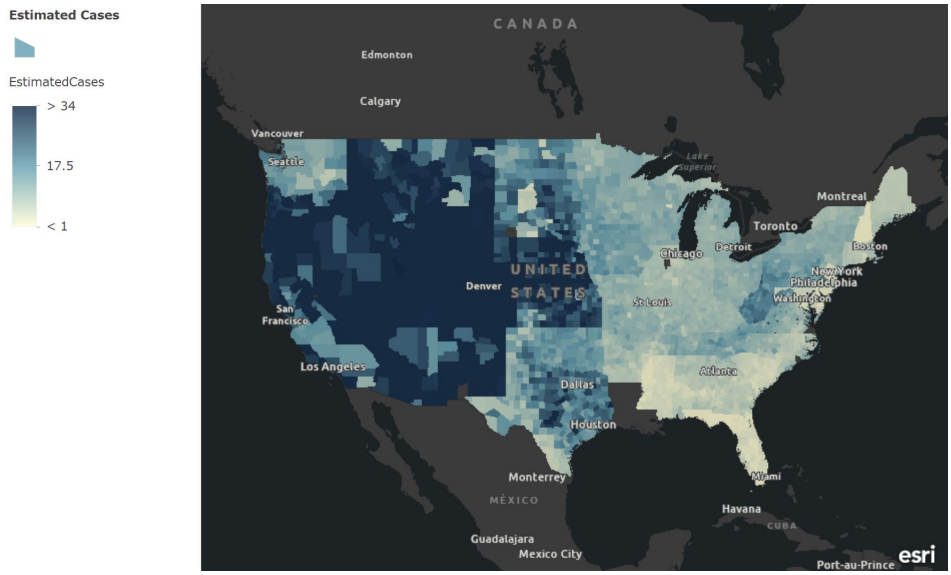

Figure 15: Estimated Cases of Melanoma in CONUS in 2020

5.4 Distribution

5.4.1 Risk distribution by level of urbanization

To grasp a thorough representation of the impact of the level of urbanization (Section 4.8), we analyze its distributions. Using Definitions 1-3, we can categorize which intervals of shade percentage are rural, suburban, and urban as displayed in the following figure (Figure 16).

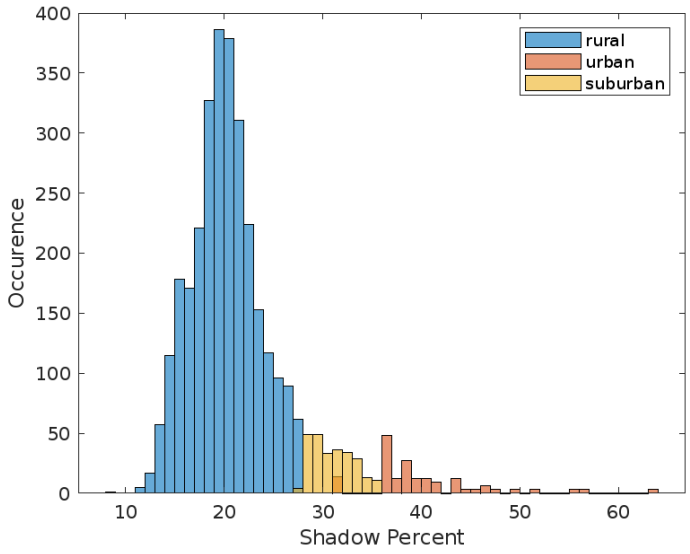

Figure 16: Occurrences of shadow percentages for rural, suburban, and urban communities

To quantitatively understand the distribution, refer to the following table (Table 9):

	Rural	Suburban	Urban
Population Percentage of US	21%	52%	27%
Mean	20.023%	31.053%	40.080%

Table 9: Distribution of US population over rural, suburban, and urban areas

We find that suburban areas (Definition 2) account for 52% of the population but only 31 % of melanoma cases. In comparison, rural regions account for 33% of the US population but only 20.023 % of melanoma cases (with a deviation of 3.929%). Therefore, we conclude that it would be most effective to prioritize suburban regions first and rural regions after.

5.4.2 Times of greatest risk

Although melanoma is not contracted immediately, since higher UVR exposure leads to higher melanoma of the skin cases, it is within reason for us to infer that exposure to certain months and hours of greatest exposure will generally produce more risk. In view of this, this section will deduce the months and hours with the greatest risk of melanoma.

To analyze months, recall Section 4.7 on UV Radiation and Section 4.8 on Level of Urbanization. As seen in Figure 6, the months with the highest risk (above the midline) are when $m = \{5, 6, 7, 8\}$ ($m = 1$ is January), or in other words, May, June, July, and August. While the specific fraction of the months may vary between counties, we proved that nearly all of the counties followed the same pattern.

Subsequently, we look at the hourly risk; specifically, the times of day when sun exposure is most potent. Recall Section 4.8 (level of urbanization) and Figure 10 which illustrates hourly shade percentage throughout a year. Although Figure 10 only represents Manhattan, when conducting similar procedures on all counties, we found 10AM-4PM (in their respective timezones) had the lowest shade percentage. Coupled with stronger UV radiation from shorter irradiance travel distance[35], these times are much more hazardous.

5.4.3 Other Distributions

Through the Mathematics Methodology and Analysis section (Section 4), we generated numerous choropleth maps to help visualize distributions. Since the distributions are straightforward to interpret, we are only going to provide a brief takeaway for each.

1. **Age Distribution (Figure 3):** distribution is very sporadic with no clear pattern.
2. **Ethnicity Distribution (Figure 4):** States with a higher percentage of fair-skinned individuals, typically northern states, are riskier. Lower California, Arizona, New Mexico, Texas , Louisiana, Mississippi, Alabama, Georgia, and Florida tend to have lower risks.
3. **Biological Sex Distribution (Figure 5):** Likewise to age, no clear patterns, however, the southeastern region has lower risks.

4. **UV Radiation (Figure 7):** Clear pattern; the more south we travel, the more UV risk associated.
5. **Elevation (Figure 13):** No spatial trends, however, states on the rocky mountain range have much higher elevations than other states.

5.5 Trends

As mentioned in sections 4.4 (Age), 4.5 (Ethnicity), and 4.6 (Biological Sex), we used linear regression to not only determine 2020's data, but also the forecasted predictions. We determined future values of demographic variables using first-order regression and found that there was not a significant difference between the years. The minimal risk values for Sex (left) and Ethnicity (Right) are shown below (Figure 17):

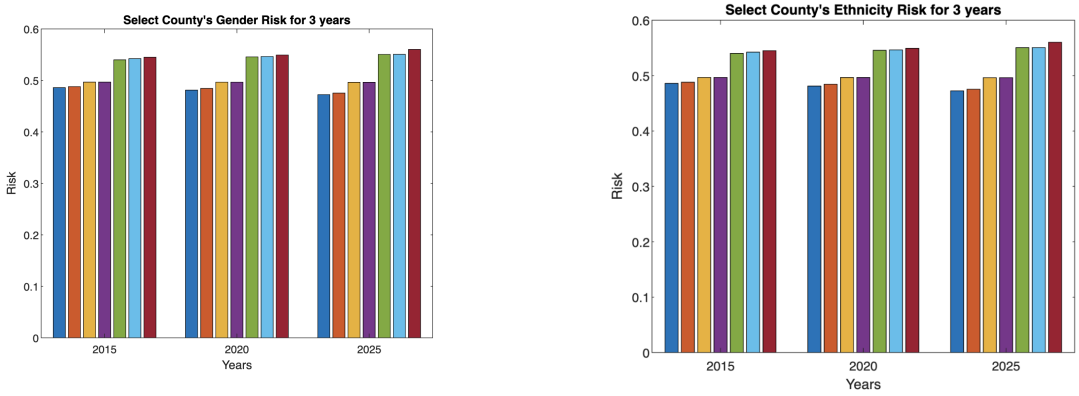

Figure 18: 2015-2025 Trends for Gender and Ethnicity

Thus, we conclude that while there are trends present, the minimal change within a 5 year duration makes the trends insignificant.

6 Recommendations

6.1 Farmers

Farmers, livestock producers, and the agriculture industry contribute to the core of melanoma statistics; many farmers work in long hours in direct sun during the sunniest times of the year. Particularly, predominantly rural states like North Dakota, South Dakota, and Montana have limited shadow coverage and a very high number of farms as supported by Figures 12 and 19 [26]. Thus, farmers need a way to mitigate the risks of significant sun exposure.

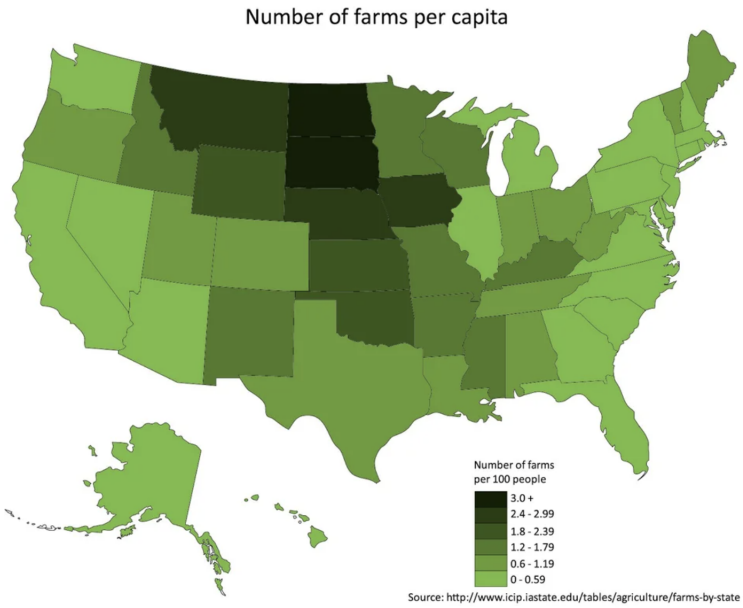

Figure 19: US Farm Distribution

We recommend that the farming industry implements more protective measures to protect their farmers. This could include encouraging farmers to wear sun-protective clothing like hats, sunglasses, and lighter-colored, long-sleeved shirts and pants, being persistent in always wearing sunscreen (even on cloudy days), and optimizing their schedule in such a way that allows some tasks to be completed during hours when the sun is not as intense (generally late afternoon and early morning rather than between 10 AM - 4 PM) as proven by our Level of Urbanization Model (Section 4.8) and shown in Figure 10. Because our model shows that the majority of farmers are located in higher-risk areas compared to other office jobs industries (Figure 12), the farming industry will benefit from having more accessible melanoma prevention events and treatment, like screening, in order to detect melanoma as early as possible. Encouraging awareness of melanoma, its warning signs, and self-examinations in this industry could go a long way.

6.2 Sun Awareness

In order to combat widespread melanoma in the US, we encourage the US government to implement sun awareness campaigns which have been found to improve safe sun habits and decrease melanoma cases. Our model identified that the following states are at the highest risk: Utah, Montana, Idaho, Nebraska, Oregon, South Dakota, and Iowa. To be most effective, we recommend a continuous melanoma campaign to be tailored toward the aforementioned states. Moreover, these campaigns should target workers in rural, high-risk occupations, such as those in outdoor industries. In Section 5.4.1, we identified that residents of rural areas experience 50% more UV exposure than urban residents, which is estimated to increase cases by more than 2% (according to our sensitivity analysis in Section 5.2).

To demonstrate the effectiveness of sun awareness, we look at multiple studies. In a study published in Oxford Academic, advertising campaigns in Australia promoting safe sun habits increased sunscreen use by up to 13.3% for children and 9.9% for adults [32]. In another study, three sun safety campaigns were analyzed for effectiveness, and they were found to have saved the

government $40 million dollars in melanoma treatment at a cost of $10 million dollars[10]. With this in mind, we recommend a 20 million dollar per year continuous campaign targeting the aforesaid states and risky profile occupations (like farmers as described in Section 6). Assuming the advertising effectiveness is the same for Australia as the United States and accounting for the difference in melanoma rates, this would save the US government 80 million dollars yearly. Using a cost-to-case reduction ratio of $16454.5 per case[10], we estimate that the campaign would decrease melanoma cases by roughly 1400 every year, a substantial amount.

6.3 Insurance

Our last recommendation is regarding insurance. Insurance policies are not preventative measures, but rather they cover any costs related to damages incurred by the policyholder. Most health insurance covers skin screenings for melanoma, but many policies don't cover melanoma biopsies and many costs related to drugs and treatment[8]. To decrease long-term costs and risk, we recommend that the US government expands the coverage for malignant melanoma treatment in existing healthcare plans like Medicare, especially to those older than 50. As described by our age risk equation (Equation 1) and Table 2, individuals above the age of 50 incur a 0.33% chance of developing melanoma, 7 times greater than that of adolescents. Even worse, as age increases, we find the associated prognosis grows exponentially. To mitigate this risk, one possible method the US government could take is Medicare-covered dermatologist wellness visits to diagnose melanoma as soon as possible. Putting heavy emphasis on prompt diagnosis detection can not only mitigate the fatality risk but also be extremely beneficial when considering costs. According to the World Health Organization, treatment for cancer patients diagnosed early is 2-4 times less expensive than the treatment of advanced-stage cancers[18]. Although early diagnosis is more beneficial to the elderly population, younger individuals may not be fully out of the picture. Research shows that melanoma prognosis for individuals in ages 15 to 39 years old has increased from 4.7 people per 100,000 to 7.7 and will continue to rise. All in all, early diagnosis is a noble cause to mitigate melanoma risk, especially considering the lethality of the disease[44].

7 Concluding Remarks

Like many other cancers, melanoma of the skin continues to pervade every city, county, and country. Despite the past decades of research to fully cure melanoma, humanity may still be decades away. Until then, we must search for and take mitigation measures—what our model aims to achieve. In tandem with our research, our model analyzes the risks and trends associated with varying demographics, UV intensity, elevation, and level of urbanization. Our estimates conclude that investing in communities with high UV index and elevation are top priority, but measures in nationwide rural and demographically risk groups should also be heavily invested in. In particular, western states lie on the rocky mountains, and rural mid-western states, are home to agricultural America. These alleviation recommendations and tactics are crucial in the fight against Melanoma.

8 Acknowledgments

We would like to sincerely thank those who supported us during long hours during our competition:

Thanks to Ms. Widener, our Math teacher, for her unwavering support and encouragement, we could not have finished this model without her invaluable guidance and infectious enthusiasm and energy.

Additionally, we would like to our mentor, Mrs. Witcraft for her incredible response time and invaluable guidance and advice throughout the entire project phase.

Our local public library staff for their endless generosity, patience, and resources, for providing space when school was not available, and for the best form of fuel: Pizza and Oreos.

Finally, we would like to thank the Actuarial Foundation for this invaluable learning opportunity and for allowing us to participate in the 2023 Modeling the Future Challenge. For providing informational webinars, free tools like ArcGIS, and weekly reminders.

References

[1] "ArcGIS Online." *Esri*, 27 December 1999, www.arcgis.com/index.html. Map. Accessed 2 Mar. 2023.

[2] "Cancer Stat Facts: Melanoma of the Skin." *National Cancer Institute*, seer.cancer.gov/statfacts/html/melan.html. Accessed 2 Mar. 2023.

[3] "Cancer Statistics at a Glance." *Centers for Disease Control and Prevention*, Nov. 2022, gis.cdc.gov/Cancer/USCS/CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcancer%2Fdataviz%2Findex.htm#/AtAGlance/. Accessed 27 Feb. 2023.

[4] "Changes Over Time: Melanomas of the Skin" https://gis.cdc.gov/Cancer/USCS Accessed 3 Mar. 2023

[5] "Choropleth Map." *The Data Visualisation Catalogue*, datavizcatalogue.com/methods/choropleth.html. Accessed 28 Feb. 2023.

[6] "Complete Health Indicator Report of Melanoma of the Skin Deaths." *Public Health Indicator Based Information System (IBIS)*, 18 Feb. 2022, ibis.health.utah.gov/ibisph-view/indicator/complete_profile/MelSkiDea.html. Accessed 2 Mar. 2023.

[7] "Contiguous United States, Continental United States, and CONUS." *NREL Transforming ENERGY*, www.nrel.gov/comm-standards/editorial/contiguous-united-states-continental-united-states-and-conus.html. Accessed 28 Feb. 2023.

[8] "Costs and Insurance Coverage for Cancer Screening" *American Cancer Society*, 20 April 2021, https://www.cancer.org/healthy/find-cancer-early/cancer-screening-costs-insurance-coverage.html.Accessed 3 Mar. 2023.

[9] "Economic Research Service." *US Department of Agriculture*, 23 Oct. 2019, https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural.aspx. Accessed 27 Feb. 2023.

[10] Doran, Christopher M., et al. "Benefit Cost Analysis of Three Skin Cancer Public Education Mass-Media Campaigns Implemented in New South Wales, Australia." *National Library of Medicine*, 29 Jan. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4732951/. Accessed 2 Mar. 2023.

[11] Everett, Jessica N. "Genetic Risk Factors for Skin Cancer." *Rogel Cancer Center*, 2023, www.rogelcancercenter.org/living-with-cancer/caregivers-and-family/genetic-risk-factors-skin-cancer. Accessed 28 Feb. 2023.

[12] "50 State Elevations (Mean Elevation)." *NETSTATE.COM*, www.netstate.com/states/tables/state_elevation_mean.htm. Accessed 1 Mar. 2023.

[13] "Google Earth" *Google*, 11 June 2001, https://earth.google.com/web/ Map. Accessed 2 Mar. 2023.

[14] Guy, Gery P., et al. "Prevalence and Costs of Skin Cancer Treatment in the US, 2002–2006 and 2007–2011." *National Library of Medicine*, 10 Nov. 2014, www.ncbi.nlm.nih.gov/pmc/articles/PMC4603424/. Accessed 28 Feb. 2023.

[15] "Historical Population Density Data (1910-2020)." *United States Census Bureau*, 26 Apr. 2021, www.census.gov/data/tables/time-series/dec/density-data-text.html. Accessed 2 Mar. 2023.

[16] Home Page of National Cancer Institute. *National Cancer Institute*, seer.cancer.gov/. Accessed 28 Feb. 2023.

[17] "How Much Does Skin Cancer Treatment Cost?" *GoodRx Health*, 15 March 2022, https://www.goodrx.com/conditions/skin-cancer/skin-cancer-treatment-cost. Accessed 3 Mar. 2023.

[18] "Key Statistics for Melanoma Skin Cancer." *American Cancer Society*, 12 Jan. 2023, www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html. Accessed 28 Feb. 2023.

[19] "Mapping the Shadows of New York City: Every Building, Every Block." *The New York Times*, www.nytimes.com/interactive/2016/12/21/upshot/Mapping-the-Shadows-of-New-York-City.html. Accessed 2 Mar. 2023.

[20] "Melanoma & Skin of Color." *Melanoma Research Alliance*, www.curemelanoma.org/about-melanoma/people-of-color. Accessed 27 Feb. 2023.

[21] "Melanoma of the Skin Cancer Risk from Birth over Time, 2017-2019." *National Cancer Institute*, seer.cancer.gov/statistics-network/explorer/application.html?site=53&data_type=6&graph_type=8&compareBy=sex&chk_sex_3=3&chk_sex_2=2&stat_type=10&race=1&hdn_age_range=300&advopt_precision=1&advopt_show_ci=on&hdn_view=0. Accessed 28 Feb. 2023.

[22] "Melanoma of the Skin Comparison of Cancer Risk, 2017-2019." *National Cancer Institute*, seer.cancer.gov/statistics-network/explorer/application.html?site=53&data_type=6&graph_type=7&compareBy=race&chk_race_6=6&chk_race_5=5&chk_race_4=4&chk_race_9=9&chk_race_8=8&series=9&stat_type=10&sex=1&age_range=300&risk_interval=99&advopt_precision=4&advopt_show_ci=on&hdn_view=1&advopt_show_apc=on&advopt_display=2#resultsRegion1. Accessed 2 Mar. 2023.

[23] "Melanoma Risk Factors." *Skin Cancer Foundation*, June 2021, www.skincancer.org/skin-cancer-information/melanoma/melanoma-causes-and-risk-factors/. Accessed 28 Feb. 2023.

[24] "Melanoma Strikes Men Harder." *American Academy of Dermatology Association*, www.aad.org/public/diseases/skin-cancer/types/common/melanoma/men-50. Accessed 27 Feb. 2023.

[25] "National Environmental Public Health Tracking Network." *Centers for Disease Control and Prevention*, ephtracking.cdc.gov/DataExplorer/. Accessed 2 Mar. 2023.

[26] "Number of Farms by State" *Iowa State University.* www.icip.iastate.edu/tables/agriculture/farms-by-state. Accessed 3 Mar. 2023.

[27] Orazio, John D., et al. "UV Radiation and the Skin." *National Library of Medicine*, 7 June 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3709783/. Accessed 27 Feb. 2023.

[28] "Radiation: The Ultraviolet (UV) Index." *World Health Organization*, 20 June 2022, www.who.int/news-room/questions-and-answers/item/radiation-the-ultraviolet-(uv)-index. Accessed 1 Mar. 2023.

[29] Raman, Ryan. "How to Safely Get Vitamin D from Sunlight." Edited by Frank Crooks. *Healthline*, 28 Apr. 2018, www.healthline.com/nutrition/vitamin-d-from-sun. Accessed 28 Feb. 2023.

[30] "Risk Factors for Melanoma Skin Cancer." *American Cancer Society*, 14 Aug. 2019, www.cancer.org/cancer/melanoma-skin-cancer/causes-risks-prevention/risk-factors.html. Accessed 28 Feb. 2023.

[31] "Skin Cancer Facts and Statistics." *Skin Cancer Foundation*, Jan. 2023, www.skincancer.org/skin-cancer-information/skin-cancer-facts/. Accessed 26 Feb.2023.

[32] Smith, Ben J., et al. "Impacts from Repeated Mass Media Campaigns to Promote Sun Protection in Australia." *Oxford Academic*, 1 Mar. 2002, academic.oup.com/heapro/article/17/1/51/550095. Accessed 2 Mar. 2023.

[33] "The State of Cancer in Colorado." *University of Colorado Anschutz Medical Campus*, 4 Feb. 2019, news.cuanschutz.edu/news-stories/the-state-of-cancer-in-colorado. Accessed 2 Mar. 2023.

[34] "Summer Fun Brings Elevated Skin Cancer Risk for Coloradans." *Colorado Health Institute*, 4 June 2015, www.coloradohealthinstitute.org/research/summer-fun-brings-elevated-skin-cancer-risk-coloradans. Accessed 2 Mar. 2023.

[35] "The Sun, UV, and You: A Guide to Sunwise Behavior." *United States Environmental Protection Agency*, www.epa.gov/sites/default/files/documents/sunuvu.pdf. Accessed 3 Mar. 2023.

[36] "Surgeon General Call to Action to Prevent Skin Cancer: Exec Summary." *US Department of Health & Human Services*, 28 July 2014, www.hhs.gov/surgeongeneral/reports-and-publications/skin-cancer/executive-summary/index.html. Accessed 28 Feb. 2023.

[37] "Ultraviolet (UV) Radiation." *US Food and Drug Administration*, www.fda.gov/radiation-emitting-products/tanning/ultraviolet-uv-radiation#1. Accessed 28 Feb. 2023.

[38] "Ultraviolet (UV) Radiation's Effects on Human Health under the Changing Climate." *European Climate and Health Observatory*, climate-adapt.eea.europa.eu/en/observatory/evidence/health-effects/uv-radiation/UV-radiation. Accessed 1 Mar. 2023.

[39] "United States Cancer Statistics (USCS)." *Centers for Disease Control and Prevention*, 13 Sept. 2022, www.cdc.gov/cancer/uscs/index.htm. Accessed 27 Feb. 2023.

[40] "US Intercensal County Population Data by Age, Sex, Race, and Hispanic Origin." *National Bureau of Economic Research*, 17 Oct. 2016, www.nber.org/research/data/us-intercensal-county-population-data-age-sex-race-and-hispanic-origin. Accessed 2 Mar.2023.

[41] "UV Index Scale." *United States Environmental Protection Agency*, 23 June 2022, www.epa.gov/sunsafety/uv-index-scale-0. Accessed 2 Mar. 2023.

[42] "UVインデックスを求めるには" ["To calculate UV index"]. *Japan Meteorological Agency*, www.data.jma.go.jp/gmd/env/uvhp/3-51uvindex_define.html. Accessed 1 Mar. 2023.

[43] "Vitamin D: The Sunshine Vitamin and Your Skin." *Henry Ford Health*, 5 July 2020, www.henryford.com/blog/2020/07/the-sunshine-vitamin-and-your-skin. Accessed 2 Mar. 2023

[44] "Increasing Incidence of Melanoma Among Young Adults: An Epidemiological Study in Olmsted County, Minnesota." *National Library of Medicine*, 1 Jan. 2012, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3538462/. Accessed 2 Mar. 2023