Matej Popovski

Data Visualization

Monday, 24 February 2025

HW2

GITHUB: <https://github.com/matejpopovski/hw2-data-visualization>

**Exploring the Paradox: Skin Cancer and Geographic Latitude**

The goal of this analysis is to investigate the correlation between geographic latitude and skin cancer incidence across the world. Given that UV radiation is strongest near the equator, the expectation is that countries closer to the equator would exhibit higher rates of skin cancer. However, the data presents a counterintuitive trend—skin cancer rates are significantly higher in northern and southern latitudes, while equatorial regions report lower incidence rates.

Using Shiny, I built a dynamic visualization tool that allows users to filter cancer rates by year, latitude, and skin cancer type. This interactive approach helps us uncover surprising trends and explore possible explanations for these unexpected patterns.

**Regression Analysis: Examining Latitude vs. Skin Cancer (First Visualization)**

The scatter plot with a regression line provides a dynamic exploration of the relationship between latitude and skin cancer rates. By adjusting the latitude range dynamically using Shiny’s interactive queries, two distinct trends emerge:

When selecting a latitude range from -40 to 0 degrees (Southern Hemisphere, closer to the equator), the regression line decreases, indicating a lower incidence of skin cancer.

However, when adjusting the latitude range to 0 to 70 degrees (Northern Hemisphere, moving away from the equator), the regression line starts increasing, revealing a higher number of reported cases.

This finding challenges the assumption that higher UV exposure directly leads to higher skin cancer rates. Instead, genetics, healthcare accessibility, lifestyle choices, and early detection efforts appear to play a much more significant role. The fact that populations in equatorial regions—where sunlight is strongest—do not exhibit the highest rates further supports the idea that sun exposure alone is not the primary driver of skin cancer prevalence.

**Global Distribution of Skin Cancer (Second Visualization: World Map)**

The interactive world map provides a geospatial perspective on skin cancer incidence across different countries. Each country is represented by a bubble, where the size corresponds to the number of skin cancer cases per 100,000 people per year.

Countries at higher latitudes, such as Australia, New Zealand, Canada, and Northern Europe, have the largest bubbles, indicating the highest incidence rates.

Meanwhile, equatorial regions display significantly smaller bubbles, further reinforcing the unexpected conclusion that higher UV exposure does not necessarily correlate with higher skin cancer rates.

This visualization highlights alternative contributing factors such as genetic susceptibility, healthcare infrastructure, early detection programs, and lifestyle habits. In developed regions, higher rates of skin cancer diagnoses may also be linked to increased healthcare access and routine screenings, rather than a higher actual occurrence.

**Sorted Table: A Tool for Comparative Analysis (Third Visualization)**

The sorted table provides a structured and numerical comparison of skin cancer incidence across different countries. Unlike the scatter plot and world map, which offer visual trends and geographical insights, the table presents a ranked list of countries with the highest and lowest reported skin cancer rates.

Filtering by year and latitude range makes it possible to identify trends over time. Sorting by reported cases allows for quick identification of outliers and high-risk regions. Displaying whole numbers ensures clarity and readability for accurate comparison.

This precise, data-driven approach is valuable for researchers, policymakers, and healthcare professionals seeking to understand geographic disparities and potential risk factors for skin cancer.

**Conclusion**

This analysis operates under the assumption that all countries accurately report their skin cancer cases, ensuring data reliability and comparability across regions. However, the findings challenge the conventional belief that greater sun exposure directly leads to higher skin cancer rates. Instead, the data shows that countries near the equator, where people are exposed to intense sunlight year-round, report significantly fewer skin cancer cases compared to those in higher latitudes.

**Questions:**

• What are some interesting facts that you learned through the visualization. Provide at least one unexpected finding?

Surprisingly, countries near the equator report fewer skin cancer cases despite higher UV exposure, while higher-latitude countries have the highest rates. This suggests that factors like genetics, healthcare access, and lifestyle choices may play a bigger role than sun exposure alone.

• How did you create the interface? Were there any data preparation steps? What guided the style customizations and interface layout that you used?

The Shiny app includes interactive filters for year, latitude, and cancer type, updating three visualizations dynamically. Data preparation involved merging cancer incidence with latitude data, handling missing values, and rounding numbers. The tabbed layout ensures easy navigation between the scatter plot, world map, and sorted table.

• What is the reactive graph structure of your application?

The app uses reactive functions, ensuring that changes in filters instantly update all visualizations. The scatter plot adjusts regression trends, the map resizes bubbles dynamically, and the table reorders countries based on selected filters, making the analysis fully interactive.

GITHUB: <https://github.com/matejpopovski/hw2-data-visualization>