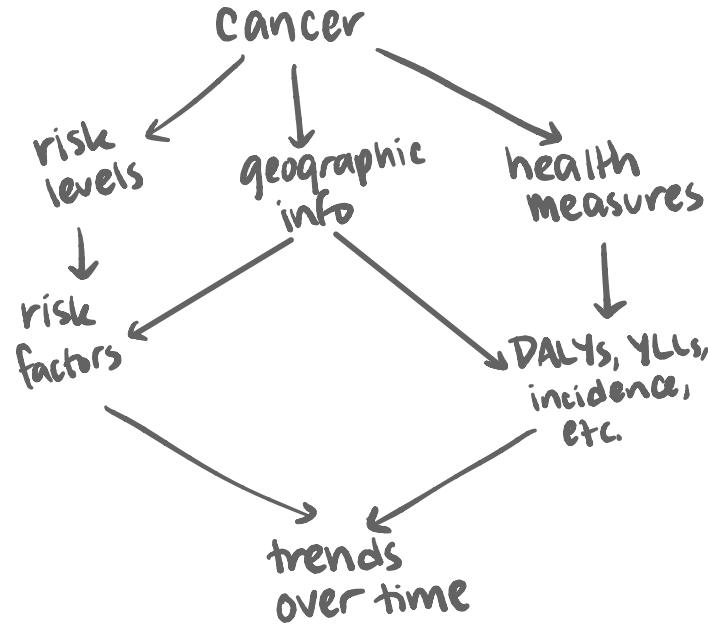


## I. IDEAS



\* Can also look at demographic effects (age, sex, income, etc.)

## 2. FILTER

### Cancer type

### risk category

- behavioral, environmental, metabolic

### specific risk factor

### geography

### year range - 1990 - 2021

### health measure

- DALYs, YLLs, incidence, prevalence

## 3. CATEGORIZE

### 1. Cross-Disease Comparison

- compare how risk categories drive different cancer types
- identify cancers most impacted by certain risks

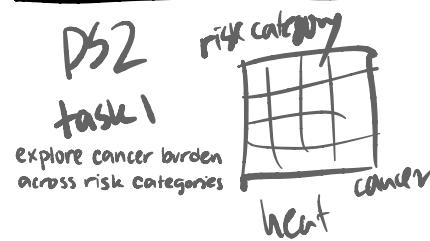
### 2. Geographic & Population Variance

- compare cancer burden across countries
- identify hotspots of environmental, behavioral, or metabolic risks

### 3. Temporal Trends

- how risks have affected cancer burden globally from 1990 - 2021
- emerging or declining risks

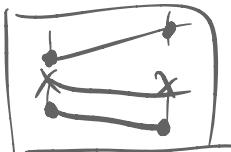
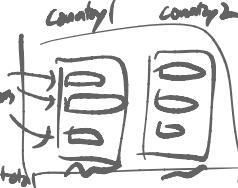
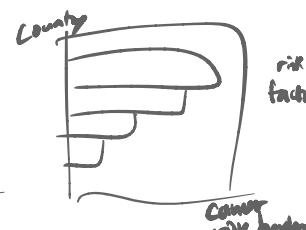
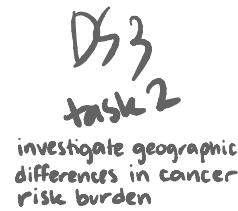
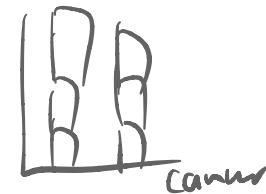
## 4. COMBINE + REFINER



### Health measure



### Health measure

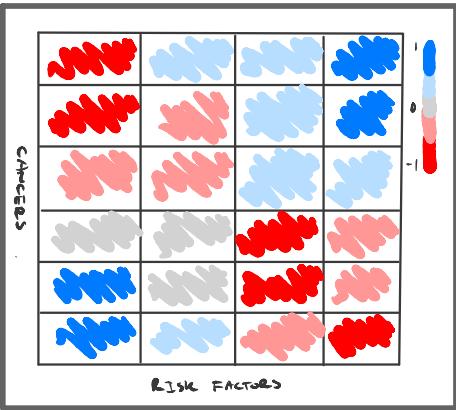


## 5. QUESTION

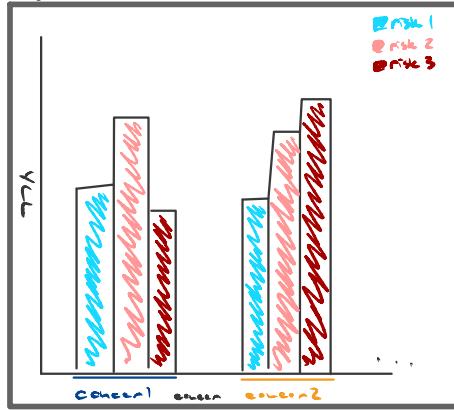
How do behavioral, environmental, and metabolic risk factors shape the global burden of cancer across diseases, regions, and time?

# LAYOUT

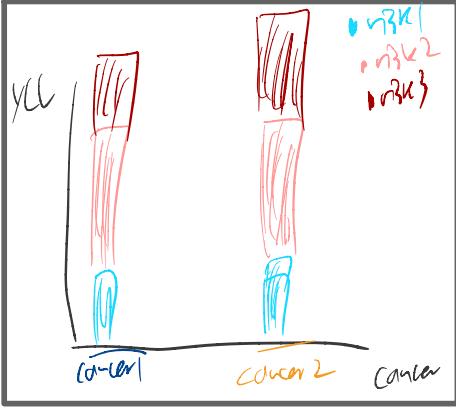
## heatmap



## grouped bar chart



## Stacked bar chart



Title: Cancer - Risk Contribution Explorer

Author: The Plot Thickens

Date: 18 Nov. 2025

Sheet: 2

Task: explore cancer burden across risk categories

## OPERATIONS

- filter by cancer type
- filter by risk category/factor
- filter by health measure
  - DALYs, YLLs, incidence, prevalence
- preview numerical results in tabular format
- tooltip to view exact numeric values
- enable multi-select comparison across cancers

## Focus

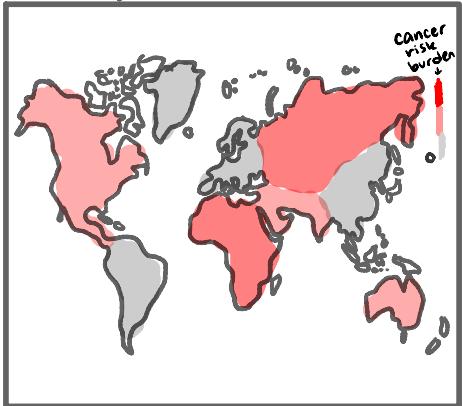
- enables high-level comparison of how major risk categories drive different cancers
- highlights which cancers are most influenced by behavioral vs. environmental vs. metabolic risks
- allows for quick visual scanning + deeper exploration
- helps identify patterns across diseases at a glance

## DISCUSSION

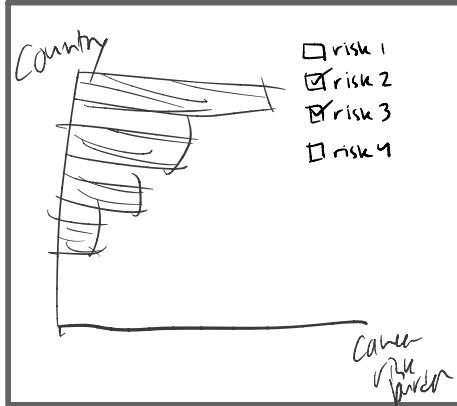
- heatmaps scale well, but require careful color choice.
- grouped bar charts are intuitive, but can become cluttered w/ multiple cancers.
- stacked bars clearly show relative proportions, but make cross-comparison harder.
- normalization is key because cancers vary widely in total burden.
- color consistency across all charts is crucial for interpretability.

# LAYOUT

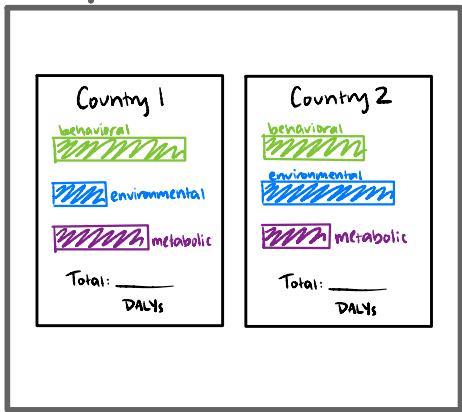
## Choropleth



## ranked bar chart



## Comparison cards



Title: Cancer-Risk Contribution Explorer

Author: The Plot Thickens

Date: 18 Nov. 2025

Sheet: 3

Task: investigate geographic differences in cancer risk burden

## OPERATIONS

- select a cancer type
- choose a risk category to visualize
  - behavioral, environmental, metabolic
- filter by geographic level
  - global, region, country
- include year range filter
- export region-level metrics in tabular/table format
- tooltip for detailed metrics

## FOCUS

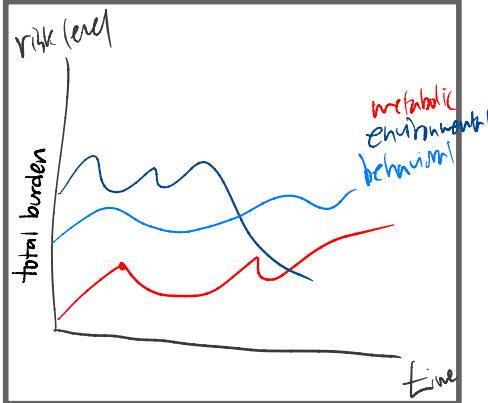
- highlights global disparities in cancer burden
- shows which regions are most affected by each risk category
- enables quick comparison of countries
  - for public health officials, policymakers, etc.
- Supports both global overview and country-level deep dive

## DISCUSSION

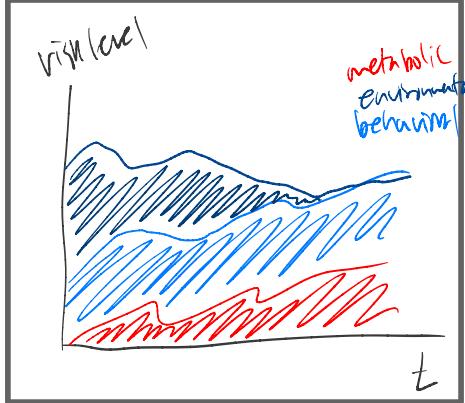
- choropleths are intuitive, but sensitive to map projection distortion.
- ranked bar charts give precise comparison, but lose geographic context.
- Side-by-side cards work well for comparing a few countries, but not many.
- missing/uneven data across countries may affect visual accuracy.
- must prevent visual bias.
- interactivity is key for map interpretability.

# LAYOUT

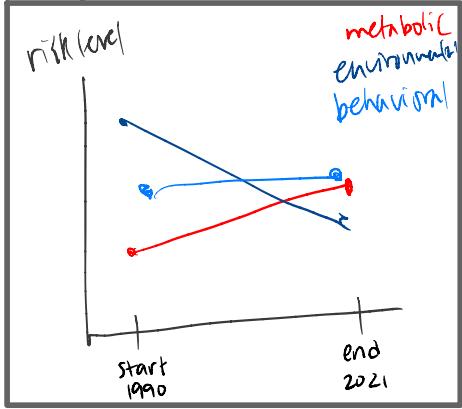
## multi-line plot



## Stacked area chart



## slope chart



Title: Cancer - Risk Contribution Explorer

Author: The Plot Thickens

Date: 18 Nov. 2025

Sheet: 4

Task: analyze temporal trends in cancer risks over time

## OPERATIONS

- filter by cancer type
- filter by risk category
  - behavioral, environmental, metabolic
- select temporal range
  - slider or multi-select
  - 1990 - 2021
- ability to overlay multiple risk categories on same plot
- toggle absolute vs. normalized values

## FOCUS

- shows how risk contributions have shifted over decades
- helps identify emerging risks
- makes long-term trends visible through slope and shape
- supports hypothesis generation

## DISCUSSION

- multi-line plots can become visually busy if many risks are shown.
- stack areas emphasize total burden, but obscure small categories.
- slope charts are minimal, but only show start/end.
- temporal smoothing helps reduce noise, but may hide important spikes.
- missing data will require careful handling.
- consistent timescale across visualizations is critical for comparison.

# LAYOUT

## Our Streamlit App

### Selections

Cancer Type

cancer ① cancer ② cancer ③

drop down  
✓ (select and edit)  
drop down  
✓ (select and edit)  
drop down  
✓ (select and edit)

Risk Categories

behavioral ②

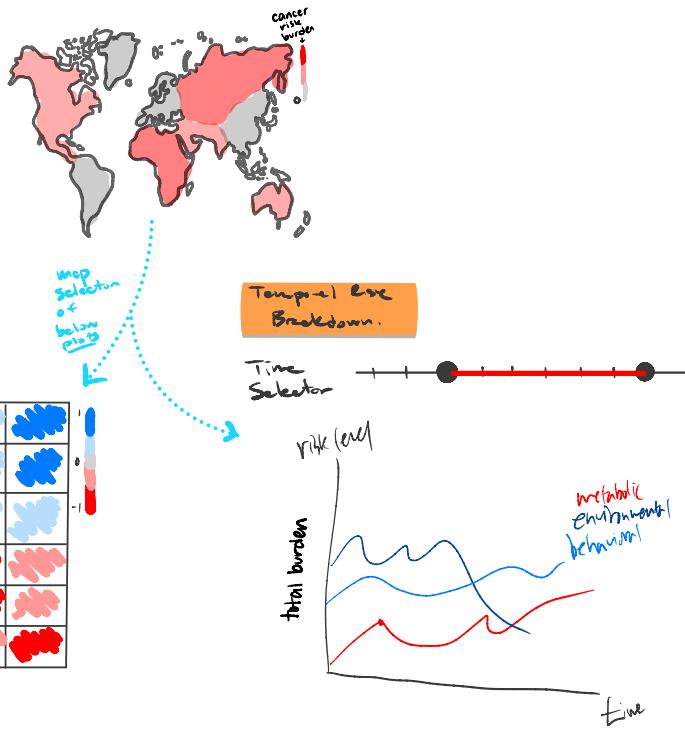
drop down  
✓ (select and edit)

Health Metrics

DALY ① prevalence ②

drop down  
✓ (select and edit)

Geographic Risk Breakdown



### Focus

See how different categories of risk factors contribute to specific cancers, and compare across countries/over time

→ e.g. visualize how environmental risk factors (air pollution, water quality) impact certain cancers while metabolic risk factors (BMI, high glucose) impact others. Maybe these trends differ across countries or have changed over the years.

Title: Cancer - Risk Contribution Explorer

Author: The Plot Thickens

Date: 18 Nov. 2025

Sheet: 5

Task: Convergence to 1 idea

### OPERATIONS

- Select a cancer type
- Choose a risk category to visualize
  - behavioral, environmental, metabolic
- filter by health measure
  - DALYs, YLLs, incidence, prevalence
- select temporal range
  - slider, 1990 - 2021
- ability to overlay multiple risk categories on same plot
- tooltip for detailed metrics
  - task 1: breakdown of risk contributions to each cancer type
  - task 2: breakdown of risk contributions by country
  - task 3: breakdown of risk contributions by year

### DETAIL

• Risk levels are quantified in YLLs (years life lost)

• We will allow user to group risk factors by category and also expand to view individual risk factors

• We may choose certain countries to focus on rather than all

• The cancer categories come from the cause of death variable, which we are subsetting to just cancers for the use of this project

3. One challenge is country data compatibility; Altair will have expectations on how to map the country names to the location on the map, and we need to make sure that the country labels from our dataset match and ensure that things populate in the visualization properly. (i.e., United States, or US, or U.S., or United States of America...)

For some countries, there could also be missing data points for certain timeframes. To deal with missing data, we will impute data from the closest previous year.

We also need to carefully communicate that the cancer variable is a subset of the cause of death variable, so we are actually seeing how certain risk factors lead to *deaths* for specific cancers. Also, risk factors are quantified by YLLs (years of life lost), which we will have to explain on our website also since it is a confusing metric.

4. We will implement these plots via Altair and Streamlit. We plan to display three graphs: a map of the world colored by cancer risk burden, a heatmap of the risk factors' contribution to different cancers, and a temporal breakdown showing how these risk factor trends have increased or decreased over time.

To make our Streamlit app interactive, we will use a combination of tool tips, dropdown selectors, and a responsive plot to filter subsequent visualizations. If the user clicks a certain country on the map, the country code will be narrowed down for the heatmap plot and the line plot. For the line plot, the user can select a time frame by using the time selector section provided above the line plot.

Together, these plots will help show how different categories of risk factors impact different cancers and how these trends differ across countries and timeframes.