Predictor of Passing

~ Working Title ~

By: Tin Nguyen , Cole Kleinhans , Kai Weiner , Jonas Bartels

Data Summary:
Our data is a collection of causes of death, the age of when they died of a given cause, their gender and the state they lived in. Additionally, there is cumulative data about the number of deaths of a given cause and the total number of people who share all the same variables (age, state, gender, cause of death).

Disclaimer:
When the CDC uses "gender" as a variable, they are reffering to sex. However, for simplicity, we will continue this convention yet we acknowledge this conflation of the normative gender and sex.

Citation
Url: https://wonder.cdc.gov/controller/datarequest/D76
Date Downloaded: 4/7/22 - 4/11/22
Authorship: CDC, Centers for Disease Control and Prevention
Exact name and version: *Personalized dataset from CDC WONDER database
Terms of use:

> *"The Public Health Service Act (42 U.S.C. 242m(d)) provides that the data collected by the National Center for Health Statistics (NCHS) may be used only for the purpose for which they were obtained; any effort to determine the identity of any reported cases, or to use the information for any purpose other than for health statistical reporting and analysis, is against the law. Therefore users will:*
>
> - *Use these data for health statistical reporting and analysis only.*
> - *For sub-national geography, do not present or publish death counts of 9 or fewer or death rates based on counts of nine or fewer (in figures, graphs, maps, tables, etc.).*
> - *Make no attempt to learn the identity of any person or establishment included in these data.*
> - *Make no disclosure or other use of the identity of any person or establishment discovered inadvertently and advise the NCHS Confidentiality Officer of any such discovery."*

Disclaimer:
Our website will not use the data to determine the identity of individuals. We will use this data for statistical reporting and analysis, with clear written disclaimers on any predictive, generated and informal information.

Suggested Citation:
Place of publication or publisher. Name of database. Full URL. Date of access.

- Ex: Centers for Disease Control and Prevention. CDC Wonder. http://wonder.cdc.gov/. April 1995.

Data prep process:
This is the process we used to generate a personalized dataset using the CDC WONDER database:

1. We generated datasets for each state containing a column with age as the variable.
    a. Generate two datasets with causes of death; one with all female instances and the other with all male instances - *CDC omitted causes that have under 10 affected individuals.*
    b. Generate two datasets without causes of death; one with all female instances and the other with all male instances - *CDC provided total number of deaths based sex & age.*

To further customize the dataset to our functionalities, we will:

2. Turn individual files into CSV files and remove unused data and text for further processing.
3. To calculate how many people of a given age died from a discluded death cause for each state: we will subtract the sum of deaths in the cause listed dataset from the same age group in the cause unlisted dataset. We will add this difference to the set with cause of death listed under the category: “Misc.”
4. We will merge our sets with cause of death listed into a single CSV.

Features/User interactions:

1) Input a variable (age, location, gender, cause of death) or combination of variables and get an output about the total number of deaths within our dataset.
2) Input a variable (age, location, gender) or combination of variables and get an output of the total number of living people who fall into the indicated categorization.
3) Input a variable (age, location, gender) or a combination of inputs and get an output of the 10 leading causes of deaths based on the highest number of deaths within the indicated categorization.
4) Input a cause of death and get the state/age where the most people die from that cause.
5) Data visualizations with age, gender, and/or location vs leading causes of deaths based on x.
6) Fun feature: Answer to “how you are going to die” using seeded randomness. User input on age, state, gender to reference with dataset along with obscure questions (ie, favorite colors, star sign) as seed. Random chance determines when and how you (the hypothetical and statistical you) will die with the chance of this happening based on the dataset's weight.

Team Contract:

1. **What are the goals of our team?**
    - Prioritize a productive, transparent and concise process just as much as creating a well-made final product.
    - Be thoughtful and considerate of each other; be through and precise with our project.
    - To have fun!
2. **What are the strengths of our team and its members?**
    - Tin: Enjoys design. Has experience with UI/UX frameworks.
    - Cole: Has (minimal) experience with git, very motivated and loves to code.

- Kai: Good team management skills, can check in on teammates and facilitate interactions.
- Jonas: a natural agile method user (always have) - pretty quick coder, loves adding extra features. Good at debugging.

How will we capitalize on the strengths of each member?

- Tin: Will focus on graphic design and organization.
- Cole: Will create the aesthetics for the website and will code a lot.
- Kai: Will check in on team members. Will help define team goals, roles, deadlines.
- Jonas: Will focus heavily on coding and help other members debug.

3. **What are the rules that will guide your team? Specifically:**
   - **When will your team meet? What time, how often, for how long, where?**
     i. Location: Anderson Commons
     ii. Time: (Tuesday/Thursday) 8:30- 9:30 as default meeting time
     iii. Frequency of meeting is TBD based on needs of the project.
     iv. Conflict time (12pm-12am):
       1. Kai: Tuesday: 3:10-4:15, Thursday: 3:10-4:15 & 4:45 -6:15
       2. Tin: Tuesday: free by 1:00 Thursday: free by 1:00
       3. Jonas: Tuesday: 1:15-3:00, 5:00-7:00 Thursday 1:15-3:00, 4:45-6:15
       4. Cole: Tuesday 1-3:30; Thursday 1-2:30
   - **What roles will members take on in your meetings? Is someone responsible for setting agendas, taking notes, facilitating discussions, etc?**
     i. Tin: “Organizer” - set agenda for meeting and assign tasks
     ii. Cole: “Driver” - typer when everyone is collaborating on a piece of code.
     iii. Kai: “Facilitator” - involve every member and guide discussions.
     iv. Jonas: “Researcher on deck” - will google answers for all our questions.
   - **How will you communicate with each other? (to share work, to ask questions, notify the group if someone is running late or if someone will miss a meeting, etc)**
     i. We will communicate through an iMessage group chat.
   - **How will you make sure communication stays respectful? (How does your team define "respectful"?)**
     i. (Definition) Respectful - willing to understand and give due regards to the feelings, ideas and thoughts of others.
     ii. We will make sure communication stays respectful by trying to be understanding of each other’s point of view. We will also practice speaking up in a calm and level way if/when we are upset.
     iii. We will listen to each other’s inputs with the utmost regards without the intention of dismissing anyone.
   - **What are the rules for dealing with a teammate who hasn't been communicating? How frequently should team members communicate / check in?**
     i. Before a meeting or when tasks are assigned, a text will be sent in the group chat to confirm participation. Conflicts are understandable, however, it is standard to communicate any conflicts in order to plan meetings and agendas accordingly.

ii. If a team member misses multiple meetings without communication or has not completed assigned tasks, the team will check-in with the member to address any conflict and clarify new tasks. As an addendum of our goals, we will try to be understanding of each member and their busy schedules.
iii. If conflict continues, we will reach out to Anya.

- **What technologies will you use to support team meetings and work? (Google Drive, Hangouts, Zoom, Facetime, etc)**
    i. Meetings will be conducted in person whenever possible. Zoom otherwise.
    ii. Files will be shared either over email, on Google Drive, or stored in our repository.
- **How will you make decisions? (Unanimous, consensus, majority rule, by assigned roles, rock-paper-scissors, etc.)**
    i. Decisions will be made through majority rule but we will always attempt to come to a compromise.
    ii. The exception to this rule is if the minority feels extremely uncomfortable with the majority decision, in that situation a compromise or cave-in to the minority must be reached.
- **How will you divide the work?**
    i. We will work together on standardizing the website's form and cohesive functionality after individual goals have been completed.
    ii. Tin:
        1. Work with Cole on ways that the data could be accessed efficiently.
        2. Create functions to sort through data to get top leading causes for a given category/ find other data.
    iii. Kai:
        1. Will work with Jonas on the cause of death predictor, the model we will use.
        2. Responsible for the seeded randomness aspect and questionnaire. Will focus on the visual end of the cause of death page, graphics and graphs displayed.
    iv. Cole:
        1. Work with Tin to access the data most efficiently
        2. Compiling general information about each state, specific age groups, gender, cause; writing functions that find all this data
    v. Jonas:
        1. Setting up the background functionality for death predictor. Statistically accurate predictions using random seed from user input. Also add the most common cause of death for the user demographic.

- **How will you ensure that everybody participates meaningfully? How will you make sure that everyone's contribution is valued?**
    i. Through open discussions, the facilitator (when needed) , and delegating individual tasks early and clearly.

- **What expectations do you have for satisfactory participation? (How much time will each group member spend per week on project activities?)**
    - We are willing to put in the amount of time and effort for exemplary work (the number of hours will depend on the project and everyone's strengths)
        - Making a thoughtful and thorough attempt with each assignment.
        - Will utilize the extension and grace-period whenever necessary but try to make deadlines.
- **What process will you follow if someone does not live up to their responsibilities and/or meet the standards for work set by the team?**
    i. Meet with the group member, explain grievances, allow them to explain their perspective. Potentially meet with the professor.
- **How will you address conflict or deal with disagreements within the team?**
    i. Open dialogue with all team members involved and a neutral mediator.
        1. Try to explain your side of the story, what you understand of the other side and how the situation makes you feel.
        2. The end goal is to reach good-will and a solution on how to move forward.