# DO NOT ADD any punctuation except commas and periods, hash mark for R comments, and single and double quote using the same exact character sex as below

# textIntroA # measureHelp

# textIntroB # causeHelp

# textIntroC # statecutHelp

# textNote.real # cutmethodHelp

# textNote.fake # SDOHHelp

# mapTab

# conditionTab

# conditionTableTab

# condtionSexTab

# rankGeoTab

# trendTab

# sdohTab

#-- Other Misc-------------------------------------------------------------------------------------------------------------------

#GitHub

#Fake Site

#Real Site

#Contact us

#Survey URL

textYear <- "YEAR: At the County and State levels of geography, YEAR is the individual year of death, with current data from 2001 to 2017. At the Community and Census Tract levels of geography, all data are displayed for the years 2013 to 2017 combined. These years are combined for statistical stability, so that for these more granular levels of geography, the displayed data are still meaningful, and not just the result of random fluctuations"

#--Home Page Text -------------------------------------------------------------------------------------------------------------

# Side Bar

textIntroA <- "The California Community Burden of Disease Engine (CCB) is a tool to explore data on burden of disease in multiple levels of geographic granularity in order to answer and generate questions, both simple and complex, about the intersection between health disparities and place."

textIntroB <- "This tool is designed for use by CDPH programs, local health departments, and community partners for epidemiologic analysis and to provide systematic scientific insight to inform public health planning, evaluation and action."

textIntroC <- "The CCB currently displays 17 years of California statewide, county, community, and census tract condition-specific mortality burden, using a range of measures, with interactive rankings, charts, maps and trend visualizations. The list of conditions is based on the Global Burden of Disease system, modified for local public health priorities. The CCB also includes a limited set of social determinants data and describes their correlations with death outcomes, as a pilot for more robust functionality in this area."

textNote.real <- "This app deployment is for preliminary internal CDPH review. Do not share these data with external partners. A very wide range of enhancements are being considered for this application. Any/all comments regarding errors, enhancements, or any other ideas about this version are most welcome."

textNote.fake <- "NOTE: THIS VERSION OF THE ENGINE IS FOR DEMONSTRATION PURPOSES ONLY - THE DATA ARE NOT REAL - THEY ARE A RANDOM SUBSET OF RANDOMLY DISTORTED DATA"

# Main Pannel

above1 <- paste0('<p><b>California Community Burden of Disease and Cost Engine (CCB):</b></p>

<ui style="list-style-type:circle">

<li style="margin-left: 40px"><b>An emerging toolset for epidemiologic analysis and scientific insight, exploring the intersection between health disparities and place</b></li>

</ui> ')

below1 <-

paste0("<ui style=align='left'>

<h4>Coming Soon:</h3>

<strong>

The CCB is a work in progress, and is intended to be an evolving toolset developing new content and functionality in response to the needs of public health practitioners. Examples of upcoming development enhancements:</strong>

<li>Hospital discharge and emergency department data</li>

<li>Cost (charges) data based on hospital discharge</li>

<li>Life expectancy estimates at the census tract, community, and county levels</li>

<li>Incorporation of 'multiple cause of death' and 'out of state' death data files</li>

<li>Expanded range and analysis of social determinants data</li>

<li>Additional displays of statistical significance</li>

<li>Enhanced user interface</li>

<li>Automated report generation </li>

<li>Our team will use the feedback gathered through this beta-testing window to prioritize future enhancements.</li><br><br>

Another great project of the CDPH Fusion Center!

<br><br>

The CCB is one of the ways the Fusion Center is working to explore the lens of place and its impact on health disparities. The CCB is an initiative of the Fusion Center implemented with participation from a crosscutting technical team, with representatives from multiple CDPH programs.

<br><br>

This platform is also a pilot component of the CDPH Ecosystem of Data Sharing, leveraging a rich multi-level data set/system for modeling and predictive analytics and demonstrating automated and integrated data processing, analytics, and visualization. The project employs nimble modular development, with the goal to share tools/resources with outside partners (counties and other states).

<br><br>

The Community Burden of Disease System (System/Application/Project/Initiative) - (CBDS/A/I/P)

<br><br>

The CCB is the California State implementation piloting the Community Burden of Disease System (CBDS). The code and system are written and structured to be useable by states and counties throughout the United States-with any state or county using their own structured input file of events (e.g. deaths), and the CBD system supplying underlying population data, social determinants of health data, and all the processing, calculations, and tools to generate a range of interactive displays of multiple rate and count measures.

<br><br>

Technical notes:

At the county level, data are displayed separately for each year, and at the community or census-tract level are displayed only for the most recent five-year period (combined). Data for some conditions with very few deaths and/or with other sensitivity considerations are suppressed in this release.

</ui>")

# DROP DOWN HELPS -------------------------------------------------------------------------------------------------

measureHelp <- paste0('<p>The current MEASURES of deaths are:</p>

<ui style="list-style-type:circle">

<li style="margin-left: 80px">Years of Life Lost (YLL)</li>

<li style="margin-left: 80px">Mean YLL</li>

<li style="margin-left: 80px">Years of Life Lost per 100,000 population</li>

<li style="margin-left: 80px">Number of Deaths</li>

<li style="margin-left: 80px">Death Rate (Deaths per 100,000 population)</li>

<li style="margin-left: 80px">Median Age at Death</li>

<li style="margin-left: 80px">Excess Relative Risk (Standard Mortality Ratio)</li>

</ui>

<br>

<p>No one measure is best--each measure provides a different view or perspective into the impact of the condition. For example,

<b>Number of deaths</b>

is the simplest, most direct measure, and has clear intuitive meaning, and, other things being equal, will be larger in areas with larger populations.

<b>Crude Death Rate</b> takes the size of the population into account by dividing the number of deaths by the number of people in the population (multiplied by 100,000 for interpretability, i.e. number of deaths per 100,000 people).

<b>Age-adjusted Death Rate</b> is the rate that would have existed if the population had the same age distribution as a reference population. This allows for comparisons between populations with differences in age distributions, accounting for the fact that age itself is generally correlated with higher mortality.

<b>Years of Life Lost</b> sums all the years of life prematurely lost across all people that die from that condition, and is influenced by the age at which people die from the condition and the number of people that die from that condition.

<b>Years of Life Lost per 100,000 population</b> divides the YLL sum by the number of people in the population, and then multiplies by 100,000 for interpretability. It is the YLL equivalent of the Death Rate

<b> Standard Mortality Ratio (SMR)</b> shows the county rate of condition divided by the rate for the condition in the State overall-this measure will highlight counties that have especially high (or low) rates of a condition <b>compared</b> to the State rate, even if the condition does not have a large number of deaths. To aid in using this measure, the vertical red line is at 1.2, corresponding to the measure being 120% higher in the selected geography than the State average. The green line is at 0.8, 80% of the State average and the grey line is at 1.0, right on the State average.

<b>Average age at death</b>

<b>Age adjusted YLL rate</b> ' )

# SMR: May want to reference https://www.cdc.gov/nchs/data/statnt/statnt06rv.pdf

causeHelp <- 'The CAUSES OF DEATH in the app are based on a hierarchal outline arrangement of about 70 conditions. The Global Burden of Disease Study condition structure forms the basis for this outline, with modifications aimed at making the conditions as relatable and useful to California Public Health priorities and programs as possible.

<br><br>

Above the "Top Level" of the outline is "All Causes", including all deaths from any cause (including those few with missing or unclear cause information).

<br><br>

The "Top Level" of the outline includes a mutually exclusive and exhaustive set of five high-level condition sets (and a "missing category"):

<li style="margin-left: 30px"> Communicable, Maternal, and Perinatal Conditions</li>

<li style="margin-left: 30px">Coronary Heart Disease</li>

<li style="margin-left: 30px"> Cancer/Malignant Neoplasms</li>

<li style="margin-left: 30px">Other Chronic Conditions</li>

<li style="margin-left: 30px">Injury</li>

At the census-tract level, only this "Top Level" set of conditions is available in the app. Because of the large numbers of deaths in these broad categories, this Top Level will be particularly useful for looking at differences within and between small geographic areas.

<br><br>

The next level, the "Public Health Level", is a separate mutually exclusive and exhaustive list of about 60 conditions (e.g. "Respiratory infections", "Trachea, bronchus and lung cancers", "Alzheimer disease and other dementias", "Suicide/Self-harm", etc.) "under" the "Top Level" categories, based on intermediate-level conditions in the Global Burden of Disease system, modified based on public health priorities and programs in the United States. Because of the greater clinical detail and public health program specificity of this level, we believe it will provide the greatest insights and the most meaningful rankings of conditions at the county and community levels.

<br><br>

Below this, there is a "Detail Level" of conditions, with a few "Public Health Level" conditions broken down into more detailed (e.g. Lymphomas and multiple myeloma, is broken down into Hodgkin lymphoma, Non-Hodgkin lymphoma and Multiple myeloma). These detailed conditions are only available in the app at the county level of geography.

<br><br>

Not shown in the app, is an underlying "Base Level". Each and every condition in the "Base Level" rolls up to one and only one "Public Health Level" condition (regardless of whether the condition is shown in the "Detail Level"). A diagram showing this hierarchy, and the long list of "Base Level" conditions is available here. Furthermore, the list of detailed ICD-10 codes that make up each "Base level" condition can be found here. We welcome comments for suggested changes to any aspect of our hierarchal list.

'

# tags$a(href="https://github.com/mcSamuelDataSci/CACommunityBurden","GitHub Site",style="color:blue"))

statecutHelp <-

"The State-based Cut-points button changes the way the Measure is broken down or

grouped-with the box checked, the cut-points are based on the State data overall

(so many/most communities in a given county might be in the highest category, if

that condition tended to be high in that county in general). If the box is unchecked,

the cut points will be based on the data in just that county, so the distribution

of the condition throughout just that one county may be easier to see and understand."

cutmethodHelp <-

"Specifies method used to determine the cut-points for the color categories. <b>Quantile</b>

divides the frequency distribution into equal categories, each containing the same fraction

of the total dataset. <b>Fisher</b> uses the Fisher-Jenks algorithm which reduces the

variance within categories and maximizes the variance between categories."

# Needs explanation on how color cutoffs were calculated, especially Fisher method

SDOHHelp <- "test"

# Maybe text on importance of SDOH on mortality, but only variable needing explanation is HPI

# ---- TABS----------------------------------------------------------------------------------------------------------

mapTab <- paste0("

These maps display the geographic distribution of disease burden among counties and communities across California. The <b>Geo Level</b> options allow the user to change the display from county, to community, to census tract. This selection is one of the key concepts behind the app, <b>Place Matters</b>. Insights into the burden of disease must be explored at multiple geographic levels, especially granular community levels. Data at the community and the census tract levels are aggregated to 5-year intervals.

<br><br>

Users can select either the state as a whole or zoom to a specific county for a better view of just that county, and its subcounty detail.

<br><br>

Users can select from various measures of mortality to assess burden of disease.

Selecting the <b>State-based cutpoints</b> option allows for comparisons based on the statewide distribution instead of just within the county.

<br><br>

The interactive map allows for zooming in and out to see streets or other geographically identifying locations.

Also, the interactive map has a pop-up which display information for the geography selected.

<br><br>

The static is better for using in an external presentation. The Place Names option displays county and community names.")

conditionTab <-

"This tab displays cause-of-death rankings for either a selected county or the whole state. The figure shows the ranking based on five different measures, and can be sorted based on any of these measures. Different insights can be gained by ranking on different measures (e.g. ranking on the number of deaths or 'age-adjusted death rates' shows the more 'typical' ranking of most systems; ranking on mean age at death shows the conditions that impact young people the most, and ranking on SMR show those conditions for which a county has particularly high rates compared to the State average). The 'levels' option allows users to select between broad or narrow categories of conditions. The 'How Many' button determines how many causes of death to display on the graph.

<br><br>

Ranking on SMR provides a special window in the potentially unique priority of a condition in the selected geography. A large SMR means the condition is especially high in that geography relative to the State average, even if the condition does not have a large number of deaths. This is an important way to detect conditions that, while perhaps not common, are unusually high (or low) in a county or community in which one is interested. To aid in using this measure, the vertical red line is at 1.2, corresponding to the measure being 120% higher in the selected geography than the State average. The green line is at 0.8, 80% of the State average and the grey line is at 1.0, right on the State average.

"

conditionTableTab <-

"This is a tabular version of the Rank Conditions tab, providing for a more granular examination of specific numbers or rates., Users can sort the table on any of the measures and can use the search window allows users to quickly find a specific condition."

conditionSexTab <-

"(This tab work in progress, awaiting further development of the tab)

<br><br>

This tab ranks causes within a selected geography separately for males and females. It can highlight conditions that appear to be a leading cause of death for one sex but not the other."

rankGeoTab <-

"This tab displays the ranked order of counties in California or the communities within a selected county for a selected condition. These rankings highlight places where a particular condition is the highest as well as highlighting geographical disparities of the condition. Years of life lost and number of deaths will tend to be highest in areas with the largest populations, whereas rate measures adjust for population size. Note that higher ranking counties or communities may not be meaningfully higher from a statistical perspective; examining the confidence intervals will help determine if there is a meaningful difference or not.

<br><br>

Note: In the current version, confidence intervals are only displayed for the crude death rate but will be available soon for all measures."

trendTab <-

"This graph displays the trend over time for a particular condition within a selected geography, separately for males, females, and the total population. Reviewing the trend over time is important for understanding which problems are improving and which are getting worse.

<br>

Note: Because the data for the communities and census tract are currently aggregated for 5 years, those data are not available currently in the trend tab."

sdohTab <-

"This tab is preliminary and under development.

<br><br>

This scatter plot displays the correlation of a selected social determinant measure with a selected condition. Each dot maps the value of the social determinant measure against the value of the condition measure for one geographic unit (county, community, or census tract). Because this association is 'ecologic' (correlation of geographic units, not of individuals), it is particularly important in this tab to look at measures that take into account the size and age distribution of the population, such as age-adjusted YLL rate and age-adjusted death rate. While correlations do not indicate causation, they are a potentially important way to understand the differential roles of some social determinants of health on disease outcomes.

<br><br>

In the current version, the colors represent the regions of the state; soon the colors could represent the rurality levels of the county/community/tract or other factors.

<br><br>

In the current version, the size of the dots is proportional to the size of the population represented by the dot. This too, could represent other things.

<br><br>

Note: Currently this tab only displays one variable, but the display and analysis in this tab will be expanded to include multiple variables simultaneously."

#--OLD NOT USED----------------------------------------------------------------------------------------------------------

# textIntroA <- "The Community Burden of Disease (CBD) is an evolving platform to view and explore detailed disease/condition burden and Social Determinants of Health data on multiple levels of geographic granularity in order to answer and generate questions, both simple and complex."

# textIntroB <- "The code and system are written/structured to useable by States and Counties throughout the United States—with any State or County using their own structured input file of events (e.g. deaths), and the CBD system supplying underlying population data, Social Determinates of Health data, and all the processing, calculations, and tools to generate a range of interactive displays of multiple rate and count measures. The list of disease conditions is based on the Global Burden of Disease system, modified for local public health priorities."

# textIntroC <- "This California State implementation of the CBD, The California Community Burden of Disease and Cost Engine (CCB), currently includes detailed death data and multiple associated measures (e.g. age-adjusted cause-specific Years of Life Lost rates) for 2001 to 2015 at the census tract, community (California Medical Service Study Areas), county, and statewide levels. The CCB also includes a small set of Social Determinates of Health, and describes their correlations with death outcomes, as a pilot for more robust functionality in this area. Other short-term road map plans include the addition of burden of disease and costs (charges) based on California Hospital Discharge data, automated report generation, and more."

textIntroOld <- "The goal of the California Community Burden of Disease and Cost Engine is to provide systematic scientific insight for allocation of Public Health resources, evaluation of Public Health interventions, and other Public Health actions. This initial version of the application displays multiple death-related measures (e.g. Years of Life Lost per 100,000 population, crude and age-adjusted death rate, standard mortality ratios) in interactive rankings charts, maps, and trend lines, for California counties, communities (Medical Service Study Areas), and census tracts for recent years. At the county level, data are displayed separately for each year, and at the community or census-tract level are displayed only for the most recent five-year period (combined). Data for some conditions with very few deaths and/or with other sensitivity considerations are suppressed in this release."