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

# textIntroA # measureHelp

# textIntroB # causeHelp

# textIntroC # statecutHelp

# textNote.real # cutmethodHelp

# textNote.fake # SDOHHelp

# mapTab

# conditionTab

# conditionTableTab

# condtionSexTab

# rankGeoTab

# trendTab

# sdohTab

#--Home Page Side Bar Text ------------------------------------------

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."

textIntroC <- " The CCB currently displays over 15 years of California condition-specific mortality burden data, using a range of measures displayed at the statewide, county, community, and census tract levels, with interactive rankings, charts, maps and trend visualizations."

textNote.real <- "This version of the app contains features that are still being tested and refined. While the contents have been reviewed extensively, the tool likely still contains 'bugs' and may generate error messages based on some selection combinations. While our team is continuously working to address any identified issues, we are confident the application in its current form will, overall, provide important and reliable insights. We look forward to and appreciate any observations you make regarding bugs, errors, or anything else of concern. We also welcome supportive comments."

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"

#--Home Page Main Panel ------------------------------------------

above1 <- paste0("<ui style=align='left'><b>

California Community Burden of Disease and Cost Engine (CCB):

</b>

<br>

<h4><b><i>

An emerging toolset for epidemiologic analysis and scientific insight, exploring the intersection between health disparities and place

</b></i></h4></ui>")

#--Warning Messages Etc. -----------------------------------------

tractWarning <- "NOTE: AT THE CENSUS TRACT LEVEL, ONLY THE 'TOP LEVEL' CONDITIONS CAN/WILL BE DISPLAYED; ALSO, RENDERING STATE-WIDE TRACT-LEVEL MAPS CAN TAKE MANY SECONDS--PLEASE WAIT"

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

levelHelp <- paste0('<li style="margin-left: 40px">The <b>Top Level</b> includes only the five most aggregated conditions (e.g. Cardiovascular diseases, Injuries) and All Causes combined.</li>

<li style="margin-left: 40px">The <b>Public Health Level</b> includes a list of about 60 conditions. This Level provides greater clinical detail and public health program specificity than the Top Level. </li>

<li style="margin-left: 40px">The <b>Detail Level</b> includes a small number further detailed conditions (e.g. Non-Hodgkin lymphoma, Opiod Use Disorders, Falls) <b>AND</b> includes all conditions in the Public Health Level. </li>

')

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

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

<li style="margin-left: 40px">Years of Life Lost (YLL)</li>

<li style="margin-left: 40px">Years of Life Lost Rate (per 100,000 population)</li>

<li style="margin-left: 40px">Age-Adjusted YLL Rate</li>

<li style="margin-left: 40px">Number of Deaths</li>

<li style="margin-left: 40px">Crude Death Rate (per 100,000 population)</li>

<li style="margin-left: 40px">Age-Adjusted Death Rate</li>

<li style="margin-left: 40px">Mean Age at Death</li>

<li style="margin-left: 40px">Standard Mortality Ratio (SMR)</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.

<br><br>

<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 Rate</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>Age-Adjusted YLL rate</b> adjusts for the age structure of the population, and is the YLL equivalent of the Age-Adjusted Death Rate.

<br><br>

<b> Average Age at Death</b> directly measures the mean age of death for a particular condition and/or for a particular geography. Younger average age at death has intrinsic greater impact on a population, and contributes to larger YLL.

<b> Standard Mortality Ratio (SMR)</b> shows the county rate of a 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 compared to the State rate, even if the condition does not have a large number of deaths.

')

causeHelp <-

'The <b>Cause</b> list contains about 70 conditions, organized based on the hierarchal structure of the Global Burden of Disease Project of the <a href=" https://www.who.int/healthinfo/global\_burden\_disease/about/en/""> World Health Organization (WHO)</a> and the <a href=" http://www.healthdata.org/gbd/about "">Institute for Health Metrics and Evaluation (IHME)</a>. We have made modifications to the WHO and IHME lists to enhance the usefulness and applicability for U.S., and specifically California, public health priorities and programs. These modifications are described on the Technical Tab of this website.

<br><br>

Depending on the geographic level selected, different levels of detail in the cause list will be available. The more granular the geographic level, the less granular the cause list will become, for statistical stability. [For our partners in Local Health Departments and in other California Department of Health Programs, fully granular data on all levels are available on request.]

<br><br>

<b>Top Level</b> conditions are a mutually exclusive and exhaustive set of five high-level conditions indicated by capital letters: A) Communicable, maternal, perinatal, and nutritional conditions; B) Cancer/Malignant neoplasms; C) Cardiovascular diseases; D) Other chronic; E) Injuries; and Z) Unknown/Missing value

<br><br>

<b>Public Health Levels</b> are a mutually exclusive and exhaustive list of about 60 conditions that are subcategories of the Top Level conditions; they are indicated by numbers associated with the capital letter of the Top Level condition under which they are subcategories. This level provides greater clinical detail and public health program specificity.

<br><br>

<b>Detail Levels</b> are available for a small number of the Public Health Level conditions; these are indicated by a lowercase letter and allow for further specificity.

<br><br>

Our full cause list hierarchy is available <a href="gbd.ICD.MapIMAGE.pdf"">here</a>. Additional information, including a link to the list of detailed ICD-10 codes as they map to all conditions, can be found on the Technical Tab of this website.

<br><br>

We welcome comments and suggested changes to any aspect of our hierarchal list.

'

stateCutHelp <-

"The <b>State-based cutpoints</b> button changes the way the Measure is broken down or grouped for display in the map and the legend. If the box is checked, the cut-points are based on the data for all 58 counties for the most recent five years. If the box is unchecked, the cut-points are based on just the county being displayed in the current map.

<br><br>

Looking at the map with the box checked allows you to see how things compare to the State overall. Looking at the data with the box unchecked allows you to see how things compare within the selected county. For example, if most communities in a given county have high rates of something compared to other counties, that county would be almost all red (indicating high rates) with the box checked, and would show a more varied distribution with the box unchecked.

<br><br>

Different insights can be gained by looking at these maps in the two different ways.

"

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."

# ---- TAB Help ---------------------------------------------------

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, that place matters! 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 map is generally better for using in an external presentation (e.g. pasting into a PowerPoint), since it is 'crisper' without the background map layer. 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 using the <b>Measure Sort Order</b> tool. Different insights can be gained by ranking on different measures (e.g. ranking on 'YLL per 100,00 population' highlights overall burden; ranking on 'mean age at death' shows the conditions that impact young people the most, and ranking on 'Standard Mortality Ratio (SMR)' show those conditions for which a county has particularly high rates compared to the State average). The <b>Levels to Show</b> option allows users to select between broad or narrow categories of conditions. The <b>How Many</b> 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 120% of 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.

<br><br>

Because the 'base' ratio for the SMR is the state rate, the SMR is not shown when the statewide 'California' geography is selected; it would be 1.0 in all cases.

<br><br>

"

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.

<br><br>

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. Users can display confidence intervals by checking the 95% CIs option. In the current version, confidence intervals are only displayed for the crude death rate and age-adjusted death rate but will be available soon for other 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 <b>Social Determinant of Health</b> 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.

"

lifeExpectancyTab <-

" This chart shows the estimated Life Expectancy for California overall and for most California counties, for males and females, from 2005 to 2018, and includes confidence intervals. Life Expectancy is a familiar and widely used measure, which summarizes in one number the 'force of mortality' in a population, and provides an extraordinarily valuable single measure to compare the overall health status between populations.

<br><br>

'Life Expectancy', or technically 'Life Expectancy at Birth', is calculated based on the number of people in each age group that die in a given year and the total size of each age group population in that same year. The method used here is complex and was programed by our colleagues in the California Department of Finance. There is more information about the specific methodology used on the technical tab.

"

# ==== 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."

causeHelpOLD <- 'The list of CAUSES in the app are based on a hierarchal outline arrangement of about 70 conditions. The basis for this outline is the Global Burden of Disease (GBD) Study condition structure, with modifications made to enhance the usefulness and applicability for U.S. public health priorities and programs.

<br><br>

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

<br><br>

The "Top Level" of the outline includes a mutually exclusive and exhaustive set of five high-level conditions:

<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 geographic 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. Because of the greater clinical detail and public health program specificity of this level, we believe this level will provide the greatest insights and the most meaningful rankings of conditions at the county and community levels.

<br><br>

Below this, the "Detail Level" includes a few "Public Health Level" conditions broken down into more detail (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 the 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 "Base Level" condition is shown in the "Detail Level"). A diagram showing this hierarchy, and the list of "Base Level" conditions is available by clicking the "Cause List Info" link. Additional information, including a link to the list of detailed ICD-10 codes that make up each "Base level" condition can be found on the Technical Tab.

<br><br>

We welcome comments and suggested changes to any aspect of our hierarchal list.

'