**INTRODUCTION**

Congratulations, you have been invited to participate in a second interview. The purpose of the second interview is for the selected candidate to demonstrate their analytical quantitative and interpretative skills. The candidate will utilize the Mortality dataset and data dictionary provided and is encouraged to obtain any additional information/data needed to help develop comprehensive conclusions about the data analyzed (examples are: <https://www.tn.gov/health/health-program-areas/statistics/health-data/death-statistics.html>, <https://wonder.cdc.gov/> ).

Please prepare a PowerPoint summary including your interpretation of the data set to present to the interviewer. The data can be presented using graphs, charts, or tables using your preferred statistical package. Your presentation should be no more than 15 slides.  Please cite your data sources.

**INSTRUCTIONS**

You are provided a Mortality Dataset to analyze, and a data dictionary for descriptive purposes. Consider this dataset to represent all deaths to Shelby County residents for this time period.

Be prepared to present and discuss:

* **Descriptive statistics and trends over a 5-year period.**
* **What underlying causes of death are of concern when you compare this subset to the whole state of Tennessee and the United States?**
* **What are the strengths of the dataset provided?**
  + Excel format is widely used and able to be adopted in several different software
  + Includes relevant variables for mortality analysis and descriptive statistics
* **What are the shortcomings of the dataset provided?**
  + Most analyses will require modifying variables into either vectors or other types for more efficient data analysis.
  + Missing/ incomplete data
  + Lack of income data
  + Incomplete definition of “Other” in cause of death
  + Age is not categorical, allows for typos and other errors to go unnoticed
    - Ex: Max age of death is 200 years old
* **What additional information would you have liked to receive to assist you in your analysis?**
  + More detail in Data Dictionary for DRace\_Ethnicity valid values and how it was categorized
  + Datasets for Mortality for Tennessee 2015-2019 and US
    - Data found but missing data for Tennessee 2018 (National Vital Statistics Reports, Vol. 69, No. 13, January 12, 2021 Pg. 60)
  + Total Shelby County Population for each year
* **What additional data did you use (if any) to support your analysis? Why did you select these relevant data sources (if any)?**
  + Tn.gov Death Statistics
  + CDC Wonder
  + CDC [National Center for Health Statistics](https://www.cdc.gov/nchs/index.htm)
  + National Vital Statistics Reports
  + https://www.tn.gov/health/health-program-areas/statistics/health-data/population.html

Please email your presentation to [Epidemiology.Section@shelbycountytn.gov](mailto:Epidemiology.Section@shelbycountytn.gov) at least one hour before the time of your interview.

Call our office at [901-222-9243](tel:901-222-9243)  with any questions or concerns.

**Data Dictionary for Exercise Dataset**

|  |  |  |
| --- | --- | --- |
| Variable Name | Description | Valid Values |
| DDODYr | Decedent’s Year of Death | 2015-2019 |
| DSex | Decedent’s Sex | M = Male  F = Female  U = Unknown |
| DRace\_Ethnicity | Decedent’s Race and ethnicity | This is a text string indicating the race and Hispanic origin |
| DAgeYrs | Decedent’s Age in Years | 1-89 represents the age of the decedent  90 represents decedents aged 90 and over  blank = age unknown, includes infants < 1 |
| DEducationCode | Decedent’s Education | 1 = 8th Grade or Less  2 = 9th-12th but No Diploma  3 = HS Graduate or GED  4 = Some College but No Degree  5 = Associate's Degree  6 = Bachelor's Degree  7 = Master's Degree  8 = Doctorate or Professional Degree  9 = Unknown |
| RankableUnderlyingdeathcause | Cause of death categorized | This is a text string indicating the cause of death category |

Dataset Title: Mortality dataset 2015-2019