

Capstone Project Proposal

Anushree Shivarudrappa

May 25 2016

1. What is the problem you want to solve?

I am considering “Death in United States” dataset for my capstone project. I am interested in finding the solutions for below listed questions

1. Main leading cause of death in US.
2. Male-to-female suicide death ratio ? Who likely take their own lives ?
3. Effect of Age & Education level on suicidal deaths.
4. Is Marital Status has effect on suicidal cases ?
5. Which age group has more accidental death ? Is younger group shows more rate ?

2. Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn't have otherwise?

Every year the CDC releases the country's most detailed report on death in the United States. Analyzing mortality data is essential to understanding the complex circumstances of death across the country. The US Government uses this data to determine life expectancy and understand how death in the U.S. differs from the rest of the world

3. What data are you going to use for this? How will you acquire this data?

I am using Death in United States dataset. This mortality dataset is a record of every death in the country for the year 2014, which includes detailed information about causes of death and the demographic background of the deceased.

4. In brief, outline your approach to solving this problem.

Loading my data, data wrangling, exploratory data analysis, statistical model to answer my data science questions and visualization of my results

5. What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck

It include my final report, R programming code. All the details will be available in GitHub repository.