Team 5 – Ivan Aguirre, minu khosla, shweta Kalavar

September 23, 2019

# Overview

## Project Title

|  |  |
| --- | --- |
|  | Tobacco Usage in USA |

## Project Description Outline

|  |  |
| --- | --- |
|  | There are numerous factors that influence the decision to start smoking or to use tobacco. These factors include age, income, education, gender. The analysis of this data is to show the percentage of people that use tobacco. |

## Research Question to Answer

|  |  |
| --- | --- |
|  | Below are the questions for analysis of Tobacco Users across USA |

* What is the Percentage of people in USA who use tobacco?
* State-wise, gender based, over time
* What is the level of Education for the Adults use tobacco?
* What is the Income level for the people who use tobacco in USA?

## Data Sets to be used

|  |  |
| --- | --- |
|  | Tobacco Use – BRFSS - <https://healthdata.gov/dataset/behavioral-risk-factor-surveillance-system-brfss-prevalence-data-2011-present> |

## Rough Breakdown of Tasks

|  |  |
| --- | --- |
|  |  |

* Extract and cleanse the Data Set
* What is the Percentage of people in the United States who use tobacco?
  + Broken down by State
  + Broken down by Gender
  + Trended over time
  + API Call to see change in average percentage of tobacco users
* What is the level of Education for the Adults who smoke?
  + Less Than High School
  + High School Graduates
  + Some College
  + College Graduate
* What is the Income level for the people who smoke in Georgia?
  + Less than $15,000
  + $15,000 - $24,999
  + $25,000 - $34,999
  + $35,000 - $49,999
  + $50,000 - $74,999
  + $75,000 or More
* Prepare a power point presentation to reflect the data analysis for people that uses tobacco in USA.