Contrast: Tobacco Use and Education Level

This visualization will be geared towards examining the relationship between different kinds of tobacco use and corresponding education level

- Education level: "< 12th Grade", "12th Grade",   "> 12th Grade"
  - Data that falls into these categories will be compared to each other
- Tobacco use: "Cigarette Use (Adults)", "Smokeless Tobacco Use (Adults)", "E-Cigarette Use (Adults)"
  - Rate of occurrence for each category as it pertains to education level

Story Intention and Takeaways

- I expect the data to reflect a negative correlation between increased tobacco use and higher education level
  - In other words, this is what I want to determine… are these things correlated or is persistent tobacco use tied to something else?
- I want our final project to examine societal biases and either validate or invalidate them
  - We expect education would lead to smarter, healthier habits
  - Perhaps biases are false
  - Maybe the opposite of what we expect is true
- Visualizations will aim to answer this question

Visualization Plan (Main Comparison Variables)

- **Education Data**: Sort Education by the three categories "< 12th Grade", "12th Grade",   "> 12th Grade"
    - Do not use "All Grades" because this does not filter the data and says nothing about education
- **Smoking Data**:
    - Topic.Desc - specifies type of tobacco use which should be the main categories of the smoking data portion of the viz
        - "Cigarette Use (Adults)", "Smokeless Tobacco Use (Adults)", "E-Cigarette Use (Adults)"
        - NOT include "Cessation (Adults)"
    - Measure.Desc -
        - "Current Smoking", "User Status", "Current Use", "Smoking Status", "Frequency of Use", "Smoking Frequency", "Percent of Former Smokers Among Ever Smokers"
    - Response -
        - "Not Current", "Current", "Every Day", "Some Days", "Never", "Former"
    - Pairs required of some Measure.Desc and Response values to indicate smoking
        - "Smoking Frequency" or "Frequency of Use" - "Every Day" or "Some Days"
        - "User Status" or "Smoking Status" - "Current"
    - Data.Value - percentage of these different smoking conditions which all pertain to conditions where smoking is occuring
        - How to display? Can I treat these as cumulative? Not really… maybe change them to rates of some sort? Or average them all for certain education/age levels since percentage is a comparable metric. These values should be combined by the
    - Key: Data filtered for this value and value included in pair

Visualization Plan (Filters)

- The data set needs to be filtered down by a few factors before appropriate analysis of education level and tobacco use can be conducted (as outlined on previous slide)

- Include a filter that allows selection of a time frame (by month)

- Age Data: Select from all age values apart from "All Ages" because the remaining categories can apply to all of the education levels
  - "Age 25 and Older", "18 to 24 Years", "25 to 44 Years", "45 to 64 Years", "Age 20 and Older", "65 Years and Older"
  - This data is what to pull other information from, it is not a variable to display in the visualization

Visualization Plan (Rough Image)

- This chart will be made into an interactive visualization in a Shiny app
- The sliding filter for setting the time frame will allow the data displayed to change with the different settings
- The exact metric used to reflect "rate of persistent use" is yet to be concretely determined
  - Likely an average of the percentage rates

EDUCATION AND TOBACCO USE IN WASHINGTON