Brevory Foster, Sijan Wu, Calvin Hathcock, Christine Chee

### Potential Data Sets

1. Prevalence and Trends Data: Tobacco Use from Data.gov
   * link to dataset: <https://catalog.data.gov/dataset/brfss-prevalence-and-trends-data-tobacco-use-four-level-smoking-data-for-2011>
   * Federal Data Set from the Behavioral Risk Factor Surveillance System that just measured the prevalence of tobacco usage
2. Hotel Booking demand
   * link to dataset: <https://www.kaggle.com/jessemostipak/hotel-booking-demand>
   * Dataset containing booking information that is used to measure the frequency of when most hotel customers book a room.
   * The dataset also contains other information such as when a room is booked, the length of stay, cancelations, etc.
3. Affordable Housing by Town in Connecticut (2011-Present) - Data.gov
   * Link: <https://catalog.data.gov/dataset/affordable-housing-by-town-2011-present>
   * Dataset containing statistics on affordable housing by town in the state of Connecticut. It includes information on the number of units, amount of assisted units, year, and percentage of affordable housing.

4. Investing factory in the world: invest factories in the world- world bank open data

* Link to dataset: <https://databank.worldbank.org/source/jobs>
* Dataset includes all the employment in service. It shows information at which country, what year, percentage of empolyment.

### Customers

1. Big Tobacco Companies
2. Hotel/Resort owners
3. Apartment companies, charities
4. Charity and companies who want less priced labors

### Problem Solved

1. It will increase their revenue by giving them the top locations in USA where there are highly profitable customers
2. It will optimise the amount of labor needed to run the hotel/resort during not-busy and busy seasons.
3. It will make it easier to see poorer communities who need assistance or communities that could use more affordable options for living.
4. It will make charities and corporations spend their money more efficiently and reasonable.

### Product Vision

1. The target customer for this data visualization are the Big Tobacco Companies. The companies would use our product because it could drive their profits through the roof by selectively marketing to their most reliable customers. It’s called Smoker’s Compliment. Big Tobacco should use our product because Big Tobacco is dying out to the various psychoactive markets chief among them being the e-cigarette market.
2. Hotel and resort owners will be the target of the data visualization. They will be able to use our product to determine the amount of labor needed during the different seasons of the year. They will be able to optimize their workers schedules along with optimizing their finances for who they should hire. Hotel and resort owners would want to use this product because they want to be able to know when and where to invest their money.
3. For apartment companies and charities for the poor. The organizations who use our product are in need of easily accessible and understandable data on Connecticut housing. The product name would be Connecticut Housing Info. Apartment companies will use our product because it will allow them to see which communities have demand for lower end and affordable housing. Charities would use our product to identify communities in need of housing assistance. Unlike Connecticut’s State Website, our product will be more user friendly.
4. Many charities can use our product to spend their money on the areas where people lack jobs. The companies who need more low-priced labor should take advantage of this chance, which means investing a factory in this area will save more money for these companies. The charities will help more people in the world where there is the highest unemployment. So my product probably can reduce the risk of crime and drugs and many other bad things.

### Major Features

1. Shows the States on a color scale of highest profitability to lowest profitability judging by prevalence of smokers and population size. They are also numbered so if the user wanted to compare they would just have to select the highest number.
2. The main features would be the calculator. The user would be able to input their own data such as number of rooms and the product will be able to calculate the recommended optimum number of labor required based on the data form the dataset.
3. The main feature would be easy to read data visualization that is interactable. This way, the user does not have to be well versed in data analytics to understand what the data means and how they could use it.
4. Show the world where it has the highest percentage of employment. The users should be able to see the darker areas in the map and know where has the cheapest labor. So they don’t need to check this website online by themself. Colors will be a better indicator.