Slide 1:

- Introduce our Project

Slide 2 Goals of Our Project:

- Our initial goals from our project plan were to determine if there was a correlation between a state's health data and the number of Planet Fitness gyms in each state.
- Regarding health data, we had the idea to either use a BMI or an exercise BMI and found the number of Planet Fitness gym locations on their website. We planned to create visualizations through plotly and write our code in a jupyter lab notebook.

Slide 3 Goals That We Achieved:

- With a more detailed measure of health, through specific questions relating to leisure activities and obesity qualifications, we were able to find correlations between the number of Planet Fitness gyms and health of specific age categories in the United States.
- We had to pivot and find a new API, which we found and used for finding state abbreviations and was our primary key, and found the health results through a nutrition physical activity and obesity behavioral risk factor surveillance system survey from the CDC.
- We chose two questions to focus on for the purpose of this project which were the percentage of adults who engage in no leisure-time physical activity and the Percent of adults aged 18 years and older who have obesity. We followed through with creating visualizations through plotly and jupyter lab notebooks.

Slide 4 Problems We Faced:

- The first problem we faced was having to find a new API once we realized that our original one from our project plan was simply not going to work. We solved this problem by finding a new API which was the state abbreviations. Working with an API in general meant that we had to sift through a lot of information, especially through the nutrition physical activity and obesity behavioral risk factor surveillance system survey from the CDC.
- Once we sifted through the data in the API, we struggled at first to find a significant correlation with the health data and number of gyms. We tested age, income, education, race, and gender and found that the strongest correlation existed with age. In general, the obesity and leisure data were very similar so it was hard to understand the correlations at first.

Slide 5 Calculations From The Data in Database:

- These are the calculations we found
- For simplicity reasons, our data is in the "Scratch data" file and then the "Analysis" tab has our tables and visualizations.

Slide 6 Table with State, Population, and Abbreviations:

- We needed to write df, health, leisure and population into our database and then merge them to create a large table with all of the values.
- We got the abbr, which are our abbreviations from the state API, and included the population, as we would use it for the number of planet fitness gyms/100k population that we will see later.
- We noticed that there are 51 states, because they include Puerto Rico and the District of columbia. This means that there was a state or two that were missing

Slide 7 Leisure Data:

- Next, we created a table with our leisure data. While we did not need all of the columns that are shown in this picture, we decided to keep them here, and then in our visualizations we only sought to include certain columns in our query. Here, important column to highlight are the

question, which was “the percentage of adults who engage in no leisure-time physical activity,” and stratification/Stratificationcategoryid1, as those are the classifications we will use later to compare income, age, gender, race, and education.

Slide 8 Obesity Weight Classification Table:

- Then, we created the same table but rather than the data associated with leisure, we looked at data corresponding to obesity classification, specifically the question “Percent of adults aged 18 years and older who have obesity." Again, some key columns from this table include stratification and Stratificationcategoryid1, as those are the classifications we will use later to compare income, age, gender, race, and education.

Slide 9 Merged Table With Primary Key (State_ID_fk):

- After creating 3 separate tables, we merged them with the primary key, which is state_id_fk, and renamed some of the columns so they were consistent across all of the tables.

Slide 10 Number of Gyms Tables:

- Next, we created a table showing the number of planet fitness gyms across all of the states from the state abbreviation API and included the obesity and leisure data here as well. Something important to note here is that the data is not specified by the stratification of inclume, education, gender, race, and age, but is by the “OVERALL”. We made sure to add population to this data table, as we would need it for our visualizations where we found the number of gyms per 1000 people in the population.

Slide 11 Visualizations We Created:

- Now, we are going to explain the visualizations we created.

Slide 12 Gyms vs. Obesity Scatterplot (by age category)

- The first scatterplot we created compared the amount of gyms vs. obesity by age category. Overall, there is a slightly positive correlation between obesity and the number of Planet Fitness gyms for the ages of 65 and older meaning that as the obesity rates rise, so do the number of gyms. This makes sense because less people tend to workout less when they get older. For the younger age groups, it surprised us that there was a positive correlation because we predicted that younger people would have lower obesity rates due to people going to the gym.

Slide 13 Gyms vs. Leisure Scatterplot (by age category)

- The second scatterplot we created compared the amount of gyms vs. leisure time by age category. For each age group, there was a slight positive correlation between leisure time and the number of Planet Fitness gyms. This made more sense to us because we predicted that Planet Fitness would want to put more gym locations in the states that had populations with more leisure time. Overall for both of our scatter plots, these correlations were very very slim. It was really interesting to conduct this exploration and it taught us a lot about converting between APIs, JSON files, databases, and visualizations.

Slide 14:

- Thank you for listening. Let us know if you have any questions!