**Research Paper:**

**Intro**

Background:

Meat is globally recognized as a prime source of protein within a human diet. Recent news and studies have questioned whether or not meat actually plays a positive role in the human diet. Meat consumption varies by a number of factors such as country, continent and type of meat, all of which vary over time. Different eating habits, cultural preferences and overall food trends may impact the quantity of meat consumed within each country. Different countries also consume different types of meat; for example, the United States does not traditionally consume sheep meat, while other countries such as New Zealand and Australia consume significant quantities of sheep meat(*Agricultural output - Meat consumption*). Wealth also may impact the amount of meat consumed; countries with higher gross domestic product(GDP) per capita have more of a fiscal opportunity to purchase meat products. Meat consumption on the basis of geography, type of meat, and time all potentially impact life expectancy.

Previous research:

A previous study was conducted to see if a combination of red meat and alcohol have negative impacts on human health at a macro and country level. This study produced results indicating that high levels of red meat consumption do appear to have a negative impact on life expectancy, more specifically in countries with higher incomes(Ranabhat, 2020). A second study explored the relationship of a low-to-no meat diet with life expectancy. This study showed that long term diets consisting of lower levels of meat may contribute to a longer life span (Singh, 2003). Another study was conducted on whether or not red and processed meat contribute to higher mortality levels, resulting in a positive relationship, especially for processed meat(Larsson, 2013). A different study analyzed mortality rates of vegetarians compared with meat-eaters, which concluded that “vegetarians have a lower risk of dying from ischaemic heart disease than non-vegetarians”(Key, 1998). Lastly, a European study was conducted to analyze the same relationship- between red meat, processed meat and mortality; finding again that processed meats have a positive association with mortality, in this case specifically due to cardiovascular disease and cancer(Rohrmann, 2013).

Study objective:

The objective of this research is to show the effect of different types of meat on life expectancy. Life expectancy is impacted by many variables, however the quality and type of food being eaten is one of the main factors. Understanding how variables impact life expectancy is key to understanding the best way to make decisions to potentially extend one’s life.

**Methods**

* **Assumptions:** 
  + All other dietary items consistent among populations
  + Quality of meat consistent among populations
* **Data**

Meat consumption data was imported as full indicator data from the Organisation for Economic Co-operation and Development (OECD) web page. Data on life expectancy, GDP per capita, and population was used from the gapminder package in Rstudio, which can also be found in its raw form on the gapminder website (<http://www.gapminder.org/data/>). These data sets each included numerous countries over differing time spans.

* **Meat Consumption Data**

The meat consumption data originally contained columns labeled: location, indicator, subject, measure, frequency, time, value and flag codes. Location contained country codes rather than country names for 35 countries and 3 non countries(WLD,OECD, BRICS). Indicator was the same for every row: MEATCONSUMP. Subject indicated the type of meat being consumed and was divided into four categories: poultry, beef, sheep and pig. Measure was divided into two measurement types: KG\_CAP, “kilograms of retail weight per capita”, and THND\_TONNE, “thousand tonnes of carcass weight (except for poultry expressed as ready to cook weight)”(*Agricultural output - Meat consumption*). Frequency was the same for every row: A. Time was from 1990-2026, conducted yearly, with actual data up to 2020 and forecasted data until 2026. Value was the actual measured data measured to the thousandth decimal. Lastly, Flag codes again had the same NA value for every row.

* **Gapminder Package**

The gapminder package consisted of six columns: country, continent, year, lifeExp, pop and gdpPercap. Country had 142 country names. Continent had 5 groups: Asia, Europe, Africa, Americas and Oceania. Year was from 1952-2007 conducted at five year intervals. LifeExp was recorded to the thousandth decimal place. Population was recorded as whole numbers. GdpPercap was to the ten thousandth decimal place.

* **Data Analysis:**

Both data sets were imported into an R script with the platform RStudio. Both the “gapminder” and the “skimr” packages were installed first. Then a collection of libraries were installed: tidyverse, lubridate, “gapminder”, moderndive, dyplr, and skimr. Once the libraries were loaded the data sets were imported as “meat\_consumption” and “gapminder”. Next, “meat\_consumption” had location data containing only country codes, so a new column was created using the mutate function called “country” and the country codes were manually mutated to the appropriate country names. Next, the data was tidied and the columns "FREQUENCY", "LOCATION" and "Flag Codes" were eliminated and then the meat data was separated by measurement type. This allowed the data to only include kilograms which was then rounded to the hundredth decimal place and converted to pounds. The meat consumption data included values forecasted out until 2026 so the data was then filtered to only include values up to 2021.

* **Plot 1-3**

The first graph that was plotted used the ggplot function and had time on the x-axis and Per Capita Consumption in Pounds on the y-axis, while color was the type of meat and the graph used facet\_wrap to separate by country. Then, some of the countries with noticeable graphs were plotted to observe more clearly; the United States, Argentina, Israel, New Zealand and Paraguay. Next, a new data set was created that spread the meat types into 4 separate columns using the spread function and this was then used to graph the same data, however this time the values for each type of meat was summed to a single total.

* **Plot 4-7**

The next step taken was to combine the two datasets. The “year” column had to be renamed to “TIME” in the gapminder data in order for the data to be joined by both year and country which produced “gapminder2.0”. This new dataset was created using the left\_join function and was named “combined\_data” and the na values were immediately filtered out. A new graph was then created to compare the meat consumption data on the x-axis with the life expectancy data on the y-axis. The color again indicated the type of meat and using the facet\_wrap function they were separated by country. A linear model was then created to observe the statistical output when life expectancy is a function of meat consumption value, country, meat type and time. Next, the data was transformed again; this time the dataset containing the spread meat types was joined to the gapminder2.0 data. ANother graph was plotted, this time looking at life expectancy vs meat consumption by country but with the meat values summed to a total. Another model was created to look at the statistics of this graph as well. The data was then changed again; the four columns were summed to create a new column with the sum value. This time the data was graphed divided by continent rather than country. Two graphs were created; one with meat type separate and the other with meat type summed from the new dataset. A model was created for each of them.

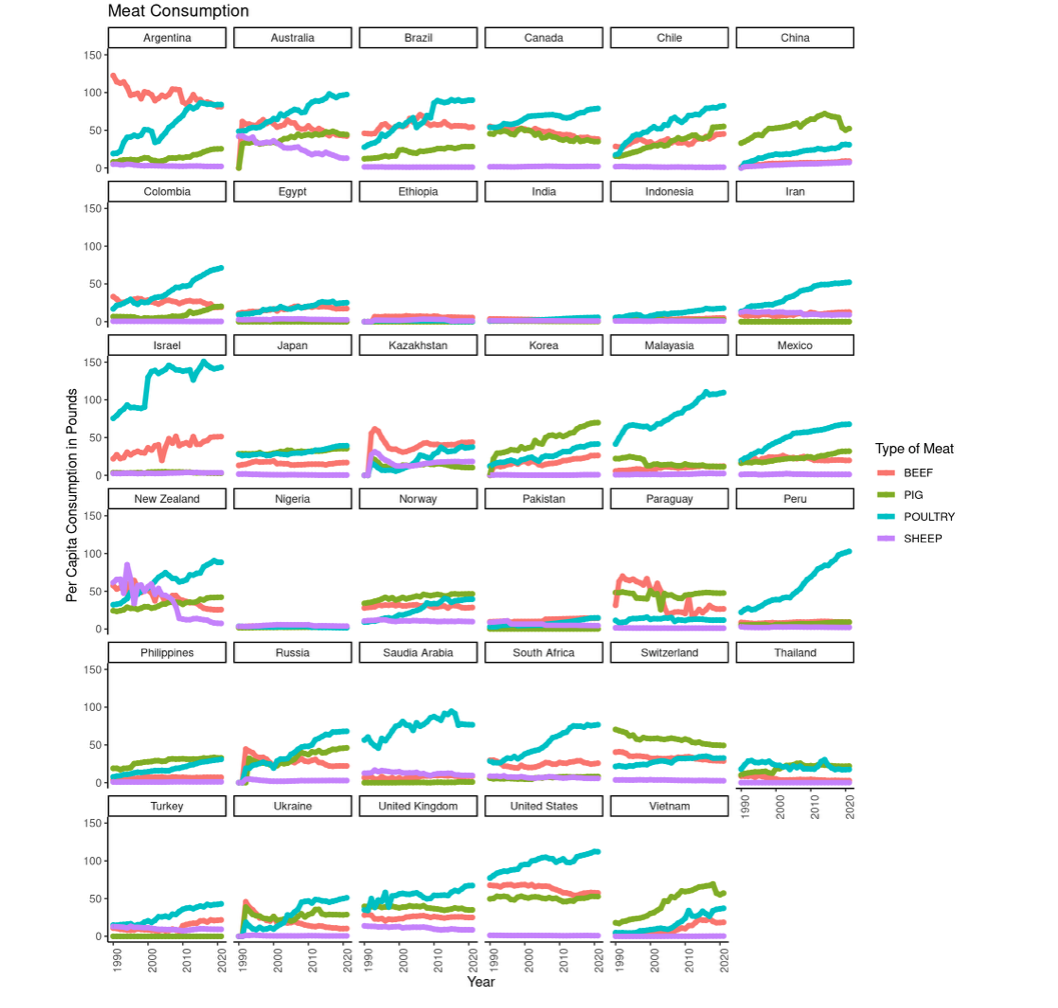
* **Plot 8-11**

The meat consumption data was then compared against GDP per capita rather than life expectancy. A function called “addUnits” was created for values with multiple zeros to be written a certain way; for the purposes of these graphs it was used to write anything with three zeros as “k” for GDP values(ex. 1000=1k). A graph of both meat by type and a summed meat value vs GDP per capita were both created with a corresponding linear model. The GDP data was then mutated to create a column dividing it by percentage. From 0-25% of GDP per capita was labelled “low”, 25-50% was considered “low-medium”, 50-75% was considered “high-medium” and the final 25% was labelled “high.” This was done for both the dataset with meat separated by type and the data set with the meat values summed which then allowed for two more graphs divided by income label rather than country or continent. These graphs showed life expectancy as a function of meat consumption again, with one showing meat type distinguished and the other showing a sum total for meat consumption. A linear model was created for each of these graphs. This time the linear models were split by percentage and viewed by each different percentage separately. A map of the r-squared values were also produced for each in order to compare them.

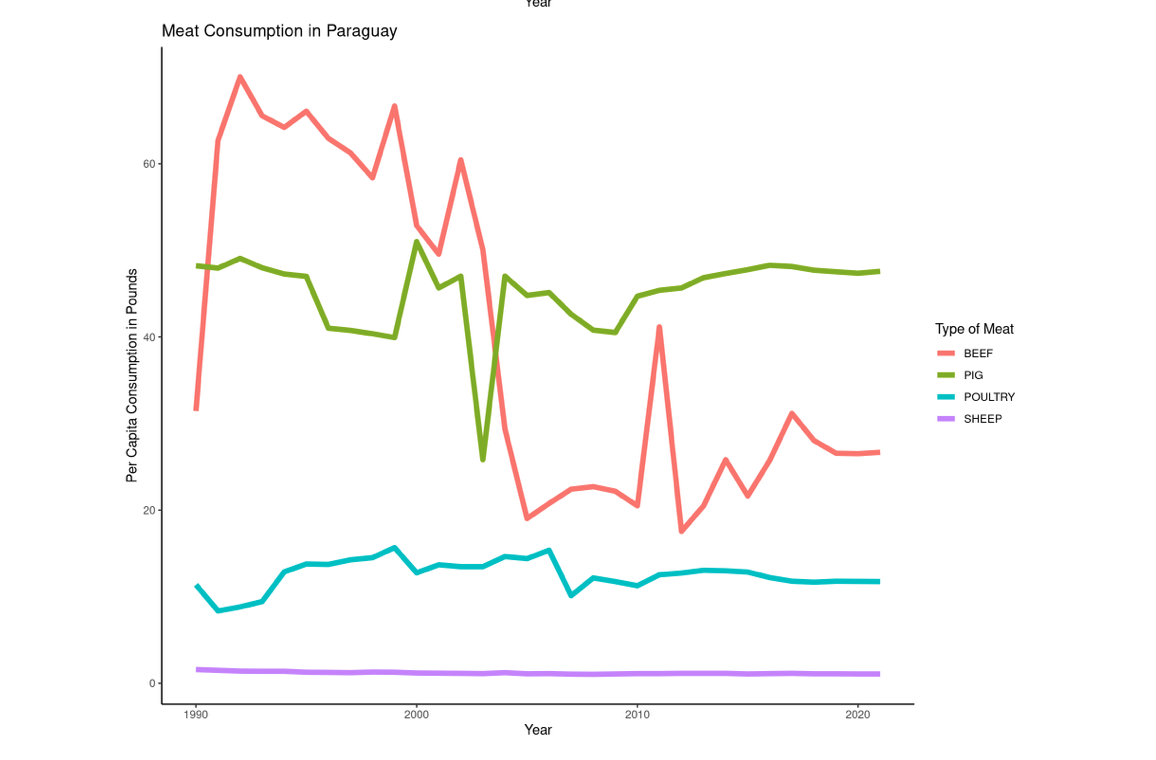
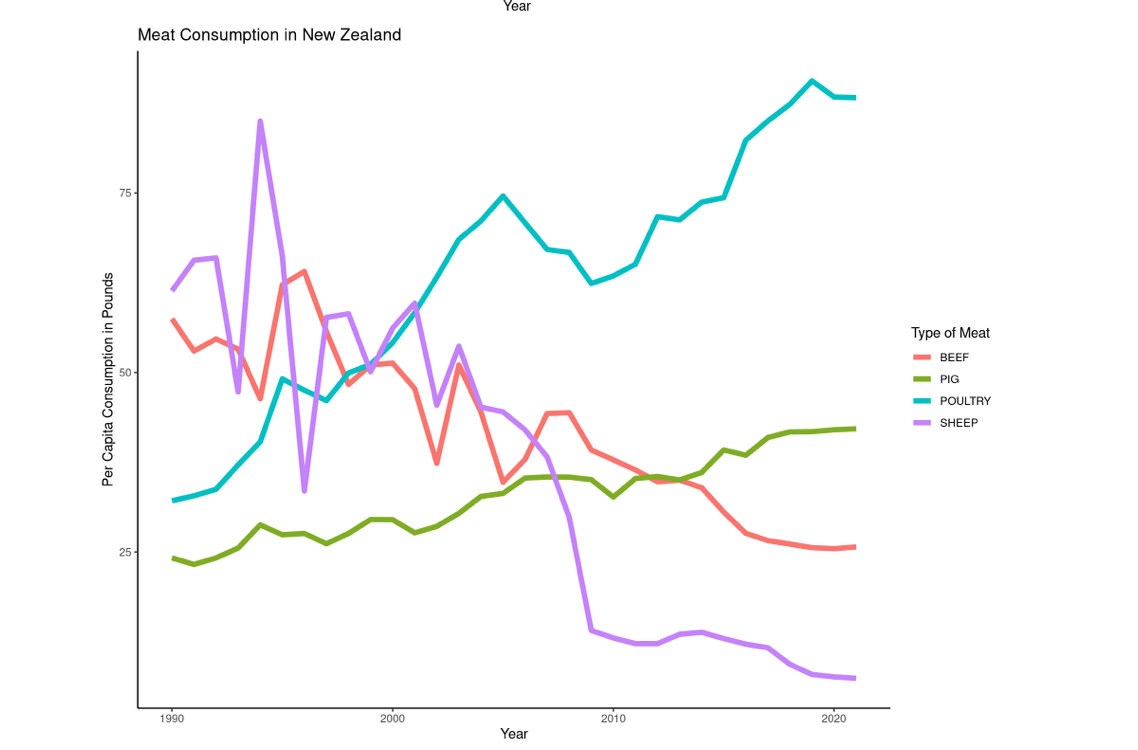
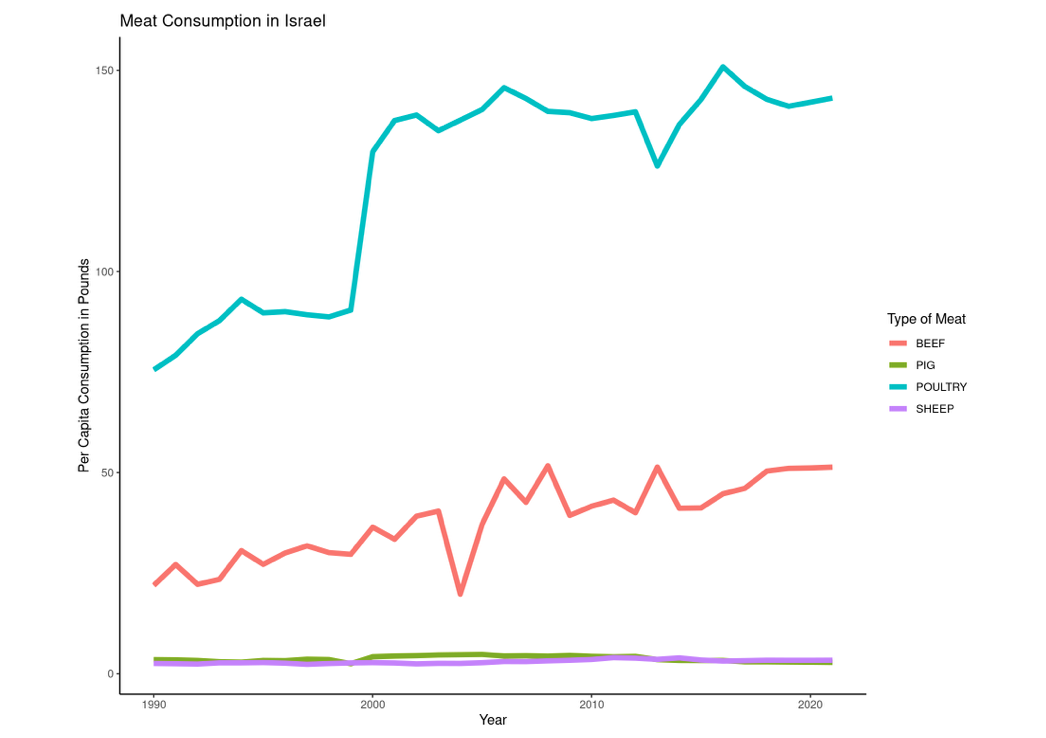
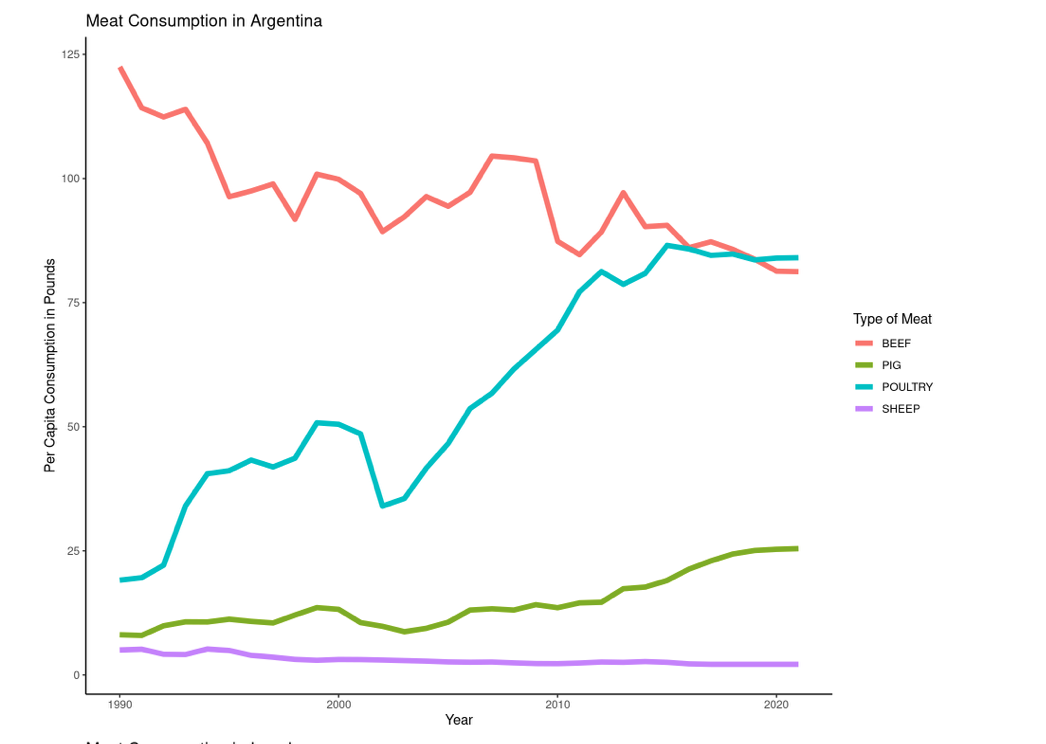
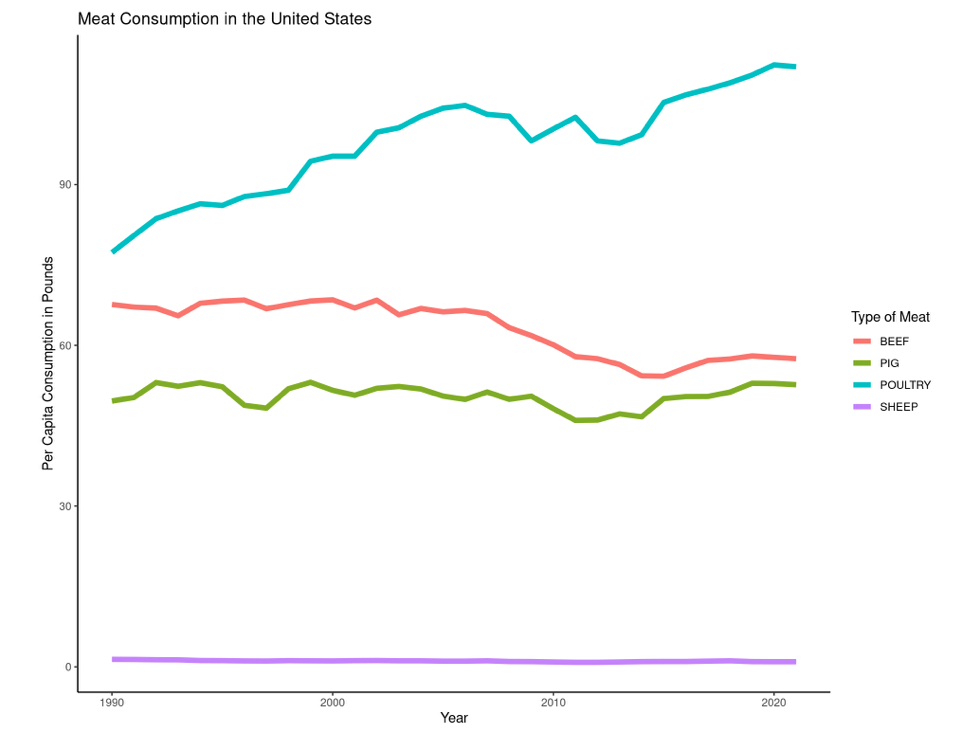
**Results**

* **Plot 1-3**

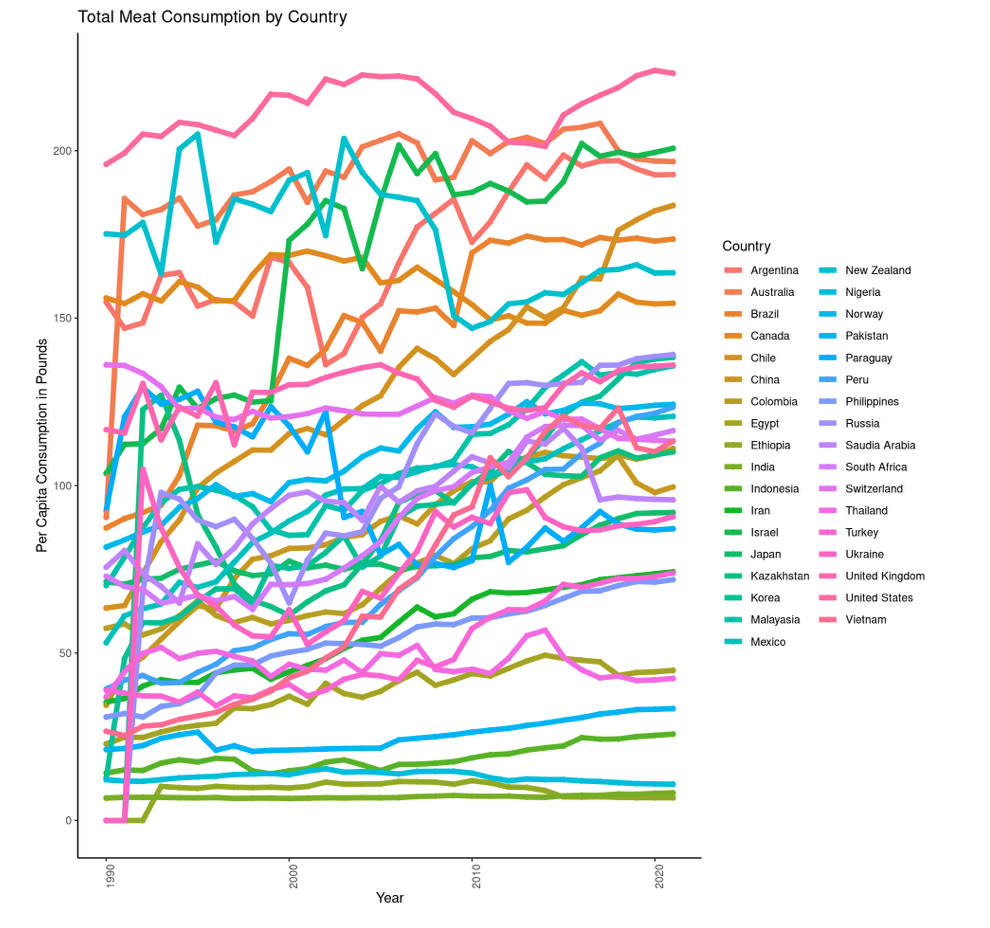
Plots 1-3 were only observing the meat consumption dataset. The first plot showed a few countries that stood out with high levels of meat consumption.



The second set of plots took a closer look at those countries that had high meat consumption.

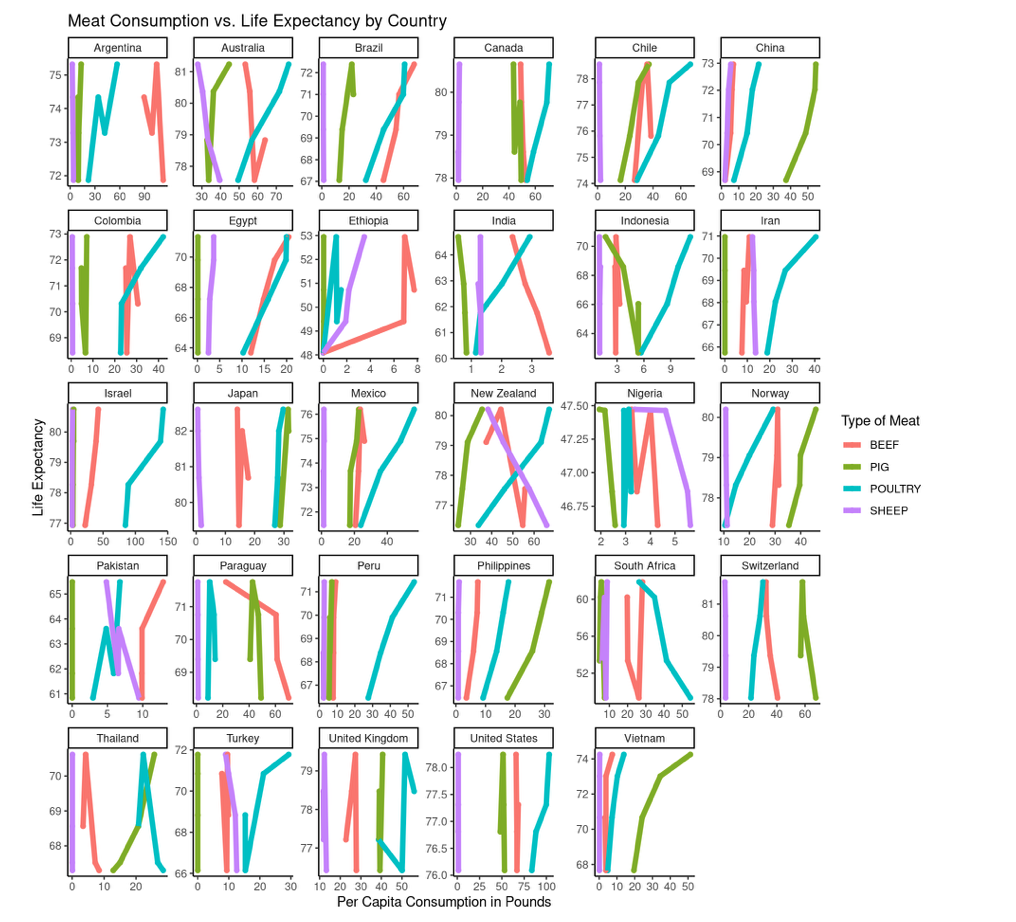


The third plot summed the meat consumption data and then graphed by country; showing that the United States is almost always the largest consumer of meat.

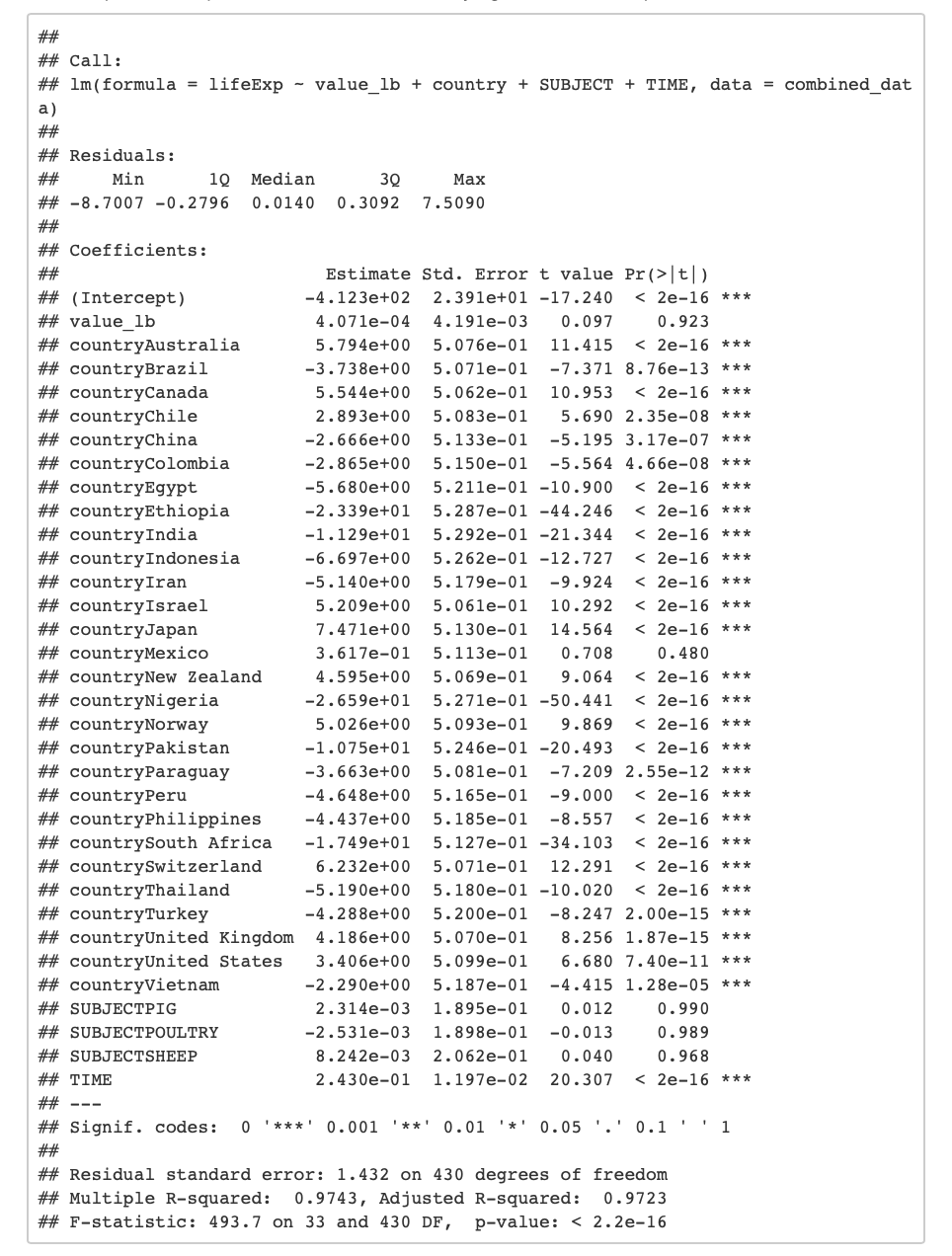
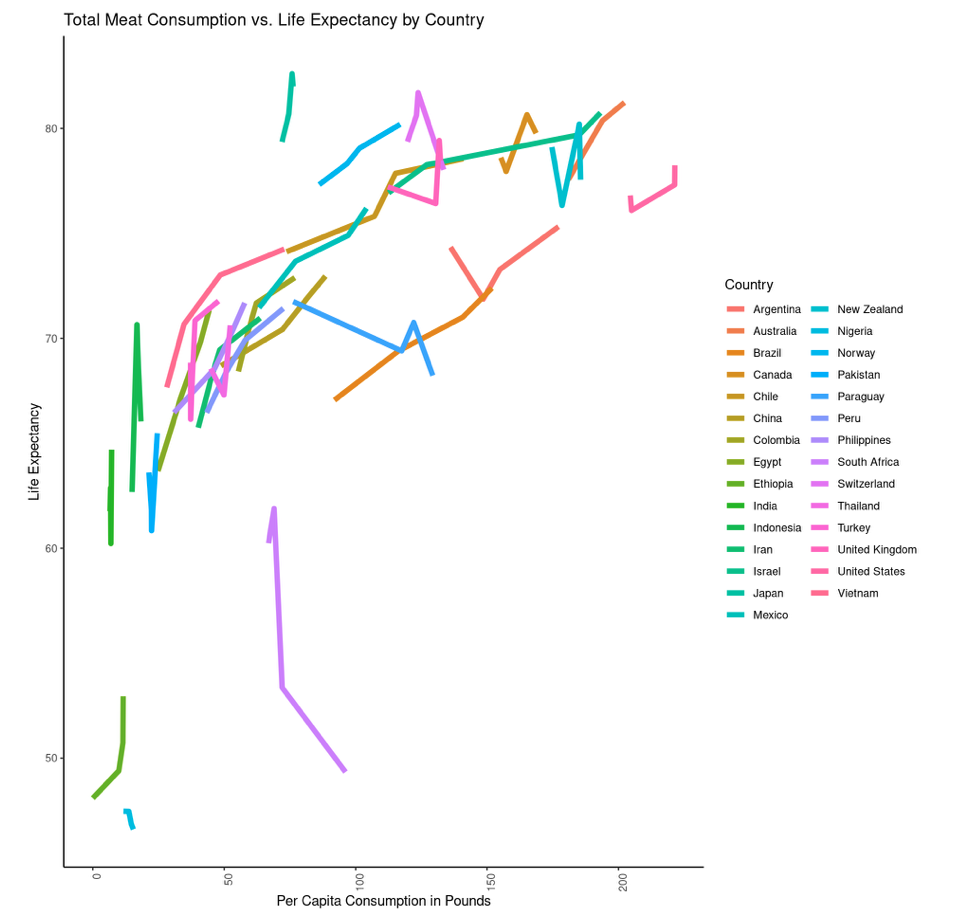
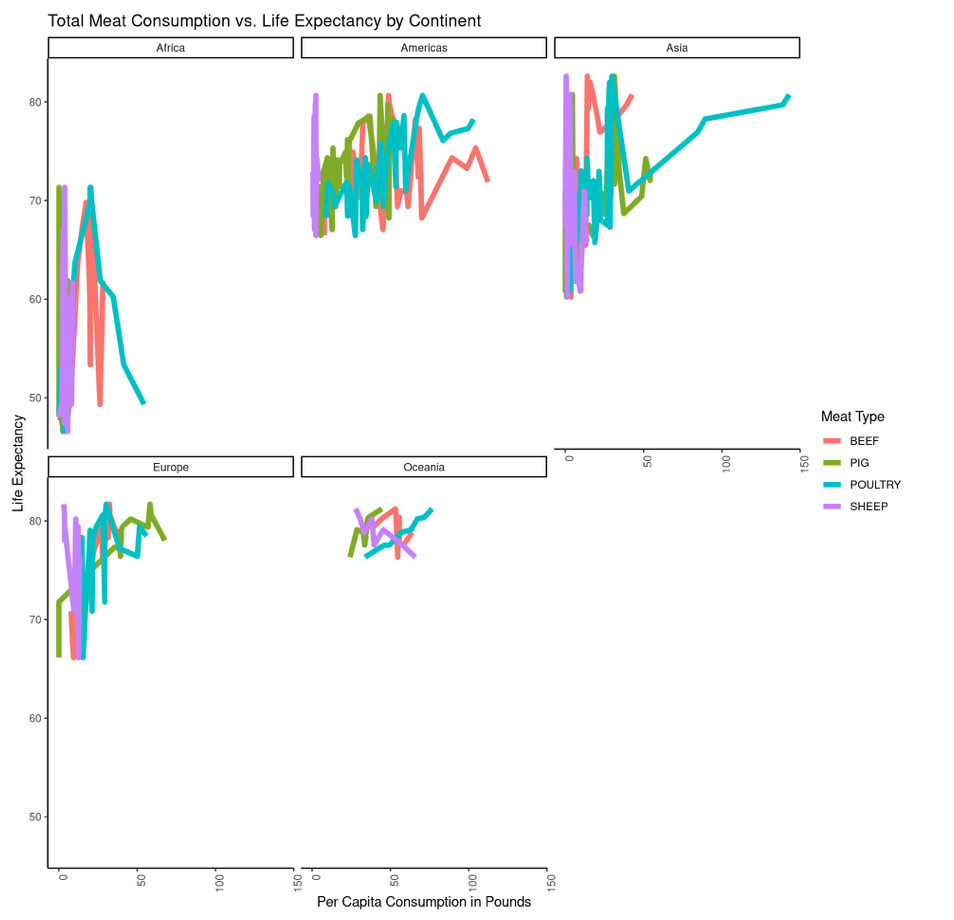
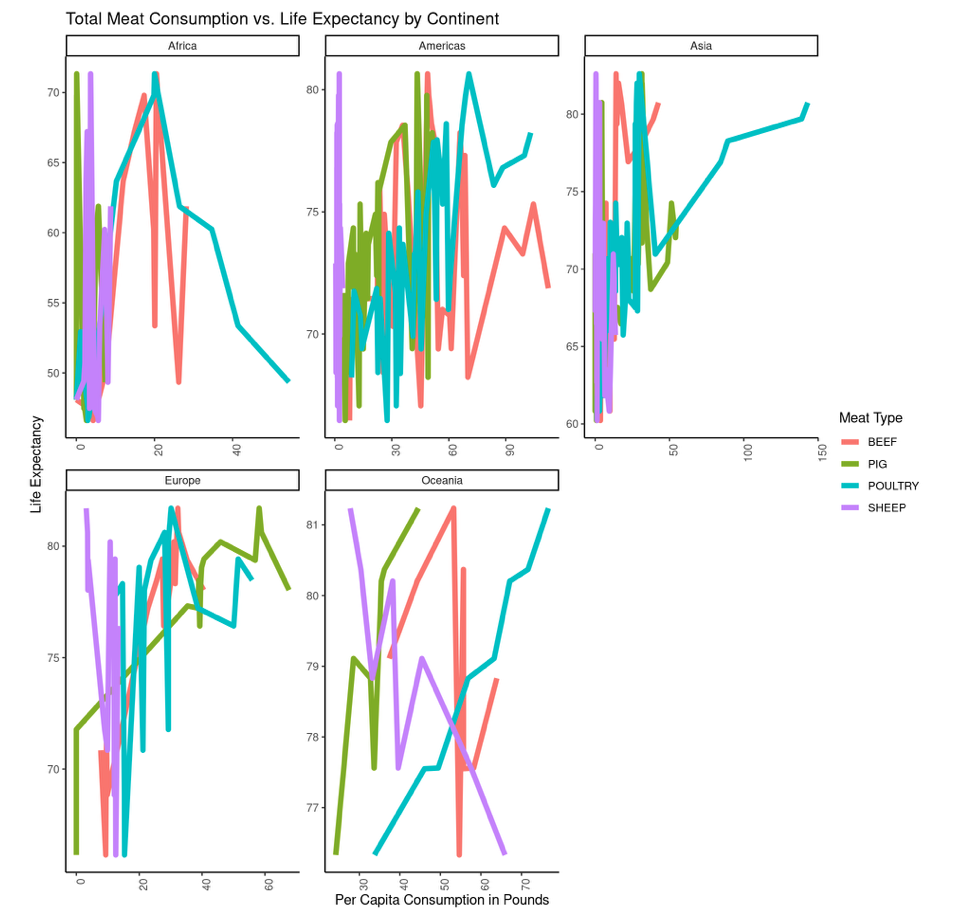
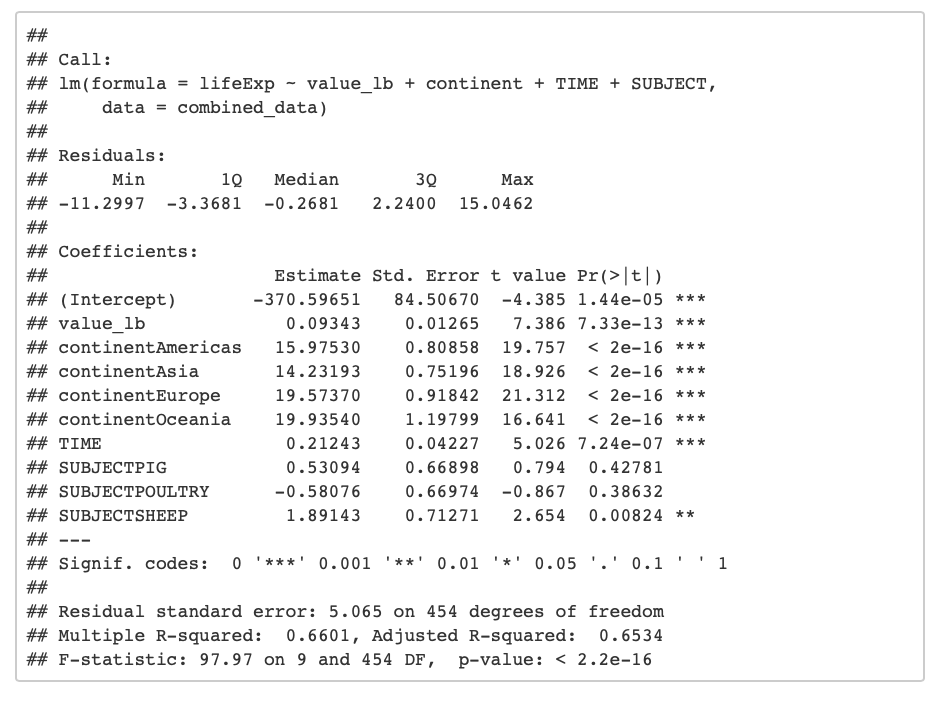
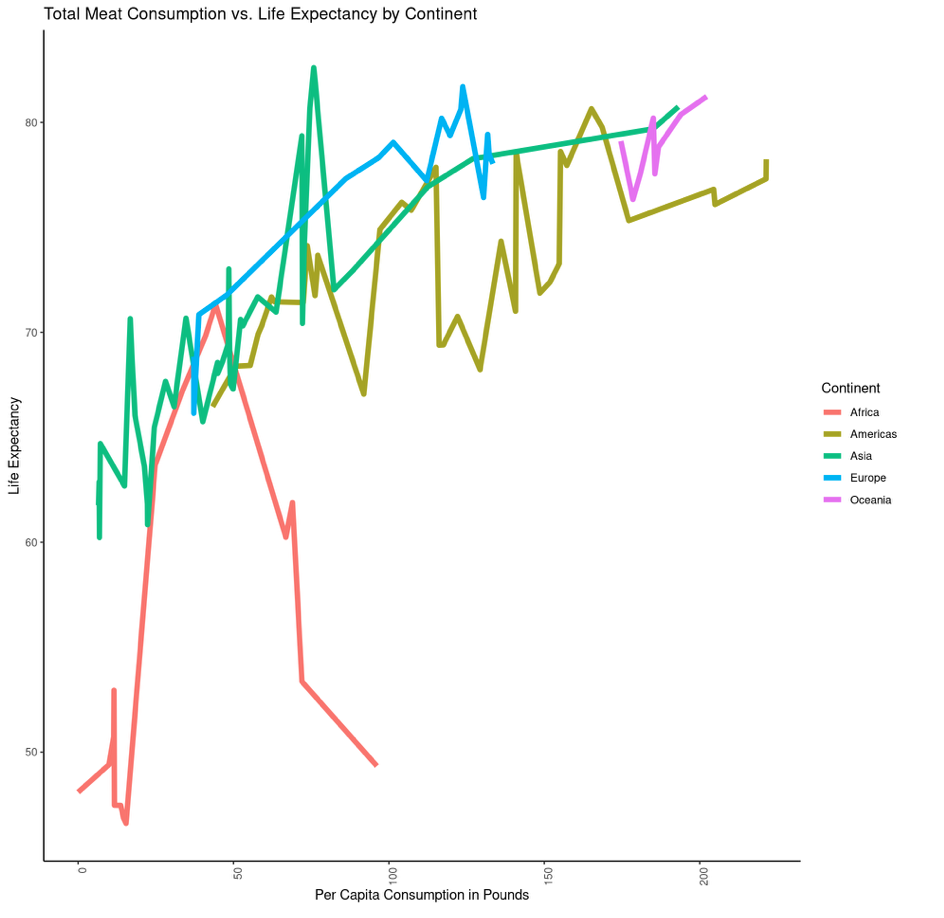
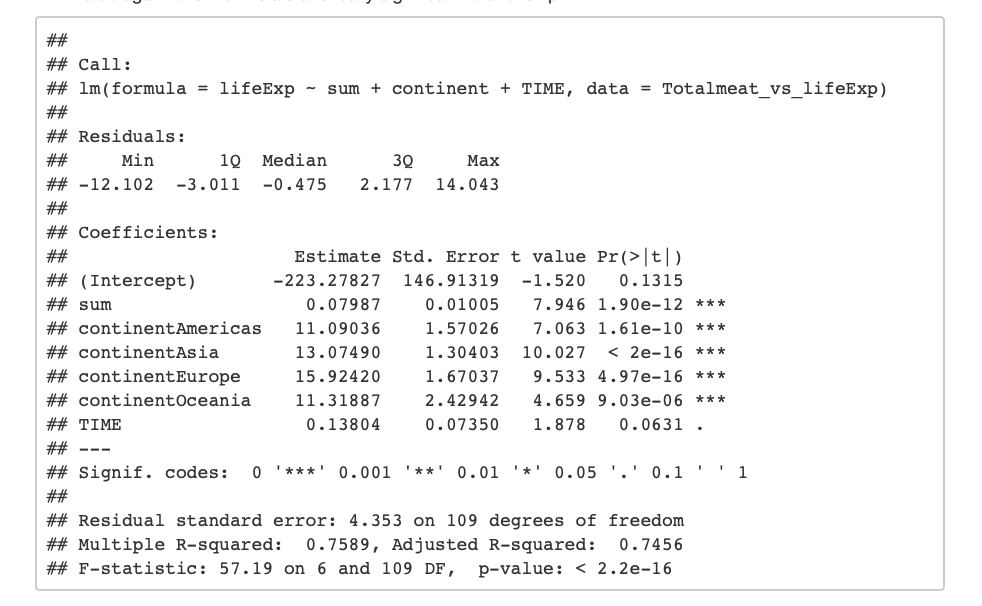


* **Plots 4-7**

The fourth plot was graphed twice because the first time it was challenging to observe any patterns.

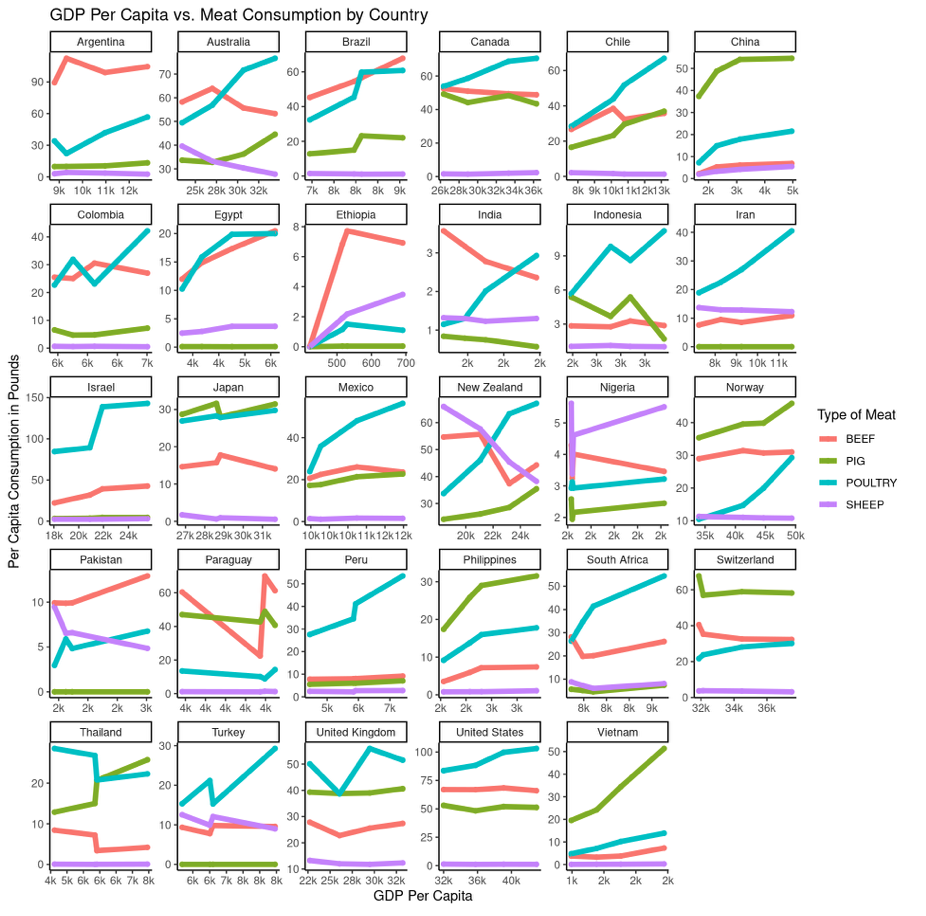
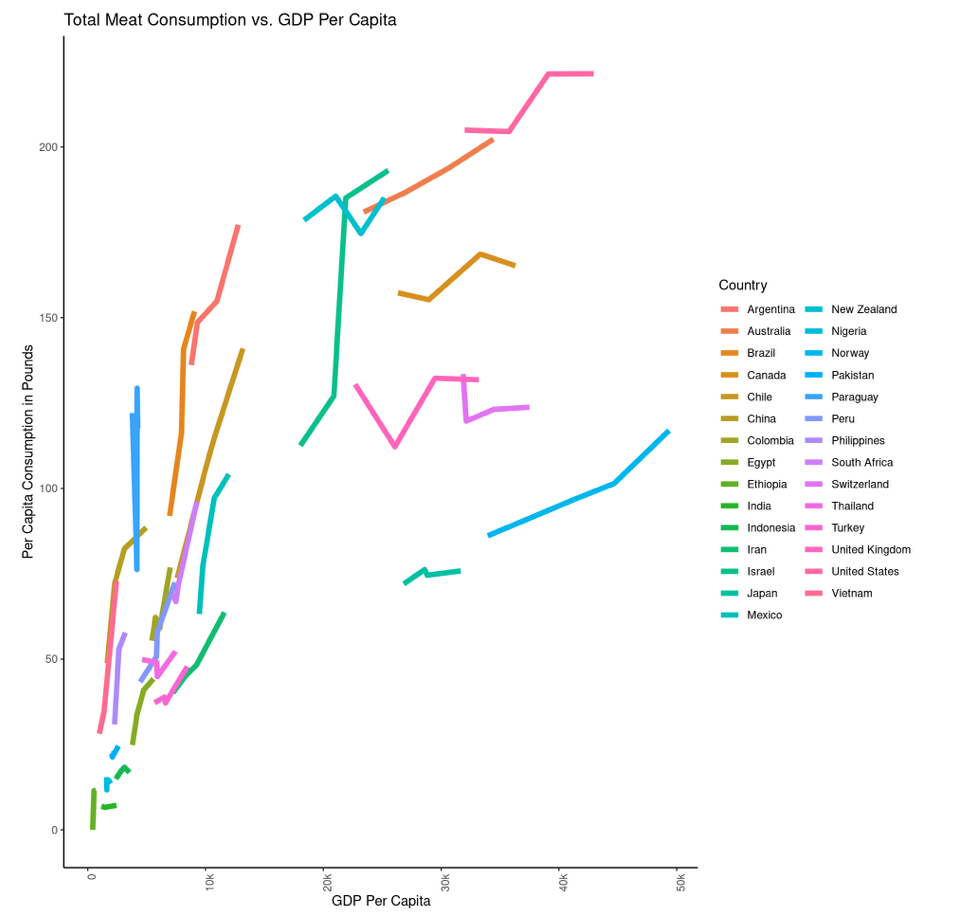
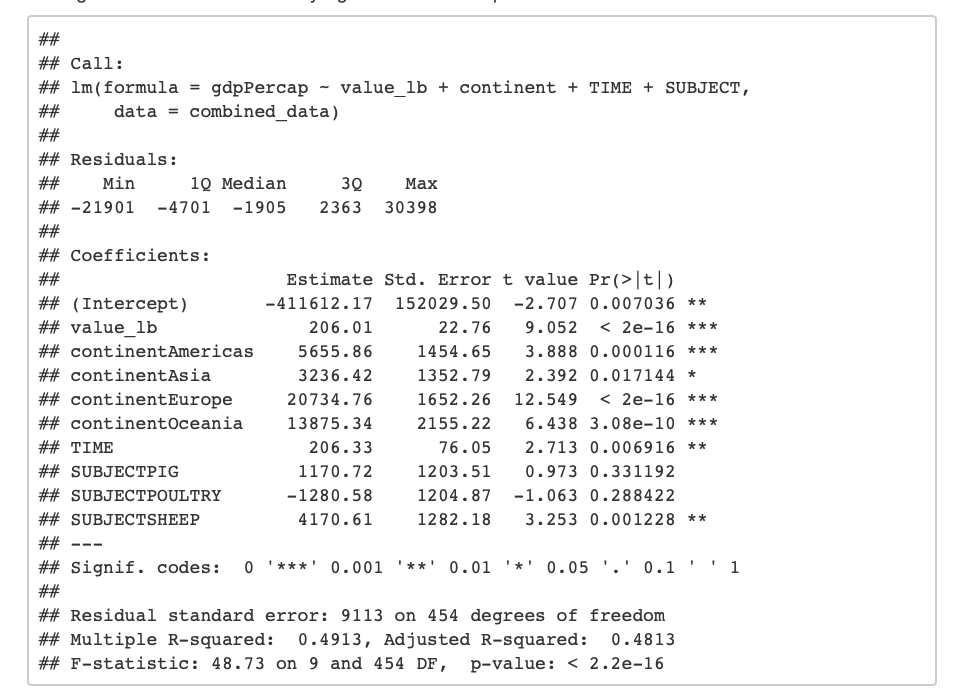
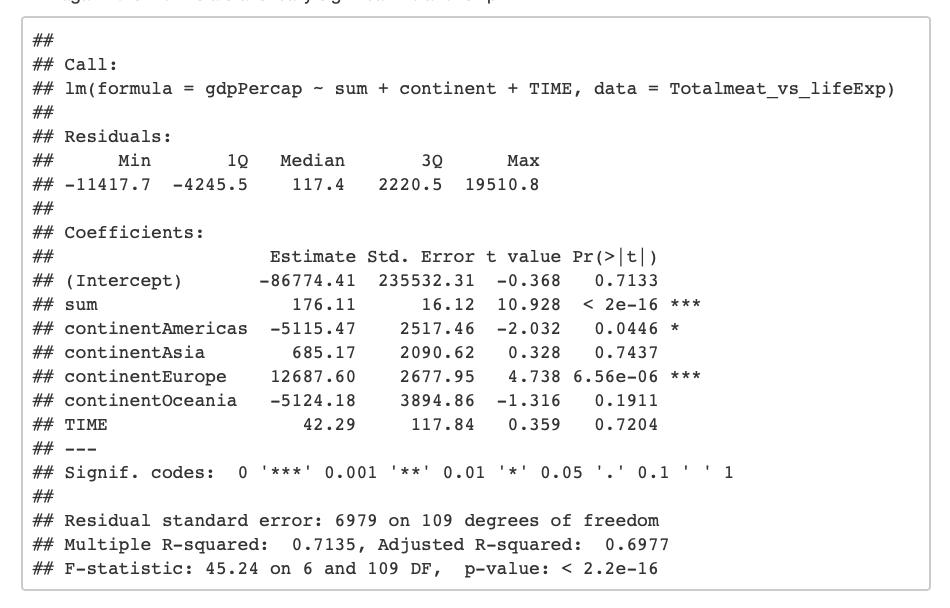
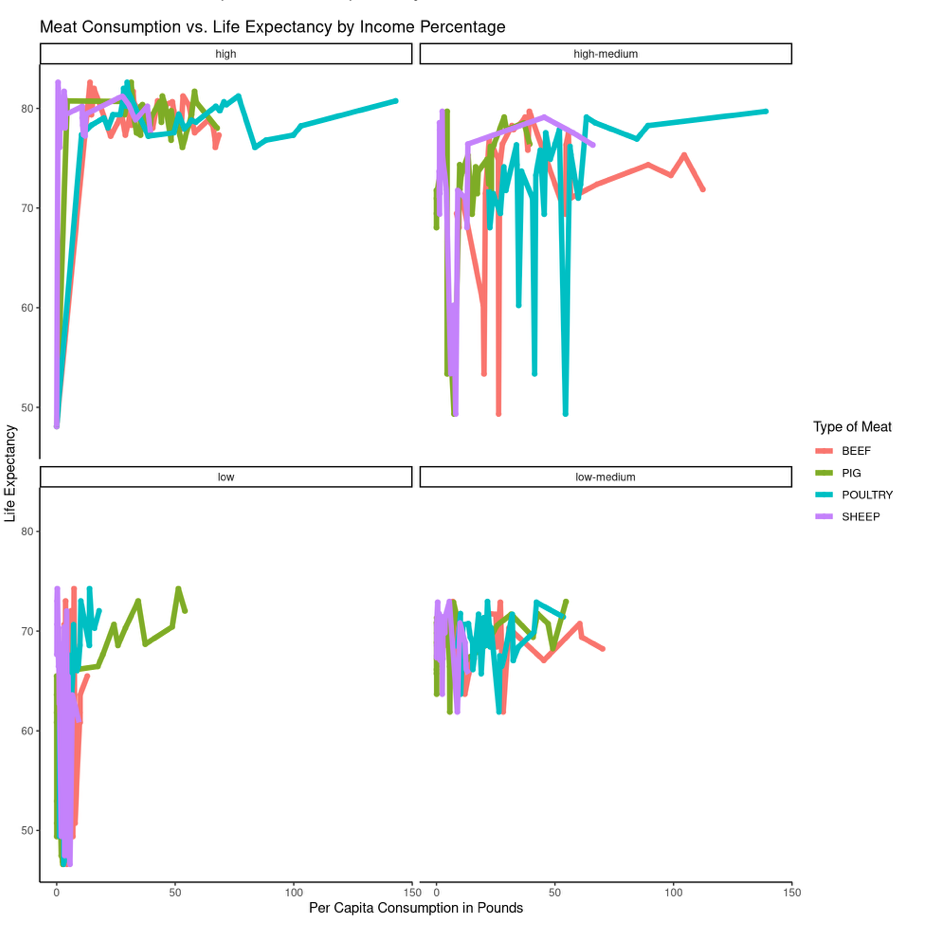
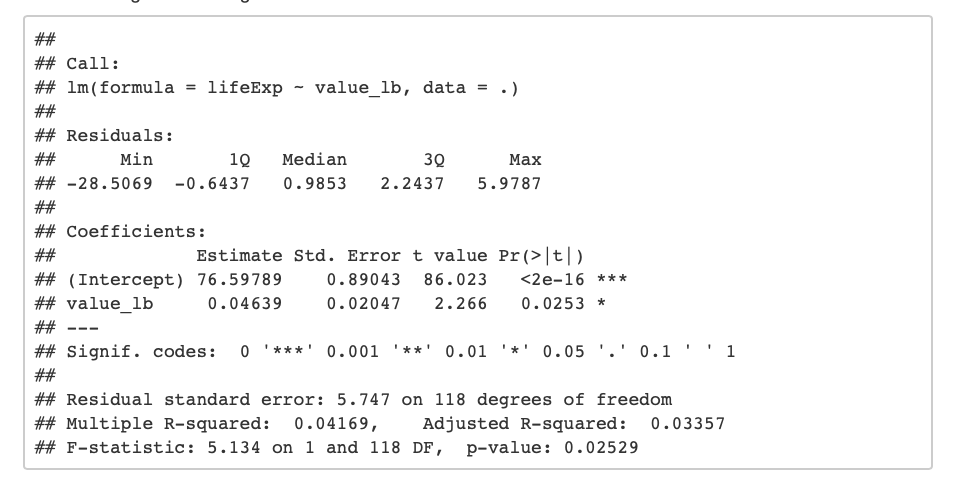
The graph indicated a higher likelihood of positive relationships between meat consumption and life expectancy rather than negative relationships.

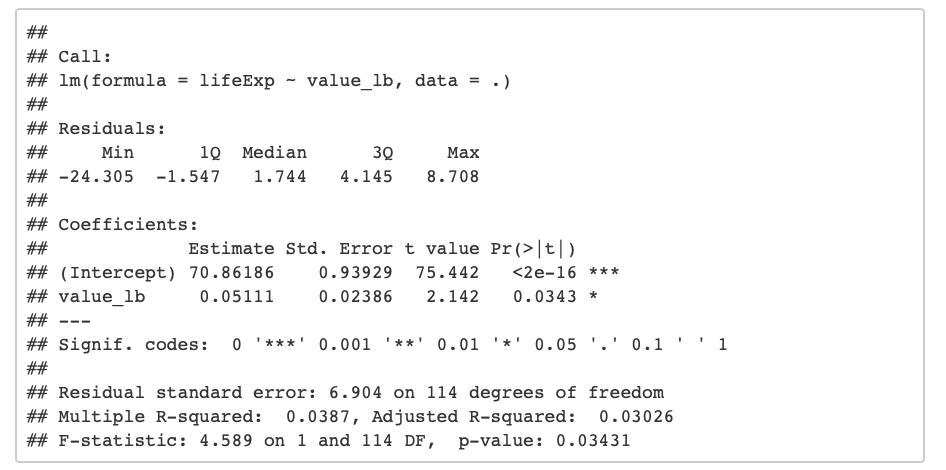
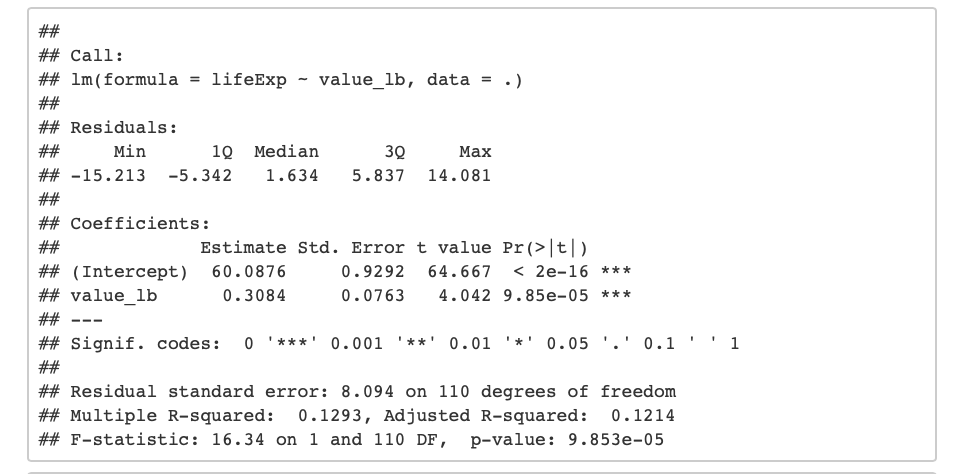
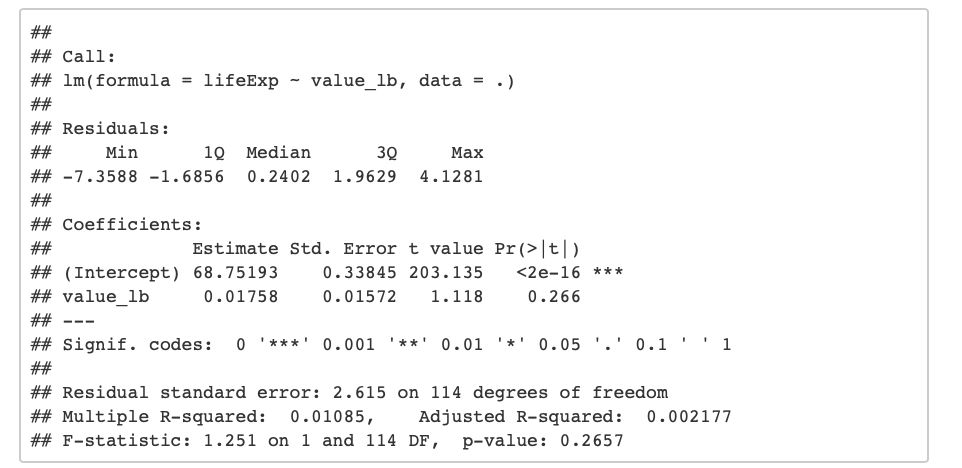
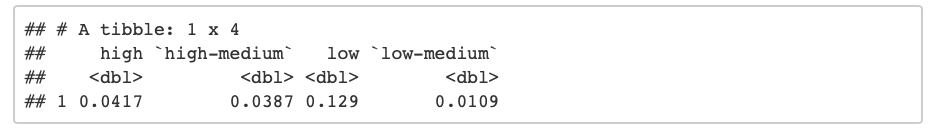
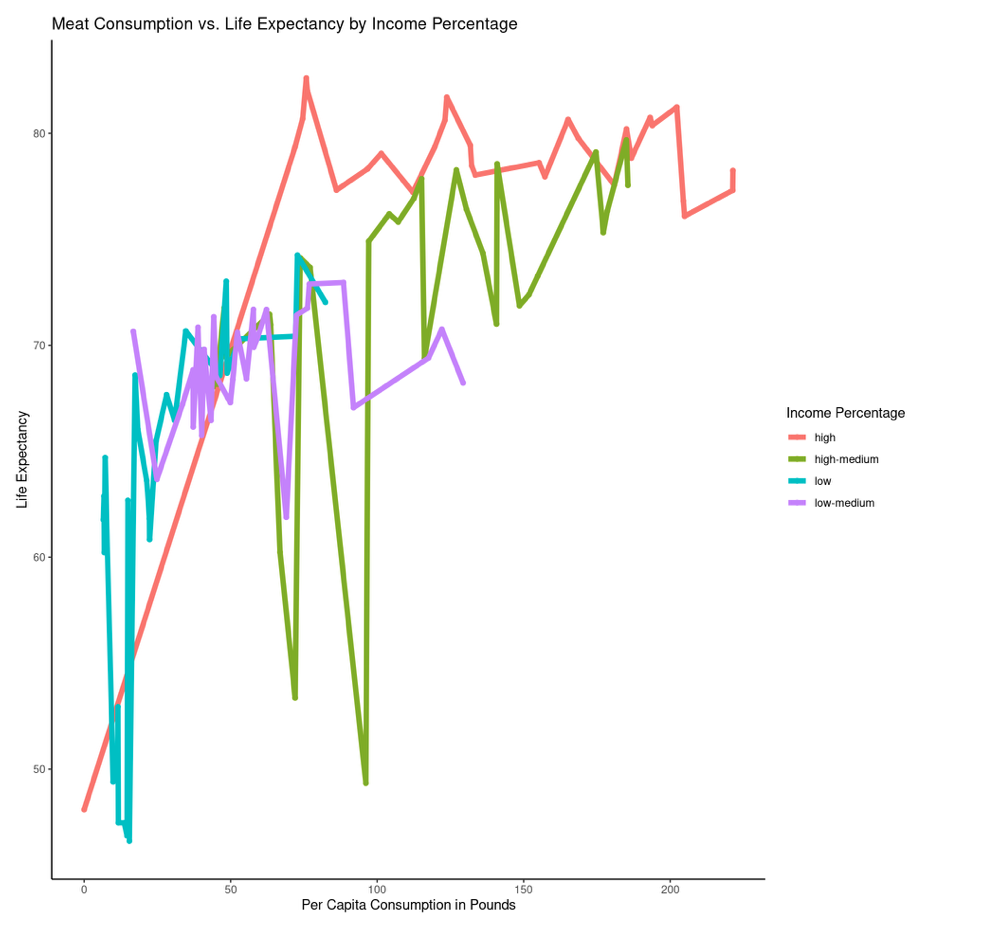
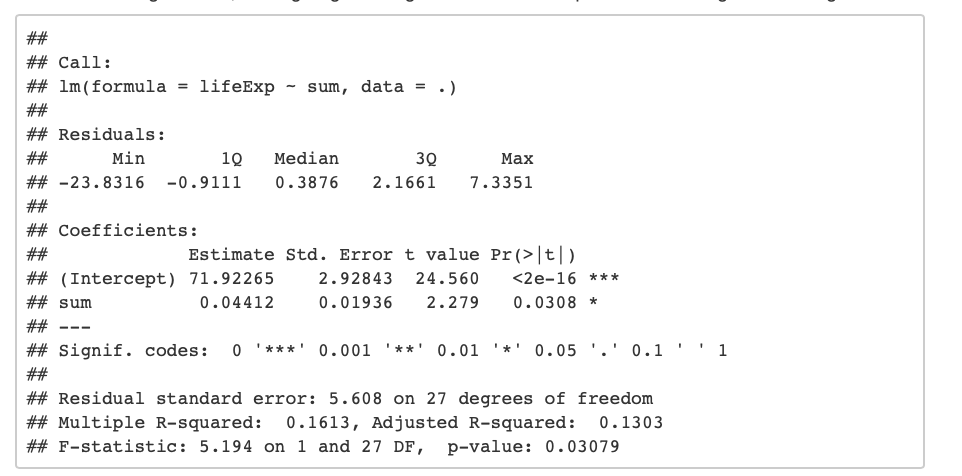
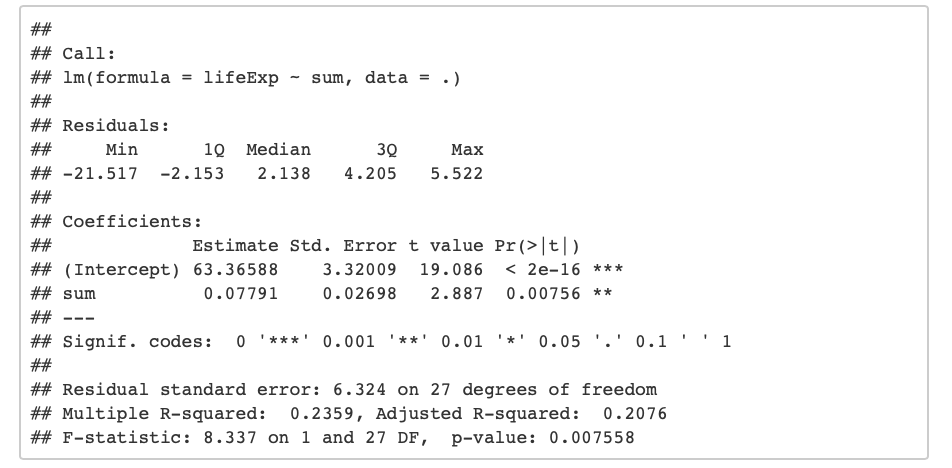
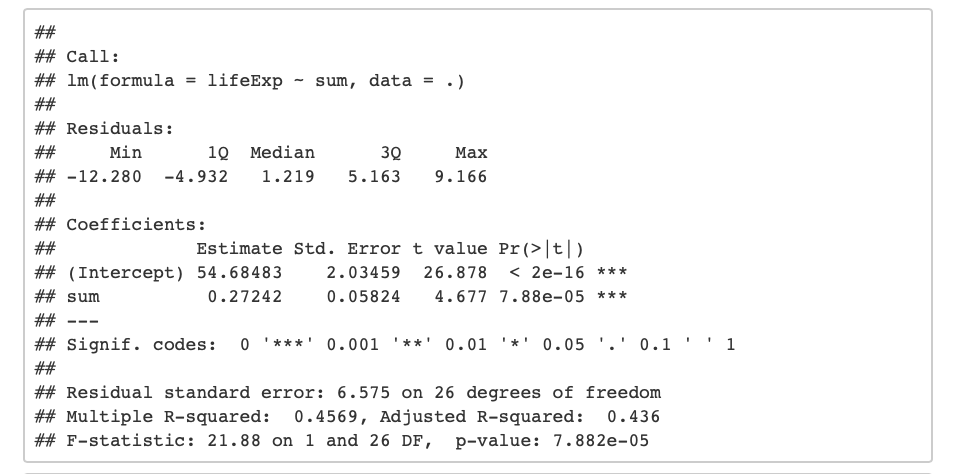
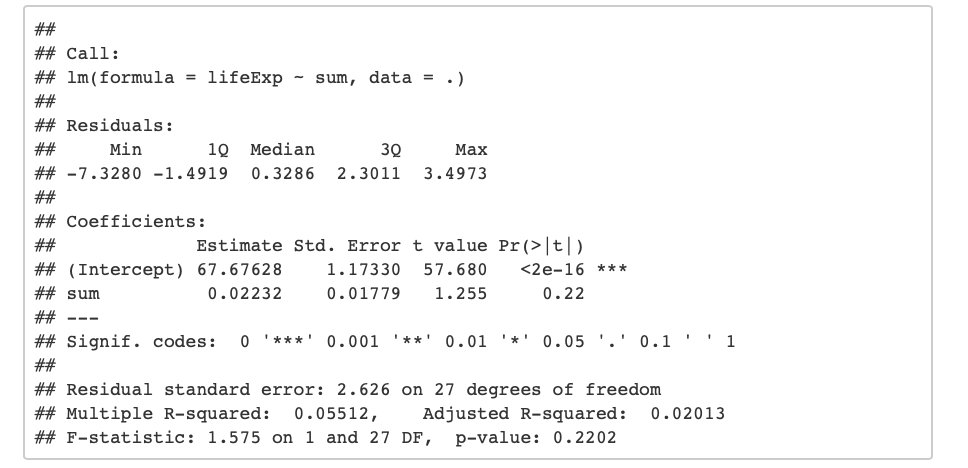
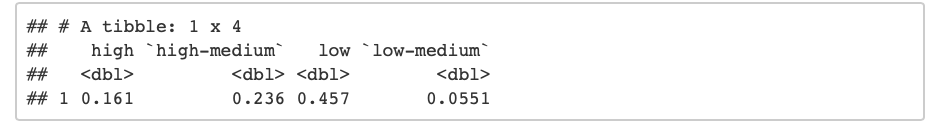
The statistical output included an R-squared value of 97% and a p-value indicating statistical significance.

The fifth graph also indicated a positive relationship may be more likely than a negative one.The statistical output again included an R-squared value of 97% and a p-value indicating statistical significance.The sixth graph was again fairly difficult to observe, so a second one was created with adjusted scales. This graph seems to show some positive and some negative relationships.The statistical output included an R-squared value of 65% and a p-value indicating statistical significance.Dividing the information by continent with meat type summed again reinforces the largely positive relationships.The statistical output included an R-squared value of 75% and a p-value indicating statistical significance.

* **Plots 8-11**

Both plots 8 and 9 showed an obvious positive relationship as expected.

The statistical output included an R-squared value of 48% and a p-value indicating statistical significance.The statistical output included an R-squared value of 70% and a p-value indicating statistical significance.Plot 10 showed some variation in the relationships between meat consumption and life expectancy by income. The statistical output included an R-squared value of .03% and a p-value indicating statistical significance.

The statistical output included an R-squared value of .03% and a p-value indicating statistical significance.The statistical output included an R-squared value of 12% and a p-value indicating statistical significance.The statistical output included an R-squared value of >.01% and a p-value indicating no statistical significance.The r-squared values were similar to the adjusted r-squared values.The final plot showed a clear positive relationship when income was in the highest bracket, up until it showed some variation around the point that meat consumption surpassed 100 lbs per capita. The other income brackets all showed fairly consistent variation.The statistical output included an R-squared value of 13% and a p-value indicating statistical significance.The statistical output included an R-squared value of 20% and a p-value indicating statistical significance.The statistical output included an R-squared value of 44% and a p-value indicating statistical significance.The statistical output included an R-squared value of 2% and a p-value indicating no statistical significance.The r-squared values were similar to the adjusted r-squared values.

**Discussion**

This study did not reveal any major trends; it appears that meat consumption has only a slight positive effect on life expectancy, however this relationship also has a number of other factors contributing to it, especially modern medicine. This study has many limitations and may produce more significant findings if compared with a more specific dataset than life expectancy, such as heart disease.

**References**

*Agricultural output - Meat consumption - OECD Data*. theOECD. (n.d.). https://data.oecd.org/agroutput/meat-consumption.htm.

Key, T. J., Fraser, G. E., Thorogood, M., Appleby, P. N., Beral, V., Reeves, G., … McPherson, K. (1998, March 1). *Mortality in vegetarians and non-vegetarians: a collaborative analysis of 8300 deaths among 76,000 men and women in five prospective studies: Public Health Nutrition*. Cambridge Core. https://www.cambridge.org/core/journals/public-health-nutrition/article/mortality-in-vegetarians-and-nonvegetarians-a-collaborative-analysis-of-8300-deaths-among-76000-men-and-women-in-five-prospective-studies/9FF29F354456308D49006B5DCBBE5BF0.

Larsson, S. C., & Orsini, N. (2013, October 22). *Red Meat and Processed Meat Consumption and All-Cause Mortality: A Meta-Analysis*. OUP Academic. https://academic.oup.com/aje/article/179/3/282/103471?login=true.

Ranabhat, C. L., Park, M.-B., & Kim, C.-B. (2020, February 12). *Influence of Alcohol and Red Meat Consumption on Life Expectancy: Results of 164 Countries from 1992 to 2013*. MDPI. https://www.mdpi.com/2072-6643/12/2/459/htm.

Rohrmann, S., Overvad, K., Bueno-de-Mesquita, H. B., Jakobsen, M. U., Egeberg, R., Tjønneland, A., … Linseisen, J. (2013, March 7). *Meat consumption and mortality - results from the European Prospective Investigation into Cancer and Nutrition*. BMC Medicine. https://link.springer.com/article/10.1186%2F1741-7015-11-63.

Singh, P. N., Sabaté, J., & Fraser, G. E. (2003, September 1). *Does low meat consumption increase life expectancy in humans?* OUP Academic. https://academic.oup.com/ajcn/article/78/3/526S/4689992?login=true.