Project 2 Final Rationale

Tiffany Wang, tw292
Adomas Hassan, ah667
Aditya Rajan, avr32

The Milk Phenomenon: Comparing Canadian and American Prices

****This project was designed to be viewed on a 13 inch Macbook screen size, please adjust the browser dimensions for best viewing!****

For this visualization, we decided to compare prices of food items across the United States and Canada over the years to see if there were any significant differences between the two countries. We split our visualization into three parts. The first part is an overview of average prices over the years, from 1996 to 2018. From that graph, we used the most recent February 2018 data to calculate how much prices in Canada differed from those of the U.S. for each item, and plotted them from greatest positive change to greatest negative change. Finally, we took a deep dive into milk prices, examining profits, costs, and production of dairy in both countries to see if we could explain the huge difference.

Please see the appendix for all data sources!

The following parts go through the story of the visualization, descriptions of the data used, and how we mapped the data to visual elements.

Story of Visualization

With our visualization, we are trying to explain why exactly milk prices are so much higher in Canada than the United States. Because it requires a bit of economics, we took the time to explain our reasoning on the visualization itself. We are copying some of the text of it here because we feel that it is a clear explanation of what we were trying to achieve.

For our project, we wanted to compare prices in America and Canada because we were curious about what consumption is like in Canada. To start off, we decided to explore prices over time of common food items to uncover any unexpected trends. This is the meat of our first visualization: comparing Canadian food prices to American food prices over time. The point of this part is to introduce the reader to the context of the visualization and highlight the huge price discrepancy in milk to set the stage for the rest of the visualization.

The second visualization shows the price percentage differences between Canada and America. The data is the most recent 2018 prices. This visualization is fairly self explanatory -- as you can see, the majority of goods are much more expensive in Canada than they are in the United

States. We chose to visualize the data this way after receiving some helpful TA feedback about how appending the images on the x axis would both make our graph more visually interesting and require less textual reading from the user, as image recognition is psychologically faster than text recognition.

The third part of our project is where we hone in on dairy and try to understand why milk prices are so different in the two countries. To try to explain the price differences, we gathered data from a variety of sources on dairy production, revenue, and cost in the United States and Canada. All data is in U.S. dollars. In all the visualizations, we represented United States data as blue and Canadian data as red to keep the visual identity of the graphs consistent. **The main assumption (that was later proved wrong) in this part was that both markets were free markets, meaning that consumers set the prices and firms have to take them as given.**

The first set of graphs focus on profit per gallon. We chose this metric because it is most commonly used for measuring financial "success", so we wanted to see how profitable farms are. Profit is a measure of how effective firms are at managing costs and production to maximize revenues. To keep comparisons consistent, we calculated profits per gallon for both countries. The user can toggle to see the relationship between profits and time or between profits and prices. This toggling functionality was kept consistent through all the sets of graphs because we wanted the user to be able to toggle one and then not toggle the other to try to identify non-obvious relationships on his/her own.

From the graphs, the user can see that profits decrease over time in the United States, but seem not to affect price fluctuations. This is consistent with our assumption that prices are independent of producer wealth. In Canada, profits increase over time - as profits increase, prices increase too. This is a bit strange because profits should not affect prices, yet the data is monotonically increasing. Clearly, this comparison was not the full story. Are profits changing because of costs, or production, or something else? We decided to decompose profits into its three pieces - revenues, costs, and production - and look at each in turn.

The second set of graphs focus on revenues and costs per gallon. We chose these two metrics because they are 2/3rds of what profit is determined by (and we wanted to save production for last)! The graphs are overlaid because we wanted the user to be able to visually see the area difference between revenues and costs.

In the United States, costs increase over time, while revenue fluctuates. In 2009, costs even overtake revenues (a product of the recession, perhaps). Like before, there is no relationship between cost and price, which is supported by our assumption. As the price of milk in the U.S. is relatively steady, this tells us that the fluctuations in revenue must be caused by a third unknown variable. Costs in Canada, on the other hand, stay steadily below revenues (although both dip in 2009). Costs increase as prices increase, which is again strange because it implies that producer costs are affecting market prices. It is unclear whether the high revenues are due to high prices

or massive production. This leads us to the final part of our visualization, which looks at production in both countries.

This last set of graphs in part 3 focus on production per capita. We saved these graphs for last because during the course of our research, we found that production was really the factor affecting the price of milk. The production trend in the United States moves in the same direction as revenues, which is to be expected. However, the production trend in Canada moves in the opposite direction as revenues - yet revenues are increasing! This is the crux of our visualization, and the answer can be found in the production over prices graph. In the United States, production increases as price increases - this is in line with our assumption from above, as farms want to produce more to maximize revenues. In Canada, however, production decreases as price increases. Less production drives prices up, allowing farms to control market prices.

After doing some research, we found that this practice is officially known as Supply Management, and it involves the government setting production quotas on dairy products. This goes back to the very beginning of the project, where you can see that milk prices in Canada stayed steady until 2002 and then started increasing rapidly. The next part of this is purely hypothesis, but it is still interesting to theorize about why this phenomenon is observed.

Our guess is that the 2002 ruling in which the government was forced to end subsidies on milk exports caused dairy producers to worry, as total export revenue would now decrease. To compensate, the government imposed dairy production quotas to drive up prices and increase domestic revenue, which would outweigh the loss of export revenue. This shift can be seen in every graph over time in this part. Clearly, this strategy was super-effective - to the detriment of the Canadian consumer.

The bonus set of graphs were appended as a way to hammer in the point about Canada under-producing and still being able to charge a lot for milk. After seeing the results above, we were curious about how many gallons of milk each farm itself was producing. Given the trends, we assumed that production per farm would be increasing in the United States and decreasing in Canada. However, the data shows that even though the number of farms has decreased, production has actually increased in both countries. In Canada especially, the small number of farms means that each one has more market power. There are less producers, which means that there is less price competition between producers, and consumers have less products to choose from so demand is relatively unaffected by price. In essence, the dairy producers in Canada are forming a joint monopoly, shifting the price-setting power from the consumers to the hands of the producers.

Part 1: Average Prices in Canada and the United States

Datasets used: canada.csv, us.csv

A. Data Description

Us.csv

This dataset contains information on the prices of various popular food commodities in the United States from January 1st, 1996 to February 2nd, 2018. The prices vary based on their decimal values. These food commodities include Round Steak, Ground Beef, Pork Chops, Chicken, Bacon, Milk, Cheese, Eggs, Bread Macaroni, Flour, Apples, Bananas, Potatoes, and Sugar

Formatting: The data is pre-processed to standardize all prices to US dollars as well as adjusted for all exchange rates. T**he data is also all represented in nominal dollars and are not adjusted for inflation. The reason we did this was because all the data (profits, costs, etc.) are all in nominal values and we wanted to stay consistent.**

Years: 1996-2018
Price: Represented in Nominal US Dollars adjusted for exchange rates
Dates: Jan 1, 1996- Feb 1, 2018
Food Commodities: Round Steak (measured per lb), Ground Beef (measured per lb), Pork Chops (measured per lb), Chicken (measured per lb), Bacon (measured per lb), Milk (measured per gallon), Cheese (measured per lb), Eggs (measured per dozen), Bread (measured per lb), Macaroni (measured per lb), Flour (measured per lb), Apples (measured per lb), Bananas (measured per lb), Potatoes (measured per lb), and Sugar (measured per lb)

Canada.csv

This dataset contains information on the prices of various popular food commodities in Canada from January 1st, 1996 to February 2nd, 2018. The prices vary based on their decimal values. These food commodities include Round Steak, Ground Beef, Pork Chops, Chicken, Bacon, Milk, Cheese, Eggs, Bread Macaroni, Flour, Apples, Bananas, Potatoes, and Sugar (same as the United States)

Formatting: The data is pre-processed to standardize all prices to US dollars as well as adjusted for all exchange rates (to be consistent with US data). The data is also all represented in nominal dollars and are not adjusted for inflation. The reason we did this was because all the data (profits, costs, etc.) are all in nominal values and we wanted to stay consistent.

Years: 1996-2018
Price: Represented in Nominal US Dollars adjusted for exchange rates
Dates: Jan 1, 1996- Feb 1, 2018
Food Commodities: Round Steak (measured per lb), Ground Beef (measured per lb), Pork Chops (measured per lb), Chicken (measured per lb), Bacon (measured per lb), Milk (measured per gallon), Cheese (measured per lb), Eggs (measured per dozen), Bread (measured per lb), Macaroni (measured per lb), Flour (measured per lb), Apples (measured per lb), Bananas (measured per lb), Potatoes (measured per lb), and Sugar (measured per lb)

B. Data Mapping into Visual Elements

Description
The method of visualizing the elements is the same for both the United States as well as Canada. The data was pulled directly from each CSV file and plotted as two line graphs in an SVG format. The SVG graphs are set side-by-side in order to visually represent the differences in prices between the two countries

We added further functionality by adding a hover feature over each individual line in which shows what food commodity the line represents. Since our project is purely focusing on milk prices, we decided to change the coloring of the Milk line to stand out among the other food commodities. To add further analysis of prices over time, we also added a few lines on the graph that represent important dates in dairy history that have determined why specific prices have either risen or dropped (i.e. 2008 financial crisis).

Method
We mapped the data from us.csv and canada.csv to a callback function that was then visually represented in a svg graph. We edited the scales (xscale and yscale) in order to put the two graphs side-by-side. The x scale was at set at intervals per year and the y scale was set at intervals per dollar. Two zscale ordinals were created in order to have two separate color schemes for the two graphs. Axis were created and added using d3 text elements in order to show what values were being represented on each axis. Tick marks and graph lines were made at intervals that best represented the data trends and also fit screen dimensions.

The hover features were created using .onmouseover and .mouseout attributes to the line features. These were tied directly to the data to visually represent text on the screen when a mouse is over the line.

The important date lines were created as svg lines that were appended to the svg graphs utilizing d3. Similar mouse over functionality was added in order to visually represent text on the screen to describe what the event was and when it took place.

Part 2: Percentage Difference in Prices
Datasets used: percentChange.csv

A. Data Description

percentChange.csv [*canada.csv, us.csv*]

This dataset uses the previous two data sets to calculate the percent difference of the prices of the same food commodities just for the year of 2018. We chose to focus only on 2018 because these prices are the most current, the most meaningful to us, and show that this is a ongoing problem, not an outdated/resolved one..

The percent difference was calculated in relation to United States prices, so we used the formula:

(canada item price - U.S. item price)/U.S item price

The percent change values vary by how much more expensive the product is in Canada - a really large value means the item is much more expensive in Canada and a negative value represents an item that is more expensive in the United States. These food commodities include Round Steak, Ground Beef, Pork Chops, Chicken, Bacon, Milk, Cheese, Eggs, Bread Macaroni, Flour, Apples, Bananas, Potatoes, and Sugar and all data is only for the year of 2018.

Formatting: Because the data is coming from the datasets in *Part 1,* all data is pre-processed to standardize all prices to US dollars, adjusted for all exchange rates, and is represented in nominal dollars without taking inflation into account. The reason for all of these adjustments are the same as above. The difference however, is that U.S Prices and Canada Prices were taken from their respective datasets and two columns were added during pre-processing: Net Difference, and Percent Difference. Net difference is the difference in price given in dollars, and percent difference is the difference in price given in a percentage of the price of the item in the United States.

Years: 2018 (only)
U.S Price: The price of the item in the United States for 2018
Canada Price: The price of the item in Canada for 2018
Net Difference: Difference in price (Canada - U.S.) in dollars
Percent Difference: Difference in price with respect to the U.S given in percent (Net Difference / U.S)
Food Commodities: Round Steak (measured per lb), Ground Beef (measured per lb), Pork Chops (measured per lb), Chicken (measured per lb), Bacon (measured per lb), Milk (measured per gallon), Cheese (measured per lb), Eggs (measured per dozen), Bread (measured per lb), Macaroni (measured per lb), Flour (measured per lb), Apples (measured per lb), Bananas (measured per lb), Potatoes (measured per lb), and Sugar (measured per lb)

B. Data Mapping into Visual Elements

Description

Our goal was to depict the percent changes in a way that would exaggerate how much more expensive (or cheaper) a product is in Canada vs the United States. Thus, we decided to use a bar graph in which bars could go either above or below the x-axis in order to represent the

percent change. We also decided to sort the prices from highest % change to lowest % change in order to exaggerate the difference in prices. The data was pulled directly from a CSV file and plotted as a bar graph in an SVG format.

We added further functionality to the graph by allowing a user to hover over a certain bar/item in order to display a tooltip for that item that had information such as: Item name, US Price, Canada Price, Net Difference, and Percent Difference while also reducing the opacity of all of the other bars in the graph. This is extremely useful because we decided to split our bar graph in two (translate positive values up and negative values down) in order to allow space in the middle for images. Because of this, it became difficult to accurately determine the percent difference value of each bar. Assigning colors to positive/negative graphs and adding the tooltip solves this issue by allowing a user to clearly determine that Milk is extraordinarily more expensive in Canada with a quick glance, and then allowing the user to hover over every bar to see by how much.

Method

First we pre-processed the data as mentioned above and then mapped the data from percentChange.csv to a callback function that was then visually represented in a svg graph. Because we knew that we would have both positive and negative values and images in between, we created one scale for x variables and 3 scales for y variables. The first y-scale would be for the y-axis above the 0% mark, the second would be for the y-axis below the 0% mark, and the last one was for all of the bars in general. We only needed one scale for x-values, which used d3.scaleBand to create even spaces for the bars with some padding for the range, and then mapped item names from our dataset to each “band”. The y-scale used the create the bar graphs had a domain of the all of the percent changes (min, max) by using d3.extent, and a range of the full height of the svg (with adjustments for margins). When actually using the scale to plot the bars, we either added or deleted 30 pixels based on whether the percent change was positive or negative in order to include space in the middle and then adjusted for our calculations later.

The hover features were created adding .onmouseover and .mouseout attributes to the bars when we first created them. These were tied directly to the data to visually represent the tooltip and reduce the opacity of all other (not-hovered) bars.

The most difficult part was definitely playing with three different y-scales (when we wanted a top and bottom axis) because we had to use scaled values from other scales in order to set the range of the two y-axis scales. There was a lot of tinkering and playing with/swapping values because of how y values are represented in svg.

The images were added by appending the svg with a link to an image for each food that we were representing. The height and width were adjusted so that the image fits directly in the center and the x values were assigned by multiplying a constant by an incrementing value so that the images are evenly spaced.

Part 3: Examining Milk

Datasets used: us-farms-income.csv, canada-farms-income.csv

A. Data Description

Us-farms-income.csv

This dataset contains information on the number of farms and cows, population of the United States, total production and profits, a profit/cost/revenue breakdown by gallons, and a production breakdown by capita and farms annually from 2000 - 2014.

Formatting: This data was heavily pre-processed from many data sources that are listed in the appendix. Because it is a combination of sources, unit conversions and data assumptions were made to reach the numbers presented. Conversions are discussed in the order that they are presented in the CSV. Some columns were not used, so the ones that were used are highlighted in blue.

Year: 2000 to 2014
Price: The price of milk in the corresponding year. This was calculated by averaging the monthly price data from us.csv.
Number of Farms: This is drawn from two PDFs containing statistics on dairy farms in the selected time range. It was difficult to find this data in a computer-readable form, so we manually copied it into a csv. There was no processing required.
Population: The number of people in the United States.
Total production: This is drawn from a UW Madison site, which had production in the selected time range in pounds. To convert to gallons, we used Excel to multiply each row by the conversion factor.
Revenues: This is drawn from USDA data containing costs and returns on dairy production. The data is directly taken from the “Production Value” row in that dataset, which contains the production value for every year per hundredweight (100 pounds). To transform the data, we converted it into production value for every gallon (by dividing by 100, then converting pounds into gallons) and then multiplied by total production.
Costs: These costs are calculated from the USDA data by adding operating costs + wages + farm overhead + rent. We performed these calculations based on the Canadian data, which had a footnote explaining how it calculated the operating costs. This is a potential point of error because it is unclear what exactly went into the Canadian calculation, so it was impossible to match the U.S. calculation perfectly. Again, these costs were per hundredweight, so we had to transform it like revenues.
Profits: This is calculated by subtracting costs from revenues.
Subsidies: This data is drawn from the Environmental Working Group, a lobbyist based in Washington D.C. These numbers are the subsidies that the U.S. government provides the milk industry. **We were originally going to factor this data into profits, but we decided not to because of potential data bias and lack of transparency about data source.** Because this data

is presented on an environmental lobbying website, it could easily be biased to be in favor of its agenda. This problem is further compounded by the fact that the data source is unclear -- they do not link to any site where they got the information from. As a result, we ultimately decided not to use this data when calculating profits.
Revenue per gallon: This is calculated by dividing revenue by total production.
Profit per gallon: This is calculated by dividing profits by total production.
Cost per gallon: Costs/total production
Production per capita: This is calculated by dividing total production by population.
Production per farm: Total production / number of farms.

Canada-farms-income.csv

This dataset is the mirror image of the United States data set, containing information on the number of farms and people across canada, total production and profits, and profit breakdowns from 2000 - 2014.

Formatting: Also listed in the order that the dataset is presented in.

Year: Same as above
Price: Same as above. The price was already converted into USD in canada.csv.
Number of Farms: This data was taken from the Canadian government's website, and no preprocessing was required.
Population: This was taken from the government website.
Total Production: Government website, but originally in hectoliters (HL) so required conversion to gallons.
Revenues: Government website, filtered for revenues from dairy farms only. This required conversion into USD using exchange.csv.
Costs: Government website, used the category "operating costs". Like mentioned before, there was a footnote explaining what went into calculating these costs, which was used to calculate U.S. costs. These also needed to be converted into dollars.
Profits: Revenues - costs
Revenue per gallon: Revenues/total production
Profit per gallon: Profit/total production
Cost per gallon: Profit/total production
Production per capita: Production/population
Production per farm: Production/number of farms

B. Data Mapping into Visual Elements

In general, Canadian data was mapped using shades of red and US data was mapped using shades of blue. The data was mapped onto a total of 10 graphs using a variety of shapes and interactivity. **All scales are d3.scaleLinear() because the data is linear and data was well separated**. Each graph is described below in order of appearance on the visualization. **Every part has a toggle button that allows the user to switch between time and price on the x axis -- this**

was kept consistent so that users could "customize" the visualization to see different relationships.

Step 1: Looking at Profit

1. **Profit over time:** This graph shows profits per gallon over time. The mapping is straightforward, as profit is on the y axis and year is on the x axis. We chose to make the circles the flags of the two countries to liven up the line graph. On hover, the profit per gallon is displayed at the specific point. We also added transitions to smooth out the hover functionality so it would not be so jarring for the user to hover over a point. We highlighted the point that was being hovered over visually by making it bigger and more opaque, making the corresponding line opaque, and making the other country's elements less opaque.
2. **Profit over price:** This graph shows profit per gallon vs. price. We chose to display it as a scatter plot because we wanted to display the relationship between the two variables. U.S. data is displayed in blue, and Canadian data is displayed in red. On hover, the price is displayed at that point. **We know that it is a bit unintuitive to display the x axis on hover, but we made that decision because the first graph already displays the profit per gallon so doing it here again would be redundant information. For all graphs over price, the hover functionality is displaying the price per gallon.**

Step 2: Counting Costs

1. **Revenues and costs over time:** This graph shows revenues per gallon and costs per gallon per year. The mapping is similar to above, but the lines have an area fill on them to best illustrate the difference between revenues and costs. On hover, the user can see the price tracking along the line to see what the revenues and costs were continuously. One issue with this functionality is that we made the selection area larger than the line area, so the mouseover captures values that are a bit below and above the actual value of the line. However, we felt that sacrificing this bit of accuracy was worth making the line much easier to hover. We also added opacity changes on hover to make the current selection more visible. We decided to plot both the U.S. and Canadian data on the same scale to make the two graphs easy to compare side by side for the user.
2. **Costs over price:** This graph is very similar to the profits over price graph, which is intentional. **We chose to make all the variables over price graphs scatter plots so that they were visually consistent throughout the third part of the visualization.** Again, the price is displayed at that point on hover. This graph is meant to illustrate the relationship between costs and prices.

Step 3: Examining Production

1. **Production over time:** This graph shows production per capita over time as a stacked bar graph. We chose to slice it per capita because of the population differences between the United States (around 300 million) and Canada (around 35 million). The Canadian bars overlay the U.S. bars so that they are easy to compare visually. On hover, you can see the exact number of gallons produced per capita. The color of the bar is determined using a

d3.scaleLinear() fill - the gradient is representative of the price of milk in that year. We added a legend at the top to explain this functionality, as we received some user feedback that it was unclear what the colors meant. **This graph is unique in that price, production, and time are all represented together:** we thought this was cool because you can see the relationship between price and production right on this graph instead of having to toggle over.

2. **Production over price:** Again, this is a scatterplot showing the explicit relationship between production and price. Although we already included it on our last graph, we chose to create another one for the sake of consistency and also to hammer in the point that production is inversely related to price in Canada.

Bonus: Less Overall Production, More Production per Farm

1. **Number of farms:** This graph shows the number of farms in the United States and Canada in 2000 and then 2014. The images were appended using svg.append(image), and the last image was clipped to the specific number by calculating the width percentage and then multiplying that by the width of the image to get the clip area. On hover, you can see the specific number of farms -- this is because graphs with pure images tend to be less informative, so we wanted to tell the viewer exactly what was going on. We added a legend at the bottom showing that each farm image represents 10,000 farms.
2. **Production per farm:** This graph shows the production per farm in the United States and Canada over time. This graph is the only non-interactive one in the visualization, which we did purposely because we wanted the message of the graph to speak for itself. The y axis was omitted because we did not want users to focus on the in-between data points, just the end points (which are labeled). The purpose of the graph is to show an increasing trend in both countries, which we think it accomplishes well.

Appendix: Data Sources

Canada.csv

This dataset contains data on average food prices of selected items across Canada from the period January 1996 to February 2018. Prices are adjusted by weight to match U.S. units for ease of comprehension, and converted from CAD to USD using the exchange rates per month from exchange.csv.

- http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=3260012&tabMode=dataTable&p1=-1&p2=9&srchLan=-1#F1

Canada-farms-income.csv

Heavily pre-processed dataset also containing information on the number of farms and milk cows across canada, total production and profits, and profit breakdowns from 2000 - 2014. In Canada, a “Supply Management” system is in use, which means that there are no government subsidies -- as a result, earned profits are equal to total profits.

- Amount of production (HL, required conversion to gallons): http://dairyinfo.gc.ca/index_e.php?s1=dff-fcil&s2=msp-lpl&s3=hmp-phl&page=histprod
- Number of cows: http://aimis-simia-cdic-ccil.agr.gc.ca/rp/index-eng.cfm?action=pR&r=219&pdctc=
- Farm income data (profits = net operating income): http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=0020035&tabMode=dataTable&p1=1&p2=-1&srchLan=-1&pattern=milk
- Number of farms: http://aimis-simia-cdic-ccil.agr.gc.ca/rp/index-eng.cfm?action=pR&r=220&pdctc=
- Canadian population:
- http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=510001

Cpi.csv

Contains CPI for food in the United States and Canada from the period January 1996 to February 2018. CPI is normalized to start at 100 for January 1996 to make applying calculations to Canada.csv and US.csv simpler.

- US CPI: https://fred.stlouisfed.org/series/CPIUFDNS#0
- Canadian CPI: http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=3260020&tabMode=dataTable&p1=-1&p2=9&srchLan=-1
    - Required conversion from percent change per month to a a classic CPI index normalized for January 1996 = 100. This was done to match the US CPI index, which is also normalized for that range. This calculation was performed in excel by treating the baseline value of January 1996 as 100 and calculating the percent change on top of the previous month's value with v(t) = v(t-1) + v(t-1)*(percent change)/100

Exchange.csv

Contains the exchange rates from CAD to USD and vice versa. This was used for pre-processing the data to transform Canadian dollars into USD so that they could be represented on the same graph.

- https://fred.stlouisfed.org/series/EXCAUS
    - The dataset only contained the conversion rate from USD to CAD, so the conversion rate from CAD to USD was calculated as the inverse of that.

percentChange.csv

Contains the February 2018 prices and percent changes for items in United States and Canada. This data is a subset of the us.csv and canada.csv data, and is used for visualization 2.

Units.csv

Contains the units for each food item chosen -- small table for reference purposes. These are taken from the descriptions in the US prices dataset.

- https://download.bls.gov/pub/time.series/ap/ap.item

Us.csv
Contains data on average food prices in the United States from the period January 1996 to February 2018.

- US prices: https://data.bls.gov/cgi-bin/dsrv?ap
    - Multiscreen data selector used to select many food items that were of interest. Area chosen was “U.S. city average” because the data contained was the least sparse out of all the other options and it provided the best coverage of the continental United States. One issue was that many food items Canada covered (butter, pork chops, coffee, tea) were not covered in this data, or had missing values, so we were forced to trim some items that we wanted to look at because of this.

US-cpi.csv, canada-cpi.csv
Divided the average prices by the CPI values/100 in each month to get the real prices. Formula based on http://www.foodsecurityportal.org/adjusting-prices-inflation

Us-farms-income.csv
Heavily pre-processed dataset containing information on the number of farms and cows, total production and profits, a profit/cost/revenue breakdown by gallons, and a production breakdown by capita and farms annually from 2000 - 2014. Information is gathered from many different sources, and unit conversions / assumptions about data accuracy were made to reach the numbers presented. Earned profits are the profits that are earned from market sales.

- Amount of production (pounds, required conversion to gallons): http://future.aae.wisc.edu/data/annual_values/by_area/99?period=complete&tab=production
- Farm income data (earned profits = value of production - operating costs - wages - farm overhead - rent (based on canadian data collection method)): https://www.ers.usda.gov/data-products/commodity-costs-and-returns/commodity-costs-and-returns/#Recent%20Costs%20and%20Returns:%20Milk
- Farm subsidies: https://farm.ewg.org/progdetail.php?fips=00000&progcode=dairy
- Number of dairy cows: https://www.statista.com/statistics/194934/number-of-milk-cows-in-the-us-since-1999/
- Number of dairy farms: https://www.hoards.com/sites/default/files/Fewer%20dairy%20farms%20left%20the%20business.pdf
- https://www.progressivepublish.com/downloads/2017/general/2016-pd-stats-lowres.pdf
- US Population:
- https://fred.stlouisfed.org/series/POPTOTUSA647NWDB

Additional Sources:
Key dates for visualization 1:
https://milk.procon.org/view.timeline.php?timelineID=000018#2000-present

Research on supply management:
https://www.dairyfarmers.ca/what-we-do/supply-management

Image Sources

Milk image in title:
https://www.stockio.com/free-icon/milk

Icons in “Story 2”:
Milk icon
https://www.flaticon.com/free-icon/milk_135635#term=milk&page=1&position=1

Chicken icon
https://www.flaticon.com/free-icon/flour_135663#term=flour&page=1&position=1

Flour icon
https://www.flaticon.com/free-icon/flour_135663#term=flour&page=1&position=1

Eggs icon
https://www.flaticon.com/free-icon/eggs_816034#term=eggs&page=1&position=32

Ground beef icon
https://www.flaticon.com/free-icon/ham_816035#term=beef&page=1&position=21

Pork chops icon
https://www.flaticon.com/free-icon/chop_816040#term=pork&page=1&position=11

Bread icon
https://www.flaticon.com/free-icon/bread_135597

Apple icon
https://www.flaticon.com/free-icon/apple_135728

Steak icon
https://www.flaticon.com/free-icon/steak_605090#term=steak%20flat&page=1&position=3

Bananas icon
https://www.flaticon.com/free-icon/banana_575393#term=bananas&page=1&position=25

Cheese icon
https://www.flaticon.com/free-icon/cheese_135652

Macaroni icon
https://www.flaticon.com/free-icon/spaguetti_135744

Bacon icon
https://www.flaticon.com/free-icon/bacon_135641

Potatoes icon
https://www.flaticon.com/free-icon/potatoes_135676

Sugar icon
https://www.flaticon.com/free-icon/sugar_714237#term=sugar&page=2&position=1

Icons in “Story 3”:

US flag circle icon
https://cdn1.iconfinder.com/data/icons/rounded-flat-country-flag-collection-1/2000/us-01.png

Canada flag circle icon
https://cdn1.iconfinder.com/data/icons/rounded-flat-country-flag-collection-1/2000/ca-01.png

Barn icon
https://www.flaticon.com/free-icon/barn_146281#term=barn&page=1&position=4