Final Project

PMAP 8551, Spring 2025

ARYA DEVKOTA

May 4, 2025

Table of contents

[Executive Summary 1](#_Toc197285869)

[Data Background 2](#_Toc197285870)

[Data Cleaning 2](#_Toc197285871)

[Individual Figures 5](#_Toc197285872)

[Figure 1: North America Case-Focus | Are Changes in Mortgage Rates Always Going to Guarantee that the Market will be More / Less Affordable? 5](#_Toc197285873)

[Figure 2: On a Global Scale, does Housing Supply Meaningfully Correlate with Home Prices? 7](#_Toc197285874)

[Figure 3: How does Renting compare to Purchasing a Home? 10](#_Toc197285875)

[Figure 4: How have Global Economic Factors Aligned with Changes in Housing Affordability Overtime (2015–2024)? 12](#_Toc197285876)

[Saving SVGs and PDFs for Vector Editor 15](#_Toc197285877)

[Final figure 17](#_Toc197285878)

# Executive Summary

The global housing market is a current trend of discussion among my generation — with most, including myself, having reserved concerns over the ability to purchase a home. When I think back to this topic, a comparison that has always stood out to me is the exaggeration of being able to purchase a family-home for merely a year’s worth of income. Ultimately, I wanted to explore this claim in more recent years, as well as see how the Affordability Index holds up against fluctuating global factors. I aimed to answer these questions through my visualizations:

1. In recent years, is there a consistent trend in how the affordability of homes responds to changes in mortgage rates? Are decreases in mortgage rates always going to garauntee that the market will become more affordable?
2. On a global scale, does housing supply meaningfully correlate with home prices? Are there trends that suggest different pricing strategies based on supply across different real-estate markets?
3. How does renting compare to purchasing a home? Is there anytime where renting is actually more expensive by the index?
4. How have global economic factors aligned with the percent change of housing affordability on a yearly basis? Is the housing affordability ratio reacting more strongly to changes in population growth and/or GDP growth?

My final infographic shows how economic factors and housing market trends can impact the lives of individuals and families around the world, as it represents the factors that could affect their chance at purchasing a house.

A few cool things that I’ve found through my visualizations are that countries with stable GDP and population growth, tend to have more consistent affordability ratios. Another analysis that can be made is that construction activity from 2015 to 2018 resulted in stable house prices for China; in general, increased construction activity seems to help stabilize house prices by increasing the supply of homes, as the USA and China display patterns where higher construction activity led to stable house prices. For Canada, the overlap between house prices and rent prices remained consistent, suggesting stable affordability for both renting and buying homes in their market.

# Data Background

This dataset, “Global Housing Market Analysis (2015-2024),” is originally from Kaggle - https://www.kaggle.com/ - It is structured as a CSV, covering various economic factors from 2015 to 2024. Ideally, it is meant for real-estate and market-trend analysis; it contains common economic values such as mortgage rates, interest rates, rent and house-price indices, GDP, and more for 20 unique countries.

# Data Cleaning

The raw file of this dataset had no particular problems with the actual values of each variable. However, column names were unconventionally formatted. Rather than spaces nor underscores, each word was separated by a period.

While this isn’t necessarily a problem for right now, it is important to stick with underscores because not only does it provide for better readability, but it is also one of the only compatible ‘special’ characters that works for a majority of coding/statistical languages and environments. This means that if I’m ever in a scenario where I need to reference my tibbles, it won’t be as much of a struggle to translate this data later on.

I went the easy, or more simple, route of using the “colnames” function. In this case, the order of your column names do matter; since I am referencing my entire tibble, each new column name corresponds to the chronological order of the columns in ‘raw\_ghm.’

I have also chose to include my custom color palettes here, since I will be referencing these palettes for multiple visualizations of mine.

I’ve also created a variable that will filter to the countries that I’d like to look at out of the twenty available; I’ve chosen six populous and well-known countries on the global stage in order to have data that is still visually readable but diverse. However, I am not filtering the tibble ‘clean\_ghm’ itself, as I need to refer back to it while creating new tibbles and values.

library(tidyverse)

library(ggplot2)  
library(patchwork)  
library(ggforce)  
raw\_ghm <- read.csv("data/global\_housing\_market\_extended.csv")  
  
clean\_ghm <- raw\_ghm   
colnames(clean\_ghm) <- c("Country", "Year", "House\_Price\_Index", "Rent\_Index",  
 "Affordability\_Ratio", "Mortgage\_Rate", "Inflation\_Rate",  
 "GDP\_Growth", "Population\_Growth", "Urbanization\_Rate",  
 "Construction\_Index")  
  
head(clean\_ghm)

Country Year House\_Price\_Index Rent\_Index Affordability\_Ratio Mortgage\_Rate  
1 USA 2015 117.4540 116.55000 9.587945 4.493292  
2 USA 2016 150.8073 51.44091 11.729189 5.662213  
3 USA 2017 123.1945 70.38604 8.506676 2.197469  
4 USA 2018 131.4234 91.46902 3.418054 4.537724  
5 USA 2019 110.4614 56.83705 9.158097 3.700762  
6 USA 2020 146.2522 71.81978 7.680612 4.233551  
 Inflation\_Rate GDP\_Growth Population\_Growth Urbanization\_Rate  
1 1.514121 -0.7520438 -0.7967074 85.98528  
2 1.880204 -0.5454003 -0.3580842 69.12727  
3 2.398940 0.9308947 0.5962449 83.55528  
4 1.608407 -1.4795873 2.3210994 88.96896  
5 1.293249 1.9614153 -0.8796402 87.27961  
6 1.701554 5.7566770 1.7129649 88.18497  
 Construction\_Index  
1 118.0892  
2 111.9805  
3 85.9739  
4 134.6718  
5 90.7024  
6 141.5862

# ----------------------------  
# Color Palettes  
# ----------------------------  
  
selected\_countries <- c("USA", "China", "Brazil", "Canada", "Russia", "Mexico")  
country\_palette <- c(  
 "USA" = "#D55E00",  
 "China" = "#e0a594",  
 "Brazil" = "#009E73",  
 "Canada" = "#fbba61",  
 "Russia" = "#61443a",  
 "Mexico" = "steelblue"  
)  
  
variablecolors <- c(  
 "House Price Index" = "#1F78B4",   
 "Construction Index" = "#adebb3",   
 "Rent Index" = "#6A3D9A",   
 "Mortgage Rate" = "darkorange",   
 "Affordability Ratio" = "darkred",  
 "Afford\_Change\_LogScale" = "darkred",   
 "Population Growth" = "#bd6e9c",   
 "GDP Growth" = "#e31a1c"   
)

# Individual Figures

## Figure 1: North America Case-Focus | Are Changes in Mortgage Rates Always Going to Guarantee that the Market will be More / Less Affordable?

First off, before starting on my visualizations, I knew that I wanted one plot to be based on North America alone, as I thought it would be fun to have a section in my infographic based on taking a closer look into the region. I chose to do Mortgage Rates, as I thought it would be more interesting to compare rates across us and our ‘neighbors.’ I also wanted to make this based on more recent years - if I’m going to look at a case-by-case focus, I might as well go all-in for this one!

I had many attempts at creating this graph until eventually reaching to this version. I knew that I wanted to display mortgage rates in comparison to the affordability ratio, however, I also knew that they were in different data formats; I couldn’t assume that the affordability ratio could be converted into a percentage without affecting the accuracy of my data.

So, I took a step back and assessed what my data actually means. Affordability ratio represents how many years of income it would take for the average household to buy a median-priced home; for example, an affordability ratio of 6 would mean that it would take the average household 6 years worth of median income to afford a home. That’s when it clicked for me; I didn’t necessarily have to scale my values, because I could create two y-axises that still contain the same ‘integers’ while maintaining the actual representation of my two variables.

This was already incredibly easy to do, as both column’s values had similar ranges. To ensure that I wouldn’t be changing the meaning of the values themselves, I added a percent sign and a multiplier symbol to ensure that my audience would understand that they are two different formats.

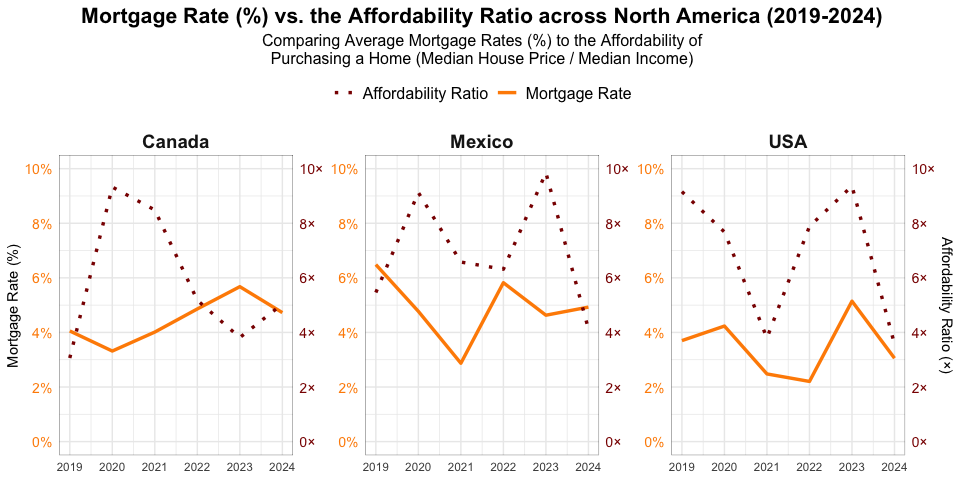
I then bolded titles and headers. As you can see, I also changed the colors of the axis values to correspond with the color of the lines in order to help viewers be able to easily refer to the correct scale; however, I purposefully did not change the color of the axis titles because I felt as if it would be distracting and unnecessary considering there is a legend to differentiate the two lines.

I also kept everything aligned to the center and faceted by country rather than displaying it all on one graph for easy readability.

Ultimately, the point of this graph is to compare fluctuations in mortgage rates to fluctuations in the affordability of homes, as this plays an important part in being able to pinpoint different factors that can affect consumers of the industry.

Being able to compare how many years worth of median income is needed to afford a median-priced home ties hand-in-hand with mortgage rates because, as mortgages are more common for the average middle-class citizen compared to down payments, any change in mortgage rates (considering you have the loan) ultimately means a change in how much you are paying off for that house.

NA\_ghm <- clean\_ghm |>  
 filter(Country %in% c("USA", "Canada", "Mexico"),  
 Year >= 2019, Year <= 2024)  
  
plot1 <-  
 ggplot(NA\_ghm, aes(x = Year)) +  
 geom\_line(aes(y = Mortgage\_Rate, color = "Mortgage Rate"), linewidth = 1.2) +   
 geom\_line(aes(y = Affordability\_Ratio, color = "Affordability Ratio"), linewidth = 1.2, linetype = "dotted") +   
 scale\_color\_manual(values = variablecolors) +   
 facet\_wrap(~ Country, scales = "free\_y") +   
 scale\_y\_continuous(  
 limits = c(0,10),  
 labels = function(x) paste0(x, "%"),  
# ----------------------------  
# I had to do my research on how to add a % symbol here without actually messing with the scale.  
# ----------------------------  
 name = "Mortgage Rate (%)",   
 breaks = seq(0, 10, 2),  
 sec.axis = sec\_axis(  
 ~ . , name = "Affordability Ratio (×)",  
 labels = function(x) paste0(x, "×"),  
 breaks = seq(0, 10, 2)  
 )  
 ) +  
 labs(title = "Mortgage Rate (%) vs. the Affordability Ratio across North America (2019-2024)",  
 subtitle = "Comparing Average Mortgage Rates (%) to the Affordability of\nPurchasing a Home (Median House Price / Median Income)",  
 x = NULL) +  
 theme\_minimal() +  
 theme(plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 panel.border = element\_rect(color = "black", fill = NA, linewidth = 0.2),  
 plot.subtitle = element\_text(hjust = 0.5, size = 12),  
 legend.position = "top",  
 legend.title = element\_blank(),   
 legend.text = element\_text(size = 12),   
 strip.text = element\_text(size = 14, face = "bold"),   
 axis.text.y = element\_text(color = "darkorange", size = 10.5),   
 axis.text.y.right = element\_text(color = "darkred", size = 10.5))  
plot1



## Figure 2: On a Global Scale, does Housing Supply Meaningfully Correlate with Home Prices?

Next, I wanted to focus on the global stage of housing prices. I knew I couldn’t compare an index with a percentage on one graph, so I decided to choose two indices to compare that I knew could provide valuable insights. I went with Home-Price Index to Construction Index. Construction Index is an indicator of how many properties and homes are being developed, meaning that it is an activity that directly affects the supply of homes.

I wanted to be able to see how these two economic factors interact with eachother, especially to be able to see whether or not the average price of homes has an overall positive resemblance to trends in construction, how ‘quickly’ the price index responds to construction activity, and if there is any outstanding differences between countries that might suggest different pricing strategies based on supply across different real-estate markets.

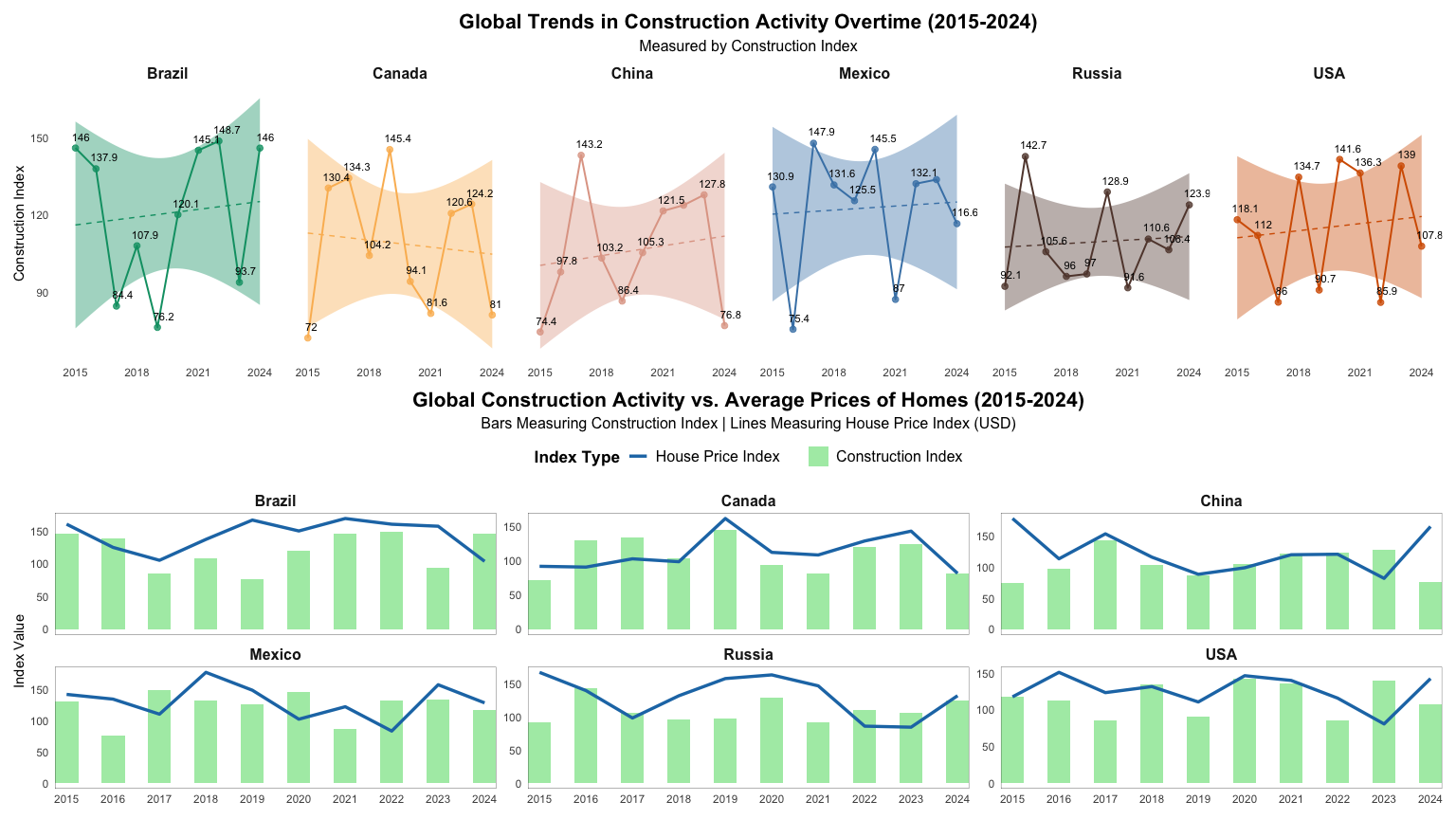
To do this, I created a patchwork plot where both graphs are faceted by countries for ease of comparison. Using the color palette I’ve created, I chose to do a smooth plot with lines and points so that way, my readers could easily compare any extremes in construction activity, as well as the trends by country based on the dashed line cutting through the center of the overall range(s). It is also easy to see outliers of the range by the points that are not covered by geom\_smooth. I also added text labels because of this; I originally wanted readers to be able to precisely compare unique values in the construction index. However, while creating my infographic in my vector editor later on, I had realized that I preferred it without any labels.

For my second graph for this plot, I tied it in with my house price index to achieve the goals I mentioned earlier. Originally, I had wanted to use the color palette I’d set for my countries, but I decided that it would make more sense in this case to use two contrasting colors representing both variables since I wanted the focus to be on comparing their relationship across global markets.

For both plots, I made sure to keep the index scale the same, as well as keep font sizes and boldness consistent. I also added subtitles that restate my values atop of my legend for clarity purposes. Finally, I also made sure to add a second title to transition to my second graph because while the two graphs are tied into my second goal for my story, they still have different purposes for their visualizations.

One thing to also note is that the color palette I’ve developed for my countries is colorblind-friendly, while still being distinct.

# ----------------------------  
# Setup for First Visual using Geom\_Smooth, Geom\_Point, and Geom\_Line  
# ----------------------------  
  
plot\_two <-   
 clean\_ghm |>  
 filter(Country %in% selected\_countries) |>  
 arrange(Year) |>  
 ggplot(aes(x = Year, y = Construction\_Index)) +  
 geom\_point(aes(color = Country), size = 2, alpha = 0.8) +  
 geom\_line(aes(color = Country), linetype = "solid", linewidth = .7) +  
 geom\_smooth(method = "lm", se = TRUE, aes(color = Country, fill = Country), linetype = "dashed", linewidth= .5) +  
 geom\_text(aes(label = paste0(round(Construction\_Index, 1))), size = 3, vjust = -1, hjust = .2, check\_overlap = TRUE) +  
 facet\_wrap(~ Country, nrow = 1) +  
 scale\_x\_continuous(breaks = seq(2015, 2024, by = 3), expand = c(0, 1)) +  
 labs(  
 title = "Global Trends in Construction Activity Overtime (2015-2024)",  
 subtitle = "Measured by Construction Index",  
 x = NULL,  
 y = "Construction Index"  
 ) +  
 theme\_minimal() +  
 theme(  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank(),  
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 plot.subtitle = element\_text(hjust = 0.5, size = 12),  
 strip.text = element\_text(size = 12, face = "bold"),  
 legend.position = "none"  
 ) +  
 scale\_color\_manual(values = country\_palette) +  
 scale\_fill\_manual(values = country\_palette)  
  
# ----------------------------  
# Setup for Second Visual using Geom\_Bar and Geom\_Line  
# ----------------------------  
  
  
plot2overlap <-   
 clean\_ghm |>   
 filter(Country %in% selected\_countries) |>   
 arrange(Year) |>   
 ggplot(aes(x = Year)) +   
 geom\_bar(aes(y = Construction\_Index, fill = "Construction Index"),   
 position = "dodge", stat = "identity", width = 0.5) +   
 geom\_line(aes(y = House\_Price\_Index, color = "House Price Index"),   
 linewidth = 1.2) +   
 facet\_wrap(~ Country, scales = "free\_y") +   
 labs(  
 title = "Global Construction Activity vs. Average Prices of Homes (2015-2024)",  
 subtitle = "Bars Measuring Construction Index | Lines Measuring House Price Index (USD)",  
 y = "Index Value",  
 x = "Year"  
 ) +  
 scale\_x\_continuous(breaks = seq(2015, 2024, by = 1), expand = c(0, 0)) +  
 scale\_fill\_manual(values = variablecolors, guide = "none") +   
 scale\_color\_manual(name = "Index Type", values = variablecolors) +   
 guides(  
 fill = guide\_legend(order = 2, title = NULL),  
 color = guide\_legend(order = 1, title = "Index Type")  
 ) +  
 theme\_minimal() +  
 theme(  
 panel.border = element\_rect(color = "black", fill = NA, linewidth = 0.2),  
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 axis.title.x = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank(),  
 plot.subtitle = element\_text(hjust = 0.5, size = 12),  
 strip.text = element\_text(size = 12, face = "bold"),  
 legend.position = "top",  
 legend.text = element\_text(size = 12),  
 legend.title = element\_text(size = 13, face = "bold")   
 )  
  
# ----------------------------  
# Combining my Plots with Patchwork  
# ----------------------------  
  
  
  
plot2 <- plot\_two / plot2overlap  
  
plot2



## Figure 3: How does Renting compare to Purchasing a Home?

Originally, this was meant to be my fourth plot. However, when I was done with creating all of my visualizations, I realized that my story would flow better with this plot being my third figure - which is why the following data mentions plot4data while being stored as plot 3.

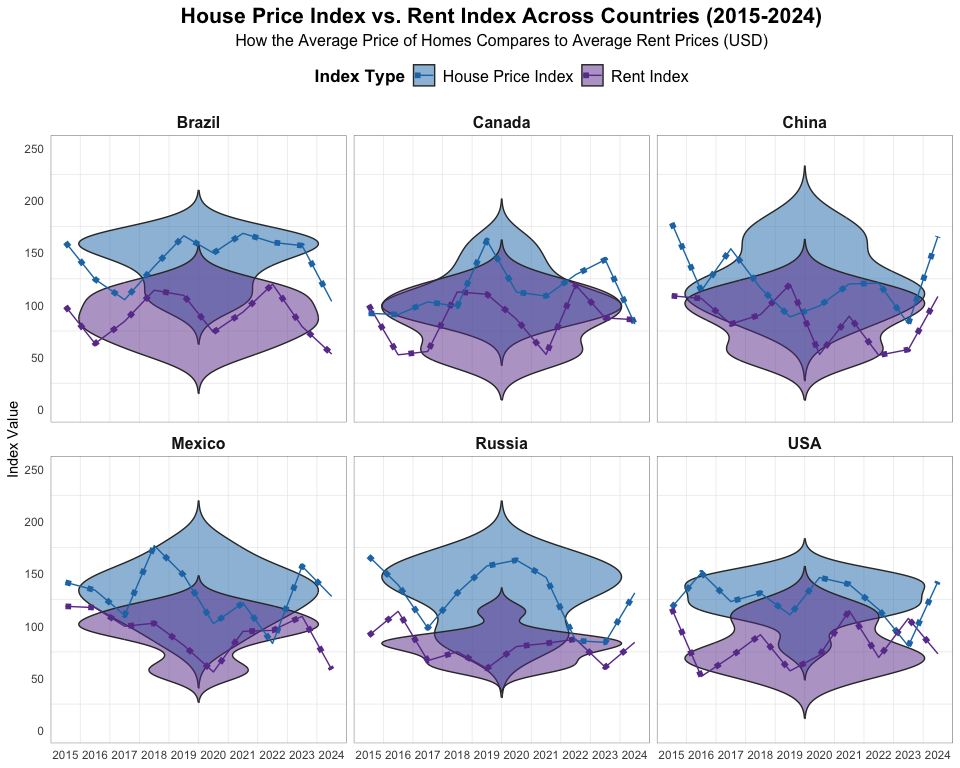
My purpose for this graph was to present trends that you might’ve not ever considered while skimming through this data. Sure, one can make an educational guess that it would be cheaper to rent an apartment compared to purchasing a home over the years. But, by how much do these indices differ? Is it safe to assume that that’s always the case, or are there years that prove this assumption to be wrong? What about country by country?

Because of all these questions for analysis that I’ve listed above, I decided to not only have lines comparing the two trends, but violin plots that are layered to visually show both indices range of values, where they actually overlap (or correlate in trend), and how stark of a difference they are across global markets. Overlapping areas mean that the distributions of house prices and rents are similar in that year, suggesting that the values of both indices for that year are close in central tendencies (mean, median), and also in shape and spread.

I first tried patchwork with this graph, where one plot contained lines, and another plot contained violins. However, I found it to be more interesting when I combined the geoms as layers. I decided to keep this graph minimal, as I will be adding captions to explain the value of certain trends within my infographic. I kept my y-axis scales consistent and with a large range to ensure that my violins could be properly displayed. I also kept my theme formatting as consistent as it could be with my previous graphs, making sure to bolden countries and titles. For any inconsistencies in font size and face, I’ll be adjusting them within Adobe Express.

I also chose to actually layer two lines, one being solid, and one being dotted. The reason for this is because while I wanted the lines to still be visible, I also didn’t want them to overpower my violins. However, if I made my solid line too thin on its own, then it almost loses its’ purpose of being there. So, in order to plot these lines without taking away from my violins and still emphasizing them, I created a mix of dotted and solid lines. The thin lines offer the illusion of transparency without having to mess with their opacity, and the dots allow for the lines to still make an impact on the graph. While it seems simple to the eye, a lot of thought process went into making sure that they were balanced.

plot4data <- clean\_ghm |>  
 filter(Country %in% selected\_countries)   
  
# ----------------------------  
# I chose to have trim set to FALSE, meaning that I didn't want for the violin plot to be cut off,  
# Because I wanted it to be able to show the full shape of the density.  
# ----------------------------  
  
  
plot3 <- ggplot(plot4data, aes(x = Year)) +  
 geom\_violin(aes(y = House\_Price\_Index, fill = "House Price Index"), alpha = 0.5, trim = FALSE, position = position\_dodge(width = 0.6)) +   
 geom\_violin(aes(y = Rent\_Index, fill = "Rent Index"), alpha = 0.5, trim = FALSE, position = position\_dodge(width = 0.6)) +  
 geom\_line(aes(y = House\_Price\_Index, color = "House Price Index"), linewidth = 2, linetype = "dotted") +   
 geom\_line(aes(y = Rent\_Index, color = "Rent Index"), linewidth = 2, linetype = "dotted") +   
 geom\_line(aes(y = Rent\_Index, color = "Rent Index"), linewidth = .5) +   
 geom\_line(aes(y = House\_Price\_Index, color = "House Price Index"), linewidth = .5) +  
 facet\_wrap(~ Country) +  
 scale\_color\_manual(name = "Index Type", values = variablecolors) +   
 scale\_fill\_manual(name = "Index Type", values = variablecolors) +   
 scale\_y\_continuous(  
 limits = c(0, 250),   
 name = "Index Value",   
 breaks = seq(0, 250, 50)   
 ) +  
 scale\_x\_continuous(breaks = seq(2015, 2024, by = 1), expand = c(0, .5)) +   
 labs(  
 title = "House Price Index vs. Rent Index Across Countries (2015-2024)",  
 subtitle = "How the Average Price of Homes Compares to Average Rent Prices (USD)",  
 x = NULL,  
 y = "Index Value"  
 ) +  
 theme\_minimal() +  
 theme(  
 panel.border = element\_rect(color = "black", fill = NA, linewidth = 0.2),  
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 axis.title.x = element\_blank(),  
 panel.grid.major = element\_blank(),  
 plot.subtitle = element\_text(hjust = 0.5, size = 12),  
 strip.text = element\_text(size = 12, face = "bold"),  
 legend.position = "top",  
 legend.text = element\_text(size = 12),  
 legend.title = element\_text(size = 13, face = "bold")   
 )  
plot3



## Figure 4: How have Global Economic Factors Aligned with Changes in Housing Affordability Overtime (2015–2024)?

To reiterate, this was meant to be my third plot. However, when I was done with creating all of my visualizations, I realized that my story would flow better with this plot being my fourth figure - which is why the variables below are listed as ‘plot3data’ despite being stored as plot4.

I thought it’d be fun to tie in my story with a bigger, more generalized picture of economic factors vs. the affordability ratio. However, I knew I couldn’t just map the affordability index onto the graph directly, since it’s scaled differently. So, instead, I decided to track the percent changes in housing affordability over time, then transform those changes into a more readable scale. I also used the lag function, as you’ll see, so that way I could tell the tibble that if there’s no previous year to calculate the percent change, to set the change to 0.

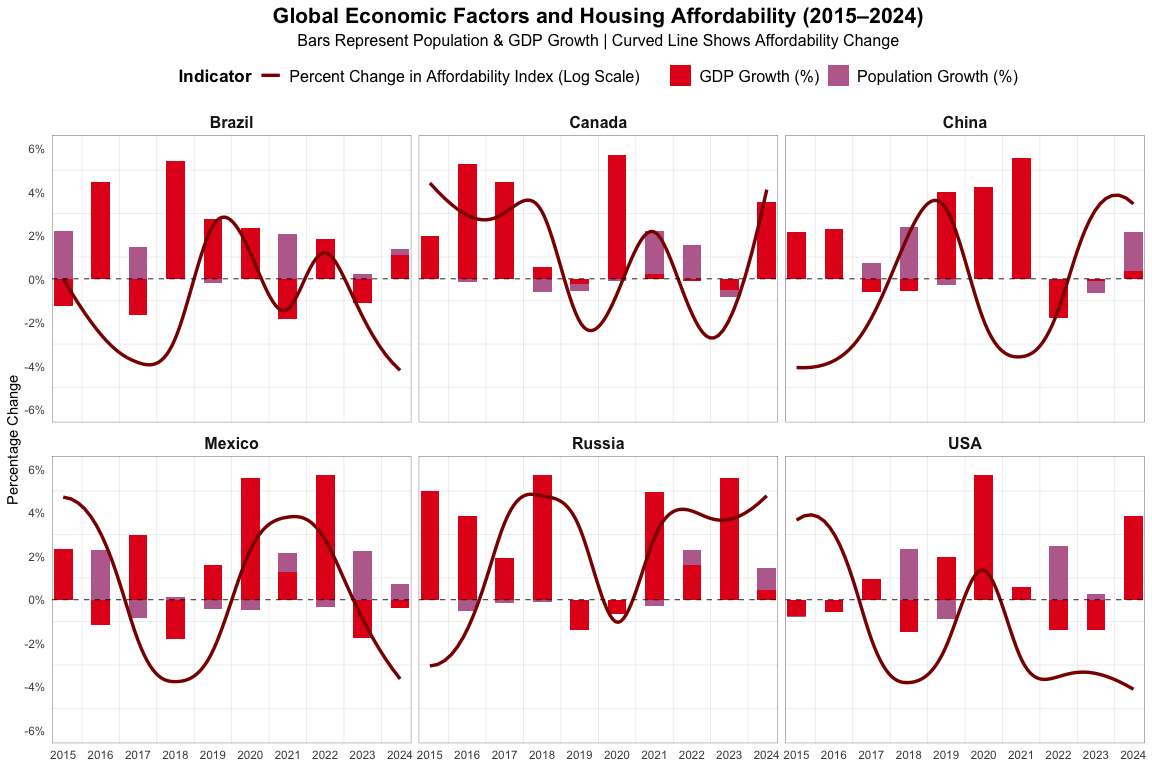
Then, because the percent changes were a lot larger than the changes in GDP growth and Population growth, I applied a log transformation so that I’m not actually changing the meaning nor accuracy of the data, but instead scaling it to the scale of GDP and Population Growth. I also added the +1 to ensure that it translates without any problem. The reason for that is because if there is a value that turns to 0, then log(0) would be undefined.

After, I pivoted it to a long tibble so that way it’s a lot easier to plot when I’ve grouped all my economic factors, or Indicators, together.

Originally, I tried to create this graph by doing three streams with ggstream. However, I realized that because ggstream isn’t using exact values and more for describing the density of the data, that it wouldn’t work well for values such as percentages. Instead, what I did was stick to bars while still maintaining the same format of my other graphs. To make this graph ‘beautiful’, I then decided to have the percent change in affordability be a curved line, since I wanted to emphasize that I wanted the percent change to be compared to the two economic indicators.

At this point, I also decided that it would look and feel a lot more easier to read if I added panel lines, spaced them out a bit more, and then have my x-values angled diagonally so that they wouldn’t collide. For the purpose of this graph, I also added a y-intercept at 0 so that it would be easy to tell between positive and negative changes over the years across all variables.

# ----------------------------  
# Calculating Percent Change of Affordability, then Pivoting Long  
# ----------------------------  
  
  
plot3data <- clean\_ghm |>  
 filter(Country %in% selected\_countries) |>  
 arrange(Year, Country) |>  
 mutate(Afford\_Change = if\_else(  
 is.na(lag(Affordability\_Ratio)),  
 0,  
 ((Affordability\_Ratio - lag(Affordability\_Ratio)) / lag(Affordability\_Ratio)) \* 100  
 )) |>  
 mutate(Afford\_Change\_LogScale = sign(Afford\_Change) \* log(abs(Afford\_Change) + 1))  
  
  
plot3data\_long <- plot3data |>  
 select(Year, Country, Population\_Growth, GDP\_Growth, Afford\_Change\_LogScale) |>  
 pivot\_longer(  
 cols = c(Population\_Growth, GDP\_Growth, Afford\_Change\_LogScale),   
 names\_to = "Indicator",  
 values\_to = "Value"  
 )  
  
# ----------------------------  
# Creating the Plot  
# ----------------------------  
  
  
plot4 <- ggplot(plot3data |> filter(Country %in% selected\_countries) |> arrange(Year), aes(x = Year)) +   
 geom\_bar(aes(y = Population\_Growth, fill = "Population Growth"), position = "dodge", stat = "identity", width = 0.5) +   
 geom\_bar(aes(y = GDP\_Growth, fill = "GDP Growth"), position = "dodge", stat = "identity", width = 0.5) +   
 geom\_bspline(aes(y = Afford\_Change\_LogScale, color = "Afford\_Change\_LogScale", group = Country), linewidth = 1.2) +  
 geom\_hline(yintercept = 0, linetype = "dashed", color = "black", linewidth = 0.3) +  
 facet\_wrap(~ Country) +   
 labs(  
 title = "Global Economic Factors and Housing Affordability (2015–2024)",  
 subtitle = "Bars Represent Population & GDP Growth | Curved Line Shows Affordability Change",  
 y = "Percentage Change",  
 x = "Year"  
 ) +  
 scale\_x\_continuous(breaks = seq(2015, 2024, by = 1), expand = c(0, .05)) +  
 scale\_y\_continuous(  
 labels = function(x) paste0(x, "%"),  
 breaks = seq(-6, 6, by = 2),  
 limits = c(-6, 6)  
 ) +  
 scale\_fill\_manual(  
 values = variablecolors,  
 labels = c("GDP Growth" = "GDP Growth (%)", "Population Growth" = "Population Growth (%)"),  
 guide = "none"  
) +  
 scale\_color\_manual(name = "Indicator",   
 values = variablecolors,  
 labels = c("Afford\_Change\_LogScale" = "Percent Change in Affordability Index (Log Scale)")) +   
 guides(  
 fill = guide\_legend(order = 2, title = NULL),  
 color = guide\_legend(order = 1, title = "Indicator")  
 ) +  
 theme\_minimal() +  
 theme(  
 panel.border = element\_rect(color = "black", fill = NA, linewidth = 0.2),  
 plot.title = element\_text(hjust = 0.5, size = 16, face = "bold"),  
 axis.title.x = element\_blank(),  
 panel.grid.major = element\_blank(),  
 plot.subtitle = element\_text(hjust = 0.5, size = 12),  
 strip.text = element\_text(size = 12, face = "bold"),  
 legend.position = "top",  
 legend.text = element\_text(size = 12),  
 legend.title = element\_text(size = 13, face = "bold")   
 )  
  
plot4



## Saving SVGs and PDFs for Vector Editor

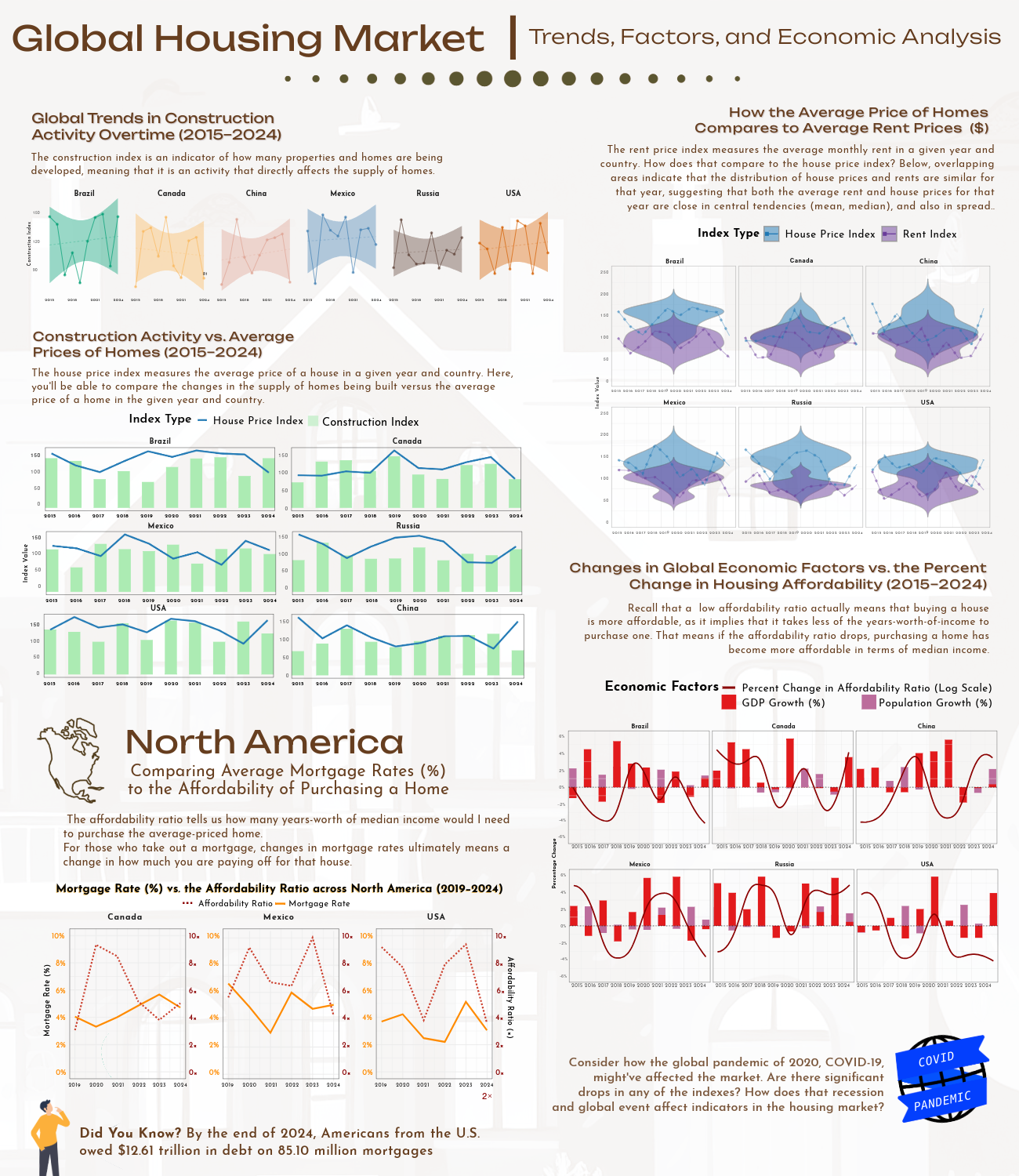
ggsave("plot1.svg", plot = plot1, width = 10, height = 5, dpi = 300, device = "svg")  
ggsave("plot2.svg", plot = plot2, width = 16, height = 9, dpi = 300, device = "svg")

ggsave("plot3.svg", plot = plot3, width = 10, height = 8, dpi = 300, device = "svg")  
ggsave("plot4.svg", plot = plot4, width = 12, height = 8, dpi = 300, device = "svg")  
  
ggsave("plot1.pdf", plot = plot1, width = 10, height = 5, dpi = 300, device = "pdf")  
ggsave("plot2.pdf", plot = plot2, width = 16, height = 9, dpi = 300, device = "pdf")

`geom\_smooth()` using formula = 'y ~ x'

ggsave("plot3.pdf", plot = plot3, width = 10, height = 8, dpi = 300, device = "pdf")  
ggsave("plot4.pdf", plot = plot4, width = 12, height = 8, dpi = 300, device = "pdf")

# Final figure



Final Project Infographic

First, I wanted the background to be something that would symbolize my infographic without interfering with the readability of my plot(s). So, I chose a free, random stock image graphic of a house, lowered the opacity, and screened it in a way that was as transparent as it needed to be. I chose to indent my visuals in a way that aligned with its closest margin, while also maintaining enough distance in the middle of the infographic. Then, I chose fonts that I felt to be ‘fresh’, engaging, and would present neatly. I also created shadows for titles the plots where I felt they might need it. If you noticed, you also might see how the color palettes transition from the top of the page to the bottom. I transitioned from pastel colors to more saturated colors by the end, because to me, it felt like as my story kept unfolding while moving down the infographic, the more it should unfold and ‘flourish’ (if that makes sense). I also chose to keep any extra facts at the bottom, along with the visualization about North America, since they focus on specific details. I tried my best to maintain all of the details of my graphs for the most part, since I refined them quite a bit in my Quarto document.

I believe my final infographic conveys truth, in which by Cairo’s philosophy comes from beauty and art, because my visualizations are aesthetically pleasing, functional, reveals patterns that are comparable across multiple demographics, years, and by economic indicators, and are also color-blind accessible. I’ve also made sure to accurately represent my data without distorting values, provided detailed explanations of variables to prevent any ambiguity and ensure that it is effective in communication, and also used geoms that would both accurately and effectively display the types of data that I was working with.