

## A Case Study on Average Home Prices in the Greater Toronto Area

### *Misleading Visualizations*

#### Introduction

---

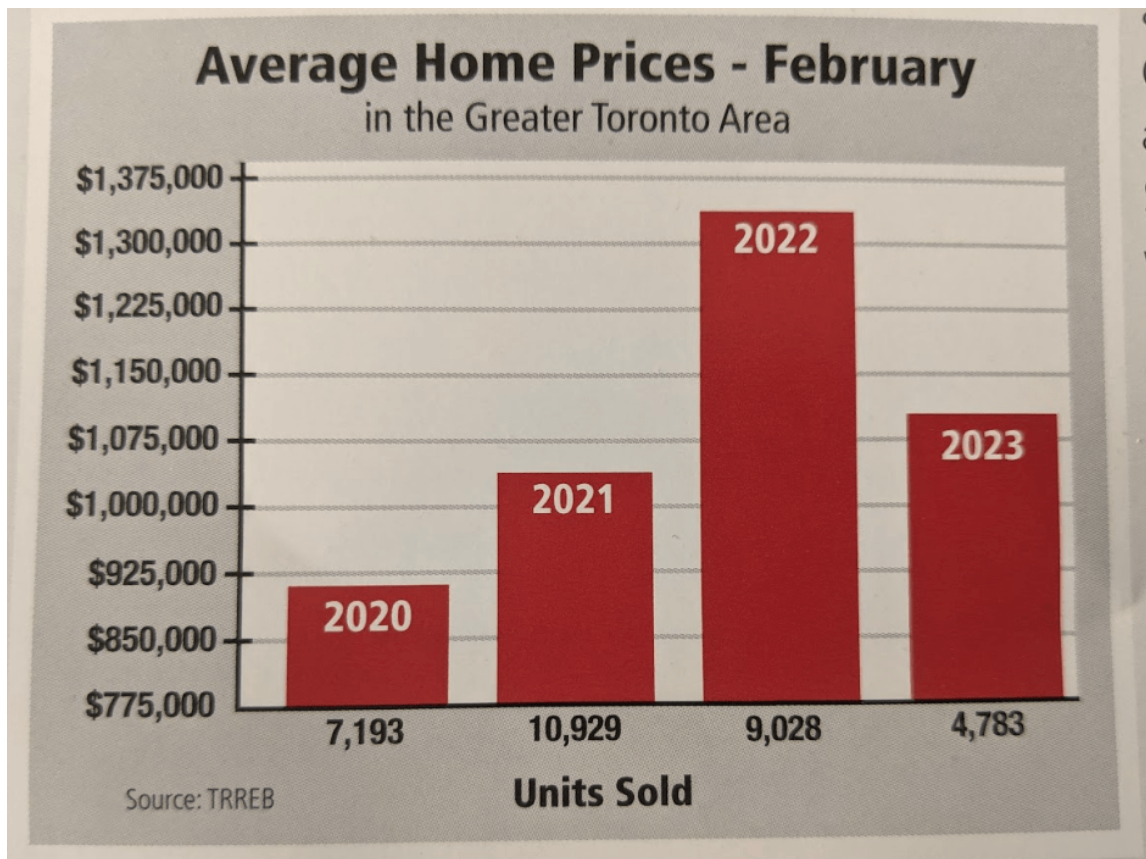


Fig 1. Average Home Prices in February in the Greater Toronto Area (2020 – 2023)

The visualization picked for the assignment is a bar graph showing average home prices in February for the Greater Toronto Area over 4 years (2020 – 2023). It also includes additional information on the number of units sold each year. At first glance, it attempts to convey two important trends in real estate: the number of homes sold and the average prices over the same period. However, the representation creates confusion due to the improper usage of bar graph, which can confuse the viewer.

## Misleading visualization

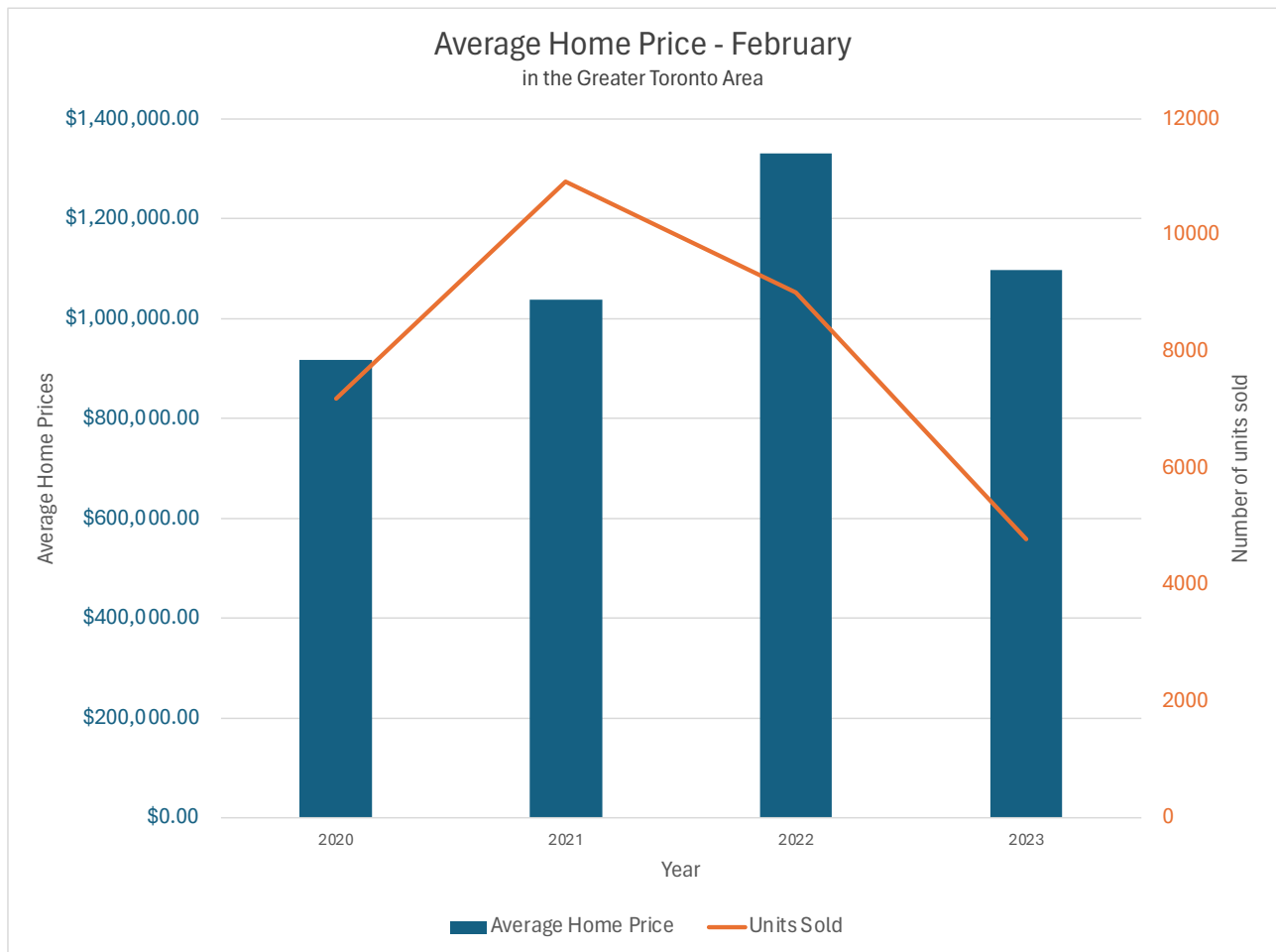
---

In general, there is too much information on the plot and although the Y – axis values are depicted with the \$ sign, the axis title is missing, which might be misleading. Let's analyze this visualization, clearly stating what is misleading about it, and propose clearer alternatives, and offer a corrected version that more effectively represents the data.

<b>Analysis</b>	<b>Fix</b>
<p><b>Inconsistent data representation and visual clutter</b></p> <p><i>Bar graphs are most effective when representing a single variable. The chart above, however, combines two metrics. The X – axis shows the number of units sold each year in the month of February and the Y – axis shows the average prices of the houses (the axis title is missing). The year is labeled on the bars. Heights of the bars can be interpreted incorrectly as representing both metrics simultaneously. This can lead to the false assumption of a strong relationship between average home prices and the units sold.</i></p>	<p><b>Separating metrics for clarity</b></p> <p>To avoid any confusion between the two metrics and allow the viewer to focus on one metric at a time. The text should clarify that each bar should represent only one metric, and a possible solution could be <b>two bars for each year</b> to compare average home prices and units sold. Another possible solution is to use <b>two separate visualizations</b> for each metric.</p>
<p><b>Axes Labeling and Titles</b></p> <p><i>The chart's X-axis currently displays the number of units sold each year, which is not an appropriate label for this axis given that the primary focus should be on comparing home prices over time. This mislabeling can confuse viewers by making it unclear whether the X-axis represents time or another metric. The Y – axis does not start with 0 for us to make accurate sense of the trend. Consequently, viewers may struggle to accurately interpret trends in average home prices across different years.</i></p>	<p><b>Proper labeling for axes</b></p> <p>To improve clarity, the <b>X-axis should be labeled with the years</b> (e.g., 2020, 2021, 2022, 2023) rather than the number of units sold. This adjustment will correctly represent the passage of time, allow viewers to understand changes more easily in average home prices over the specified years and focus on the temporal trend of home prices without confusing it with units sold.</p>
<p><b>Misleading trends</b></p> <p><i>The X – axis does not accurately categorize the units sold; this can distort the viewer's perception of the data. Moreover, the number of units sold are printed beneath the bars. If they are trying to show the relationship between home prices and units sold, it is not effectively communicated.</i></p>	<p><b>Adding a line chart for complementary data</b></p> <p>We can use a bar chart to display average home prices per year and a <b>line chart to show the trend in units sold per year</b>. This approach clearly distinguishes between the two metrics and helps viewers understand their individual trends without confusion.</p>

## Updated Visualizations: A more effective approach

We can use separate visuals for the two metrics, average home prices and units sold. The revised graph, created using Microsoft Excel, represents a significant improvement over the original version for several reasons.



*Fig 2. Fixed Visualization with uniform X – axis and separate Y – axes for clarity*

The corrected X – axis in *Fig 2* just the year information accurately reflects the temporal dimension of the data and eliminates any confusion caused with the *Fig 1*. *Fig 2* more effectively allows viewers to identify and analyze trends in the average home prices and the units sold over time. It is easier to see if the prices are increasing, decreasing, or remaining stable and understand how market dynamics evolve each year.

The use of dual axis for each metric provides a comprehensive view of the relationship between the two variables which can be studied further based on the influencing factors. This allows viewers to see how changes in average home prices correlate with changes in units sold. The legend eliminates any ambiguity.

The usage of bar chart and a line plot clearly separates the two metrics. This can be depicted with 2 different graphs as well. This separation has reduced visual clutter, and each type of data is best suited to its nature; quantitative data with bars and trends with a line.

## Conclusion

---

The original chart, while visually striking, misrepresents the data by truncating the Y-axis, combining two variables in a single visual, and poor axis labeling. By separating the metrics into distinct visualizations—using a bar chart for home prices and a line chart for units sold with dual Y-axes—the revised charts offer clearer insights. This approach eliminates ambiguity, accurately represents trends, and facilitates a better understanding of the relationship between the two variables. These changes help viewers better understand trends in the real estate market in Greater Toronto Area and make informed interpretations and decision – making based on the visualization.