

I was retained by Ralls & Wille to review data for a client, Constantine Panousopoulos, and to produce graphical summaries of the data. The data relates to properties in Santa Cruz County, their assessed values for taxation purposes, protests filed to challenge those assessed values, and the results of those protests. I was specifically tasked with building plots that can help to identify temporal variations in several metrics having to do with whether or not the protests were granted, the dollar amounts that were granted or not granted, and percentages of the total property values that these dollar amounts reflect.

Data:

164 protest documents related to properties owned by Panousopoulos were compiled into a Microsoft Excel document and the shared with me. The relevant variables included information regarding the tax year of the protest, the date of the protest, six categories of property value (full cash value and limited property value for the originally assessed values, the owner's desired values, and the ultimate decision values), the percentage of change from the requested value to the decision value, as well as observations about whether the protests were granted in full, granted in part, or denied. Other variables about the relevant properties were also recorded in the excel sheet, but those variables were not used in the analysis.

Data cleaning and preparation was conducted with the help and consultation of Ralls & Wille. This process began when I observed that some of the relevant fields had non-numerical values that could not be plotted. When a row of data included fields with non-numerical entries, those fields were converted into a numerical values if possible, or if not possible, the row was removed from the analysis. Attached to the end of this report is a log of the steps that were taken to arrange the data into formats that could be used for plotting. Below is a description of the kinds of changes that were made.

For rows where the protest decision was labeled "N/C" or "no change", the dollar value of the protest decision was set to the original assessed value. For older protests that did not specify FCV or LPV, both of those fields were set to be the same (implying that there was no substantive difference between FCV and LPV). For older protests that separated values for "land," "improvements," and/or "total," only a total (combining land and improvements) was used to allow comparison to modern protests. In protests that did not challenge the assessed value of the property (such as protests challenging only the property type and corresponding tax ratio rate), the protests were designated to be omitted from the plots. For protest documents that did not include a decision value, those rows were also designated

to be omitted from the plots. For protest documents that did not include a new requested value by the owner, but did show a reduction, the decision/reduced value was copied into the requested value column. Because there were two instances in which the decision LPV increased from the originally assessed value, the outcome designation of “No Change” was altered to “No Change or Worse.” For protest documents that did not include a “tax year,” the tax year was presumed based on the date of the protest filing (in other words, one year was added to the filing year to determine the tax year). For some of the rows, two documents were provided for the same protest/property/year, in which case data was used from either (1) the document that showed an actual decision, or (2) the document that showed the larger change in value.

After generating draft graphical summaries, outliers in the data became apparent. For example, in Row 38, Column L (owner’s requested LPV), the value “8442810” was incorrectly written (the correct amount per documentation was “2442810”). Similarly, it was apparent that a row of data was missing for the 2020 tax year due to the number of results that were displayed. Correction of these data entry errors are noted in the cleaning log at the end of this document as well.

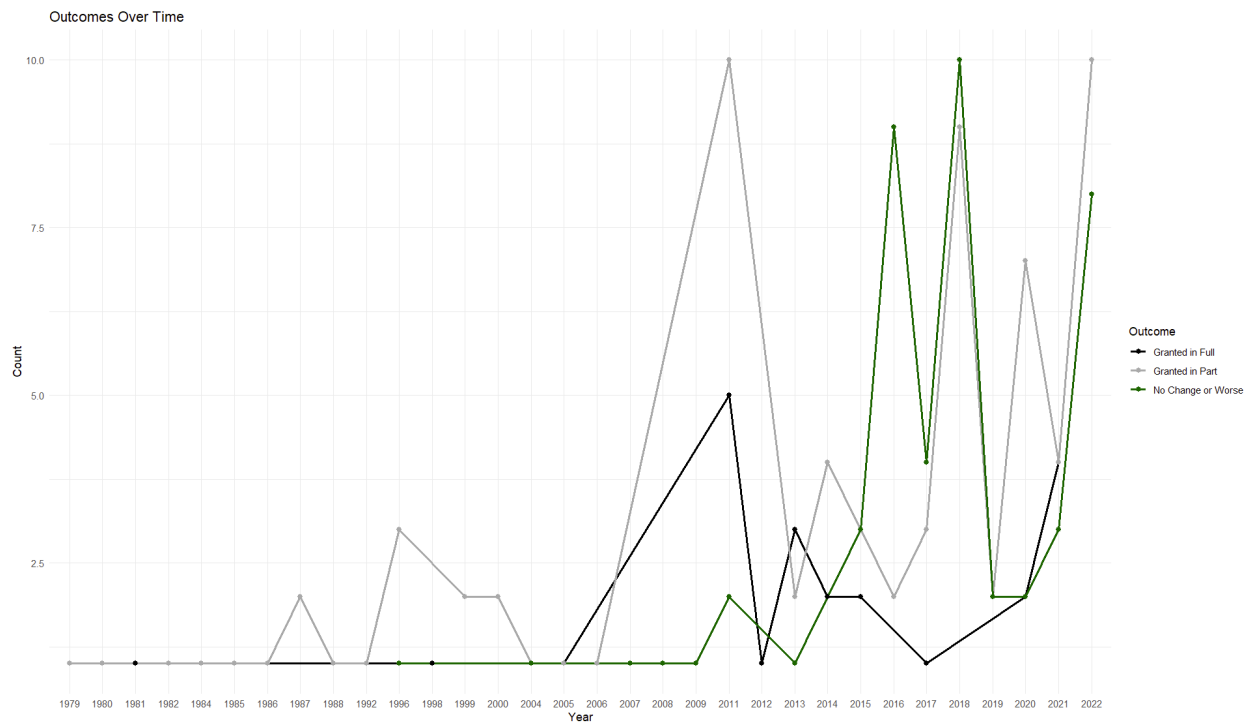
A second set of data was also obtained from three Microsoft Excel documents titled “Notice of Value 2014-2015-2016,” “Notice of Value 2017-2018,” and “Notice of Value 2018-2025.” These documents included information regarding the original assessed values for a series of properties. In that data set, there were also tables that indicated the properties were not located in Santa Cruz County, and therefore irrelevant to the issues at hand. This data set included many more properties than the first data set, because only a few properties’ values were protested in any given year. At this time, I have not produced any summaries of this larger dataset.

Process:

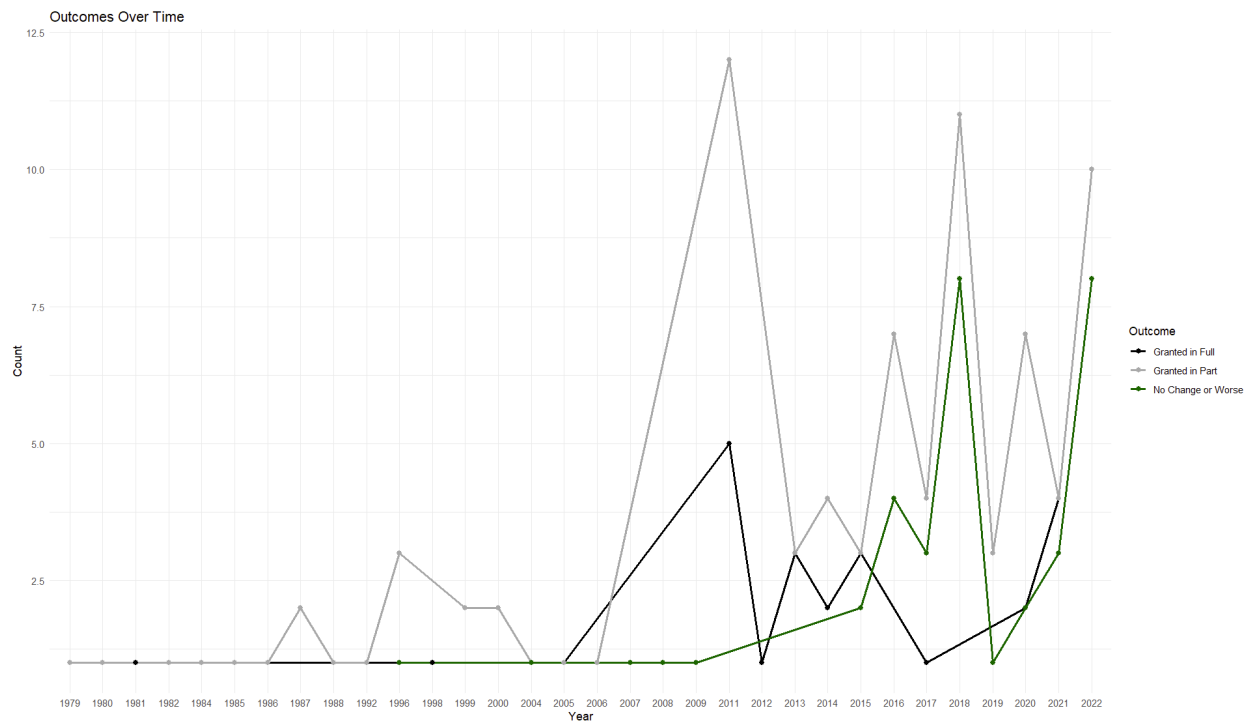
Using the data set generated from the protest documentation and the programming language “R,” I generated various graphical summaries (“plots”) of the information. Each plot was generated twice, once with FCV and once with LPV. In the plots, the x-axis always represents the tax year (as opposed to the year a protest document was filed). The y-axis was specific to the variables analyzed, as described below.

The first variable summarized was a categorical description of the decision of the protest, which was either “Granted in Full,” “Granted in Part,” or “No Change or Worse.” I plotted the same variables in line graphs and stacked bar graphs. In these plots, the y-axis represents the raw number of protests that were granted in full, granted in part, or denied (no change).

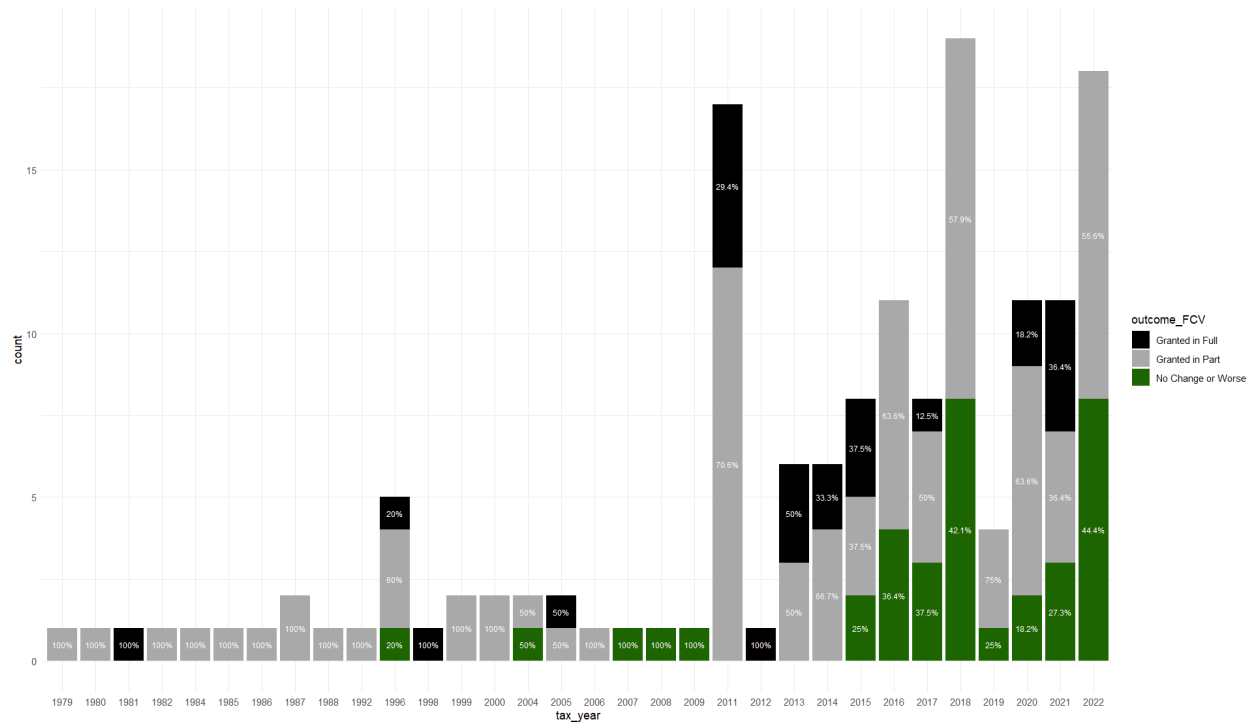
“FCV decision summary line.png”



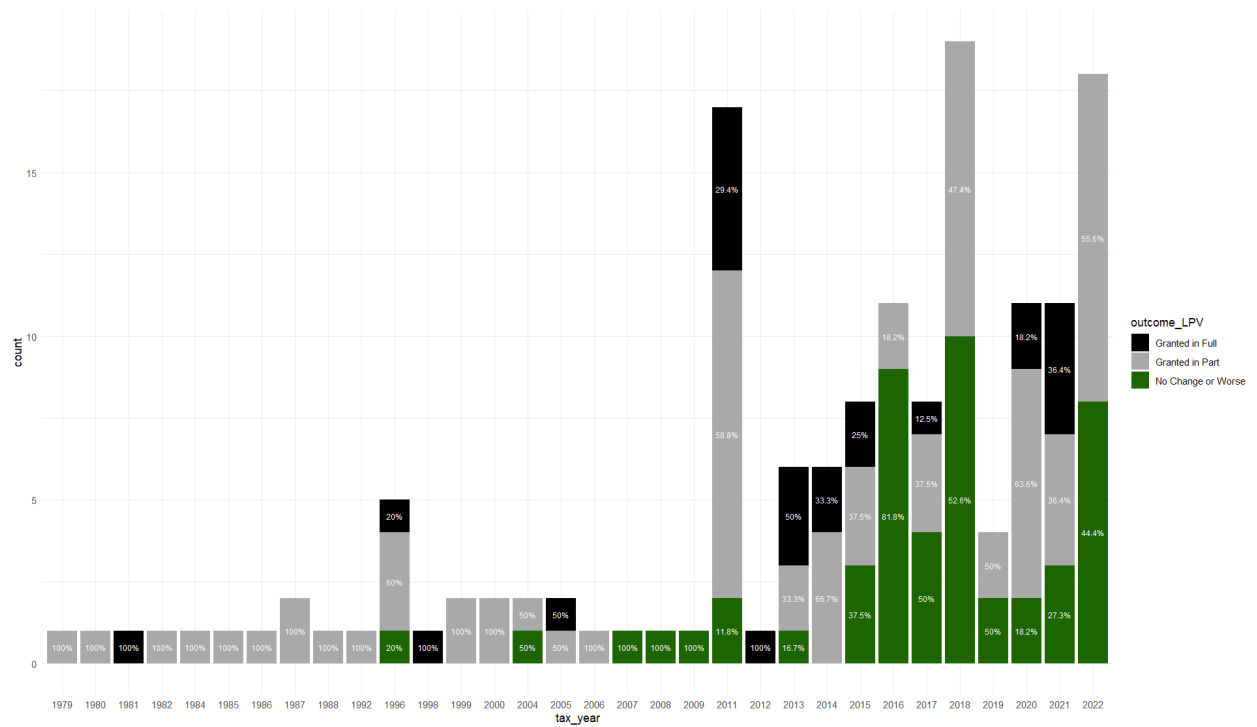
“LPV decision summary line.png”



“FCV decision summary stacked.png”

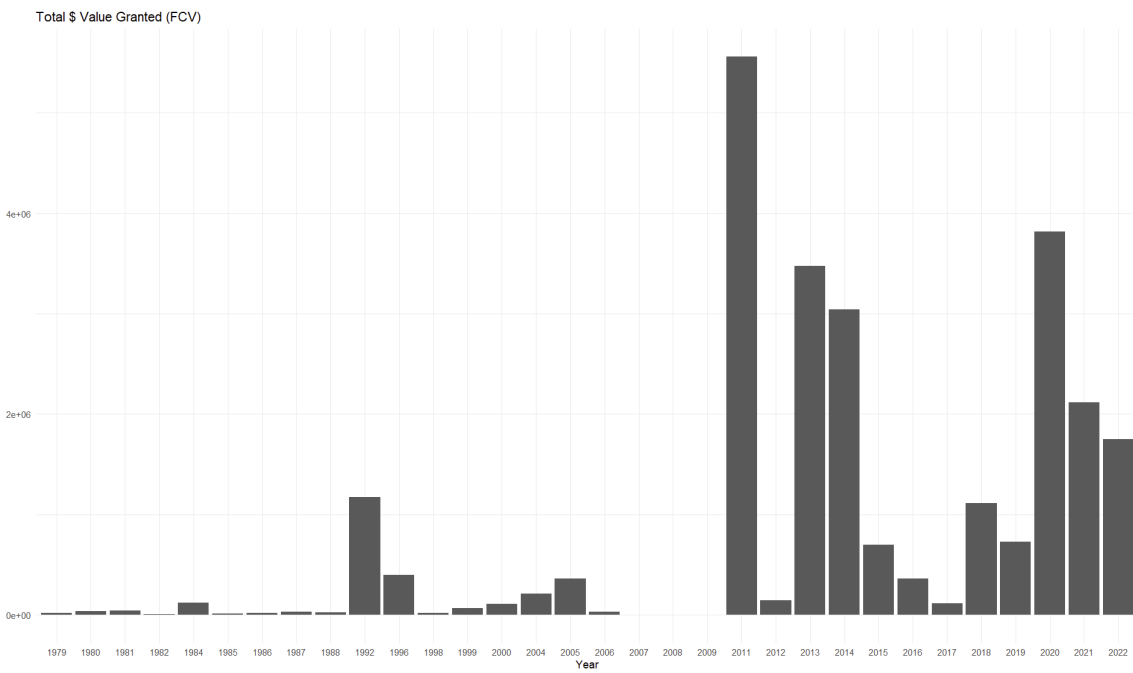


“LPV decision summary stacked.png”

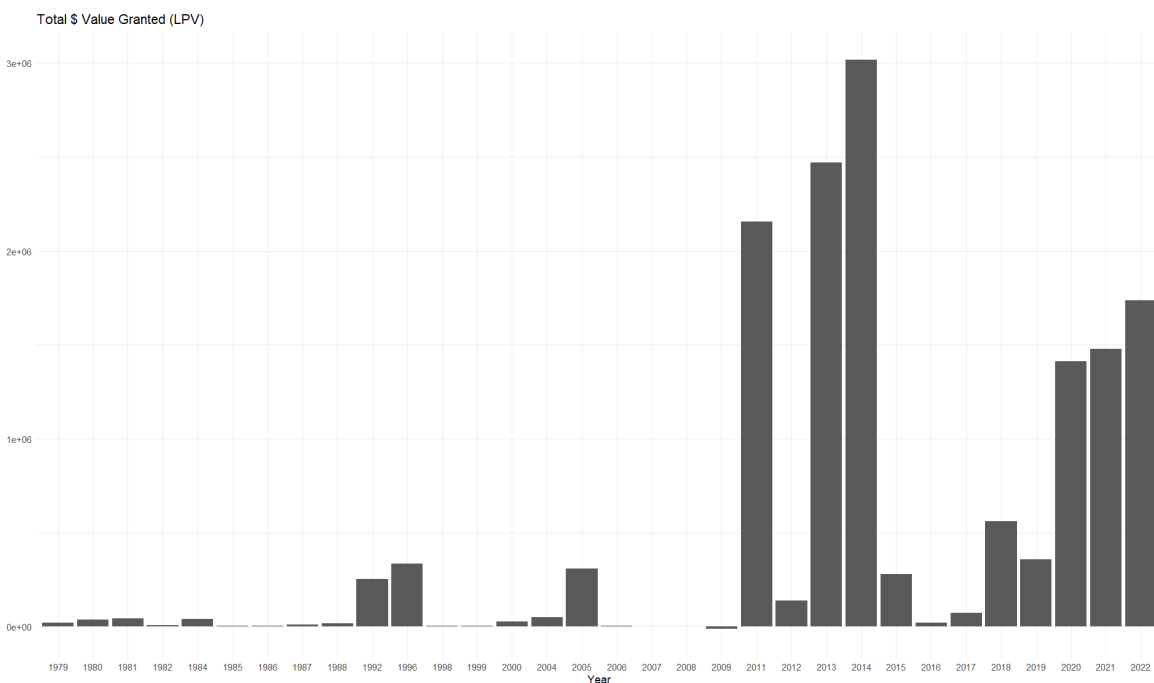


In addition to plotting the number of protests that were granted or not granted, as above, I also produced bar plots of the total dollar value reduction granted each year, both in terms of FCV and LPV.

“FCV total dollar granted.png”

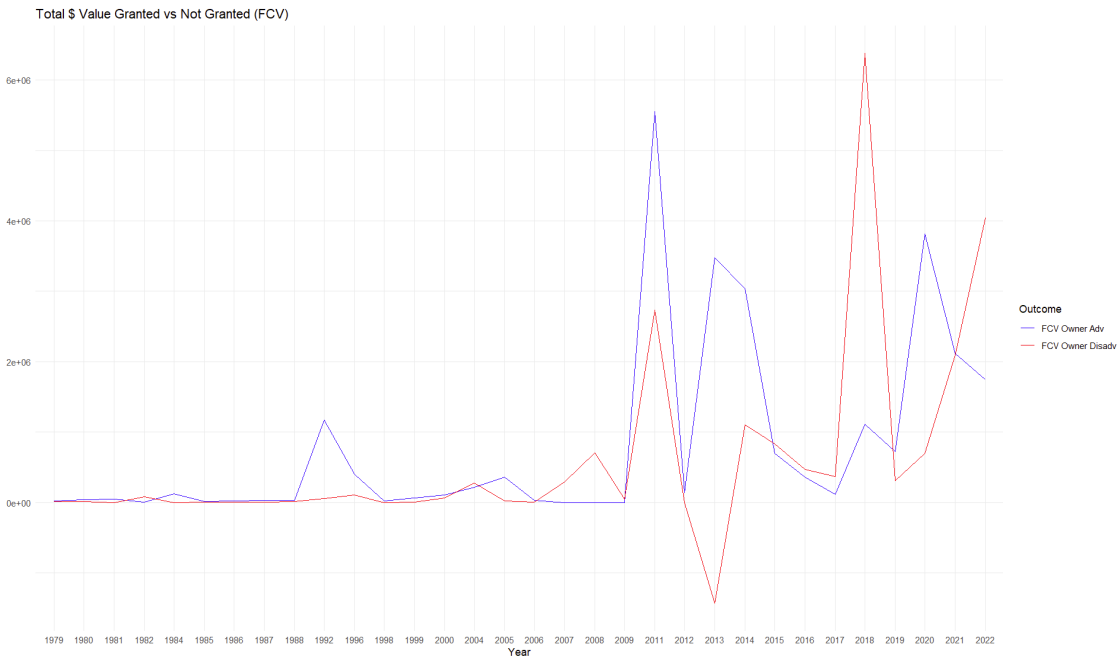


“LPV total dollar granted.png”

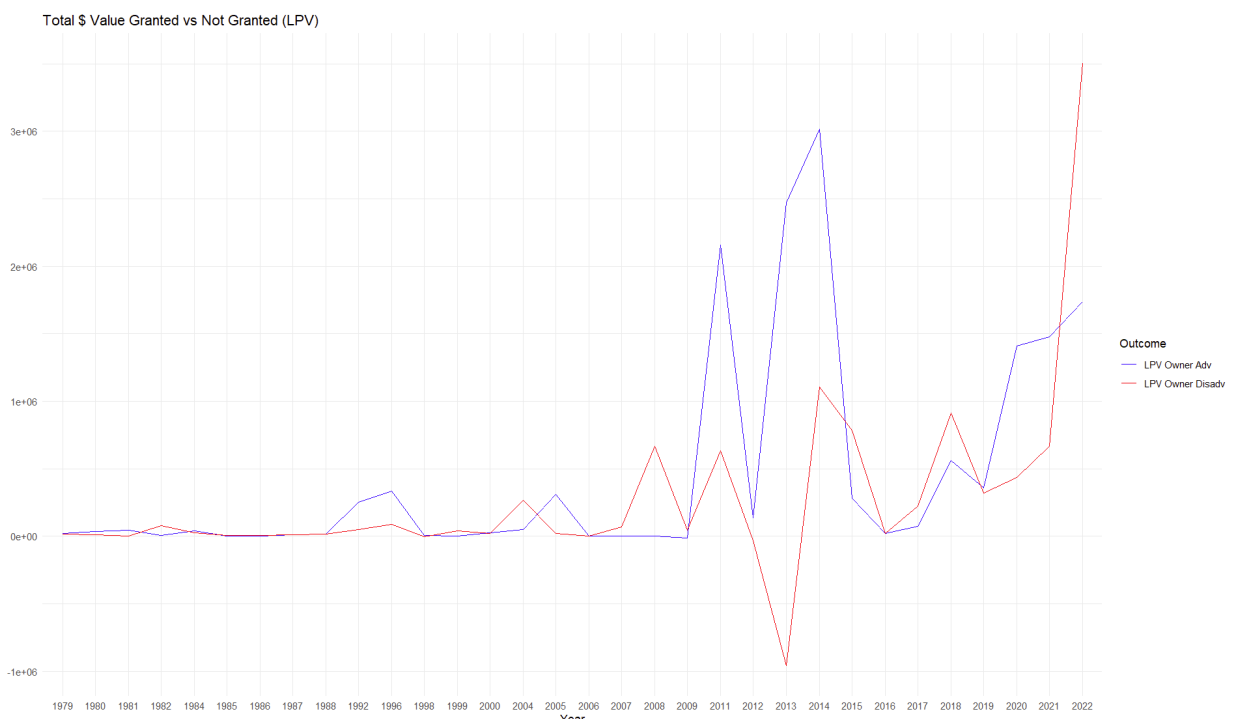


The next series of plots address the total dollar value granted and the total value not granted in the protest decisions. For simplicity, I labeled these variables as “Owner Adv” or “Owner Disadv,” but the values represent dollar-denominated accepted changes and denied changes to the property values, respectively.

“FCV total dollar granted versus not granted.png”



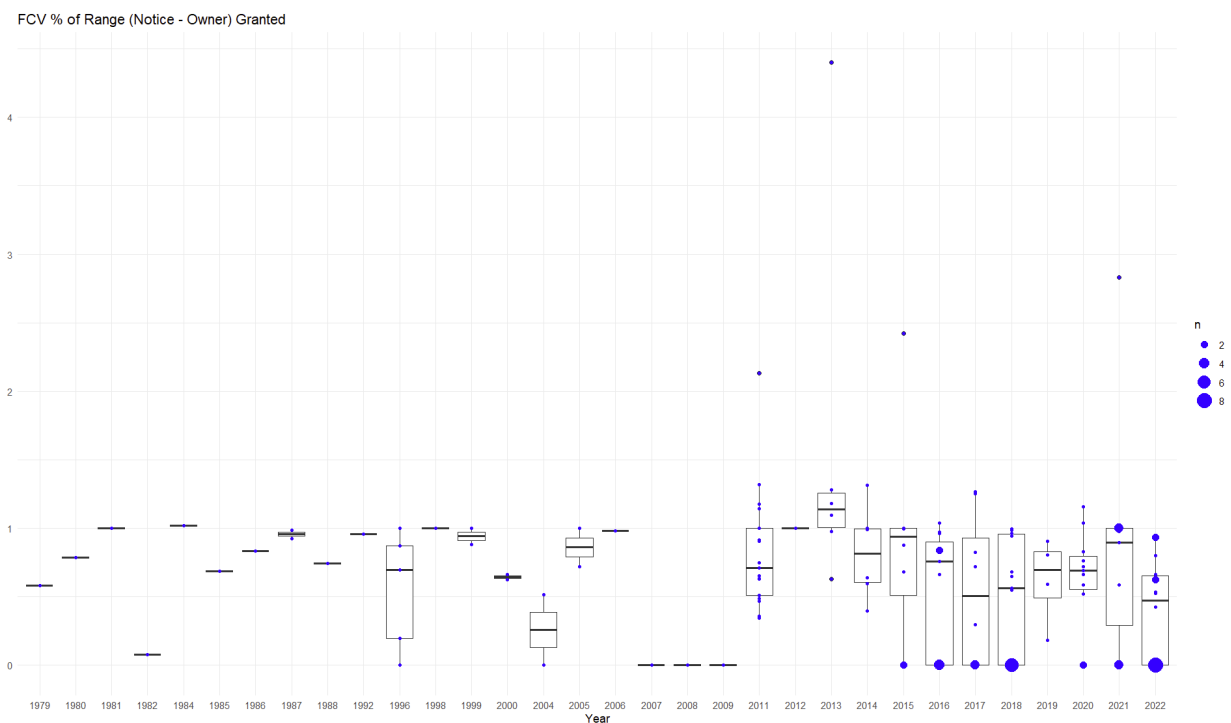
“LPV total dollar granted versus not granted.png”



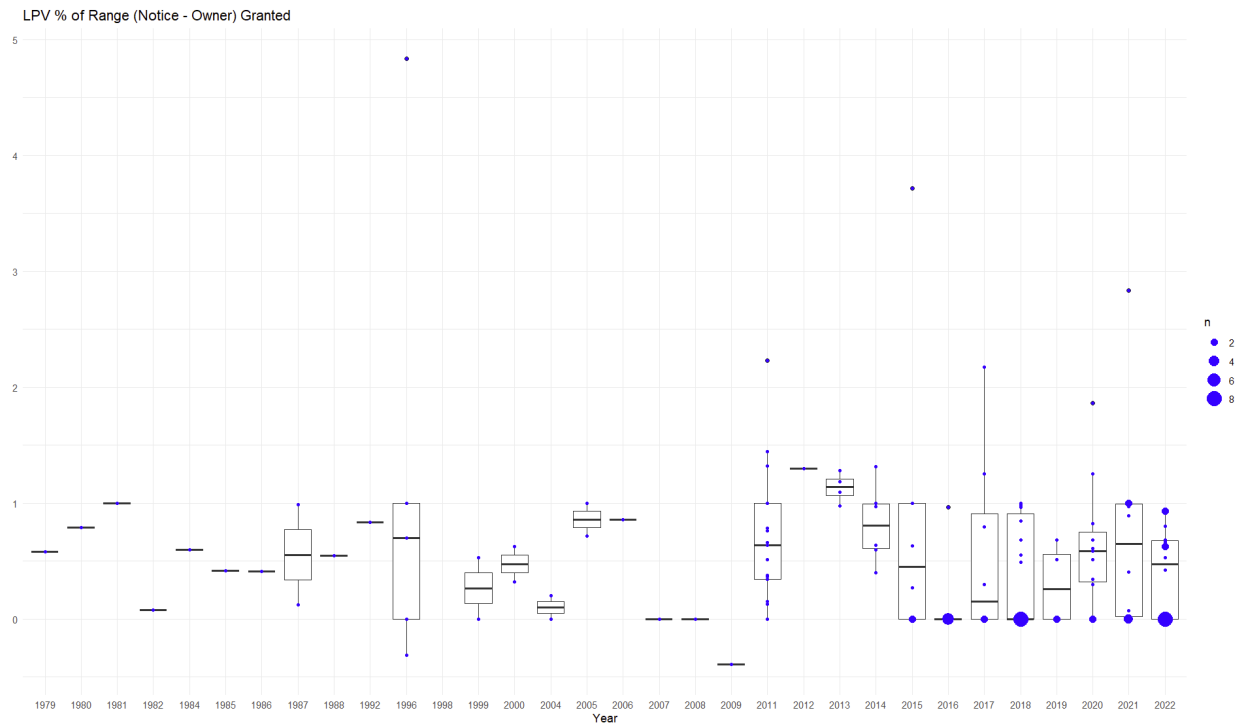
Next, I produced a series of box plots (also known as box-and-whisker plots) to summarize the proportional (or percentage) change resulting from each protest. A box plot provides a visual summary of key statistical measures such as the median, quartiles, and potential outliers. In the graphics below, each dot represents a protest's percentage of change (and larger dots indicate multiple protests landing on the same spot). Each box represents the middle 50% of where the data lies (i.e., the data falling between the 25th and 75th percentile, also known as the inter-quartile range). The horizontal line inside the box is the median line (the middle value when ordered) of the protest data for the corresponding year. The lines extending above and below a box are the “whiskers,” which represent the other 50% of where the data lies (with the exception of outliers, which are reflected in dots that fall outside of the box and whiskers of the plot).

In the pair of box plots below, I plotted what percentage of the property owner's request was granted in a decision. A 0 here reflects that 0% of the request was granted, while a 1 reflects that 100% of the request was granted. In a small number of cases, the number exceeds 1, which reflects that the decision ended in a property value below the request of the owner.

“FCV percentage of range granted.png”

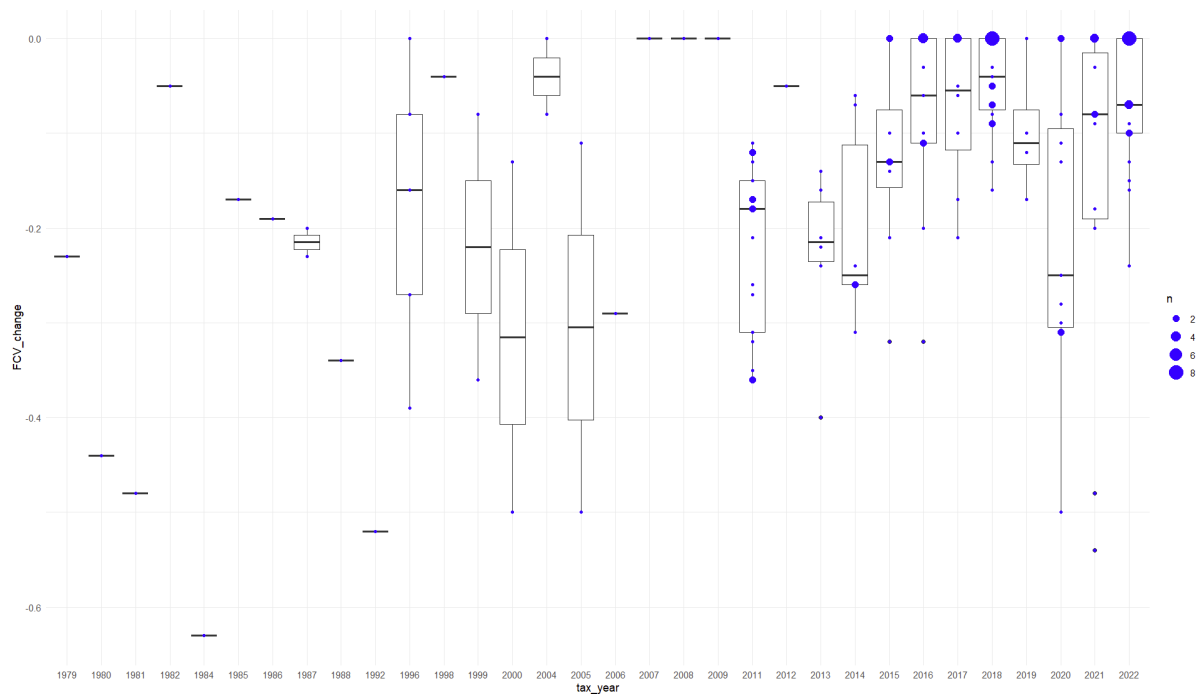


“LPV percentage of range granted.png”

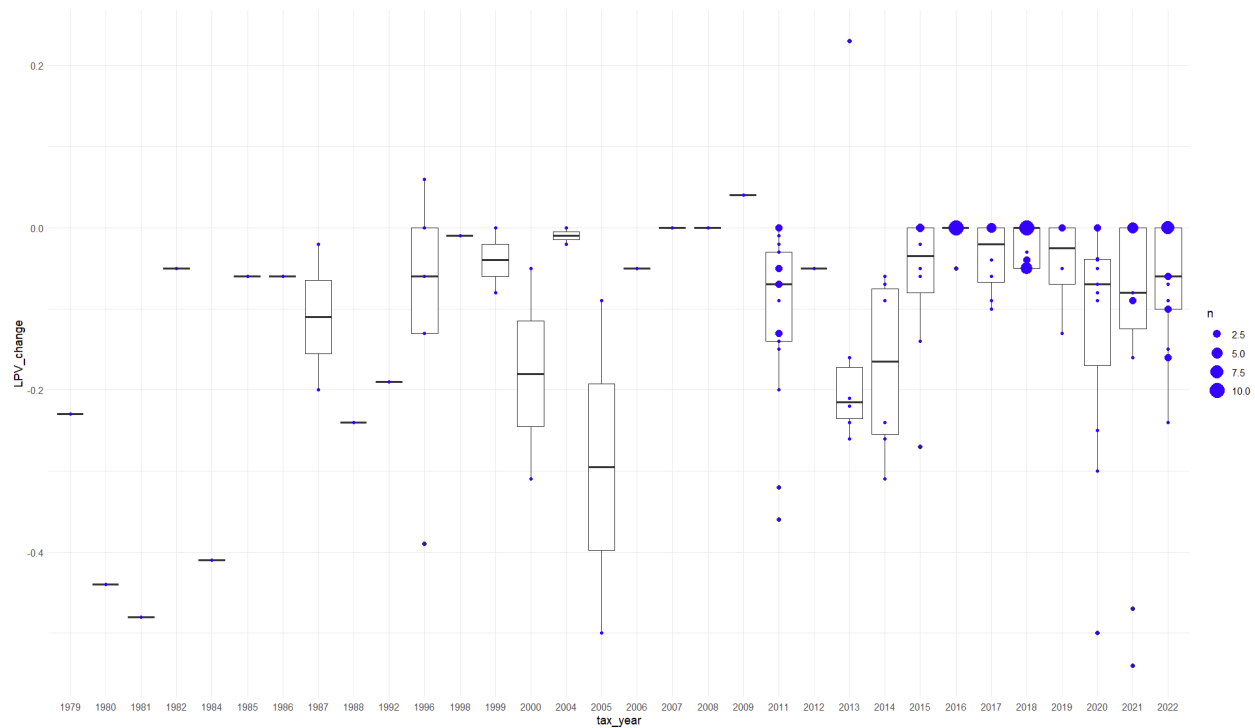


In the next pair of box plots, I plotted the percentage of actual change occurred between the original notice of value and the final decision. Because this plot represents decreases in value, this plot shows negative values (below zero).

“FCV percent change.png”



“LPV percent change.png”



Conclusions:

My expertise as a data scientist involves the processing and visualization of data in ways that reveal their underlying trends or patterns. I am capable of explaining how these graphs were generated and what these graphs represent in terms of variation (or changes) over time. Understanding what conclusions one should draw from these graphs, however, would require additional context beyond the scope of the data that I visualized. I am not an expert in property tax valuations, and once the work of generating a plot is complete, I am as capable as anyone else of observing the graph and drawing inferences or conclusions from its graphical representations.

Yotam Shmargad, Ph.D.

3/21/2025

Date

Data Cleaning Log:

Only protest documents related to properties owned by Panousopoulos at the time of the protest were included in the data set.

If a protest form did not include a “tax year,” the correct tax year was presumed based on the date of the protest. For example, a protest filed in April of 2025 would necessarily be for the 2026 tax year.

Row 2: Land and improvements combined. The same figure was placed under both full cash value and limited property value, because that distinguishing variable did not exist at that time.

Row 3: Land and improvements combined. The same figure was placed under both full cash value and limited property value, because that distinguishing variable did not exist at that time.

Row 4: Land and improvements combined. The same figure was placed under both full cash value and limited property value, because that distinguishing variable did not exist at that time. Note: In the decision values, land and total values were omitted, and only a new improvements value was listed. To resolve this issue, the originally assessed land value was added to the decision total of the improvements to get the total of \$47,000.

Row 5: Land and improvements combined. The same figure was placed under both full cash value and limited property value, because that distinguishing variable did not exist at that time.

Row 12: This row was removed because the protest addressed the taxation ratio, not the values of the property.

Row 14: This row was removed because the protest did not request a specific new value and the protest was denied in any event.

Row 20: This row was removed because there was no data provided as to the assessor’s decision.

Row 22: This row was removed because the protest addressed taxation ratios, not the values of the property.

Row 31: This row was removed because the protest did not request a new value nor was there any assessor’s decision provided.

Row 31: The row of data included the wrong year in the “date filed.” The correct year was 2005, because of the “date received” on the original documentation and because this was

for the 2006 tax year. Also note, the correct parcel number was only included in the “basis for this petition” section.

Row 35: Column V – Because the LPV actually rose to the owner’s detriment, the outcome for this row are listed as “No Change or Worse.”

Row 36: This row was removed because no assessor’s decision was provided.

Row 38: Two different protest documents and decisions were recovered. They show different values for the decisions. The lower decision values were included in the data set. Later, it was discovered that Row 38, Column L, incorrectly listed “8442810.” The correct value per documentation is “2442810.”

Row 48: This row did not include a requested new value, but nonetheless reflects that a decrease was granted. To accommodate, the row was treated as granted in full, and the decision value was copied into the requested value.

Row 52: Two different protest documents and decisions were recovered. They show different values for the decisions. To accommodate, the lower decision values were included in the data set.

Row 53: This row was removed because no assessor’s decision was provided.

Row 55: This row was removed because no assessor’s decision was provided.

Row 56: This row was removed because no assessor’s decision was provided.

Row 61: Because the LPV actually rose above the original notice, the column regarding outcomes was changed to reflect “No Change or Worse” to encompass any adverse ruling.

Row 66: Two documents were obtained pertaining to this protest. One had no decision. The other had an official decision. Data was used from the document that included an official decision.

Row 68: Two documents were obtained pertaining to this protest. One had no decision. The other had an official decision. Data was used from the document that included an official decision.

Row 129: This row was removed because no assessor’s decision was provided.

Row 130: This row did not have data for the owner’s new requested value. To accommodate, the decision value was copied into the requested value. Columns Q, U, and V were also changed to reflect this change.

Row 133: This row did not have data for the owner's new requested value. To accommodate, the decision value was copied into the requested value. Columns Q, U, and V were also changed to reflect this change.

Row 135: This row was removed because no original value and no new requested value was provided. The "basis for decision" does indicate that the protest was denied in full, but due to the other missing data, the row is removed.

Row 137: This row was removed because the decision included a clear error, raising the assessed value from around \$800,000 to around \$8,000,000.

Row 138: This row did not have data for the owner's new requested value. To accommodate, the decision value was copied into the requested value. Columns Q, U, and V were also changed to reflect this change.

Row 149: This row was removed because no decision value was provided.

Row 151: This row was originally compiled from online data. There is a scan of the original petition, but it does not have a decision value listed and a post-it note covers part of the data. Upon review, it is clear there is at least one data error in the excel sheet—the original Notice of Value for LPV is listed as 446499, but the scanned document shows 682423. The scanned document's data needs to be added to the data set. This row will be updated upon receipt of a complete copy of the protest and decision from the County Assessor.

Row 163: During further review of protest documents obtained by the Government, a document showing the date filed and parcel number for this row was identified. That information has been added to the excel sheet.

Row 164: During further review of protest documents obtained by the Government, a document that was not reflected in the excel sheet was identified, regarding the 2020 tax year. All data from that protest has been added into row 164.