

# Analyzing the Dutch Housing Market:

**Trends, Patterns, and Correlations**

**Piotr Kucharzak**



# **Key Questions through Data Analysis:**

## **1. The Dutch Housing Market: An Analysis of Prices Over Time**

- What are the current price trends for houses in the Dutch market, and how have they changed over time?
- How has the price per square meter of houses in the Dutch market evolved over the years of build houses.
- How have prices for houses built in the last 50 years changed, and what insights can we gain from this analysis?

## **2. Understanding Housing Trends in the Netherlands**

- What do the energy labels of houses in the Dutch market tell us about the state of the housing?
- Can we identify any patterns or trends in the types of houses that buyers are looking for in the Dutch market?
- How does the location of a house impact its price in the Dutch market, and what insights can we draw from this analysis?

## **3. Statistical Analysis of the Dutch Housing Market**

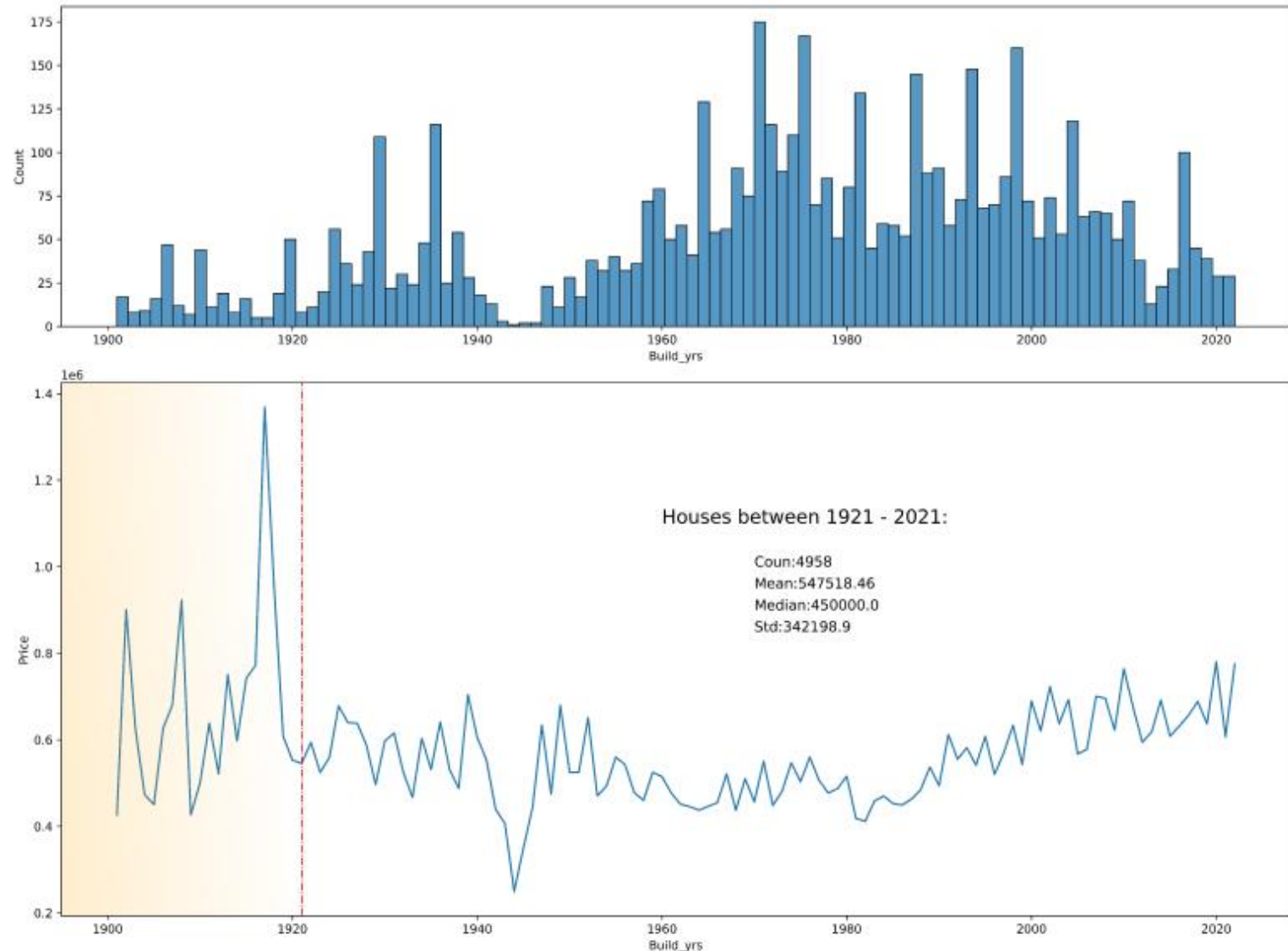
- Can we identify any statistically significant differences in the factors that influence house prices in the Netherlands?

# Houses Price vs Houses Build\_yrs

\* To provide a more accurate representation, the analysis will be focused on the **House buildied between 1921 and 2021**. This approach will ensure that the analysis is based on a more consistent and accurate representation of the property prices over time.

\* The analysis shows that the property prices between 1921 and 2021 have a **median value of 450.0 thousand €**, indicating that half of the properties in the dataset have prices below this value and the other half have prices above it. However, the **mean price of 547.518.46 €** is much higher than the median, suggesting the presence of a few very high-priced properties that are significantly impacting the average value.

\* **The standard deviation of 342.198.9** means that there is a large amount of variability in the data, with values spread out from the mean. Specifically, the standard deviation is a measure of how much the individual data points deviate from the average value (mean) of the dataset. In this case, the high standard deviation indicates that there are many data points that are far from the average price, indicating a large range of prices for the properties in the dataset.





Comparision Price\_per\_(m2) vs Houses build years



\* Based on observation, it appears that there was a period of decreasing Price\_per\_(m2) for houses built prior to 1970 where also Dataset contain value gaps between 1940-1950, followed by an increase in Price\_per\_(m2) for more recent constructions.

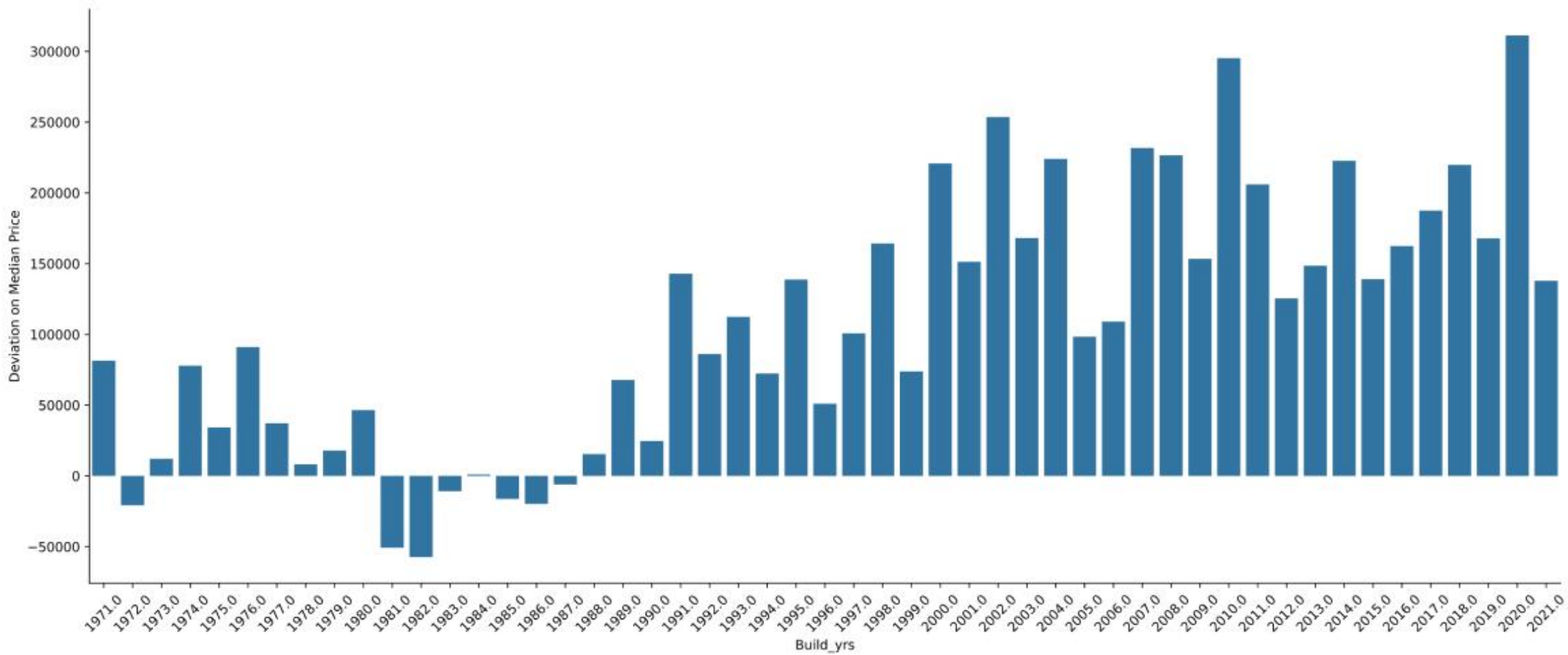
This could be due to a variety of factors such as changes in housing demand, economic conditions, or housing regulations.

In next analysis i will focus only on Houses build between 1970 -2021.

- 1940 Nazi German invades, Rotterdam heavily bombed. The Velvet Glove: German occupation for most Dutch is relatively mild.
- 1945 Canadian force liberate Netherlands.
- 1948 Marshall Plan: Massive US economic assistance to Europe including Netherlands.
- 1950s An economic miracle: Netherlands part of western Europe's economic recovery.

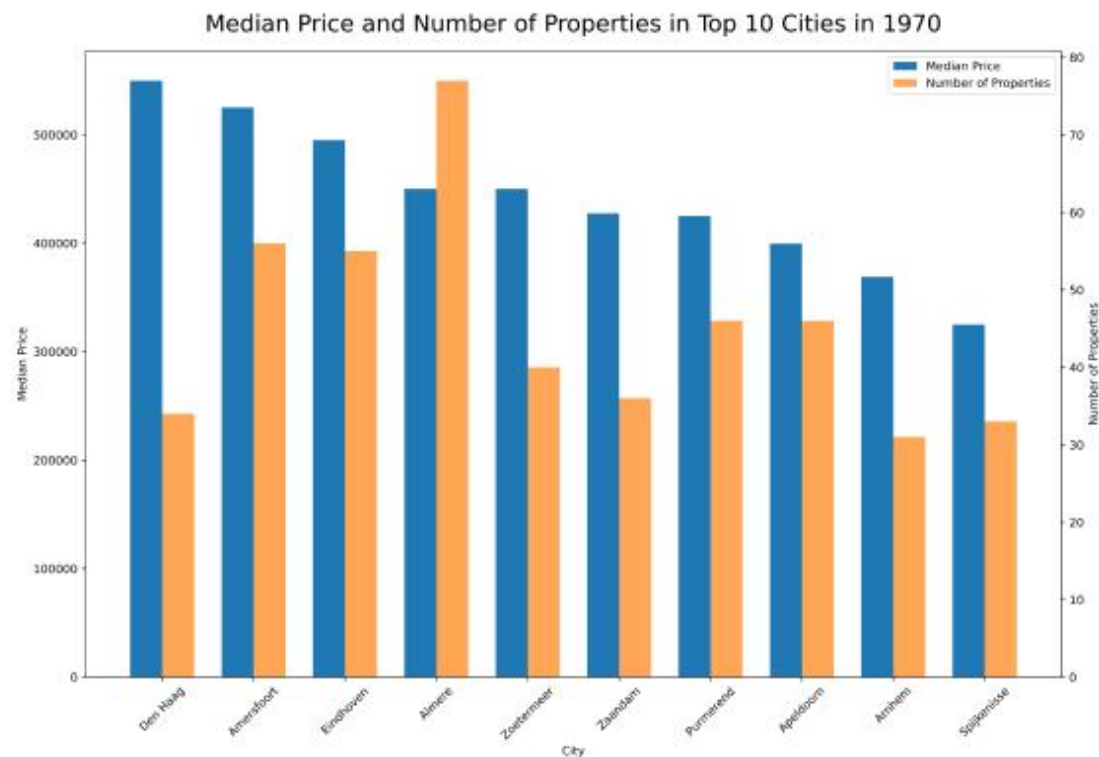
\* The housing market has consistently shown a pattern of increasing prices for **houses built after 1987**. This is likely due to the fact that houses constructed during this time period generally have better overall condition and quality, which is reflected in their higher market value. Additionally, newer houses often have modern features and amenities that are in high demand among buyers, further contributing to their appreciation in value over time. As such, the trend of constant growth in housing prices for houses built after 1987 is a reflection of their perceived value and desirability in the real estate market.

Deviation Chart



# Overview of Property Prices in Top 10 Cities since 1970:

- The dataframe shows the top 10 cities with the highest median property price since 1970.
- Den Haag had the highest median property price of €549,750, with a range of €335,000 to €945,000.
- Amersfoort, Eindhoven, and Almere also had median property prices above €450,000, with ranges of €375,000 to €1,450,000.
- Zoetermeer, Zaandam, Purmerend, Apeldoorn, Arnhem, and Spijkenisse rounded out the top 10 with median property prices ranging from €325,000 to €427,500.
- The number of properties sold in each city varied widely, with Almere having the highest number at 77 and Arnhem having the lowest at 31.
- The mean property price was higher than the median in most cities, suggesting that the distribution of prices may have been skewed towards higher-end properties.
- The maximum property price was much higher than the median in some cities, indicating the presence of luxury properties or outliers in the data.





# The summary statistics provide valuable insights into the data:

From the results, we can see that the most common house type is 'Eengezinswoning', with 2630 houses in the dataset. The 'Villa' category has the highest median price of 995,000.0, which is nearly double the median price of the most common house type, 'Eengezinswoning' (439,000.0). The 'Herenhuis' and 'Bungalow' categories also have relatively high median prices of 675,000.0 and 565,000.0, respectively.

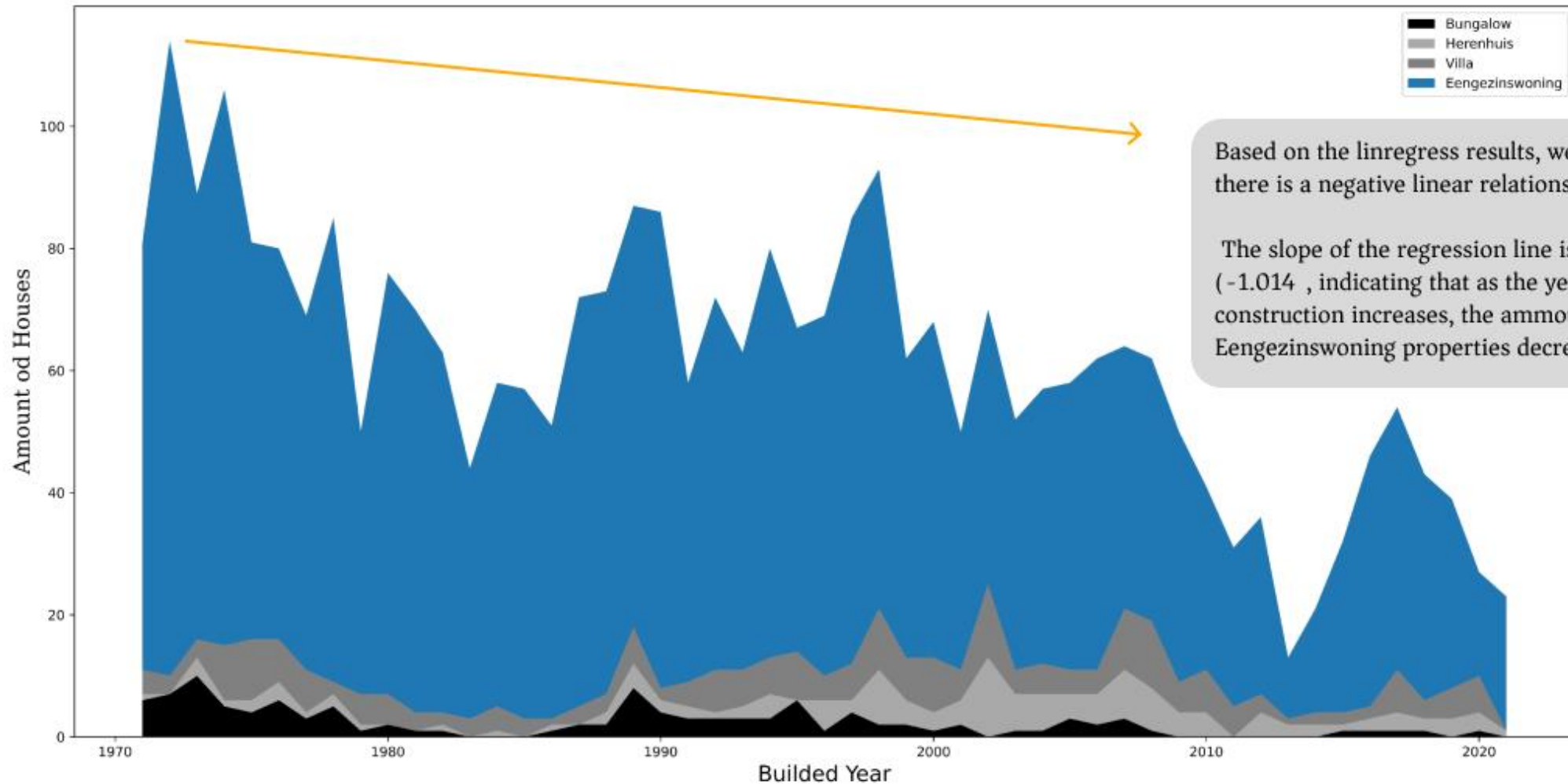
It's important to note that there is a large difference in the number of houses between the different categories, which could impact the results. For example, the 'Grachtenpand' and 'Woonboot' categories only have one house each, so their statistics may not be very representative of the population as a whole.

This information can be used to identify potential areas for further analysis and decision-making. For example, it may be worthwhile to investigate why 'Villa' houses are more expensive and whether this trend is consistent across different areas or regions. Additionally, the data could be further analyzed to identify any patterns or relationships between house type, price, and other variables, such as location or size.

	House_types	N	Mean	Median	Variance	SD	SE	95% Conf. Interval
1	Eengezinswoning	2630	4.723441e+05	439000.0	2.934514e+10	171304.2389	3340.3366	[465794.1941, 478894.104]
5	Villa	255	1.217320e+06	995000.0	4.374369e+11	661390.1459	41417.8575	[1135753.4951, 1298885.8068]
3	Herenhuis	141	7.070035e+05	675000.0	5.779665e+10	240409.3347	20246.1183	[666975.8826, 747031.2096]
0	Bungalow	114	6.050526e+05	565000.0	5.824615e+10	241342.3868	22603.7855	[560270.4557, 649834.8075]
4	Landhuis	49	1.225338e+06	975000.0	4.006864e+11	632997.9572	90428.2796	[1043519.9011, 1407156.3846]
6	Woonboerderij	22	9.793636e+05	990000.0	1.093793e+11	330725.3998	70510.8921	[832728.2089, 1125999.0639]
2	Grachtenpand	1	8.250000e+05	825000.0	NaN	NaN	NaN	[nan, nan]
7	Woonboot	1	6.250110e+05	625011.0	NaN	NaN	NaN	[nan, nan]

\* Based on the chart provided, it appears that there has been a **significant decrease in the number of "Eengezinswoning"** houses sold over time. This downward trend may suggest that fewer "Eengezinswoning" houses are being built which mean less general new buildings in market.

## Trend of House\_types over time since 1970



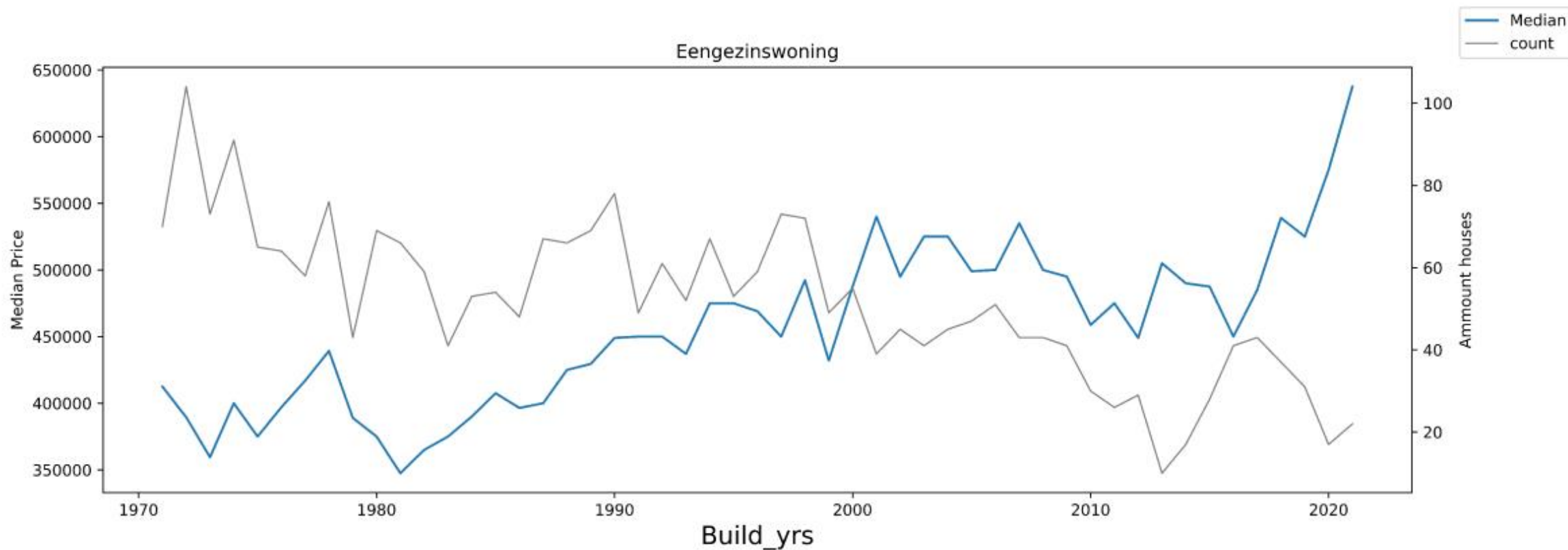
Based on the linregress results, we can see that there is a negative linear relationship .

The slope of the regression line is negative (-1.014 , indicating that as the year of construction increases, the ammount of Eengezinswoning properties decreases.



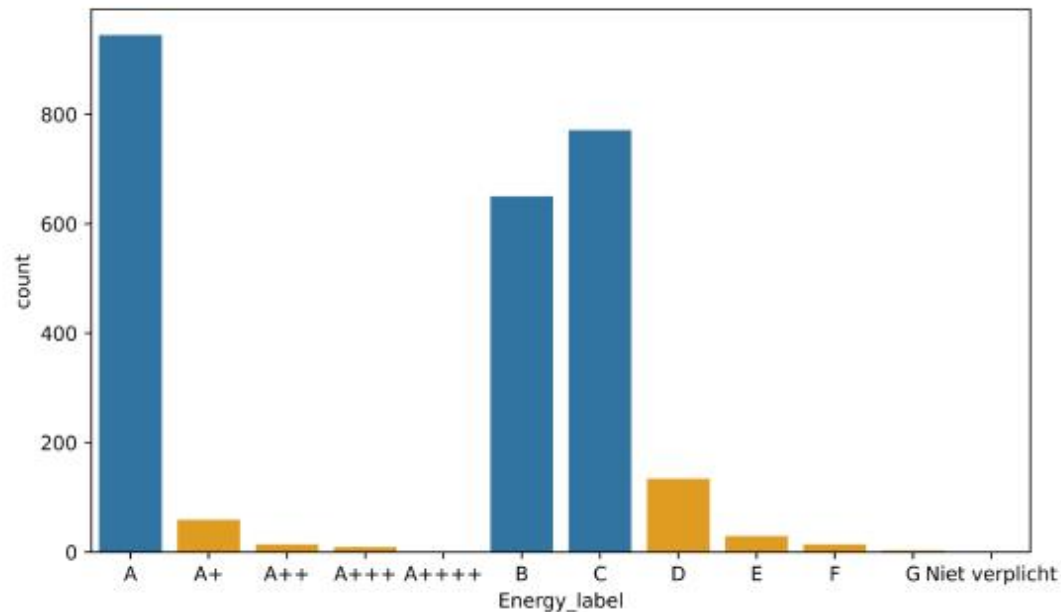
# Comparing Median and Total Amounts for Eengezinswoning

\* **Toper "Eengezinswoning"**: Based on the data in the chart, we can conclude that as the house age decreases, the number of Eengezinswoning houses also decreases. However, the median price for these houses tends to increase. This suggests that there may be a higher demand for newer Eengezinswoning houses, leading to an increase in their value. It is also possible that newer Eengezinswoning houses are built with more modern amenities and features, and in higher cost which can also contribute to their higher median price.

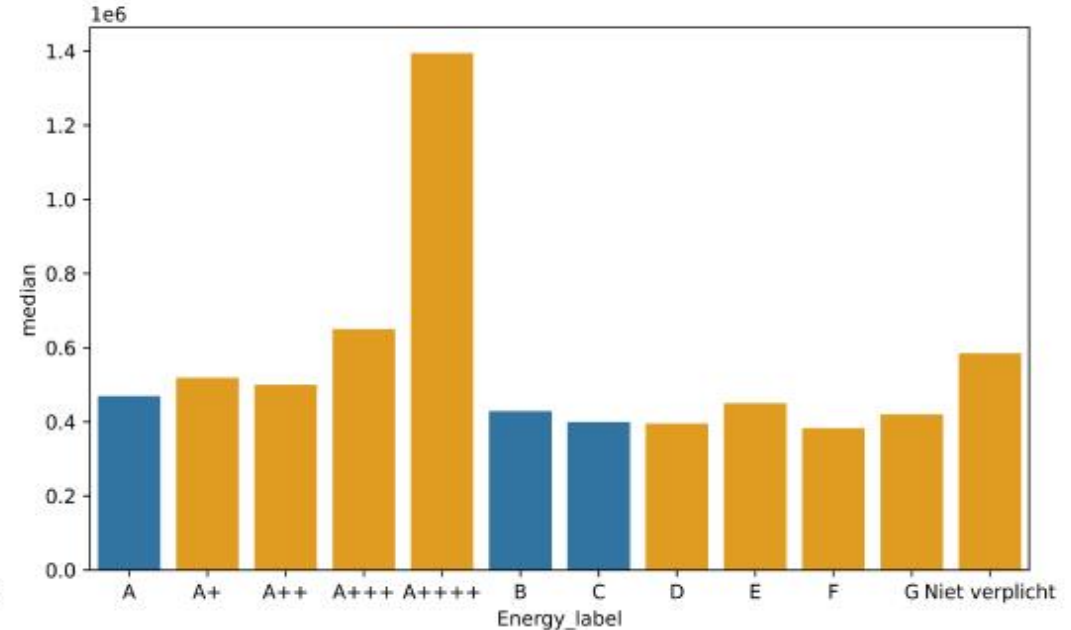


# Analyse Eengezinswoning by Energy\_label:

COUNT Prices in Energy Labels



MEDIAN Prices in Energy Labels



The top three Most frequent values in Energy\_label:

A with 35.93%  
C with 29.32%  
B with 24.71%

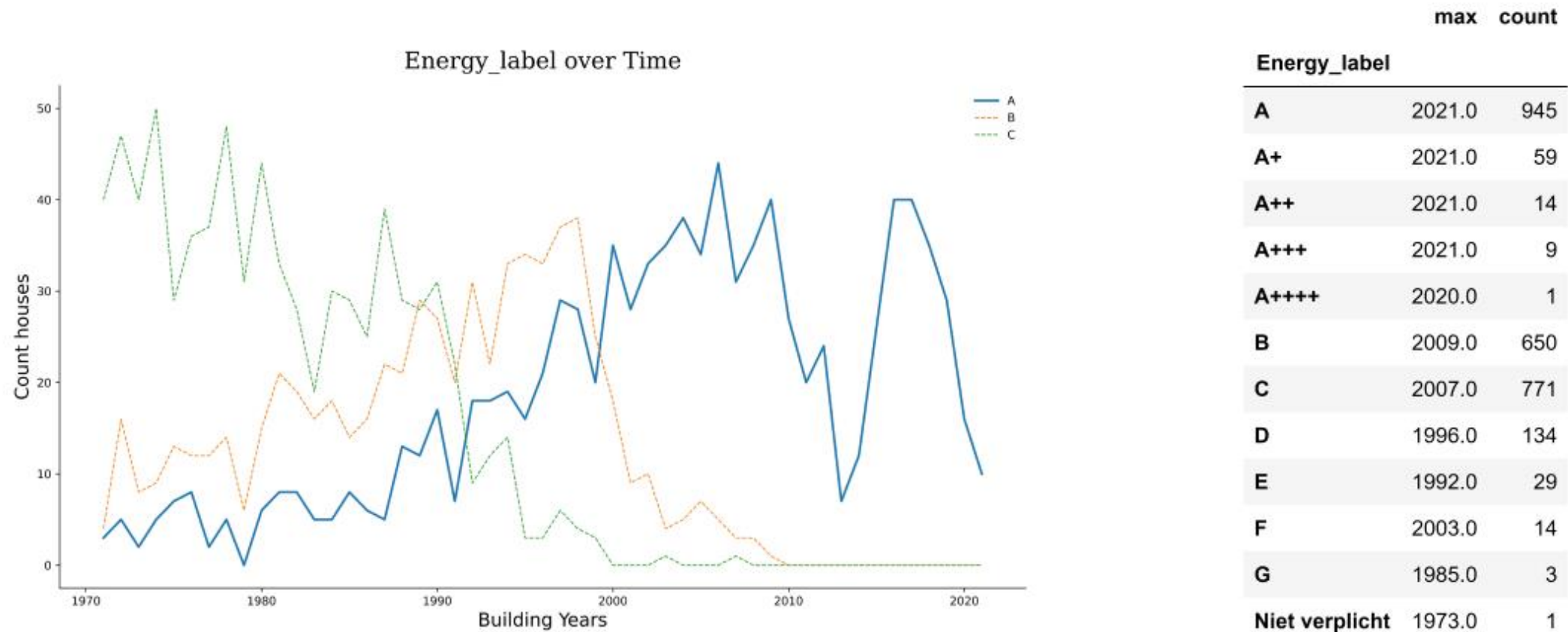


Overall, it seems that homes with better energy labels tend to have higher median prices, and homes with worse energy labels tend to have lower median prices. This is consistent with the idea that energy-efficient homes may be more valuable which is generally the most officiant and les energy consuming and desirable to buyers.

# Analyse Eengezinswoning Energy\_labels over time:

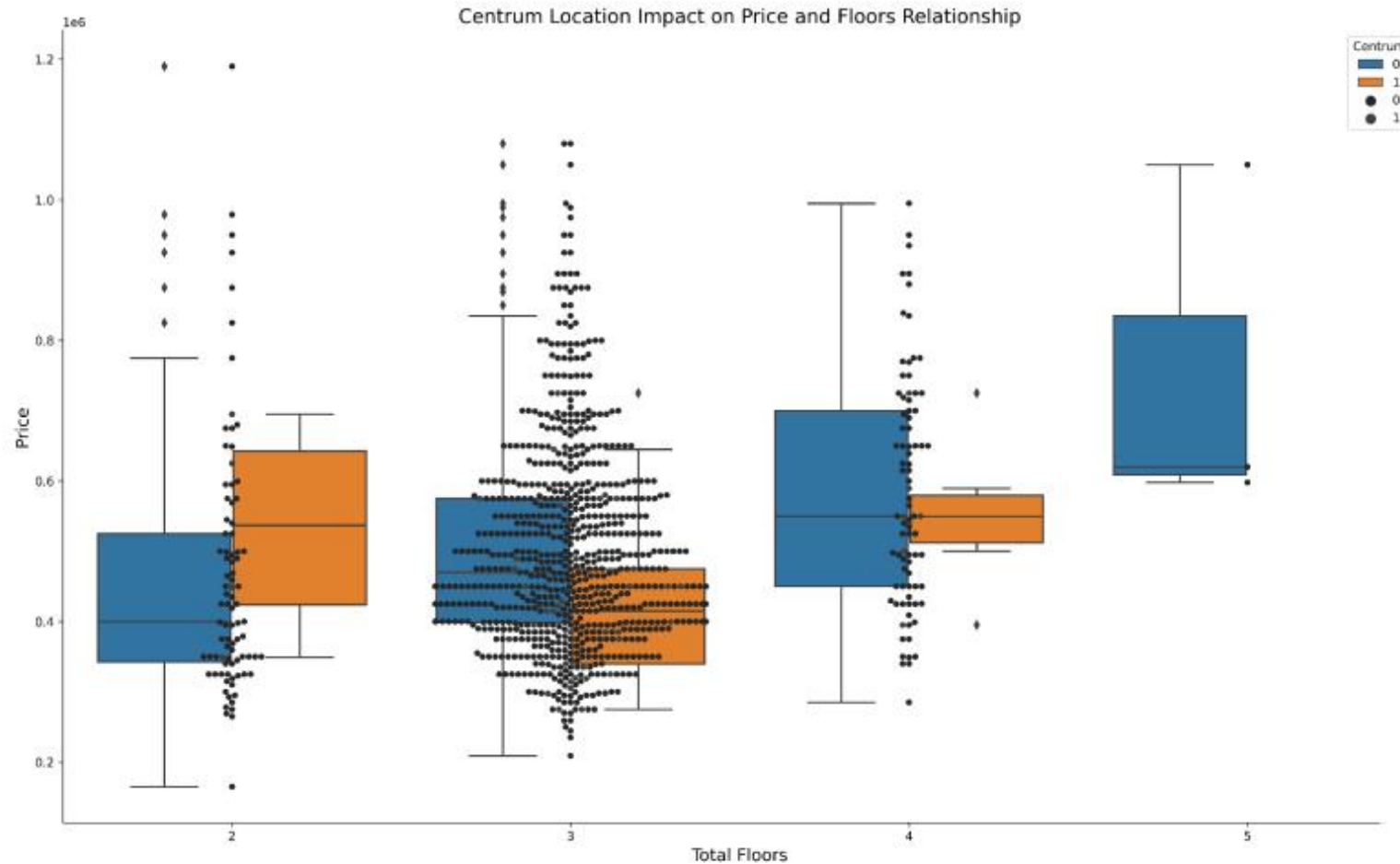
The majority "Eengezinswoning" switch between 1990-2000 from lower into a high energy labels (in this case, 'A', 'A+', 'A++', 'A+++', 'A++++', it likely means that there have been changes in building regulations or incentives for energy-efficient housing in Netherland. This can have a potential implications for the house market:

Higher demand for energy-efficient homes and Increased competition for older homes





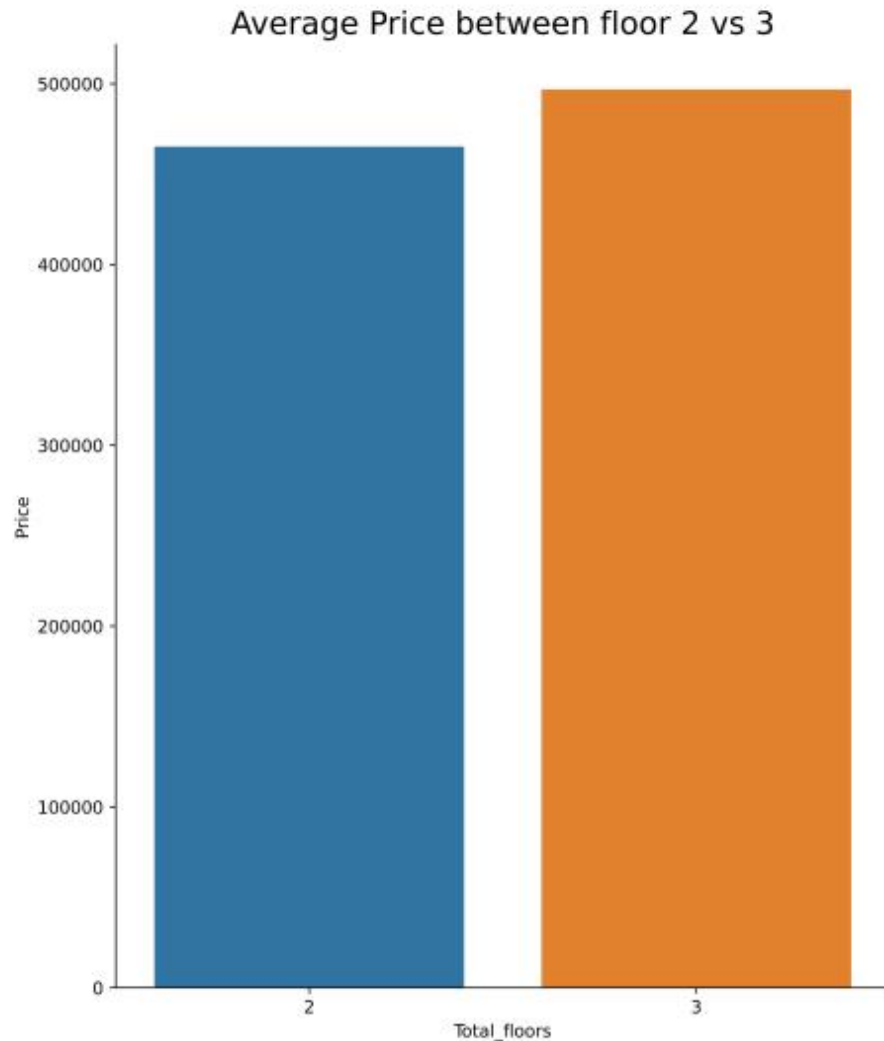
# Price and Floors Distribution by Centrum Location:



**81%** houses belong into house with 3 floors where **95%** of observations belong into houses out of centrum.



# Independent T-test Results Comparing Mean House Prices of 2-floor and 3-floor Eengezinswoning Houses with Energy Label 'A':



- The mean house price for houses with 3 floors is higher than those with 2 floors by €31,691.89.
- There is a statistically significant difference between the mean house prices of the two groups, with a p-value of 0.0588.
- However, we cannot reject the null hypothesis of equal means at the typical threshold of 0.05.
- The Cohen's d value is -0.2140, which indicates a small effect size.