Analysis of AMD’s and Intel’s Processors

Statistical Analysis performed by **Faycal Kilali** at **2023, April.**
This is a project based on the data of most **CPUs from AMD and Intel** released up to 2021, with predictions up to 2031.

Source of dataset: https://data.world/fiftin/intel-processors

The 4 datasets used contain information on all CPUs in the market, including specifications and benchmarks for each CPU. This data can be useful for understanding trends in the CPU market, comparing different CPUs, comparison of innovation in the CPU industry between different attributes and competitors, and to help with making informed purchasing decisions.
Although, the source states that the data was collected automatically, there may be certain outliers or *bugs* in the collection that resulted in data collection loss, which may bias the analysis performed in relation to the population of CPUs of AMD and Intel. The rest of the statistical analysis was performed by me as explained in the details and shown in the code included.

Explanation of columns and interpretations:

- Cores and threads: These columns represent the number of physical and logical cores in the CPU, respectively;
- Name: The name of the CPU model;
- Launch date: The date when the CPU was released;
- Lithography: The size of the transistors in the CPU, in nanometers;
- Bus_speed: The maximum speed at which the CPU can communicate with other components, measured in gigabytes per second;
- Base_frequency: The clock speed of the CPU under normal operating conditions, in gigahertz;
- Turbo_frequency: The maximum clock speed of the CPU when under heavy load, in gigahertz;

- Configurable_tdp_up_frequency: The frequency at which the CPU can operate with its configurable TDP up, in gigahertz;
- Cache_size: The amount of cache memory available on the CPU, in megabytes;
- TDP: The thermal design power of the CPU, in watts;
- Configurable_tdp_up: The maximum power consumption of the CPU when configured with TDP up, in watts;
- Price: The price of the CPU in US dollars;
- Product_line: The product line of the CPU (e.g. Core i9, Ryzen, and so on);
- Socket: The type of socket the CPU uses;
- Memory_type: The type of memory supported by the CPU;
- URL: The URL of the product page for the CPU;
- Vertical_segment: The market segment the CPU is intended for (e.g. consumer, server);
- Max_memory_size: The maximum amount of memory the CPU can support, in gigabytes;
- Status: The current status of the CPU (e.g. discontinued, available);
- Max_temp: The maximum temperature the CPU can operate at, in degrees Celsius;
- SKU: The stock-keeping unit (SKU) of the CPU;
- Package_size: The size of the CPU package, in millimeters;
- Fullname: The full name of the CPU, including the product line and model number.

Data cleaning

The data was cleaned due to missing values in a few select columns that we'd like to include in our analysis. The columns of the datasets that were cleaned are: base frequency, turbo frequency, tdp, max_temp, launch_date, cores, threads by dropping their **empty** rows, based on the analysis we are using. A few outliers in launch_date were also filtered for the AMD dataset as AMD was founded at 1985, up to today, 2023. Some outliers for some reason are outside that range, hence were excluded.

This was performed in order to include those attributes in our analysis.

Explanation of methods and tools to be used:

To answer the questions mentioned above, we will use various data visualization techniques such as line plots, violin plots, heat maps, bar plots, categorical plots, and any other plot that may improve our analysis. We will also use summary statistics to gain insights into the dataset as well as the correlation coefficients, and t-tests to compare means. The analysis will be performed through multiple libraries in Python.

All predictions will use linear regression, specifically, the least squares method for this project.

Here are the inquiries we'd like to explore

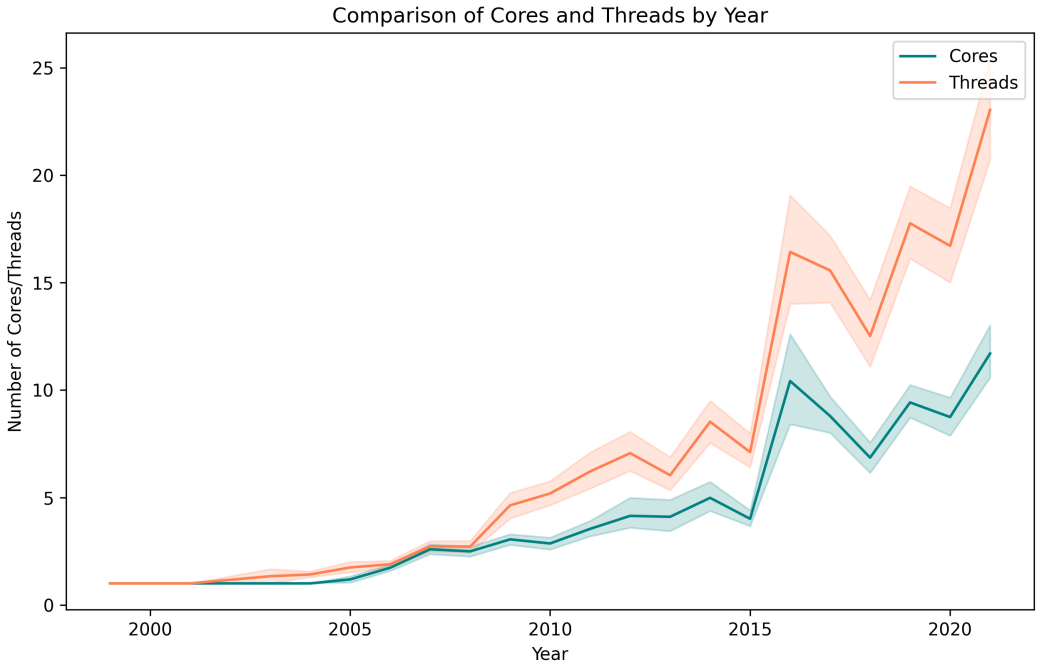
The following inquiries are intended to provide insights into the connections among the attributes of the CPU, which is the central component of any computer hardware. These insights can be valuable for any work involving computerized hardware. Moreover, the findings may help make better-informed decisions when acquiring hardware and predicting potential changes in the CPU's attributes in the future.

- How do the number of cores and number of threads vary across the years?
- How do the base clocks and turbo clocks vary across the years?
- How does TDP and temperature vary across the years for all CPUs?
- How do AMD and Intel compare in terms of these specifications?
- Are there any notable spikes or dips in these specifications at certain points in time?
- Are there any notable trends or changes over time that are unique to either company?
  Are there any significant differences between the two companies in terms of these specifications?
- What is the correlation between TDP and temperature, both in general and between the two companies?
- Is there a connection between cores, base clocks, number of threads, TDP, and temperature? If so, what is the correlation between them?
- Based on the data, what can we predict to be the means of the number of CPUs/Threads, and Max Temperature / TDP, which shows the efficiency of a unit across the period 2021-2031?

TDP & Temperature

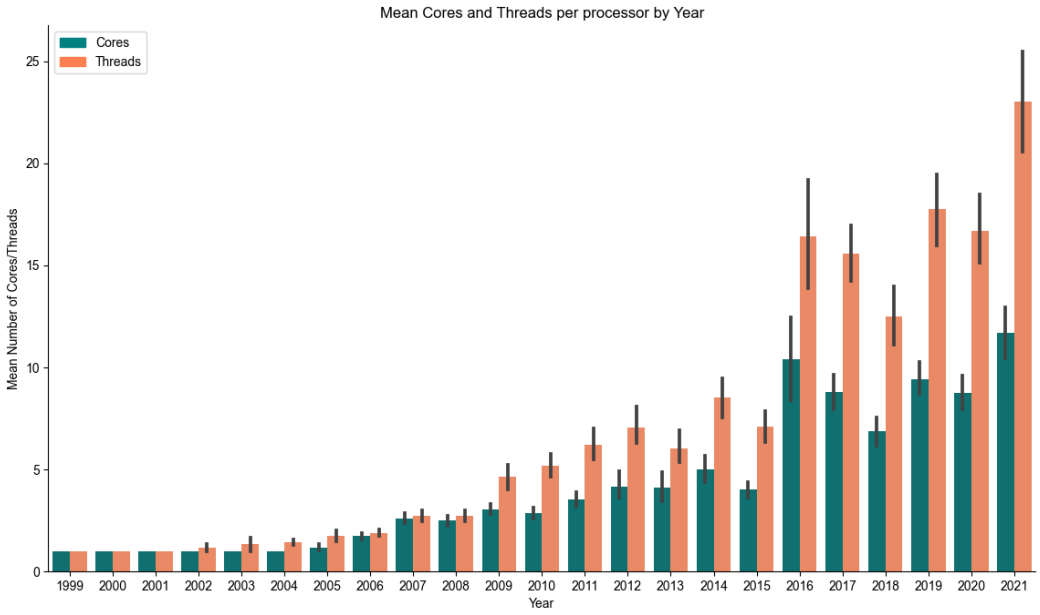
- What is the overall distribution of TDP (thermal design power) and temperature values across all CPUs in our analysis?
- How does the distribution of TDP and temperature values compare between AMD and Intel CPUs?
- Are there any significant differences in TDP or temperature values between the two companies across the years?

Analysis



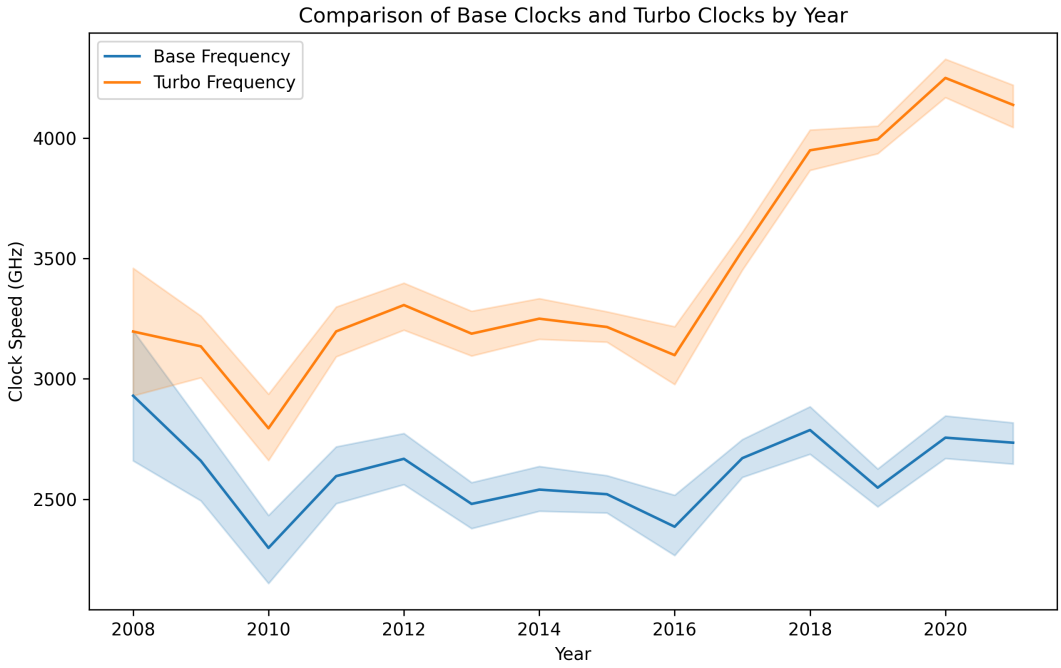
- We can see that the general trend for Cores and Threads is increasing faster as time passes by, and so is the variance between the years;
- The number of threads increases even faster than the number of cores, suggesting that newer CPUs are more likely to be multi-threaded than previous generations, with more than two threads per CPU;

- There is a significant jump in number of cores and threads around the year 2015-2020, suggesting a technological innovation has occurred that resulted in a large jump in the number of cores and threads manufactured within those years having more cores and threads per CPU unit;
- From the earlier years, the majority of the CPUs were single-threaded, which is why the curves of the number of cores and threads is almost coinciding.

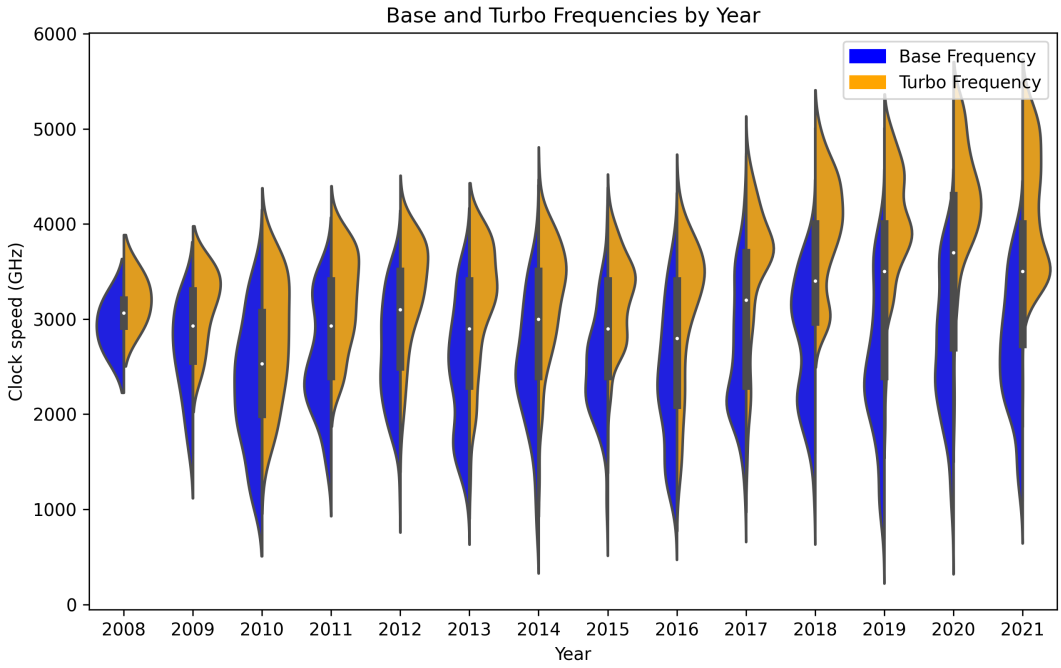


Distribution wise, we can see some of our observations even more clearly from this plot. Specifically, the number of cores and threads coincides up until 2002, where the shift happens. Suggesting an innovation in the CPU Industry, making the consumer CPUs in the dataset contain more than one thread per core. Additionally, We see the extreme rise in the threads along the years 2016 to 2021, suggesting a further push towards more threads per core.

- Interestingly, the number of mean cores per CPU produced in 2016 is more than the next 4 years up until 2021. Suggesting that innovation was focused on other areas in the CPU industry during that period which likely resulted in a pull-back of number of cores per CPU.

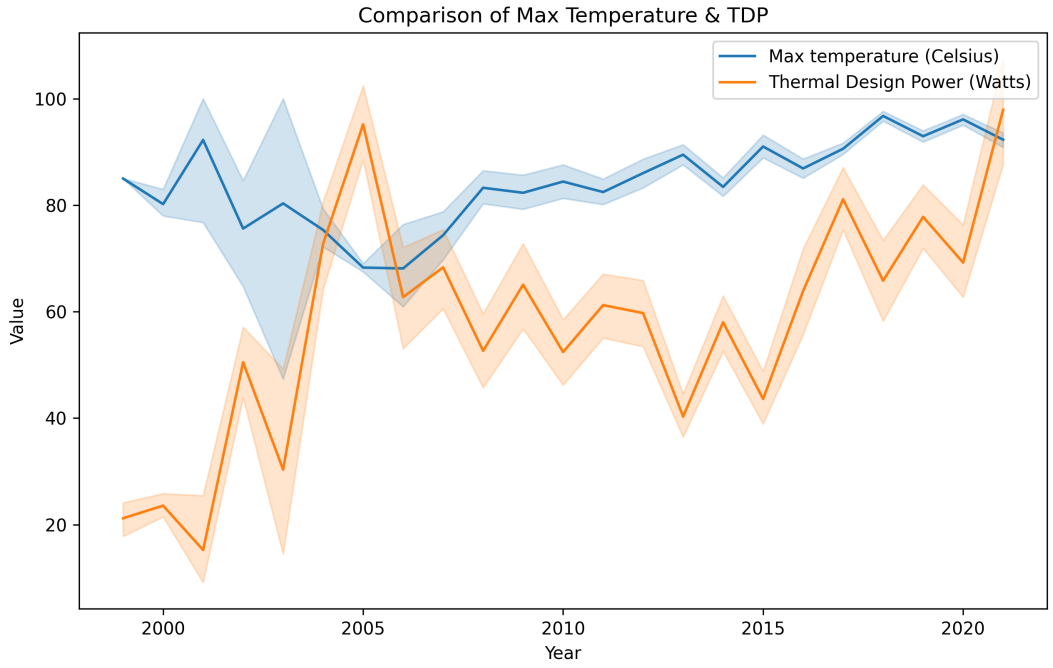


- The standard deviation for both the base frequency and turbo frequency decreases as the years pass by, suggesting continuous development in the industry for more predictable clock speeds;
- There was an interesting dip at the year 2010 for both the base frequencies and turbo frequencies of the CPUs manufactured during that period;
- The overall trend is increasing, with significant increases for the turbo frequencies in comparison to the base frequencies as the years pass by.



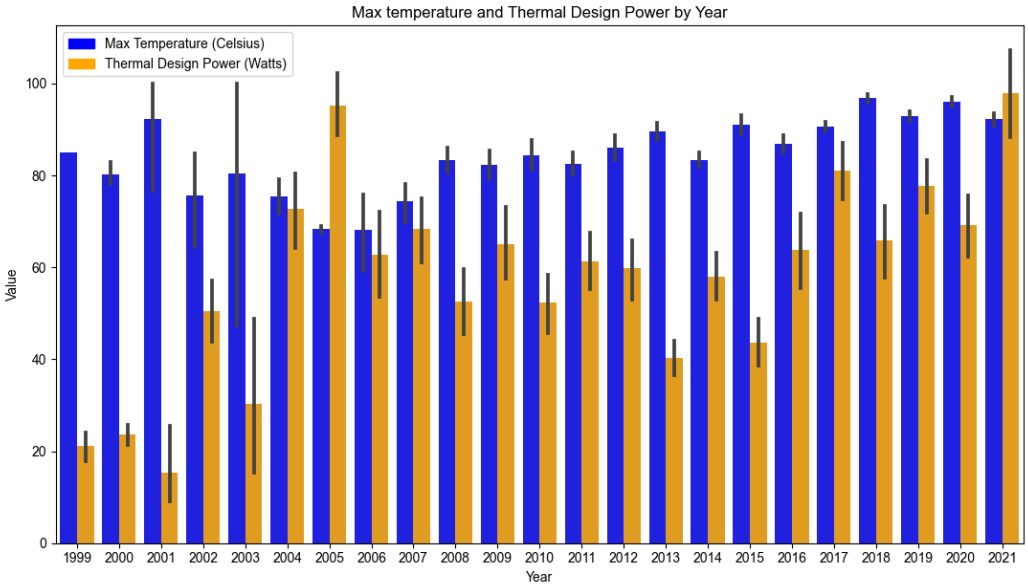
We can see the distributions of the frequencies between the base frequencies and turbo frequencies are deviating from the mean in opposite directions in terms of the magnitude of their difference as the year passes. That is, turbo frequency clock speeds are increasing faster than base frequency clock speeds as the years pass by.

The overall trend is increasing, with generally the mean being higher than the previous per year. There is an interesting dip from 2009 to 2010, suggesting that the innovation resulted in a greater disparity between the frequency ranges for both turbo and base frequencies, but with the mean overall actually ending up *lower* than all the other years within the period.



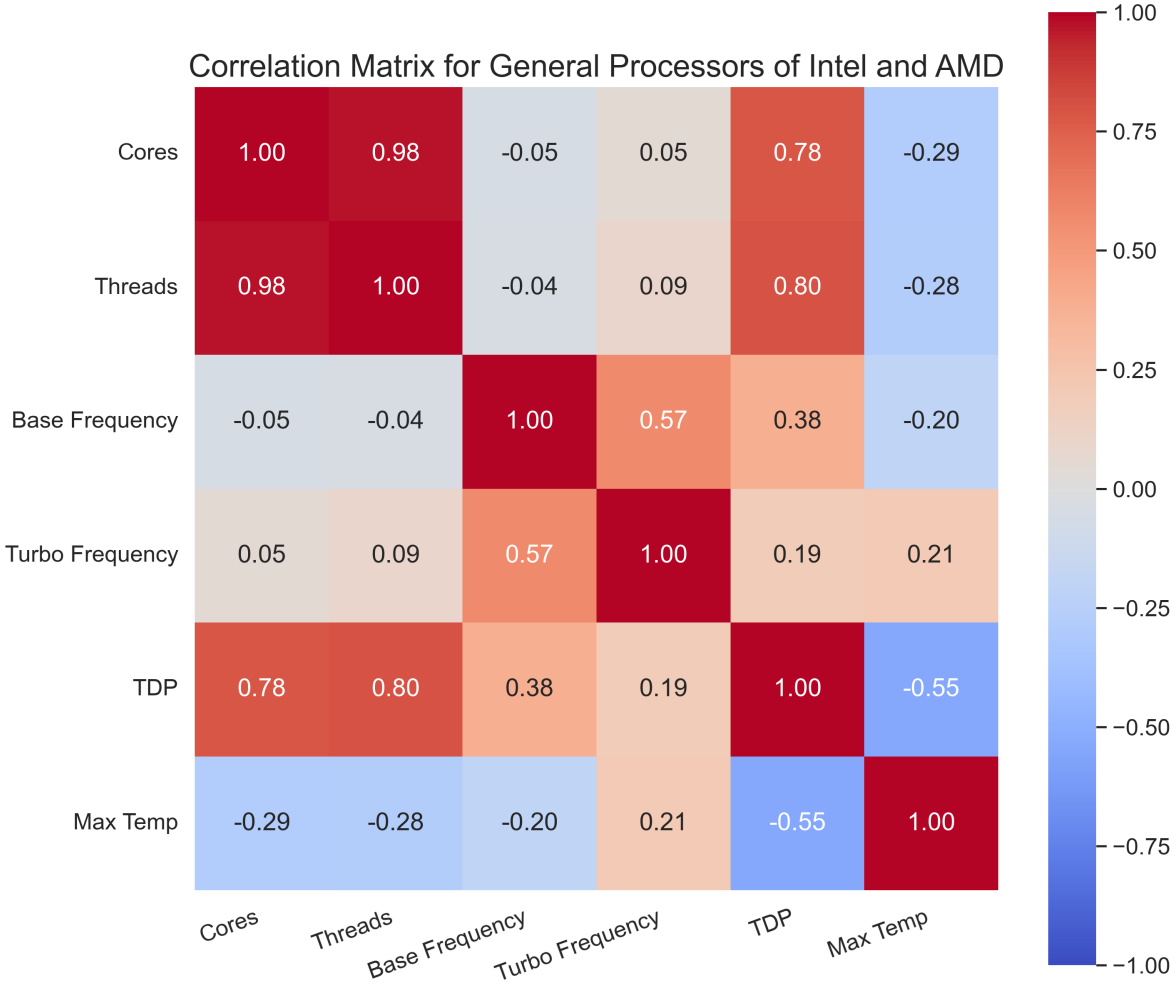
Generally, the heat output of a CPU has been increasing over the years, requiring appropriate coolants and other innovations for cooling to maintain optimal functionality. The biggest jump occurred at 2005 for the TDP, suggesting the CPUs manufactured around that period produces an immense amount of heat in comparison to their counterparts over the next few years and over the preceding few years.

The max temperature overall stayed relatively stable, and both the TDP and Max Temperature of the CPUs only intersected three times across the years. Twice around 2003-2007 and once around 2020, with the TDP overtaking the max temperature, requiring even stronger cooling conditions in order to prevent complete shutdown of the CPUs produced in that period.



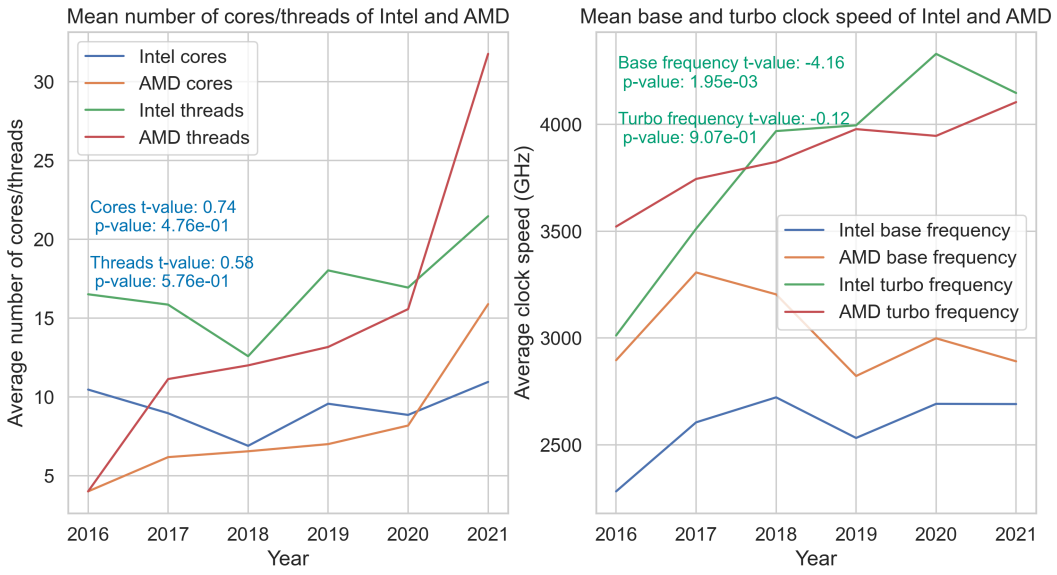
We can see that the majority of the CPUs around 1999 to 2001 had a low TDP to max operational temperature ratio, with a massive shoot up from 2004 and 2005 in terms of TDP to Max Temperature, where the first overtake happened at 2005. The general trend is an overall increasing TDP as the years progress and a rather steady max temperature per CPU unit.

This suggests that the CPUs' manufactured across the years generate more heat and are more likely to reach their maximum temperature barring advances in coolants. With the newest generations likely requiring the newer, more innovative coolant methods in comparison to the previous years to be operational.



- Cores and Threads are very strongly positively correlated through a linear relationship, as we can see with the correlation coefficient of 0.98, suggesting that the more cores there are, the more threads there will be and vice versa in most cases;
- The relationship between Cores/Threads with TDP is also very strongly positively correlated, suggesting that as more cores/threads are manufactured in a unit CPU, the TDP generally increases at a very strong correlation rate and vice versa;

- The max temperature is negatively correlated to Cores and Threads, that suggests the more cores and threads there are, the lower the max operational temperature will be for a unit CPU;
- There is a lightly weaker correlation, but still a positive correlation, between the frequencies and the TDP, whereas the higher the clock speeds the more heat is generated by the CPU;
- There is a moderately strong positive linear relationship between base frequency and turbo frequency of a rate of 0.57, suggesting a moderately strong positive linear relationship.



When it comes to the cores and threads, the general trend is an overall mostly linear increase in the number of cores and threads for both AMD and Intel, with a significant trend change at the year 2021 where AMD threads undergo exponential increase.

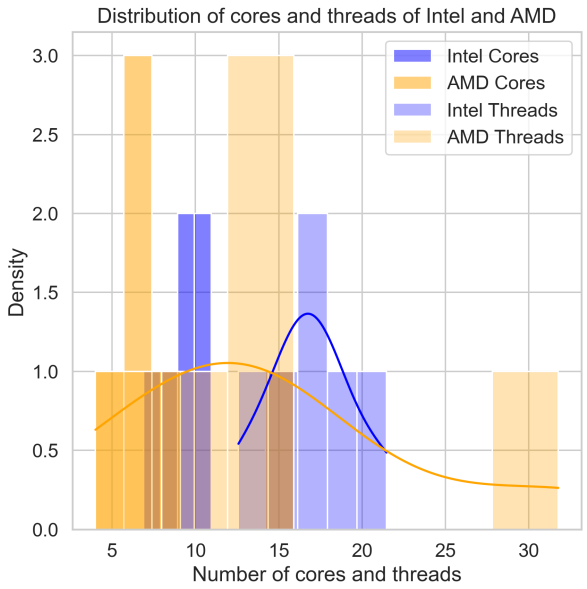
There is an overall larger amount of cores for Intel up until 2021 per unit CPU, where AMD overtakes Intel. Likewise, for the number of threads per unit CPU.

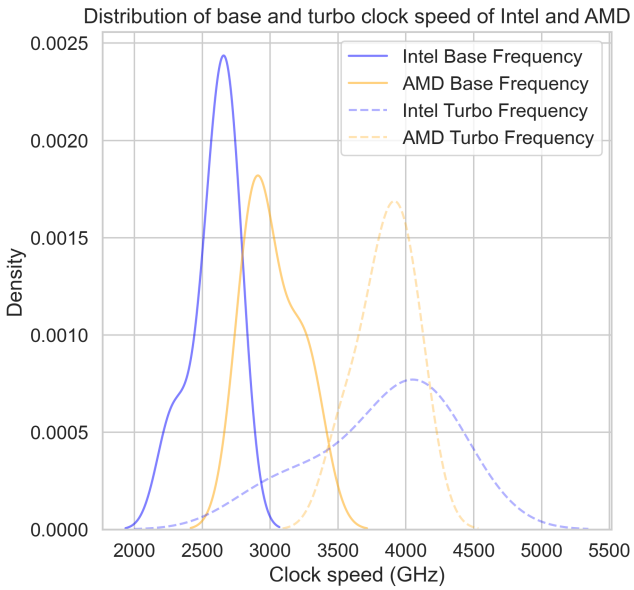
The t-values and their associated p-values for the Cores and Threads suggest there is no significant difference between the two attributes being compared of the two companies. Furthermore, the p-values are higher than the common threshold of 0.05, which suggests that the difference in means occurred due to chance alone.

For the frequencies, the trend for the turbo frequencies appears to be mostly an almost linear increase, and for the base frequencies it oscillates throughout the period.

Intel mostly has a lower base frequency than AMD but the trend suggests that they will likely intersect at some point in the future as AMD's base frequency appears to be decreasing throughout the years. Additionally, the turbo frequencies of Intel have overtaken AMD over time, up until 2021, where they intersect once more. Suggesting that the two companies in terms of innovative advantage are reaching an *equilibrium*.

The t-value comparison suggests that the base frequencies have a significant difference and that it is not due to chance. Whilst for the turbo frequency, the t-test suggests that there is no significant difference between the two companies in terms of turbo frequency, and any differences are likely due to chance.

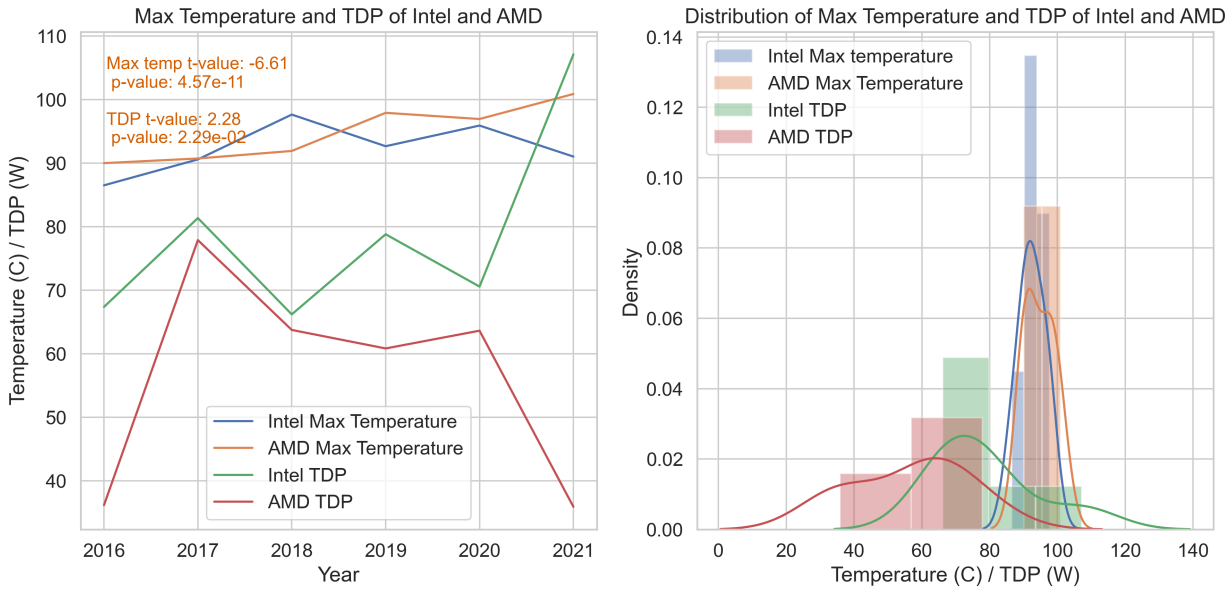




The distribution of the number of cores and threads shows that the majority of AMD cores and threads lay around the 10-15 region, and the majority of the Intel ones lay about the 15-20 region. Additionally, the density for the AMD cores and threads is mostly around 8 and 12-16, respectively. Whilst for Intel its 10 and between 15-22.

There is an outlier for both companies, with none of the CPUs analyzed having cores/threads between 23 and 27 due to no density there.

The distribution of their base and turbo frequencies shows that the Intel base frequency mostly lays about 2500, and ranges from 2000 to 3000. Whilst for AMD its mostly at 3000 and ranges between 2500 to 3500.
Additionally, Intel's turbo frequencies are mostly around 4000 and ranging from 2500 to 5500, with AMD ranging from 3000 to 4500 for their turbo frequencies, a much more limited range. This suggests that AMD has a finer range for their frequencies of their units.



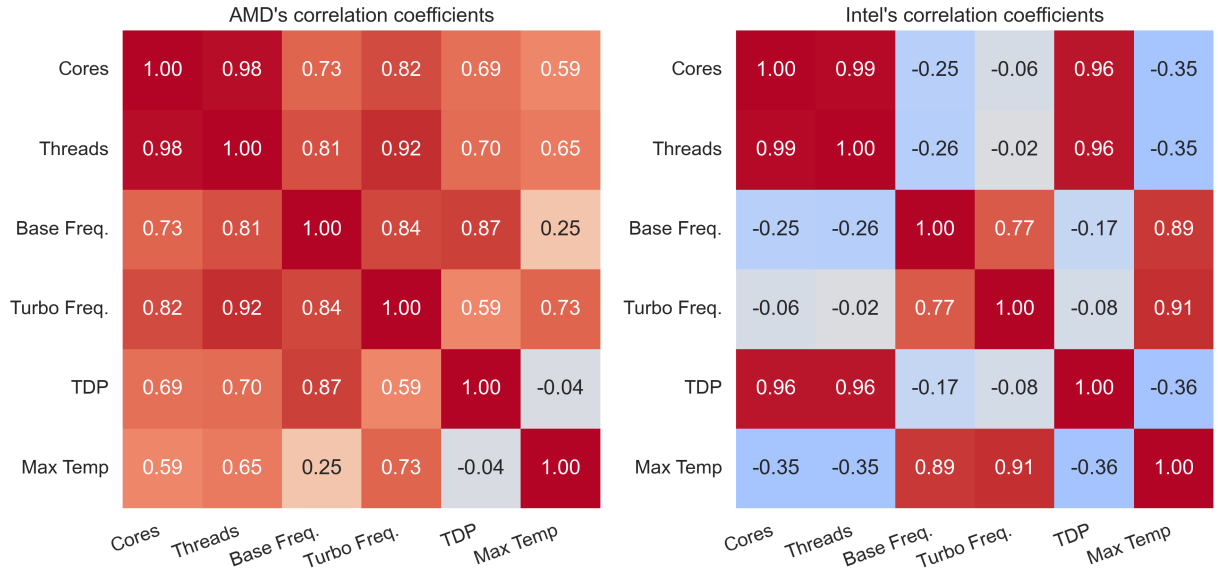
The overall trend shows that the max temperature is mostly increasing linearly for both companies with the TDP for AMD mostly oscillating up until 2021 where it dips heavily, and for Intel instead it shoots up significantly at 2021.

The t-test for max temperature suggests that there is a very strong and statistically significant relationship between the max temperatures' of AMD and Intel that are being compared. The relationship appears to be an inverse relationship.
The t-test for TDP suggests a strong statistically significant relationship between the TDP of Intel and AMD, and that the relationship is a positive relationship. However, this is not as strong as the max temperature relationship.

The distribution of Max Temperature and TDP of Intel and AMD shows us that the majority of Intel max temperature falls at around 90 Celsius with about 14% of their CPUs at this value, and the rest ranging from 80 Celsius to about 100 Celsius. AMD has

a very similar curve, slightly shifted to the right, starting from about 82 Celsius and ranging to about 105 Celsius, mostly dense around the 90 to 93 Celsius range.

The TDP for Intel mostly falls between 60 to 80 watts, but ranges from 40 to 140. Whilst for AMD it mostly falls around 30 to 80, and ranges from about 1 to 100 for most of its density.



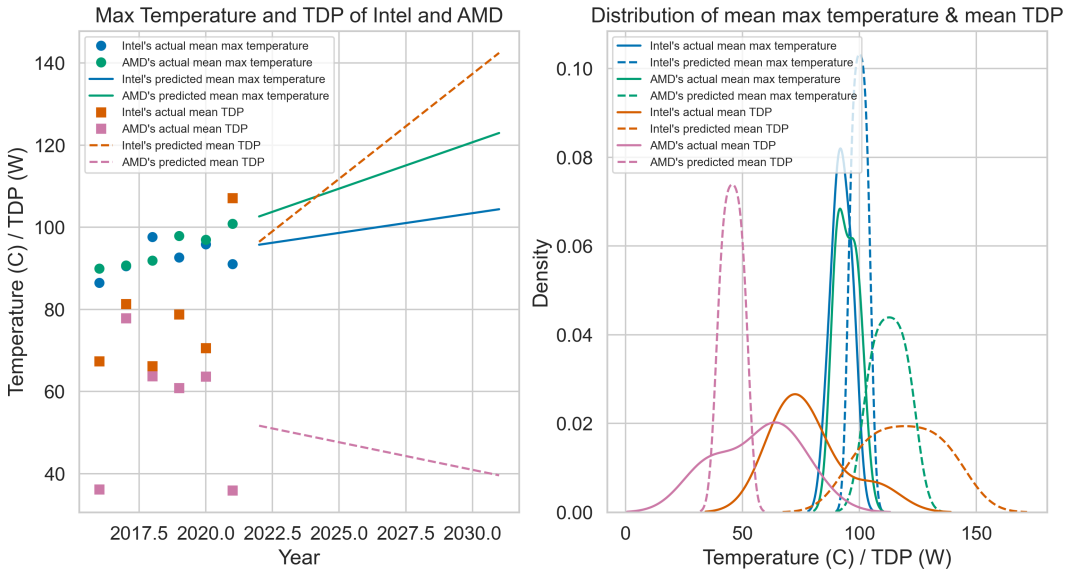
Interestingly enough, the correlations for Intel are very similar to the general processor correlations that we analyzed earlier, while for AMD there are some significant differences. Because the Intel analysis here would be essentially very similar (but just a lot stronger) in comparison to the general analysis we performed earlier, we'll leave it to the reader to compare.
There's one thing to note though before I do leave it to the reader. Observe that, in general, there is a negative correlation between the frequencies and the TDP, suggesting that as frequencies increase the heat generation dissipates which suggests that Intel may be innovating their heat dissipation technologies at a faster rate than AMD.

For the AMD analysis, there are some very significant correlations to observe. The first observation is that the frequencies and number of cores/threads are actually very strongly positively correlated rather than negatively correlated. Additionally, the heat

generation and the frequencies is also very strongly correlated, much stronger than the general processors.

Predictions

The following predictions are intended to attempt to predict what the future will hold in terms of some of those attributes, trend wise.



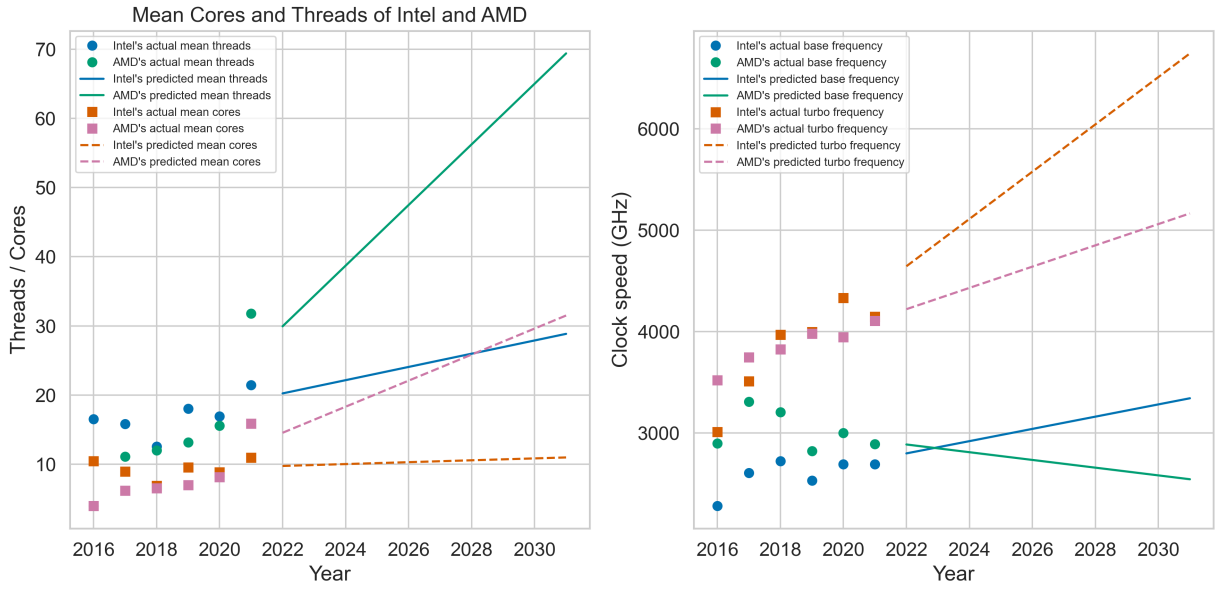
Our modelling shows us that Intel's main TDP will increase to about 140 W by the year 2030, and in the other hand, AMD's TDP will actually drop to about 40 W by 2030. Suggesting that the two companies are going the opposite directions in terms of their priorities of innovating heat generation.

The predicted max temperatures are both positively trending upwards, with AMD trending upwards faster than Intel regardless of its TDP trending downwards which is contradictive in most cases as we saw for AMD the correlation between TDP and max Temperature is very strongly positively correlated.

Distribution wise, our model predicts that AMD's curve will shift from a much wider range covering 0 to 100 W to a much narrower range to the left, from about 40 to 60 W. Whilst for Intel, the shift in TDP is towards the right — towards a higher watt consumption — from a range of about 40 to 130 to a range of about 80 to 160 W. The curve also gets smoother with less jumps.

For the max operational temperatures, Intel’s curve shifts slightly to the right to a new range of about 90 to 120 Celsius, and AMD shifts to the right even more to a new range of about 90 to 140 Celsius.

Overall, the trends are pretty clear in the predicted changes, with only a single contradiction that requires further analysis.



Our prediction shows that AMD’s mean threads is increasingly with a fairly positive slope, giving an estimate of roughly 65 threads by the year 2030. Whilst for Intel, the increase is far more conservative — about 28 threads per CPU unit produced in 2030 for Intel.

The mean cores are again, of a similar trend. With AMD having about 30 cores per CPU unit manufactured in the year 2031, while for Intel its roughly only about 11 manufactured in average per CPU unit in 2031.

Those trends are likely exaggerated significantly, it is likely that the connection in the predicted dates is somewhere in between Intel’s conservative prediction and AMD’s prediction that our model is producing.

For the frequencies, AMD has significant increases in Turbo Frequency predicted per year, with up to 6500 GHz by the year 2031. For Intel its roughly 5200 GHz, both of which are quite sensible. The base frequencies are more interesting to look at, the prediction is that AMD’s base frequency will drop lower than Intel’s by the year 2031, to

around 2500 GHz, while Intel will be about 3200 GHz, showing that the current innovations of Intel favor Base Frequencies while AMD's favor Turbo Frequencies.

Conclusions and Summary

In *summary*, there has been a significant increase in the number of cores and threads over the years, with more multi-threaded CPUs being produced in recent years. The innovation that led to this increase occurred around 2015-2020. The standard deviation for base and turbo frequencies decreased over time, suggesting an industry focus on more predictable clock speeds. The heat output has been increasing over time, with the biggest jump occurring in 2005. The relationships between the number of cores/threads and TDP are strongly positive, while the relationship between the number of cores/threads and max temperature is negative. The relationship between frequencies and TDP is also positive, and the base frequency and turbo frequency are moderately positively correlated. AMD and Intel have had similar trends in the number of cores and threads produced over time, with a significant trend change in 2021 when AMD threads underwent exponential increase. Overall, newer CPUs generate more heat and require better cooling mechanisms.

Conclusions

- **How do the number of cores and number of threads vary across the years?**

  The number of cores and threads in CPUs has been increasing over time, with the variance between the years increasing as well. The number of threads has been increasing faster than the number of cores, indicating that newer CPUs are more likely to be multi-threaded than previous generations, with more than two threads per CPU. There is a significant jump in the number of cores and threads around the year 2015-2020, suggesting a technological innovation has occurred that resulted in a large jump in the number of cores and threads manufactured within those years having more cores and threads per CPU unit. From the earlier years, the majority of the CPUs were single-threaded, which is why the curves of the number of cores and threads is almost coinciding.

- **How do the base clocks and turbo clocks vary across the years?**

  The overall trend is increasing for both base and turbo clocks, with significant increases for the turbo frequencies in comparison to the base frequencies as the

years pass by. Interestingly, the standard deviation for both the base frequency and turbo frequency decreases as the years pass by, suggesting continuous development in the industry for more predictable clock speeds. There was an interesting dip at the year 2010 for both the base frequencies and turbo frequencies of the CPUs manufactured during that period.

- **How does TDP and temperature vary across the years for all CPUs?**

  The TDP of CPUs has been increasing over time, requiring appropriate coolants and other innovations for cooling to maintain optimal functionality. The biggest jump occurred at 2005 for the TDP, suggesting the CPUs manufactured around that period produces an immense amount of heat in comparison to their counterparts over the next few years and over the preceding few years. The max temperature overall stayed relatively stable, and both the TDP and max temperature of the CPUs only intersected three times across the years. The general trend is an overall increasing TDP as the years progress and a rather steady max temperature per CPU unit.

- **How do AMD and Intel compare in terms of these specifications?**

  AMD and Intel show some similarities and differences in terms of specifications. Both companies generally follow the trend of increasing cores, threads, base and turbo clocks, and TDP over time. However, there are some differences in the magnitude of these increases, with AMD showing a more significant increase in the number of cores and threads in comparison to Intel. In terms of clock speed, Intel CPUs generally have higher base clocks, while AMD CPUs generally have higher turbo clocks. There are also some differences in the variance of these specifications, with AMD CPUs generally having a higher variance in the number of cores and threads, and Intel CPUs having a higher variance in clock speeds.

- **Are there any notable spikes or dips in these specifications at certain points in time?**

  There are some notable spikes and dips in these specifications at certain points in time. For example, there is a significant jump in the number of cores and threads around the year 2015-2020. Additionally, there was an interesting dip at the year 2010 for both the base frequencies and turbo frequencies of the CPUs manufactured during that period. The biggest jump in TDP occurred in 2005,

suggesting the CPUs manufactured around that period produced an immense amount of heat relative to that period of time.

- **Are there any notable trends or changes over time that are unique to either company?**

  There are some trends or changes over time that are unique to either company. AMD generally shows a more significant increase in the number of cores and threads in comparison to Intel, while Intel CPUs generally have higher base clocks, and AMD CPUs generally have higher turbo clocks. AMD CPUs also generally have a higher variance in the number of cores and threads, while Intel CPUs have a higher variance in clock speeds.

- **Are there any significant differences between the two companies in terms of these specifications?**

  Yes, there are some significant differences between the two companies in terms of the specifications analyzed.

- **What is the correlation between TDP and temperature, both in general and between the two companies?**

  As the TDP increases, so does the heat output of the CPU, which can lead to an increase in the max temperature of the CPU if proper cooling methods are not in place.

- **Is there a connection between cores, base clocks, number of threads, TDP, and temperature? If so, what is the correlation between them?**

  Yes, there is a connection between cores, base clocks, number of threads, TDP, and temperature. As mentioned in the analysis, cores and threads are strongly positively correlated, as are cores/threads and TDP. The max temperature is negatively correlated to cores and threads, suggesting that the more cores and threads there are, the lower the max operational temperature will be for a unit CPU. There is a lightly weaker positive correlation between the frequencies and the TDP, whereas the higher the clock speeds, the more heat is generated by the CPU.

- **Based on the data, what can we predict to be the means of the number of CPUs/Threads, and Max Temperature / TDP, which shows the efficiency of a unit across the period 2021-2031?**

The analysis of the dataset using linear regression allowed us to predict the means of the number of CPUs/Threads and Max Temperature/TDP for the period 2021-2031. Which is, again, detailed in the analysis section.

- **What is the overall distribution of TDP (thermal design power) and temperature values across all CPUs in our analysis?**

  The analysis mentions that the overall trend is an overall increasing TDP as the years progress and a rather steady max temperature per CPU unit. The majority of the CPUs around 1999 to 2001 had a low TDP to max operational temperature ratio, with a massive shoot up from 2004 and 2005 in terms of TDP to Max Temperature, where the first overtake happened at 2005. This suggests that the CPUs manufactured across the years generate more heat and are more likely to reach their maximum temperature barring advances in coolants.

- **How does the distribution of TDP and temperature values compare between AMD and Intel CPUs?**

  The analysis we've performed details those aspects.

- **Are there any significant differences in TDP or temperature values between the two companies across the years?**

  Yes, there are differences in TDP and temperature values between the two companies across the years. The analysis we've performed shows that there are differences in the number of cores and threads, clock speeds, TDP, and max temperature between the two companies across the years.