Capstone III Project Proposal: Moore’s Law

By Sean Dickerson

1. **What dataset or datasets do you plan to use? What are the features, rows, and data types of each?**

I plan on using a CSV dataset from Kaggle called *“CPU and GPU Performances Dataset”*. The file contains various GPUs, CPUs, and their specifications such as transistor count or frequencies. Listed below are the columns along with their variable types.

|  |  |  |
| --- | --- | --- |
| **Variable** | **Quantitative or qualitative** | **Variable type** |
| Product | Qualitative | Nominal |
| Type | Qualitative | Binary |
| Release Date | Quantitative | Continuous |
| Process Size (nm) | Quantitative | Continuous |
| TDP (W) | Quantitative | Continuous |
| Die Size (mm2) | Quantitative | Continuous |
| Transistors (million) | Quantitative | Continuous |
| Frequency (MHz) | Quantitative | Continuous |
| Foundry | Qualitative | Nominal |
| Vendor | Qualitative | Nominal |
| FP16 GFLOPS | Quantitative | Continuous |
| FP32 GFLOPS | Quantitative | Continuous |
| FP64 GFLOPS | Quantitative | Continuous |

1. **What research or business questions do you want to answer?**

Moore’s Law states that roughly every two years, the amount of transistors on a microchip doubles. First inquiry I want to explore is if there’s any correlation between frequency and transistor count.

Second, is there any correlation between frequency and GLFOP counts?

1. **What are your hypotheses going in?**

My first hypothesis is there is indeed a correlation between frequency and transistor count. Higher transistors would allow chips to run at faster frequencies more efficiently which would allow for faster performance. I’m expecting correlation coefficients of at least 1 to confirm the results.

My second hypothesis is that there isn’t a significant difference in GFLOPs depending on what the frequency is used for the chip. I’m conducting my testing at a 90% significance level.

1. **How will you use your data to test your hypotheses?**

Scatterplots and Pearson correlation coefficient will be used for checking the correlation. I’m going to average the frequency clocks if various chips are using the same amount of transistors and conduct a t-test for the samples.

1. **Who will find your findings valuable, and how will they use them?**

Businesses and customers that are purchasing components will want to see how long they can expect their hardware to last before needing to upgrade. They should also find the results of the frequency/GFLOPs test useful for better purchasing decisions when looking at raw spec sheets.