Decoding the Future of the Semiconductor Industry Through an Era of Innovation and Global Risks

Ryan Reiser

CSCI 303: Introduction to Data Science

Instructed by Professor Morgan Cox and

Professor Wendy Fisher

Colorado School of Mines

Golden, CO USA

rreiser@mines.edu

November 26th, 2024

Project Abstract

This report presents a comprehensive analysis of the semiconductor industry, analyzing data mostly within technological progress. This report uses the following insightful and highly related Kaggle datasets to explore the industry's current dynamics and predicts its trajectory: "AI Global Index" from Kateryna Meleshenko on Kaggle [1], "Semiconductor Shortage" by Ram Jas on Kaggle [2], "Nvidia Daily Stock Price Data" by Julia Zwittlinger on Kaggle [3], and "ASML: Leading Semiconductor Innovator" by Zeeshan Younas on Kaggle [4].

Key findings of this report reveal a rapidly growing demand for semiconductors driven by artificial intelligence, particularly GPUs, alongside a critical vulnerability rooted in Taiwan's dominance in chip manufacturing. Modeling approaches, including Random Forest Regression, k-Means clustering, and Artificial Neural Networks, provided insights into industry patterns, predicting future trends with high accuracy and uncovering latent market structures.

This analysis additionally qualitatively investigates the impending end of Moore's Law as a potential innovation catalyst, geopolitical tensions as a driver for technical progress, and ethical challenges such as widening economic disparities and labor disruptions. Future research would further investigate global supply chain resilience, the effects of other major countries on the industry, and likely future technological progress. This report highlights the semiconductor industry's dual role as a driver of technological progress and a critical geopolitical asset, as well as ways to address its opportunities and risks in a future defined by rapid innovation.

Section 1: Review Project Overview

The following paragraph summarizes the initial ideas expanded upon to create this final project:

The sector of engineering that interests me most is the Semiconductor industry. One dataset titled "ASML: Leading Semiconductor Innovator" from Kaggle provides mass data on stock information related to ASML, which is an extremely important company whose Lithography machines are necessary for the created of essentially all modern CPUs. Thus, by understanding this dataset, we can get a greater understanding for the Semiconductor industry in its whole. This data can be used to predict how stocks relating to the Semiconductor industry will look in the future.

Section 2: Proposal

The following paragraph shows the initial proposition for this final project:

This project expands upon the review project overview to examine how the semiconductor industry is affected by many dependent variables, including ASML stock data, Nvidia stock data, as well as other factors such as the amount spent on research in a given country. The project objective is to use predictive analytics to view how these independent variables affect critical dependent variables that explain the complicated state of the semiconductor industry, such as ASML and Nvidia stock data into future months.

Section 3: Literature Review

Within the last year, artificial intelligence has exploded in global popularity. Artificial intelligence stands out as one of the most important innovations in recent human history with how much it promises to change nearly every industry in the coming decades. However, above everything else, AI appears to be most affecting the already very geopolitical semiconductor industry, specifically in Graphical Processing Units, or GPUs. Figure 1 below shows Google Trends data [5] showing exactly how the recent growth of AI has affected NVIDIA, the leading company for GPU manufacturing, a subdivision of the semiconductor industry.

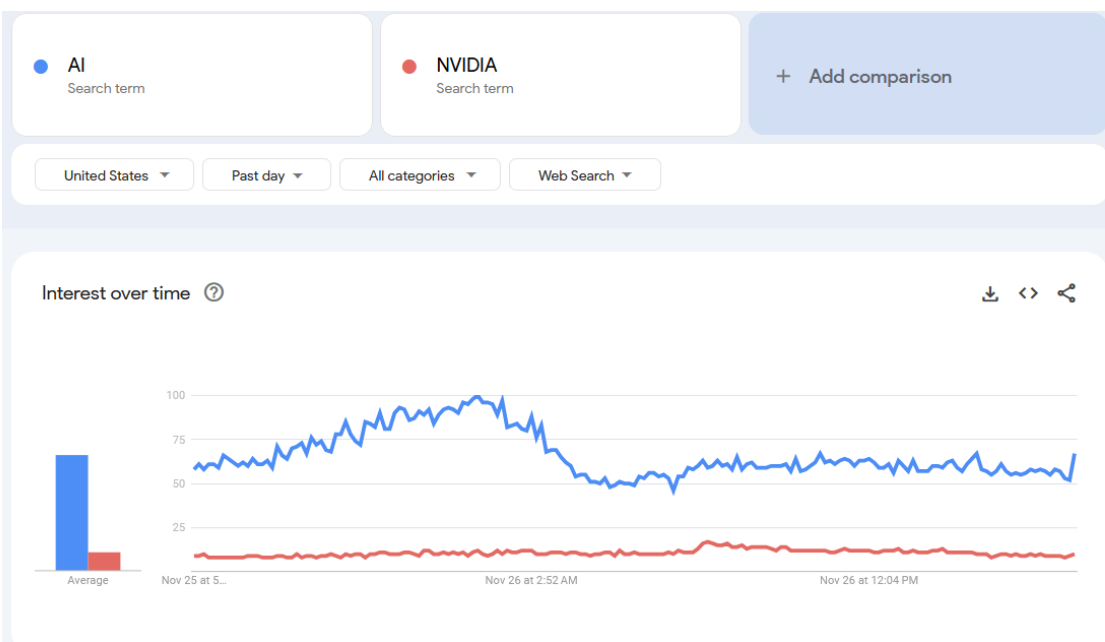

Figure 1: Interest Over Time: AI vs. NVIDIA

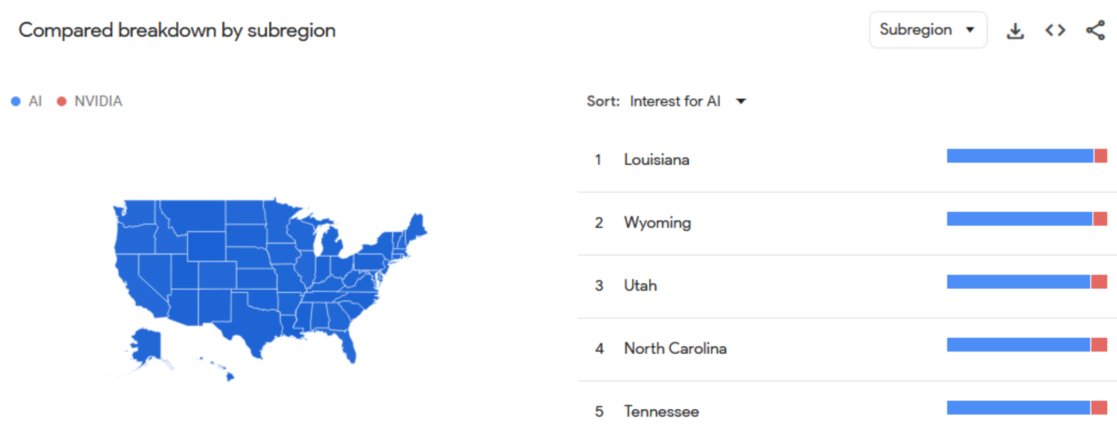

Figure 2: Compared Breakdown by Subregion: AI vs. NVIDIA

According to Toru Baji's *GPU: the biggest key processor for AI and parallel processing,* "AI requires a massive parallel operation to train many-layers of neural networks. With CPU alone, it was impossible to finish the training in a practical time. The latest multi-GPU system with P100 makes it possible to finish the training in a few hours" [6], showing just how important the GPU market is to the AI revolution. Predicting future trends for the semiconductor industry, it is critical to consider how the AI and GPU boom will affect future production and sales and semiconductors.

This new surge in AI popularity will undoubtedly continue to shift the semiconductor market into the future, but it is still very unknown how. Due to the naturally high prices of semiconductors, the global economy plays a significant role in chip sales, which could be significantly altered depending on how AI is used and regulated into our future. According to Jason Furman and Robert Seaman's *AI and the Economy*, "Artificial intelligence has the potential to dramatically change the economy. On the one hand, the potential for increased productivity growth is welcome given the decades-long slowing in productivity growth in the United States and other advanced economies. On the other hand, the potential for AI-induced labor distributions could potentially exacerbate existing problems in the labor force, including the decades-long decline in male labor force participation rate" [7]. This point shows how important it is to our findings to additionally research how global economies may shift into the future due to innovations in AI.

Interestingly, despite the immense growth of American and European countries such as NVIDIA and ASML within the semiconductor and AI industries, the company that continues to monopolize the semiconductor industry more than any other is Taiwan Semiconductor Manufacturing Company, which produces over 60% of the world's semiconductors and over 90% of the most advanced semiconductors, according to *Taiwan's dominance of the chip industry makes it more importan*t by The Economist [8]. This grasp that TSMC has over the industry is a major contributor to the geopolitical significance of Taiwan. Denny Roy's novel *Taiwan: A Political History* highlights the major modern political tensions between Taiwan's independence and China's desire for unification [9]. This is important in our prediction of the semiconductor industry, as a possible takeover of Taiwan from China could result in the global supply of semiconductors being massively disturbed.

Of course, it is always possible that future innovations under the semiconductor industry could lead to the manufacturing process becoming cheaper and more advanced, changing the way that the industry could lead. Semiconductors are already built so small that future concerns arise of humans reaching the molecular limit of how small we can make circuits. According *to What's Next? [The end of Moore's law]* by R. Stanley Williams, Moore's law states that the number of transistors on an integrated circuit doubles every two years, meaning that in 2025, humans are likely to reach the molecular limit of how small transistors can be made. This could lead to a wall preventing further semiconductor innovation, or, as Williams writes, "The end of Moore's law may be the best thing that happens in computing since the beginning of Moore’s law. Confronting the end of an epoch should enable a new era of creativity by encouraging computer scientists to invent new biologically inspired paradigms, implemented emerging architectures, with hybrid circuits and systems that combine the best of scaled silicon CMOS with new devices, physical interactions and materials” [10]. It is critical to consider this industry's unique wall for technological progress into predictions for the future of the industry.

Section 4: Data Collection and Preprocessing

For this project, the objective is to view the semiconductor industry from a wholistic perspective to see why its past trends have occurred and to use models to predict the future of the industry, as well as understand the industry from a multitude of viewpoints without overwhelming the models or readers of my analysis. Because of this it was chosen to analyze “AI Global Index” from Kateryna Meleshenko on Kaggle, “Semiconductor Shortage” by Ram Jas on Kaggle, “Nvidia Daily Stock Price Data” by Julia Zwittlinger on Kaggle, and “ASML: Leading Semiconductor Innovator” by Zeeshan Younas on Kaggle. All these datasets are highly rated on Kaggle and provide immense amounts of data related to my topic.

The first preprocessing step taken was to load, read, and display all raw data into a Jupyter Notebook as shown in Figure 3:

[75]:
```
#Loading data into file
semiShortage = "SemiconductorShortage.csv"
nvidiaStock = "NvidiaStock.csv"
asmlStock = "ASMLStock.csv"
ai = "AIData.csv"

#Reading and displaying the raw data
semiShortageData = pd.read_csv(semiconductorShortage)
nvidiaStockData = pd.read_csv(nvidiaStock)
asmlStockData = pd.read_csv(asmlStock)
aiData = pd.read_csv(ai)
semiShortageData.head()
nvidiaStockData.head()
asmlStockData.head()
semiShortageData.head()
```

[75]:

	DATE	Producer Price Index	Export Price Index	Import Price Index	Employed peoples in semiconductor industry	Export Price Index.1	Relative Importance Weights	Import Price Index.1	Export Price Index.2	Industrial Production
0	1/1/1985	100.5	.	.	670.8	.	5.5335	.	NaN	81.4300
1	1/2/1985	100.8	.	.	665.6	.	5.3861	.	NaN	81.4319
2	1/3/1985	100.9	169	141.6	660.9	223.9	5.3103	69.2	NaN	82.1699
3	1/4/1985	100.6	.	.	653.9	.	5.1255	.	NaN	81.9731
4	1/5/1985	100.1	.	.	647.2	.	5.0821	.	NaN	82.1311

Figure 3: Data Collection Display

The next preprocessing step taken was to clean and standardize the data. After reviewing all the datasets, it was found that only the Semiconductor Shortage Dataset was missing a few values, so it was decided to handle these values by filling them with the column means. All columns were converted to numeric and standardized to the clean data, as shown in Figure 4:

```python
# Define the scaler
scaler = StandardScaler()

# Function to clean and standardize
def clean_and_standardize(data, cols):
    # Convert columns to numeric, forcing errors to NaN
    data[cols] = data[cols].apply(pd.to_numeric, errors='coerce')

    #Gandle NaN values by filling with column mean
    data[cols].fillna(data[cols].mean(), inplace=True)

    # Standardize the cleaned data
    standardized_data = pd.DataFrame(scaler.fit_transform(data[cols]), columns=cols)
    return standardized_data

# Standardizing semiShortageData
numerical_cols_ai = ['Producer Price Index', 'Export Price Index', 'Import Price Index',
                     'Employed peoples in semiconductor industry', 'Relative Importance Weights',
                     'Industrial Production']
semiShortageData_standardized = clean_and_standardize(semiShortageData, numerical_cols_ai)

# Standardizing aiData
numerical_cols_semi = ['Talent', 'Infrastructure', 'Operating Environment', 'Research',
                       'Development', 'Government Strategy', 'Commercial', 'Total score']
aiData_standardized = clean_and_standardize(aiData, numerical_cols_semi)

# Standardizing nvidiaStockData
numerical_cols_nvidia = ['Open', 'High', 'Low', 'Close', 'Volume']
nvidiaStockData_standardized = clean_and_standardize(nvidiaStockData, numerical_cols_nvidia)

# Standardizing asmlStockData
numerical_cols_asml = ['Open', 'High', 'Low', 'Close', 'Adj Close', 'Volume']
asmlStockData_standardized = clean_and_standardize(asmlStockData, numerical_cols_asml)

# Optionally, print standardized data
print(semiShortageData_standardized.head())
print(aiData_standardized.head())
print(nvidiaStockData_standardized.head())
print(asmlStockData_standardized.head())
```

Figure 4: Data Preprocessing

Finally, made bar and scatter plots were made for each set of numerical data as partially shown below. There were no notable outliers in the dataset, thus no outlier manipulation was performed. Even if outliers were found in the dataset, it would likely be worth keeping them in an analysis around an industry as delicate as the semiconductor industry. Additionally, there was no removing of any duplicate data from my analysis because it was recognized that it would harm any analysis related to this specific industry. There was use for further feature engineering at this point because there is already so many useful data columns.

One discovered challenge was standardization of the data across so many different data sets, though the standardization that was performed should do a proper job of cleaning the data. Even more difficult was outlier detection. Attempts were made to find outliers in my visual analysis of the data, but these attempts were unable to find anything significant, likely due to the sheer size of these datasets. Though, this brings reassurance that any outliers will likely not have any major impact on my models and conclusions.

Section 5: Exploratory Data Analysis

For the exploratory data analysis, it was important to take a deep dive into the semiconductor industry to see how notable factors within the industry affect each other, such as the stock prices of the most important modern semiconductor manufacturing companies, the amount of people employed within the industry, and producer price index.

After loading in and cleaning our data sets, descriptive statistics are run to begin to see the likely variables that contribute to the overall state of the semiconductor industry using the .describe() function, as can be seen in Figure 5. There are many more statistics in the actual python file than are listed in the figure, though many notable ones were captured here:

```
       Producer Price Index  Export Price Index  Import Price Index  \
count            239.000000          239.000000          239.000000
mean              68.032636           70.071967           76.325941
std               10.992814           11.840853           13.282132
min               55.100000           59.000000           57.300000
25%               58.700000           61.500000           64.100000
50%               64.800000           64.600000           75.200000
75%               77.700000           76.800000           86.600000
max               89.300000          102.000000          101.200000

       Employed peoples in semiconductor industry  \
count                                  239.000000
mean                                   432.743515
std                                     89.585594
min                                    359.800000
25%                                    369.850000
50%                                    385.400000
75%                                    454.900000
max                                    714.500000

       Relative Importance Weights  Export Price Index.2  \
count                   239.000000            239.000000
mean                      3.986199            131.856485
std                       1.599349             15.408438
min                       2.224600             99.800000
25%                       2.510600            121.100000
50%                       3.646300            135.600000
75%                       4.972650            142.450000
max                       8.396000            163.500000

       Industrial Production
count             239.000000
mean              104.436920
std                 5.743083
min                93.281600
25%               100.150400
50%               102.189800
75%               109.281550
max               116.123100
```

Figure 5: Descriptive Statistics

Next, we begin to visualize the data using bar and scatter plots. This gives us a visual understanding of each statistic alone which is much easier for a human to understand than the raw descriptive statistics. These help to visualize industry trends such as the close similarity between producer price index, export price index, import price index, and employment in the industry. Below, Figures 6 and 7 show some of the data visualizations performed in this section of the analysis, though the charts were difficult to fully capture. Again, there are many more charts in the actual python file than are listed in these figures:

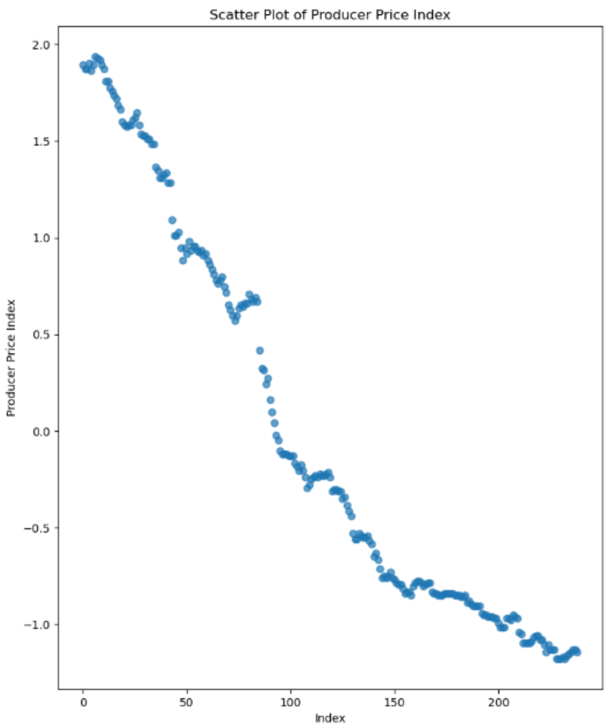

Figure 6: Scatter Plot of Producer Price Index

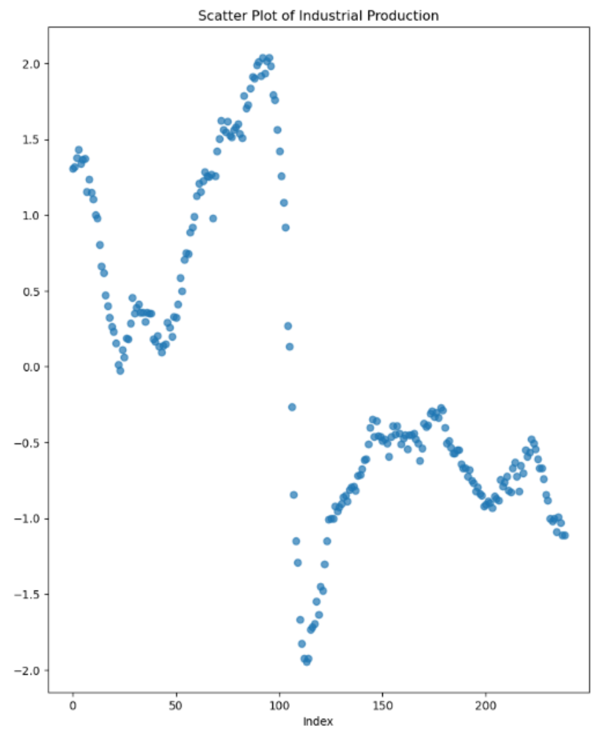

Figure 7: Scatter Plot of Industrial Production

Next, we can investigate the calculated correlations between variables. With this analysis we can see exactly how much these variables of industry significance affect each other. It is particularly insightful to see how strong of a correlation factors like Producer Price Index has with Employed peoples and Import Price, while these factors seem to have a negative correlation with Nvidia and ASML stock prices. Figure 8 below presents a correlation matrix supporting these insights. Again, there are many more charts in the actual python file than just listed in this image:

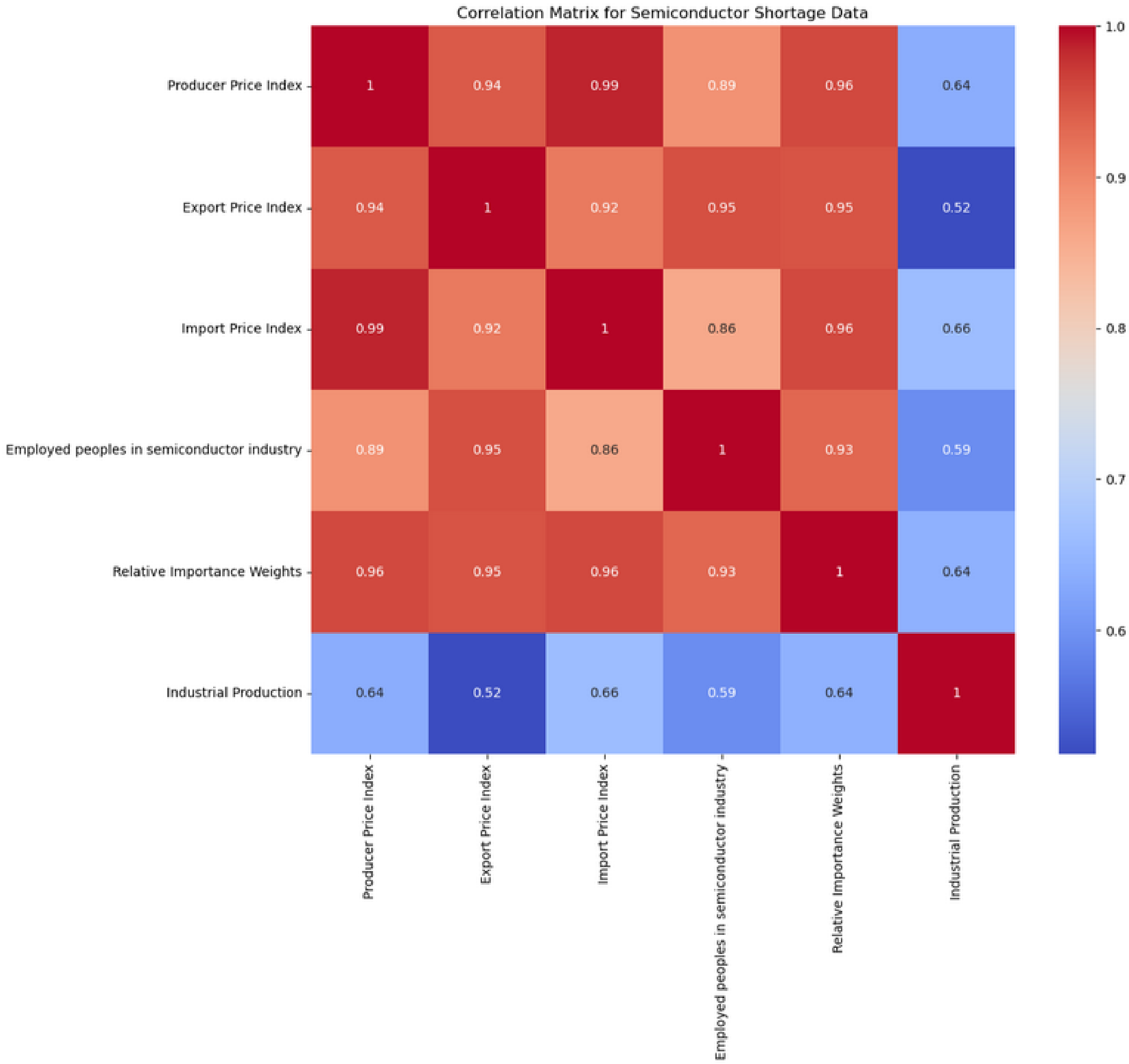

Figure 8: Correlation Matrix of Semiconductor Shortage Data

Finally, we conclude the exploratory data analysis by looking at visualizations of the notable company stock prices overtime. Below, Figures 9 and 10 present layered time-series data of stock prices of the notable semiconductor companies Nvidia and ASML. These visualizations give us an understanding of how the industry has been doing over the last few decades so that we can begin to understand what it will look like into the future. Interestingly, the stock prices between Nvidia and ASML seem completely unrelated despite being in similar industries and fields of work.

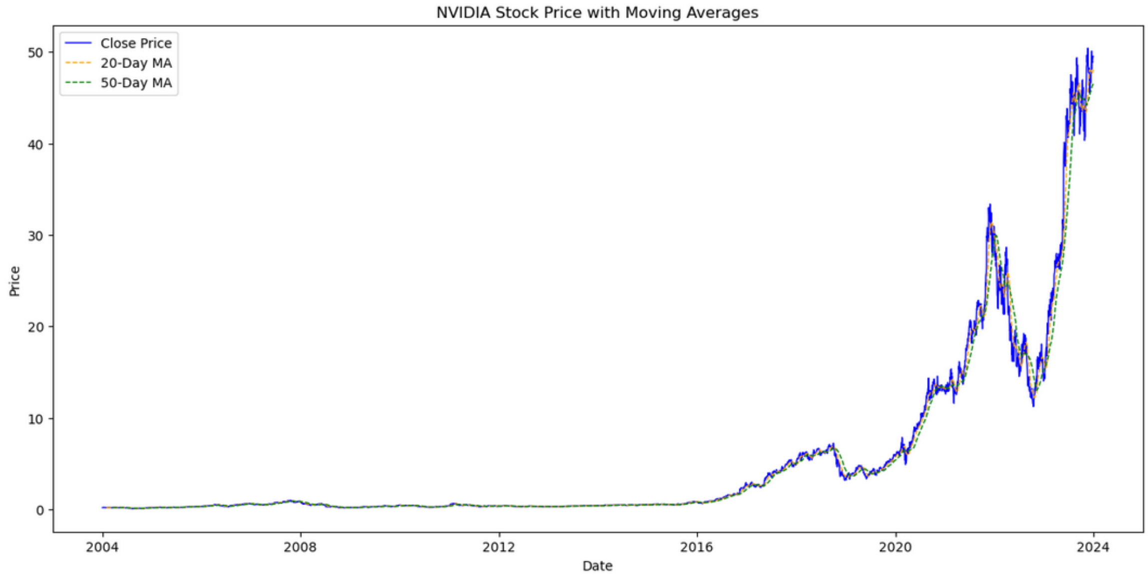

Figure 9: Nvidia Stock Price Time-Series Visualization

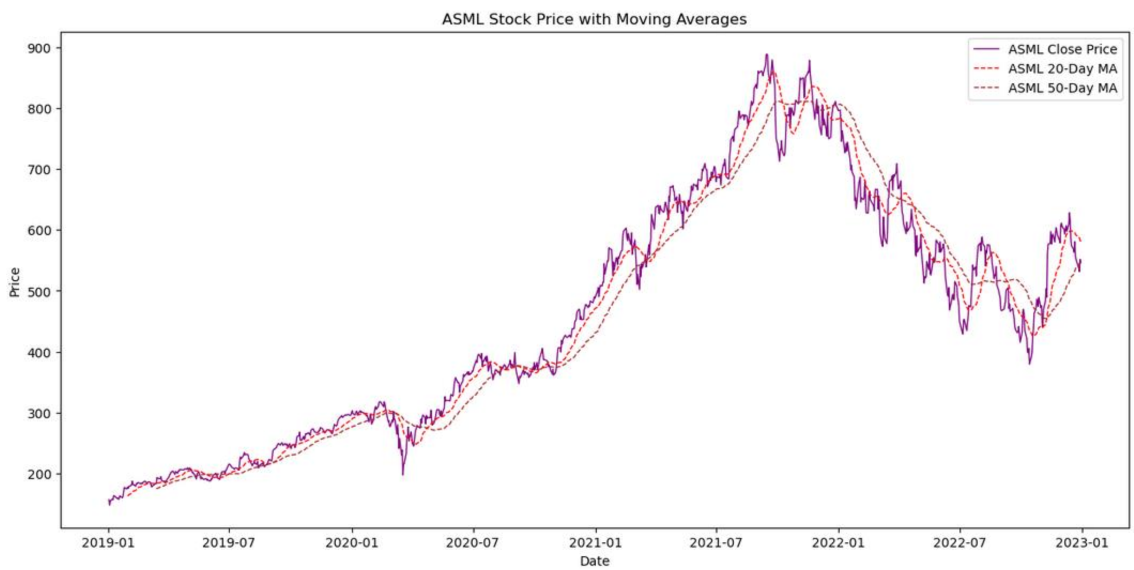

Figure 10: ASML Stock Price Time-Series Visualization

Section 6: Model Development

The first model chosen to use was a random forest regression due to its capability to manage complex, non-linear relationships. In an industry as complex and rapidly changing as the semiconductor industry, random forests are well-suited for capturing complex interactions while remaining resistant to over-fitting. The random forest model is highly understandable because it provides feature importance scores, allowing us to pinpoint the most influential factors in predicting metrics like stock prices. For this model, n_estimators was tuned to balance model performance with computation time, max_depth was tuned to avoid overfitting, and min_samples was split to control the minimum number of samples required to split a node to help generalize the model best.

The second model chosen was a k-Means clustering model to find patterns within the stock price data. k-Means models are computationally very efficient and well-suited for exploring and visualizing structures within large datasets, providing easy-to-understand insights into the data. The only hyperparameter tuned for this model was k to adjust the number of clusters. To find the optimal number of clusters, a silhouette score analysis was used to ensure that the model is as accurate as possible.

The final model chosen to use was an artificial neural network because of its strengths in capturing non-linear, complex data interactions. ANNs are particularly useful with time-series data which is useful for analyzing stock price data overtime. ANNs typically require large datasets to function properly, but my datasets are already very large. For this model, the number of hidden layers and neurons per layer were tuned to balance complexity and generalization. Additionally, optimization method was tuned to randomly search across hyperparameters to prevent over-fitting.

Section 7: Results and Evaluation

This project uses three different modeling approaches to analyze the semiconductor industry dataset: a Random Forest Regression, a K-Means Clustering, and an Artificial Neural Network. To evaluate the results of the models, each model's performance was tested using metrics of means square error, mean absolute error, and R-squared. Or, in the case of the k-Means model, we use silhouette score to measure the quality of clustering. These metrics provide insights into the strengths and weaknesses of each model, as mentioned below.

1) Random Forest Regression

```
Random Forest Regression MSE: 0.03769785416666993
Mean Absolute Error: 0.1295208333333441
Mean Squared Error: 0.03769785416666993
R-squared: 0.9997295727518876
```

Figure 11: Random Forest Regression Results

The Random Forest Regression model performed extremely accurately, achieving an R-squared score of 0.9997, as shown in Figure 11. This indicates that the model explained nearly all variance in the target variable, with very low error rates, which we can tell from the resulting MSE of 0.0377 and MAE of 0.1295. These metrics reflect the Random Forest’s ability to capture complex, non-linear relationships in the data. It is important to note that, while random forest regressions are very with handling large datasets, they are also computationally difficult to run, and more complex models will take a long time to predict results.

2) K-Means Clustering

```
Number of clusters: 2, Silhouette Score: 0.8186050106816524
Number of clusters: 3, Silhouette Score: 0.8330148741495828
Number of clusters: 4, Silhouette Score: 0.8087615940825517
Number of clusters: 5, Silhouette Score: 0.8174757385560414
Number of clusters: 6, Silhouette Score: 0.7969279256117164
Number of clusters: 7, Silhouette Score: 0.8020875579959896
Number of clusters: 8, Silhouette Score: 0.8094178651324756
Number of clusters: 9, Silhouette Score: 0.794529489457818
```

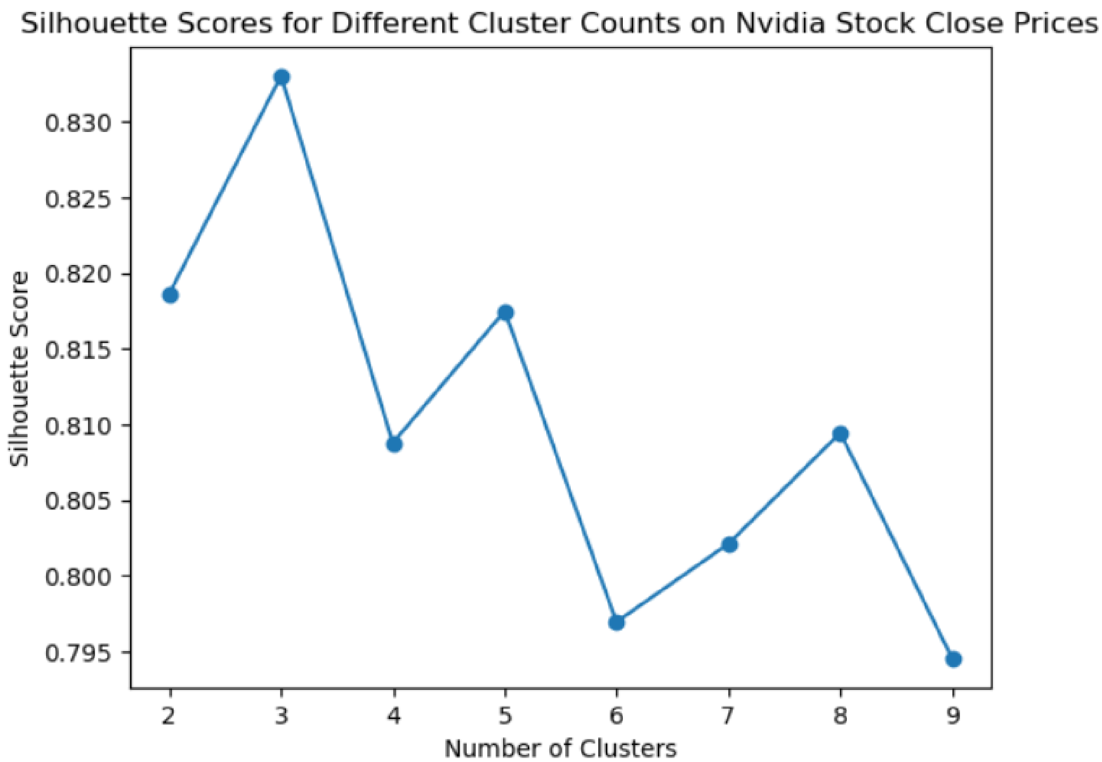

Figure 12: K-Means Clustering Results

K-Means was evaluated using the Silhouette Score, which measures the quality of clustering. The highest score of 0.8330 at 3 clusters, as shown in Figure 12, suggests that this configuration provided well-defined and meaningful groupings within the data, highlighting potential market segments or behaviors in the semiconductor industry. This is also shown in the visualization which visualizes the peak silhouette score at 3 clusters. k-Means clustering is good at discovering hidden patterns within data but is susceptible to initial cluster centroids and will struggle at capturing non-spherical cluster shapes.

3) Artificial Neural Networks

```
Scaled Test Loss: 0.0810086727142334
Mean Absolute Error: 0.1122315524317243
Mean Squared Error: 0.10439509097849399
R-squared: 0.9988984077903325
```

Figure 13: Artificial Neural Networks Regression Results

The ANN showed strong predictive capability, with a high R-squared of 0.9989 and an MAE of 0.1122, as shown in Figure 13, which is lower than the Random Forest model, suggesting high precision in individual predictions. However, the MSE was slightly higher, indicating that while the ANN captured subtle patterns, it may be more affected by outliers. ANNs are strong with handling complex data patterns, and in this case have the lowest MAE among the three models suggesting high prediction accuracy, but also have very high computational costs which could be a greater problem among very large and highly complex datasets.

Concluding Table

Model	MAE	MSE	R^2
Random Forest Regression	0.1295	0.0377	0.9997
k-Means Clustering	Not Applicable	Not Applicable	Peak Silhouette Score: 0.8330
Artificial Neural Network	0.1122	0.1044	0.9989

Figure 14: Model Conclusion Table

As shown in Figure 14, all the models were quite successful in their prediction evaluations. The RF Regression and ANN were very close in their metrics. ANN had slightly higher MSE, indicating some errors might be larger, while also having a lower MAE, indicating that it made less errors on average. While we cannot directly compare the k-Means clustering to the other models, a peak silhouette score of 0.8330 shows that the clusters were ell separated and cohesive.

Section 8: Ethical Considerations

1. **Data Privacy and Transparency:** All the data used in this analysis was collected with the consent of the databases, though it is not entirely clear where the databases sourced their data collections from.
2. **Fairness and Bias:** Given the quantity of different datasets used in this analysis and the preprocessing techniques used, there is limited potential bias. Additionally, semiconductor data is not an evaluation topic that has much room for bias, as most of the data, such as stock values, are rather linear in nature.
3. **Model Interpretability:** The interpretability of the models, particularly the neural network, is crucial. Using easy-to-understand visualizations and storytelling techniques, we can establish high model interpretability in the analysis.
4. **Data Quality and Quantity:** The quality and quantity of the data can significantly impact model performance. Fortunately, this analysis uses a multitude of thorough datasets from differing sources for evaluation.

Ethical Limitations

1. **Model Assumptions:** Each model has its own assumptions. For instance, linear regression assumes a linear relationship between variables, while k-means assumes spherical clusters. These assumptions can impact the results of the evaluation.
2. **Generalizability and Interpretability:** The generalizability and interpretability of the findings is limited by the specific dataset used. Though the results of the analysis are useful, it is important to understand that findings may not generalize in other contexts or time periods.
3. **Computational Cost:** The computational cost of training and deploying complex models like neural networks can be significant, especially for large datasets. This dataset does not particularly have any high computational times, but if one tries to evaluate this analysis on a low-end computer, they may struggle.

In summary, by addressing these ethical considerations and limitations, the project can be conducted responsibly, and the findings can be used to make informed decisions regarding the semiconductor industry. It is important to have a deep understanding of how data analyses affect factors such as data privacy and data fairness to ensure that our work as data scientists does not put others at risk.

Section 9: Conclusion

The semiconductor industry stands at the forefront of modern technological progress, serving as an anchor for future advancements in artificial intelligence, computing, and global economic development. Yet, trying to predict this critical industry's future trajectory becomes extremely difficult due to the sheer wealth of factors affecting semiconductor production, such as technological innovation, geopolitical pressures, and economic trends, all of which are themselves extremely difficult to predict. This comprehensive analysis seeks to explore datasets such as the AI Global Index, Semiconductor Shortage numerical data, Nvidia stock prices, and ASML stock prices to uncover critical insights and better understand the industry's role in our future.

Through exploratory data analysis, we find connection between variables such as producer price indices, employment, and global trade metrics, highlighting how these economic drivers collectively shape the semiconductor market. For instance, while the industry's growth is linked to advances in artificial intelligence, as evidenced by the rising demand for GPUs, geopolitical dependencies, such as Taiwan Semiconductor Manufacturing Company's dominance, additionally present both opportunities and vulnerabilities. With TSMC producing over 60% of global semiconductors and 90% of the most advanced chips, any disruption in Taiwan could have catastrophic effects across industries reliant on these technologies. While it is not entirely feasible to capture the geopolitics of Taiwan as an independent variable for this study, attempts are still made in this analysis to acquire semiconductor workforce data through measuring factors such as global employment statistics in the semiconductor industry, which have a similar impact on the state of the industry.

One of this study's most important findings comes from the juxtaposition of technological and geopolitical factors. The looming possibility of the end of Moore's Law, which was once seen as a barrier, now appears to be a catalyst for innovation, suggesting that overcoming physical limitations in chip design could unlock entirely new technical industries to shift future progress. From a geopolitical perspective, Taiwan's significance, alongside rising tensions with China, presents not just an economic risk but a potential global crisis, with implications for supply chains far beyond the semiconductor sector. Finally, while artificial intelligence and grand-scale automation promise growth, they simultaneously introduce vulnerabilities to the health of the industry and society as a whole, as unchecked reliance on AI-centric advancements may deepen inequalities, disrupt labor markets, and create monopolistic pressures.

Modeling outcomes further supports these findings. The Random Forest Regression model demonstrated near-perfect predictive accuracy, revealing the nonlinear relationships between key factors like research spending, stock price movements, and global semiconductor demand. Meanwhile, K-Means clustering examines distinct market behaviors, identifying how geopolitical events and technological innovation segment the industry. The Artificial Neural Network further captured complex patterns that point to the sector's rapid evolution and the growing divergence between leading semiconductor firms like Nvidia and ASML.

Future research on this industry should directly address these challenges. Developing predictive models that incorporate real-time geopolitical data could provide early warning systems for supply chain disruptions. Exploring alternative manufacturing locations, materials, and sustainable production methods would be critical in mitigating risks. Additionally, investigating the regulatory frameworks necessary to balance innovation with ethical considerations regarding AI could help navigate the delicate interplay between technological growth and societal impact.

Ultimately, the semiconductor industry is not just a technological marvel but is also deeply connected to technological progress, as well as global economic and geopolitical tensions. This analysis reveals that AI-driven demand for GPUs, reflected in Nvidia's soaring stock prices, has become a dominant force, but it also highlights a dependence on the Taiwan Semiconductor Manufacturing Company. This reliance creates a vulnerability, where geopolitical tensions, particularly in Taiwan, threaten to disrupt the global supply chain. Modeling further uncovered that research investments and innovation cycles are key drivers of future industry performance, emphasizing the strategic importance of sustained research and development. Simultaneously, unknowns like the possibility of the end of Moore's Law call attention to how difficult it can be to actually model predictions of such a complex industry. This project also found ethical challenges, such as the risks of increased economic inequality and labor market disruptions through unchecked technological advancements, specifically in AI. Addressing these challenges would require future research into diversifying semiconductor manufacturing and implementing regulatory frameworks to balance innovation with equity. The stakes going into the future of this explosive industry are monumental, and as a society we need to ensure the semiconductor industry is used as a tool to drive progress without sacrificing global stability.

Bibliography

[1] K. Meleshenko, “AI Global Index Dataset,” Kaggle, https://www.kaggle.com/datasets/katerynameleshenko/ai-index (accessed Oct. 23, 2024).

[2] R. Jas, “Semiconductor Shortage Dataset,” Kaggle, https://www.kaggle.com/datasets/ramjasmaurya/semiconductor-shortages19852021 (accessed Oct. 23, 2024).

[3] J. Zwittlinger, “Nvidia Daily Stock Price Data 📈,” Kaggle, https://www.kaggle.com/datasets/jvanark/nvidia-daily-stock-price-data (accessed Oct. 23, 2024).

[4] Z. Younas, “ASML: Leading Semiconductor Innovator,” Kaggle, https://www.kaggle.com/datasets/zeeshanyounas001/asml-leading-semiconductor-innovator (accessed Oct. 23, 2024).

[5] “Google Trends.” Oct. 16, 2024 (accessed Oct. 16, 2024).

[6] T. Baji, “GPU: The biggest key processor for AI and Parallel Processing,” SPIE Digital Library, https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10454/1045406/GPU--the-biggest-key-processor-for-AI-and-parallel/10.1117/12.2279088.short (accessed Oct. 16, 2024).

[7] J. Furman and R. Seamans, “AI and the Economy,” The University of Chicago Press Journals, https://www.journals.uchicago.edu/doi/full/10.1086/699936 (accessed Oct. 16, 2024).

[8] “Taiwan’s dominance of the chip industry makes it more important,” The Economist, https://www.economist.com/special-report/2023/03/06/taiwans-dominance-of-the-chip-industry-makes-it-more-important (accessed Oct. 16, 2024).

[9] D. Roy, *Taiwan: A Political History*. Ithaca, New York: Cornell Univeristy, 2003.

[10] R. S. Williams, "What's Next? [The end of Moore's law]," in *Computing in Science & Engineering*, vol. 19, no. 2, pp. 7-13, Mar.-Apr. 2017, doi: 10.1109/MCSE.2017.31. keywords: {Computers;Nonvolatile memory;Scientific computing;Moore's Law;Random access memory;Accelerators;Technological innovation;Moore's law;memory driven;nonvolatile memory;photonic interconnect;accelerator;ASIC;brain inspired;scientific computing},(accessed Oct. 16, 2024).