
NVIDIA

Introduction

NVIDIA is a global leader in graphics processing units (GPUs), high-end performance computing, and later expanding into the AI industry, data centers, and autonomous vehicles. As a first mover in the AI space, NVIDIA has gained a significant competitive advantage and edge over its competitors. However, maintaining market dominance remains a key issue, with competitors such as AMD and Intel rapidly advancing in AI and GPU production as well as many AI startups threatening to disrupt the industry. NVIDIA must keep up with the fast-changing industry standard and R&D to stay ahead of the relentless technological progress being made in the world.

Value Proposition

NVIDIA has consistently been the industry standard for parallel computing and high-performance GPUs, especially relating to AI.

Their influence in cloud computing is unparalleled in the space, and they continue to expand into different sectors such as robotics, AI computing, and cloud web services. The main competitors for NVIDIA are AMD and Intel, but their AI-driven GPUs have been able to outperform their competitors since they have taken the first mover's advantage. NVIDIA has secured a first-mover advantage in AI computing, allowing its cutting-edge H100 and A100 GPUs to outperform rivals in training and inference tasks (Shah, 2024). Referring to Appendix 1, the company has invested heavily in AI, robotics, and cloud web services, adding to its dominance. In doing so, NVIDIA has secured a strong technological edge over competitors, which is their key competitive advantage, as it allows them to continue being the top player in the GPU space.

According to Appendix 2's AAA analysis, NVIDIA shows strong industry-leading practices, currently holding 92% of the data centre GPU Market. One of its greatest differentiation factors is the CUDA software ecosystem, which acts as a competitive advantage. This asset builds a strong developer community by simplifying GPU programming and allows developers to write in familiar languages like C, C++, and Python (Wills, 2024), leveraging the power for new developers to enter the complex field of AI, deep learning, and simulations. CUDA also provides an advantage in producing a strong network effect that drives innovation since it is a self-reinforcing ecosystem. As developers create open-source libraries, tutorials, and repositories, it provides more resources for new developers to contribute, creating a positive feedback loop. Through continuous innovation, strong developer support, and the CUDA ecosystem, NVIDIA is not just a leading hardware manufacturer but a foundational platform for the future of computing (Barron's, 2025).

Differentiation

NVIDIA can differentiate itself from its competitors. NVIDIA strategically differentiates itself through both vertical and horizontal differentiation. In terms of vertical differentiation, NVIDIA

offers superior GPU performance in areas such as gaming, AI, and data centers (Wills, 2024). Customers are willing to pay a premium for this quality, justifying the higher pricing compared to its competitors. The H100 and G100 GPUs offer unmatched speeds for machine learning training and inference, making them the gold standard for AI computing (Datacrunch, 2024). The introduction of those chips was revolutionary, shattering all GPU benchmarks at the time of release (Moorhead, 2022).

NVIDIA's Omniverse platform provides horizontal differentiation for consumers. While competitors such as AMD may offer GPUs with similar raw performance, NVIDIA offers unique features like Deep Learning Super Sampling, ray tracing, and deep integration with AI frameworks. Many developers also prefer the CUDA ecosystem for its simplicity and easy barrier to entry into the community. This reinforces the horizontal differentiation for consumers with its unique features over other brands.

First Mover Advantage

First-mover advantage describes the benefits a company gains by being the first to enter a specific market segment. In 2017, NVIDIA exemplified this by being the first entrant into the GPU chip market for artificial intelligence (AI) and machine learning applications. Recognizing the increasing demand for enhanced processing power, NVIDIA introduced the Volta architecture, specifically designed to accelerate AI workloads. For example, the Tesla V100 GPU, based on Volta, can deliver over 120 teraflops of deep learning performance, a significant leap over the previous GPUs (NVIDIA, 2017). NVIDIA's strategic foresight of the market's demands positioned it at the forefront of AI infrastructure, with its hardware and software solutions, such as CUDA and the NVIDIA Deep Learning Accelerator, allowing researchers and developers to launch AI models with an increased level of speed and efficiency. Now, advanced AI models like ChatGPT use NVIDIA's GPUs, specifically the A100 and H100 Tensor Core GPUs, to handle their extensive computational requirements (Insights of Nature, 2023). By securing early entry into the AI GPU market and continually innovating, NVIDIA has maintained its competitive advantage, benefiting from substantial intellectual property protections and setting industry standards for AI hardware acceleration.

Barriers to Entry & Deterrence Strategies

The GPU market is largely characterized by blockaded entry, as NVIDIA's structural advantages can naturally deter most competitors. This is evident by analysts' estimates that in 2024 NVIDIA held approximately "92% share of the data center GPU market" (ExtremeTech, 2024). Although now in 2025, competition has increased, and companies like AMD, Intel and Juniper do remain in the space, they struggle to compete at NVIDIA's scale, suggesting some level of accommodated entry (Gartner, 2024).

NVIDIA's competitors face significant structural barriers to entry, which strengthens NVIDIA's position in the AI and GPU markets. One key barrier is the economies of scale that NVIDIA

benefits from through its established production capabilities, which enable cost advantages, making competition on price challenging for new entrants. Additionally, NVIDIA has strong brand loyalty, particularly in AI and gaming, which further solidifies its market dominance. High switching costs also deter customers from moving away, especially given the deep integration of NVIDIA's CUDA platform, which significantly outperforms traditional CPUs in processing AI and machine learning workloads and is deeply integrated into many AI development pipelines. Ultimately, due to current market conditions and the increasing urgency felt in the AI development market, the risk of slowing AI development by switching chip providers creates high switching costs for customers and makes it unlikely that major players will abandon NVIDIA. As Seitz (2024) points out, the time-to-market trade-off involved in switching to a different provider poses a significant risk for AI developers, making NVIDIA's position even stronger.

Alongside structural barriers, NVIDIA employs strategic deterrence strategies to maintain its dominance. Exclusive partnerships with firms like OpenAI, Microsoft, and Apple increase rivals' costs and limit market opportunities. OpenAI's recent push in 2025 to develop its own AI chips underscores its dependence on NVIDIA and highlights the current reliance on NVIDIA (TipRanks, 2025). Additionally, NVIDIA's aggressive capacity expansion strengthens its market position, making it harder for competitors to gain market share. Its full-stack ecosystem, combining hardware and software, further strengthens its dominance and discourages market entry. Overall, NVIDIA's strong reputation as an industry leader serves as a deterrent, signaling to potential competitors the difficulties of challenging its market position.

Technology

For NVIDIA to maintain its competitive advantage (CA), it must figure out which strategy/strategies will make its technology the dominant one. NVIDIA currently dominates the GPU space in terms of revenue across all industries and in terms of market share in the data center industry (Fernandez, 2025), the discrete desktop GPU market (Shilov, 2024), and the gaming market (Wilson, 2024). These factors have already solidified NVIDIA as the leading GPU manufacturer, but how can they remain this way? For them to maintain their CA, they should find ways to make the network effect and positive feedback loop work against their competitors. The network effect is a phenomenon where a good or service increases in value the more that people use it, creating a positive feedback loop.

NVIDIA's Network Effects:

1. **NVIDIA's Ecosystem**
    a. NVIDIA has fostered a strong community of developers, creating a cycle where these developers develop applications that then attract more developers to use NVIDIA, resulting in more consumers for the company (Lazo, 2025).
        i. An example of a killer application that attracts more developers is CUDA. Compute Unified Device Architecture (CUDA) allows developers to use

NVIDIA GPUs for tasks beyond just graphics rendering, allowing for faster computing for non-graphics rendering tasks (Krumholz, 2024).

2. **Dominance and Power in the Industry**
    a. NVIDIA GPUs make up 92% of the data center market. In 2024, the data center market more than doubled year-over-year, with no signs of slowing down, meaning that as the data center market grows, the demand for NVIDIA GPUs grows as well (Fernandez, 2025).

Sustaining a Competitive Advantage

Strategies that exhibit a sustained competitive advantage often contain four characteristics:

1. Heterogeneous.
    a. Something different than what competitors do.
2. Inimitable.
    a. Something entrants/competitors cannot replicate.
3. Appropriable.
    a. Something from which the firm can capture value.
4. Foresight.
    a. Something whose value the firm recognizes before its competitors/entrants.

To understand if NVIDIA possesses a sustained competitive advantage, we must analyze its GPU strategy under the four specified categories.

1. **Heterogenous:** Technological Leadership in the GPU Space (dcfmodelling, 2025).
    a. NVIDIA chips are almost solely used in the AI space, as well as dominating the discrete GPU space, outlining the heterogeneous nature of their products.
    b. NVIDIA's spending on R&D ($12.91 Billion USD) (macrotrends, 2025) further outlines their technological leadership and constant urge to innovate, leading to differentiated products that the market desires.

2. **Inimitable:** Complex Barriers to Entry & Patents.
    a. NVIDIA chips are not easy to copy, as they are highly advanced and technical pieces of technology, often at the forefront of computational innovation.
        i. They even invented the first consumer-facing GPU back in 1999 (NVIDIA, 2025).
    b. NVIDIA has an extensive array of patents, they have over 26,000 active patents (dcfmodelling, 2025).
    c. NVIDIA also sinks roughly USD 10 billion annually into AI chip design, something not easily achievable given the cost and expertise required to create new technology (dcfmodelling, 2025).

3. **Appropriable:** NVIDIA provides state-of-the-art GPUs, that are used in numerous industries (dcfmodelling, 2025).
    a. NVIDIA's value can be understood by its revenue breakdown by segment
        i. Data Centers—$115.19 Billion (USD) (BullFincher, 2025).
        ii. Professional Visualization - $1.88 Billion (USD) (BullFincher, 2025).
        iii. Gaming - $11.35 Billion (USD) (BullFincher, 2025).
        iv. Total Revenue (FY 2025) - $130.5 Billion (USD) (BullFincher, 2025).
    b. NVIDIA's data center segment makes up a large majority of its revenue (88.27%) (BullFincher, 2025), outlining its predominant value to the data center and cloud computing industries.

4. **Foresight:** First Movers into the GPU Space.
    a. NVIDIA's high R&D spending and dominance in the AI chip space highlight its ability to capture value before it is fully recognized.
        i. They've historically shown foresight, as NVIDIA was the first company to invent a consumer-facing GPU (NVIDIA, 2025).
    b. NVIDIA is extremely strategic with its product development and deployment as it prioritizes its focus on emerging technologies such as AI.
        i. NVIDIA holds a 95% market share in the AI chip industry.
        ii. NVIDIA's GPUs are market-leading.
        iii. NVIDIA's technological leadership is world-class (dcfmodelling, 2025).

Financial Analysis

Nvidia attained a revenue of $130 billion in the 2025 fiscal year, up 114% from the previous year, of which $73 billion was retained as profits. Their revenue streams can be broken down into four segments. Data centers were responsible for 88% of their revenue, at $115 billion, making it by far their highest-grossing segment. Their gaming and AI PC segment composed 9% of their revenue, at $11 billion. Professional visualization and automotive and robotics made up the final 3% of their revenue, with $1.5 billion and $1.3 billion in revenue, respectively (appendix 3). Their annual revenue is up 114% from the previous year, largely due to the 142% increase in data center revenue (Nvidia, 2025). This upward trend is expected to continue, with an industry growth rate of around 33.4% presenting an opportunity for Nvidia to further increase revenue by maintaining market share (GPU Global Market Report, 2025). Their gross margins are very high but are predicted to decrease from 75% in Q4 2025 to 70.6% in Q1 2026. Operating expenses are generally around 12.5% of revenue, with 10% going to R&D and 2.5% to SG&A.

Recommendations

Nvidia currently holds an estimated 92% market share in the data center GPU market (Fernandez, 2025), which can largely be attributed to their technological lead, with GPUs offering significantly improved processing speeds, scalability, and energy efficiency compared to competitors. This is supported by their strong software ecosystem, which optimizes hardware to

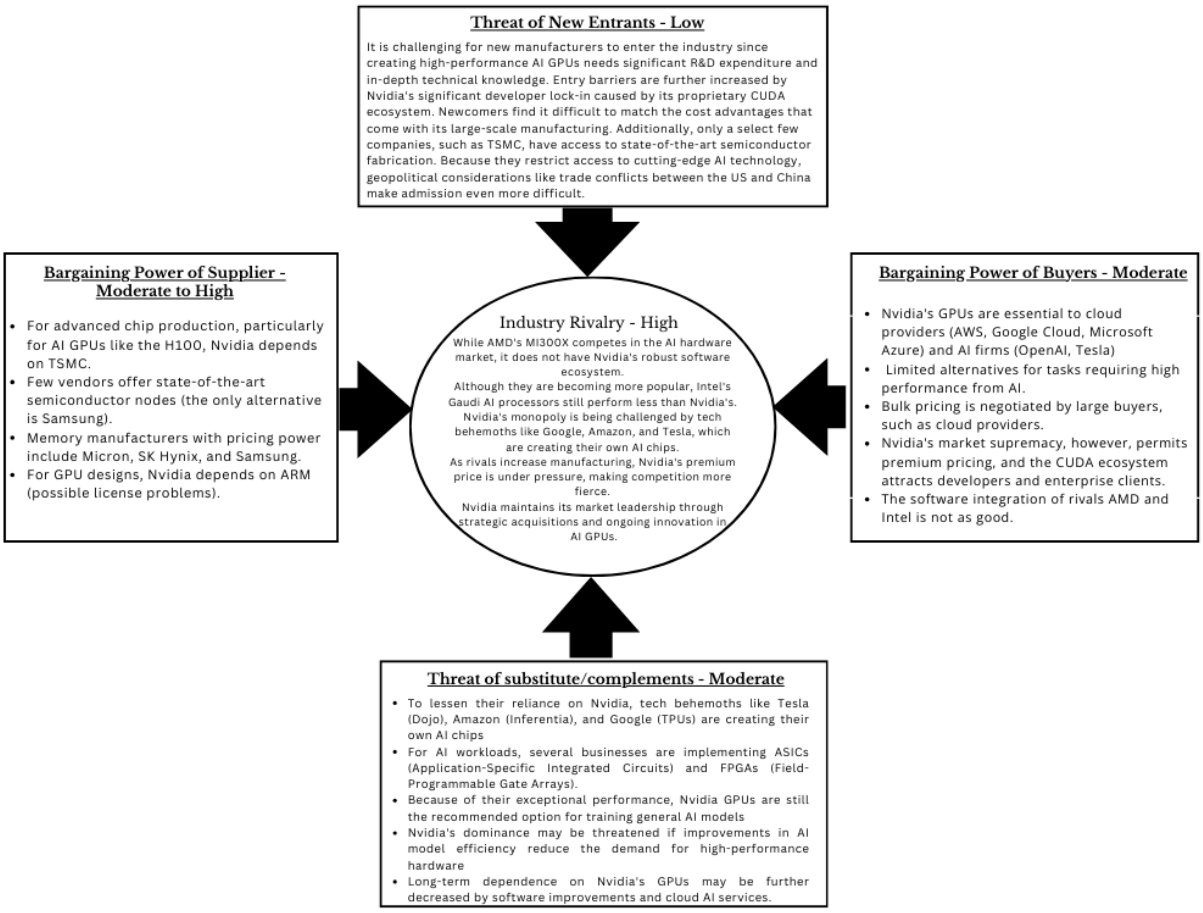
further improve performance. With Nvidia's competitors and current customers developing competing products, Nvidia risks losing market share in data center GPUs. As this accounts for 88% of Nvidia's yearly revenue, Nvidia must retain its market share to maintain high profitability. This can be achieved by increasing their investments in R&D by 25% to maintain their technological advantage in the long run. Currently, Nvidia expects a 9.3% quarter-over-quarter increase in revenue from Q4 2025 to Q1 2026. Even if Nvidia manages to maintain this revenue following releases of competing products, their profits the following year will be similar to if they increase R&D investments, at $89 billion and $87 billion, respectively, as doing so would ensure they do not lose share in the data center GPU market. Even if market growth falls as low as 20%, Nvidia's profits will remain incredibly high at $79 billion, while if market growth exceeds expectations to 50%, they will see profits as high as $97 billion (appendix 3). This means the investment carries little risk while ensuring their competitive advantage is maintained in the long run. Reinvesting profits into advancements in data center GPUs will reinforce technological leadership and market dominance, ensuring sustained profitability in an increasingly competitive landscape, without sacrificing significant short-term gains.

Aside from increasing investment in data center GPUs, NVIDIA should also proactively diversify its partnerships beyond its current chip manufacturers due to the rising geopolitical and supply chain disruptions between the USA and Taiwan with the tax increase on imports, thereby reducing their costs. Although its technological achievements have been largely made because of its dependency on TSMC in Taiwan, future collaboration with alternative manufacturers like Samsung would reduce the dangers of external instabilities in the economy. To improve its position in the industry, NVIDIA could also look at strategic M&As. Purchasing startup semiconductor companies and smaller AI chip companies would boost their innovation and provide access to a wider spectrum of technological capabilities. Their market reach will be improved by their alliance with international hardware and software suppliers, maintaining their dominance in the AI and GPU markets.

Cloud computing infrastructure is becoming more and more necessary for AI-driven workloads, and NVIDIA's market contribution by offering high-performance AI chips for data centers is a major factor in the company's revenue growth. Additionally, its Omniverse platform plays a vital role in encouraging high switching costs for organizations adopting its ecosystem. By making it difficult and expensive for companies to switch to other supplies, sustaining investments in these areas would improve their value proposition and also guarantee customer retention in the longer run. They can build a complex, high-value ecosystem that reinforces its competitive edge through these offerings.

Appendices

Appendix 1: Porter's 5 Forces



Appendix 2: Triple A Framework

Assets	Activites	Advantages
92% of the data center GPU market	Builds upon its already dominant footprint in data centres, where demand for NVIDIA GPUs grows alongside AI adoption	Most data centres powered on NVIDIA AI chips leading to 88% of NVIDIA's profits
NVIDA owns over 26000+ patents relating to software and hardware innovations	Build upon its patents and breakthrough with more R&D for more market dominance	Full-stack software integration, global supply chain management, and deep IP portfolio make it nearly impossible for new entrants to compete.
Strongest GPU chips in industry	Continue to reinvest its profits into developing strong hardware	Leading positions in gaming GPUs, discrete desktop GPUs, and AI infrastructure
CUDA is proprietary and widely adopted, acting as a deep moat that competitors can't easily replicate.	Continues to improve CUDA's capabilities and integration into academic, enterprise, and research environments.	High switching cost for software from competitorss
Developer ecosystem and community	Continuously adds supports for new use cases like gene AI, simulation and LLM deployment	NVIDIA has fostered a strong community of developers that provide a positive feedback loop. This reinforces the network effect, increasing the value of the platform with each new user.
Educational programs, tutorials, and frameworks for developers	Build upon more languages and systems for more developer supports	A large, loyal base of developers that build on NVIDIA's tools, which in turn attracts more developers and consumers
First movers advantage in AI	Maintain investments in AI and R&D to continue market dominance	Even competitors like AMD and Intel struggle to replicate NVIDIA's AI software suite.
High brand reputation and loyalty	Continue its marketing strategy offering premium products and services	Higher premium prices for consumers who are used to NVIDIA's services

Appendix 3: Financial Analysis

	Data Centers			Gaming and AI PC			Professional Visualization			Automotive and Robotics		
	Revenue $	Revenue %	Yearly growt	Revenue $	Revenue %	Yearly growt	Revenue $	Revenue %	Yearly growt	Revenue $	Revenue %	Yearly growth
Revenue	115200	88.27789145	142%	11400	8.73583301	9%	1900	1.45597217	21%	1700	1.30271194	55%
	2025	**2026 no change**	**2026 increased R&D**									
(Units: Million of USD)			2026	2026	2026							
Total Revenue	130497	$172,000.00	$159,670.15	$174,646.15	$194,230.15							
Gross Margin	75%	70.60%	70.60%	70.60%	70.60%							
COGS	$32,624.25	$50,568.00	$46,943.02	$51,345.97	$57,103.66							
R&D	$12,914.00	$17,200.00	$19,958.77	$21,830.77	$24,278.77							
SG&A	$3,419.00	$4,300.00	$3,991.75	$4,366.15	$4,855.75							
Operating expenses	$16,405.00	$21,500.00	$23,950.52	$26,196.92	$29,134.52							
Operating income	$81,453.00	$99,932.00	$88,776.60	$97,103.26	$107,991.96							
Net Income	$72,880.00	$89,414.07	$79,432.79	$86,883.05	$96,625.71							
	0.89474912											
Cash	$43,210.00											

(Nvidia, 2025)

2026 no change forecast:

- Revenue calculated using projected quarter-by-quarter growth of 9.33% from Nvidia's projections, and assuming revenue is constant after Q1 2026 with increased competition
- Assumes 70.6% gross margin as predicted for q1 in 2026
- Assumes consistent proportion of revenue as operating expenses (historically 10% for r&d, 2.5% for SG&A)
- Assumes same ratio of net income to operating income (0.89)

2026 Increased R&D forecast:

- Revenue was determined by assuming a constant market share at 92% and a market growth of 33.4% for data center GPUs. It was calculated using 9.33% revenue growth as predicted by Nvidia for q1 in 2026 for all other revenue streams
- Assumes 70.6% gross margin as predicted for q1 in 2026
- Includes a 25% increase in R&D, from 10% of revenue to 12.5%, but assumes a consistent proportion of revenue as SG&A (2.5%)
- Assumes same ratio of net income to operating income (0.89)

Bibliography:

Barron's. (2025, March 21). *Nvidia is going big on quantum computing, and it isn't going it alone*. Barron's. https://www.barrons.com/articles/nvidia-quantum-computing-jensen-huang-e35343f5

Bradshaw, T., & Acton, M. (2025, March 20). *Nvidia to spend hundreds of billions on us supply chain, says chief*. Nvidia to spend hundreds of billions on US supply chain, says chief. https://www.ft.com/content/3fd3a717-2fbf-42ef-bb08-5baecdeb1985?utm_source=chatgpt.com

Bullfincher. (2025). *NVIDIA Corporation Revenue Breakdown By Segment*. NVIDIA Corporation Revenue Breakdown by Segment | Bullfincher. https://bullfincher.io/companies/nvidia-corporation/revenue-by-segment

DataCrunch.io. (2024, June 5). *NVIDIA H100 vs A100 GPUs – Compare price and performance for AI training and inference*. DataCrunch.io. https://datacrunch.io/blog/nvidia-h100-vs-a100

ExtremeTech. (2024, February 5). Analysts estimate NVIDIA owns 98% of the data center GPU market. Retrieved from https://www.extremetech.com/computing/analysts-estimate-nvidia-owns-98-of-the-data-center-gpu-market

Fernandez, J. (2025, March 4). The leading generative AI Companies. IoT Analytics. https://iot-analytics.com/leading-generative-ai-companies/

Gartner. (2024). NVIDIA alternatives in data center and cloud networking. Retrieved from https://www.gartner.com/reviews/market/data-center-and-cloud-networking/vendor/nvidia/alternatives

Harrington, M. J. (2025, January 26). Nvidia at a crossroads - A value investor's analysis of intensifying competition. NVIDIA At A Crossroads - A Value Investor's Analysis Of Intensifying Competition. https://www.barrons-independent.com/nvidia-at-a-crossroads-a-value-investors-analysis-of-intensifying-competition/

Insights of Nature. (2023). What does it actually take to run ChatGPT? Retrieved from https://medium.com/insights-of-nature/what-does-it-actually-takes-to-run-chatgpt-6fe9b891c691

Krumholz, M. (2025, February 24). *The exponential edge: How network effects are shaping AI and Nvidia's future*. LinkedIn. https://www.linkedin.com/pulse/exponential-edge-how-network-effects-shaping-ai-nvidias-krumholz-iujvc

Lazo, M. (2025, January 19). *The Network Effect Multiplier: Lessons from NVIDIA's Trillion-Dollar Rise*. Datalearningscience.com; Data Learning Science. https://datalearningscience.com/p/the-billion-dollar-secret-of-network

Moorhead, P. (2022, March 22). *NVIDIA H100 GPU performance shatters machine learning benchmarks for model training*. Moor Insights & Strategy.

NVIDIA. (2025, March 18). Nvidia announces financial results for fourth quarter and fiscal 2025. NVIDIA Newsroom. https://nvidianews.nvidia.com/news/nvidia-announces-financial-results-for-fourth-quarter-and-fiscal-2025

NVIDIA Corporation: History. (2018). NVIDIA. https://www.nvidia.com/en-us/about-nvidia/corporate-timeline/

NVIDIA Research and Development Expenses 2010-2025 | NVDA. (2025). Macrotrends.net. https://www.macrotrends.net/stocks/charts/NVDA/nvidia/research-development-expenses#:~:text=NVIDIA%20annual%20research%20and%20development,a%2039.31%25%20increase%20from%202022

NVIDIA. (2017). NVIDIA launches revolutionary Volta GPU platform, fueling next era of AI and high performance computing. Retrieved from https://www.globenewswire.com/news-release/2017/05/10/1149756/0/en/NVIDIA-Launches-Revolutionary-Volta-GPU-Platform-Fueling-Next-Era-of-AI-and-High-Performance-Computing.html

Review, H. B. (2024, October 7). Nvidia (NVDA) : Porter's five forces industry and Competition Analysis (2nd edition). NVIDIA (NVDA) : Porter's Five Forces Industry and Competition Analysis (2nd Edition). https://www.hivelr.com/2024/10/nvidia-nvda-porters-five-forces-industry-and-competition-analysis-2nd-edition/

Shilov, A. (2024, June 6). *Nvidia's grasp of desktop GPU market balloons to 88% — AMD has just 12%, Intel negligible, says JPR*. Yahoo Finance. https://finance.yahoo.com/news/nvidias-grasp-desktop-gpu-market-135746789.html

Shilov, A. (2023, September 13). *Jensen Huang's Unusual Nvidia Management Style Doesn't Include Long-Term Plans or Reports*. Tom's Hardware. https://www.tomshardware.com/news/nvidia-ceo-shares-management-style-always-learn-make-no-plans#:~:text=Tech%20Industry-,Jensen%20Huang's%20Unusual%20Nvidia%20Management%20Style%20Doesn't,Long%2DTerm%20Plans%20or%20Reports&text=No%20plans%2C%20no%20reports%2C%20no%20clear%20hierarchy.&text=Nvidia%20this%20year%20became%20the,that%20Huang%20manages%20his%20company

Shah, A. (2024, June 10). *Nvidia shipped 3.76 million data-center GPUs in 2023, according to study*. HPCwire.

TBRC. (2025, January). Data Center Graphics Processing Unit (GPU) market share report 2025. The Business Research Company. https://www.thebusinessresearchcompany.com/report/data-center-graphics-processing-unit-gpu-global-market-report

The4. (2025). *NVIDIA Corporation (NVDA): VRIO Analysis [Jan-2025 Updated]*. DCF, SWOT, CANVAS, PESTEL, BCG Editable Templates. https://dcfmodeling.com/products/nvda-vrio-analysis#:~:text=NVIDIA%20Corporation%20(NVDA)%20%2D%20VRIO%20Analysis%3A%20Brand%20Reputation&text=NVIDIA%20reported%20%2426.91%20billion%20revenue,in%20discrete%20graphics%20processing%20units

TipRanks. (2025, February 10). OpenAI looks to reduce reliance on Nvidia (NVDA) by making its own AI chips. *Markets Insider*. https://markets.businessinsider.com/news/stocks/openai-looks-to-reduce-reliance-on-nvidia-nvda-by-making-its-own-ai-chips-1034336546

Wilson, H. (2024, December 23). *NVIDIA Facts and Statistics (2025)*. Investing.com. https://www.investing.com/academy/statistics/nvidia-facts-and-statistics/

Wilis, T. (2024, February 22). *Nvidia's CUDA software gives it a key edge in AI computing*. Investor's Business Daily. https://www.investors.com/news/technology/nvidia-stock-nvda-cuda-software-ai-artificial-intelligence/