

데스크탑 PC CPU, GPU 성능 변화와 게임 사양의 변화 분석

3학년 B반 202144070 강지승

제출일: 2025년 12월 07일

GitHub Repository: [<https://github.com/gelatinjelly/Hardware-Performance-Trend-Analysis/>]

1. 프로젝트 개요

최근 데스크탑 PC의 CPU·GPU 성능은 세대가 바뀔 때마다 큰 폭으로 향상되었고, 동시에 최신 게임들은 보다 높은 그래픽 품질과 물리 연산을 요구하게 되었다.

하지만 일반 사용자 입장에서 “내가 새로 사려는 PC로 앞으로 나올 게임들을 충분히 돌릴 수 있을까?”를 판단하기 쉽지 않다.

이 프로젝트의 목적은 다음과 같다.

1. CPU·GPU 성능의 세대별 변화 추세를 정량적으로 분석한다

2. PC 게임들의 요구 사양을 대규모로 수집·정규화 하고, 시대별 경향을 파악한다.

3. 2025년 기준 주요 기대작들의 요구 사양을 실제 PassMark 점수와 연결하여, 앞으로 몇 년간 무난하게 게임을 즐기기 위한 적적 PC 성능 기준을 제안한다.

이를 위해 Wikipedia, PassMark, PCGamingWiki 등에서 웹 스크레이핑을 통해 데이터를 자동 수집하고, 파이썬 기반으로 전처리·변환·시각화를 수행하였다.

2. 데이터 수집

2.1 CPU·GPU 스펙 및 성능 데이터

2.1.1 Wikipedia - CPU / GPU 목록

- Intel Core / AMD Ryzen 주요 마이크로아키텍처 및 제품군 페이지 다수에서 Model, Release/Launch Date, MSRP 컬럼이 포함된 테이블을 자동 수집.
- GPU 역시 NVIDIA GeForce GTX/RTX 시리즈, AMD Radeon RX 시리즈 관련 페이지를 대상으로 스크레이핑.
- pandas.read_html()과 BeautifulSoup를 조합해, 페이지마다 제각각인 테이블 구조를 자동으로 인식하도록 구현:
 - Model / Processor / SKU / Name 등 다양한 헤더 이름을 공통 “Model”로 매핑
 - Release / Launch / Date 컬럼을 인식하여 출시일 추출
 - Price / MSRP / USD 컬럼에서 가격 정보 추출
- 멀티 인덱스 헤더, 병합 셀(rowspan) 등이 섞여 있어, 브랜드 컬럼을 앞에서부터 ffill() 하는 등 구조 적응형 로직을 사용.

2.1.2 PassMark - CPU/GPU 벤치마크 점수

- CPU:
 - https://www.cpubenchmark.net/high_end_cpus.html (멀티코어 점수 + 가격)
 - <https://www.cpubenchmark.net/singleThread.html> (싱글코어 점수)
- GPU:
 - https://www.videocardbenchmark.net/high_end_gpus.html (3D 성능 점수 + 가격)
- HTML 내 ul.chartlist > li 구조에서
 - .prdname → 모델명
 - .count → 점수
 - .price-new / .price-neww → 가격를 추출하여 Name / Multi_Score / Single_Score / G3D_Score / Price 형태의 DataFrame으로 저장.
- 네트워크 오류나 4xx 응답에 대비해, 동일 페이지를 로컬 HTML 파일로 저장 후 재파싱하는 백업 코드도 구현하였다.

2.1.3 데스크톱 전용 필터링

- 수집된 CPU·GPU에는 모바일, 서버, 임베디드용 제품도 섞여 있기 때문에, 본 프로젝트의 목적에 맞춰 데스크톱용 모델만 남기도록 필터링했다.
- CPU:
 - 제거 대상: Mobile, Laptop, Tegra, Snapdragon, Xeon, EPYC, Atom 등과 U / H / HS / HX / T /

GE와 같이 모바일·저전력 SKU에 흔히 쓰이는 접미사.

- 예외 처리:
 - Ryzen 5 5600G 같이 실제 데스크톱 APU인 모델은 유지
 - Pentium G5400 처럼 데스크톱 G 시리즈는 살리고, G5400T 같은 저전력 T 모델만 제거
- GPU:
 - 제거 대상: Mobile, Laptop, Max-Q, Integrated, Quadro, FirePro, Tesla, GRID 등 노트북/워크스테이션/서버 GPU, 및 960M 같이 숫자+M 패턴.

그 결과, 데스크톱용 CPU·GPU의 출시 연도·MSRP·벤치마크 점수를 갖춘 정제된 데이터셋을 확보하였다.

2.2 게임 요구 사양 데이터

2.2.1 PCGamingWiki 게임 목록 수집

- MediaWiki API를 사용하여 Category:Games의 모든 멤버를 페이지네이션을 통해 반복 요청.
- cmlimit=500 으로 한 번에 최대 500개씩 가져오고, continue.cmcontinue 값을 이용 카테고리 전체가 끝날 때까지 연속 호출.
- 이렇게 얻은 수만 개 규모의 게임 타이틀을 pc_games_list.csv 로 저장.

2.2.2 개별 게임 페이지에서 시스템 요구사항 추출

- 각 게임 제목에 대해 다시 API 요청:
 - prop=revisions, rvprop=content 옵션으로 위키텍스트 원문을 가져온다.
- 위키 본문에서 {{System requirements ...}} 템플릿 블록을 정규표현식으로 찾고, 그 내부를 파라미터 단위로 파싱: | minCPU = ..., | minRAM = ..., | recGPU2 = ... 등
- 키 이름을 통일하여 아래와 같은 필드를 추출:
 - 최소 사양: OS (Min), CPU (Min), GPU (Min), RAM (Min), Storage (Min)
 - 권장 사양: OS (Rec), CPU (Rec), GPU (Rec), RAM (Rec), Storage (Rec)
- 위키 마크업 [[링크]], <ref>,
 등은 모두 제거하여 순수 텍스트만 남김.
- 중간에 작업이 끊겨도 이어 할 수 있도록,
 - 이미 OUTPUT_FILE에 기록된 게임 타이틀은 completed_games로 인식해 건너뛰는 resume 기능을 구현했다.
- System requirements 템플릿이 없는 경우, 페이지 없음, 내용 없음 등 상황별로 ☒ / ☐ / ☒ 상태를 로그로 남겼다.

2.2.3 게임 사양 데이터 정리

- 중복된 게임 행 제거 (Game Title 기준).
- 최소/권장 사양이 모두 완전히 비어 있는 “깍뎀기” 행 제거.
- 수집된 결과를 pc_games_specs_fixed.csv → 정제 후 pc_games_specs_normalized.csv로 저장.

3. 데이터 전처리·변환 과정

3.1 CPU·GPU 스펙 정규화

3.1.1 모델명에서 날짜·가격·잡음 제거

- 일부 위키 테이블은 모델명 옆에 (January 2024) 같은 날짜나 \$399 같은 가격이 붙어 있는 형태.
- remove_date_safely() 함수로 다음을 제거:
 - (Jan 1, 2020), January 2020, Q1 2020, 2020-01-01 등 각종 날짜 패턴
 - US \$169, \$1,599, 2049 RMB 등 가격 표기
 - (OEM), (FB), Tray 등 옵션/판매 방식 정보
- 잔여 괄호, 쉼표, 중복 공백까지 제거하여 순수한 모델명만 남기도록 했다.

3.1.2 출시일(Release Date) 형식 통일

- 위키에 기록된 날짜는 형식이 제각각이므로 `unify_date_format()`으로 모두 YYYY-MM-DD로 변환:
 - 분기: Q1 2025 → 2025-01-01, Q3 2024 → 2024-07-01
 - 반기: H1 2025 → 2025-01-01
 - 일반 날짜: Jan 1, 2022, 2023-03-15 등은 `pd.to_datetime()`에 위임
 - 변환 불가(Unknown/TBD 등)는 None 처리
- 변환 후 최신 출시일 순으로 정렬하고 `_Final.csv` 로 저장.
- 일부 최신/특수 모델(예: Ryzen 9 9955HX3D, RTX 4090 D 등)은 문서/뉴스를 참고하여 수동으로 CPU_Data.csv, GPU_Data.csv에 출시일을 보정하였다.

3.1.3 성능 점수 통합

- CPU:
 - 멀티코어 점수와 싱글코어 점수는 서로 다른 PassMark 페이지에서 수집되므로 Name 기준 outer merge 후 Multi_Score / Single_Score를 하나의 표로 통합.
 - 모델명에서 @ 3.60GHz 같은 클럭 정보는 `remove_clock_speed()`로 제거하여 위키의 모델명과 매칭 가능한 공통 이름으로 통일.
- GPU:
 - G3D_Score는 한 페이지에서 함께 제공되므로 추가 병합 없이 사용.

3.1.4 데스크톱용 위키 목록 + PassMark 성능 데이터 병합

- CPU:
 - 위키 Model과 PassMark Name 에 대해 공통 매칭 키(Match_Key) 생성:
 - ▶ 소문자 변환 → 공백·하이픈 제거 → intel, amd 문자열 제거
 - Match_Key 기준 Left Join 후, 점수가 없는 행은 제거하여 Wiki_CPU_Cleaned.csv 완성.
- GPU:
 - 정제된 모델명을 기반으로 `create_match_key()` 적용:
 - ▶ nvidia, geforce, radeon, gpu, oc 등 브랜드/수식어 제거
 - ▶ 영문자+숫자만 남겨 키 생성
 - 마찬가지로 Left Join 후 GPU_Final_List.csv 생성.

3.2 게임 사양 데이터 정규화

3.2.1 RAM·저장공간 단위 통일 (GB 기준)

- `parse_size_to_gb()` 함수를 통해
 - 512 MB, 4GB, 1 TB 등을 모두 GB 기준 숫자로 변환.
 - TB → 1024GB, MB/KB는 1024를 기준으로 환산.
- 최소 RAM, 최소 저장공간, 권장 RAM에 각각 적용하여 RAM (Min) GB, Storage (Min) GB, RAM (Rec) GB 열 생성.

3.2.2 운영체제(OS) 분류

- `normalize_os()` 함수로 텍스트를 아래의 카테고리로 분류:
 - Win 9x/ME, Win XP, Win Vista, Win 7, Win 8, Win 10
 - 그 외는 Other 혹은 Unknown
- 이를 통해 OS 세대별 게임 출시/사양 트렌드를 비교 가능하게 만들었다.

3.2.3 사양 등급 그룹화

- RAM 구간: ~2GB, 2~4GB, 4~8GB, 8~16GB, 16GB+
- 저장공간 구간: ~1GB, 1~10GB, 10~50GB, 50GB+
- pd.cut()으로 그룹화 후, 구간별 게임 개수를 교차표로 집계하여 히트맵 시각화에 활용했다.

3.3 2025년 기대작 요구 사양 매칭

3.3.1 타겟 게임 선정 및 요구 사양 정의

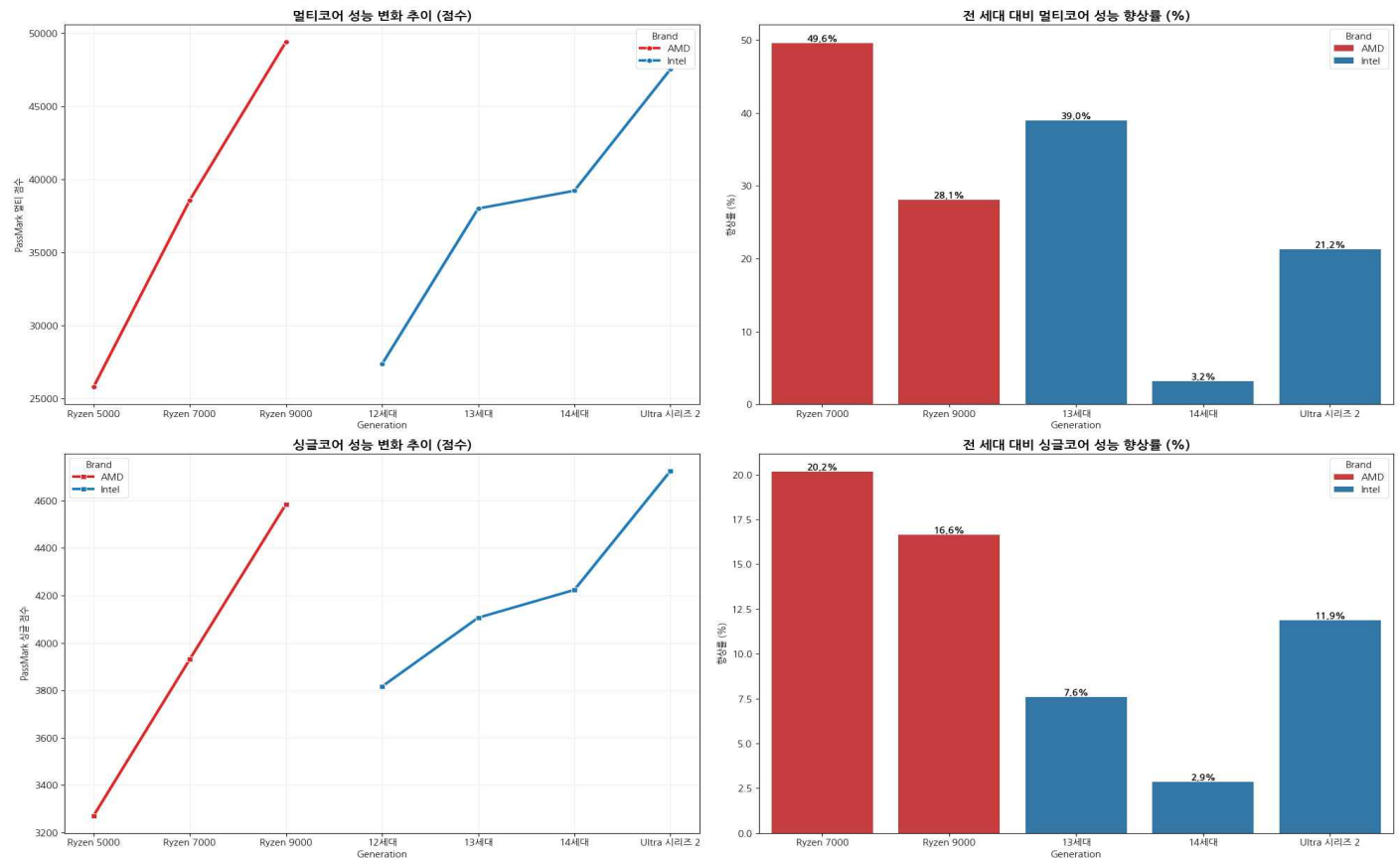
- 2025년 기준 PC 게이머들이 관심을 가질 만한 작품들을 선정:
 - GTA 6 (루머 기반 PC 사양)
 - 몬스터 헌터 와일즈
 - 문명 7
 - 킹덤 컴: 딜리버런스 2
- 각 게임에 대해 최소/권장/울트라(있는 경우) 사양을 직접 정리:
 - 예: GTA 6 권장 → CPU: Core i9-10900K / Ryzen 9 5900X, GPU: RTX 3080 / RX 6800 XT, RAM: 32GB, Storage: 150GB 등.

3.3.2 요구 사양을 PassMark 점수로 변환

- get_max_score() 함수로, CPU·GPU 모델 문자열을 CPU_Data.csv, GPU_Data.csv 내 Model과 부분 일치 검색.
- 찾은 행의 Multi_Score(CPU), G3D_Score(GPU) 중 최대값을 해당 게임·사양의 점수로 사용.
- 이렇게 만들어진 df_reqs에는 게임별로
 - CPU_Score, GPU_Score, RAM, Storage 정보가 담기며, 이후 바 차트로 시각화 및 “목표 사양 계산”에 사용된다.

4. 시각화 및 분석 결과

4.1 CPU 세대별 성능 변화 (멀티/싱글 코어)



AMD의 멀티코어 성능은 Ryzen 5000 → 7000에서 약 49.6% 라는 매우 큰 폭의 상승을 보였고, Ryzen 7000 → 9000 구간에서도 28.1%의 향상률로 꾸준한 성장세를 유지했다.

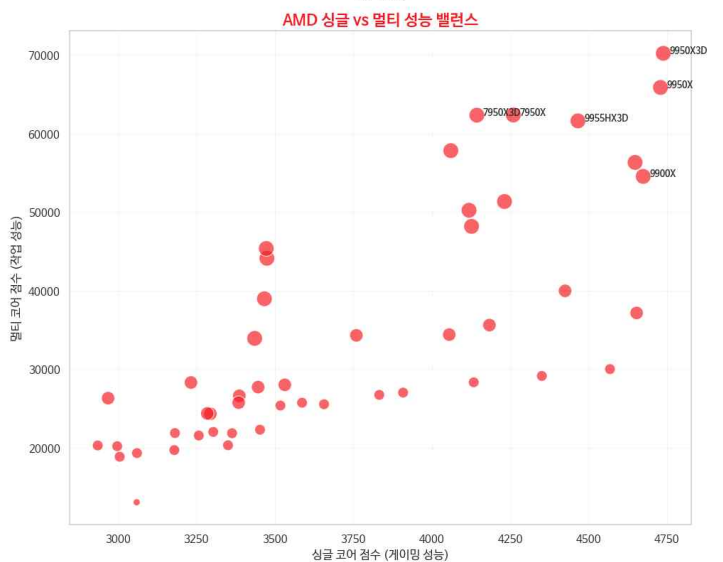
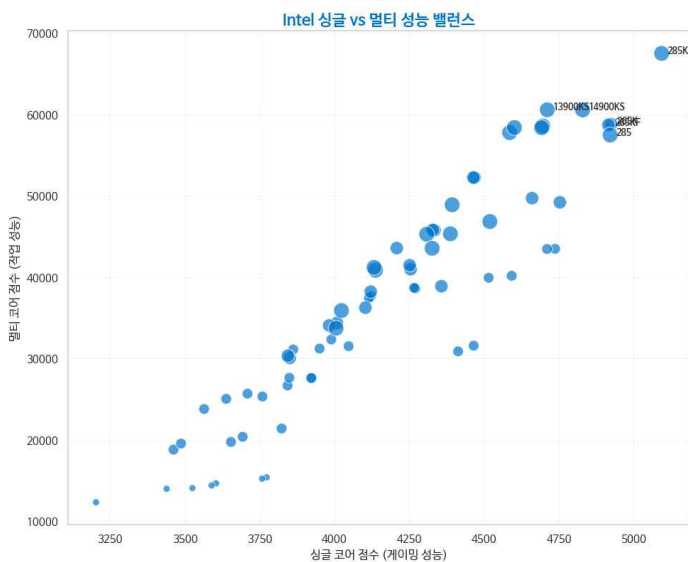
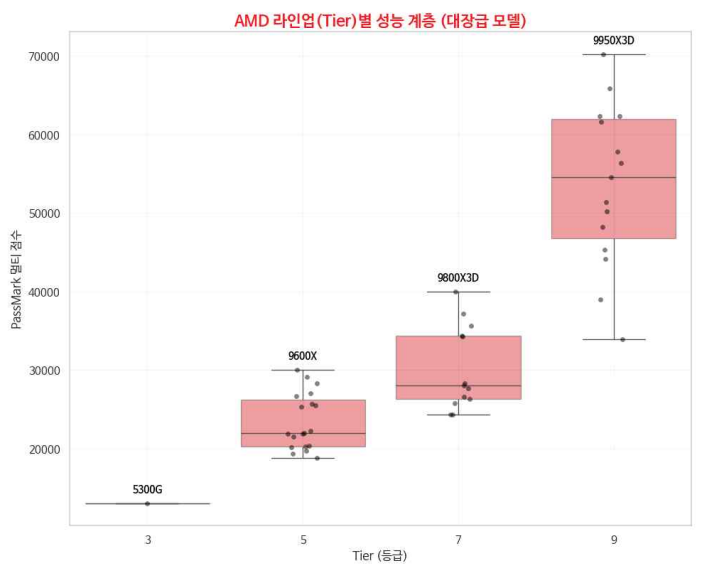
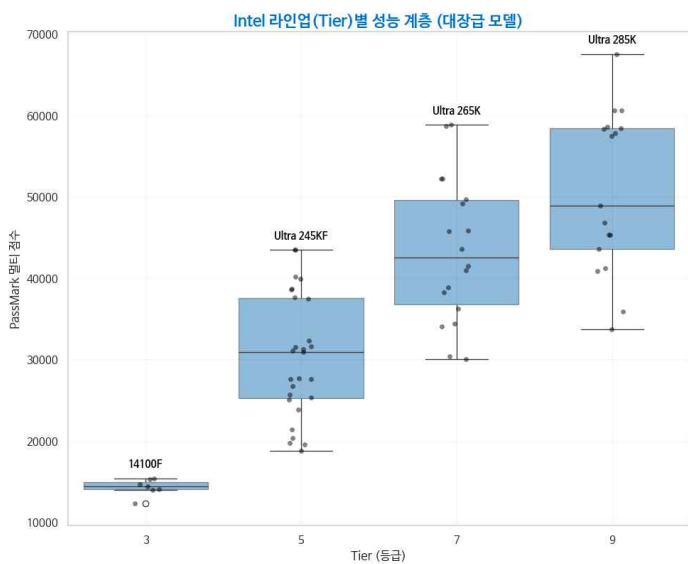
Intel의 경우 12세대 → 13세대는 멀티코어 약 39.0% 향상으로 경쟁력이 크게 상승했지만, 13세대 → 14세대 향상률이 3.2%에 그쳐 정체 구간을 보여준다.

이후 Ultra 시리즈 2에 이르러 다시 21.2% 수준의 향상률을 보이며 성능을 끌어올렸다.

싱글코어에서도 AMD는 Ryzen 5000 → 7000은 +20.2%, Ryzen 7000 → 9000은 +16.6%로 꾸준한 단일 코어 성능 향상이 관찰된다.

반면 Intel은 세대 간 향상폭이 AMD에 비해 상대적으로 낮지만, 최신 세대 기준으로는 여전히 높은 절대 성능을 유지하고 있다.

4.2 CPU 라인업·티어별 밸런스 분석



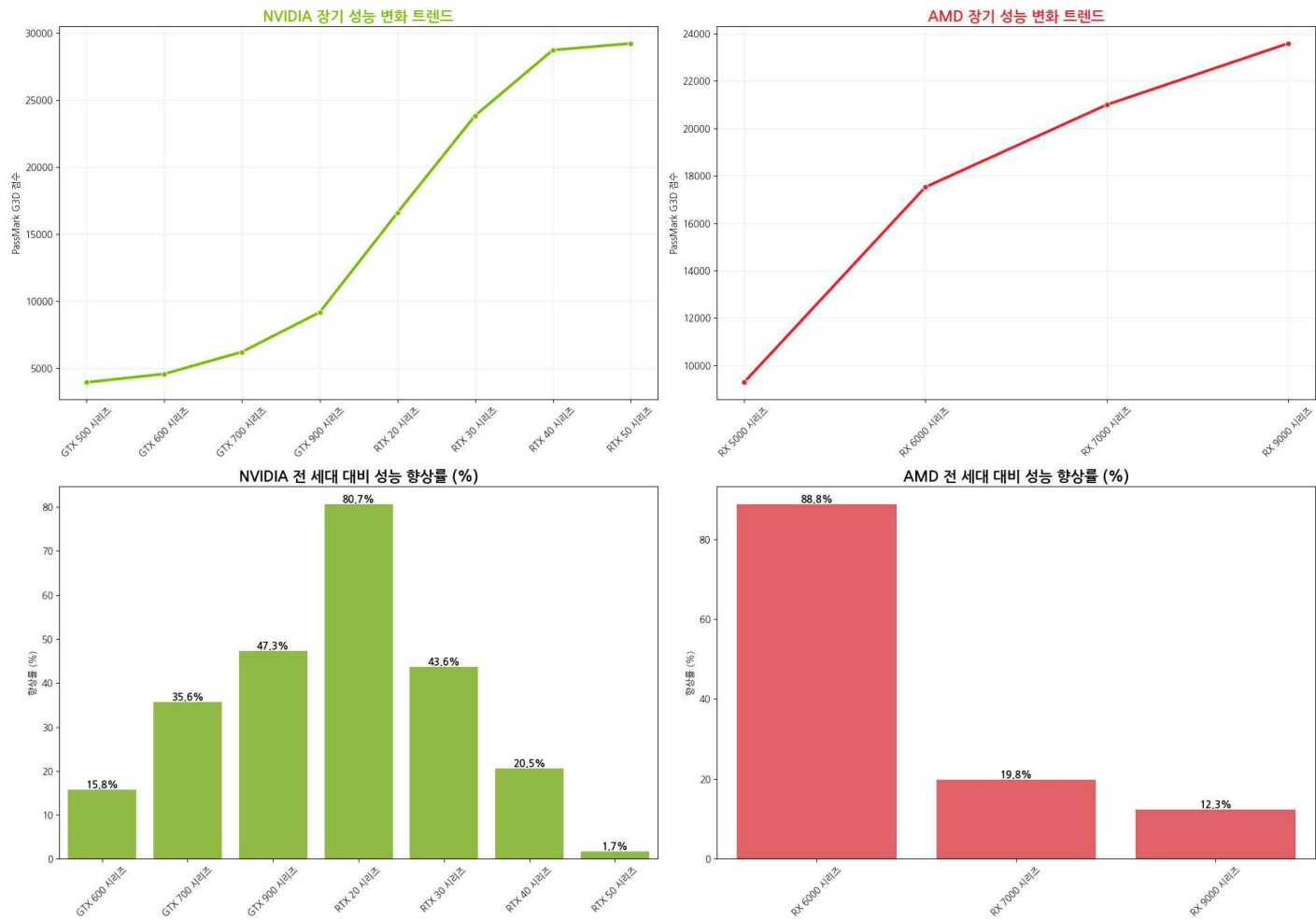
티어별 박스 플롯을 보면, Intel은 라인업이 올라갈수록 성능 분포가 꾸준히 상향 이동하며, 각 티어 내에서도 비교적 좁은 분산을 보여 제품 포지셔닝이 명확하다.

AMD는 X3D 모델이 포함된 상위 티어에서 멀티코어 및 게임 성능이 모두 매우 높게 분포하고, 같은 티어 내에서도 모델별 편차가 다소 큰 편이다.

싱글 vs 멀티 산점도에서는 양 브랜드 모두 싱글·멀티 성능이 함께 올라가는 우상향 구조를 보이지만, AMD X3D

계열은 싱글 점수 대비 멀티 점수가 특히 높은 “게임+멀티” 특화 포인트가 나타난다.
Intel의 K/KF 상위 SKU는 멀티 성능뿐 아니라 싱글 코어 점수도 높은 영역에 위치해 고주사율 게임 + 작업용으로 적합한 균형을 보여준다.

4.3 GPU 세대별 성능 변화

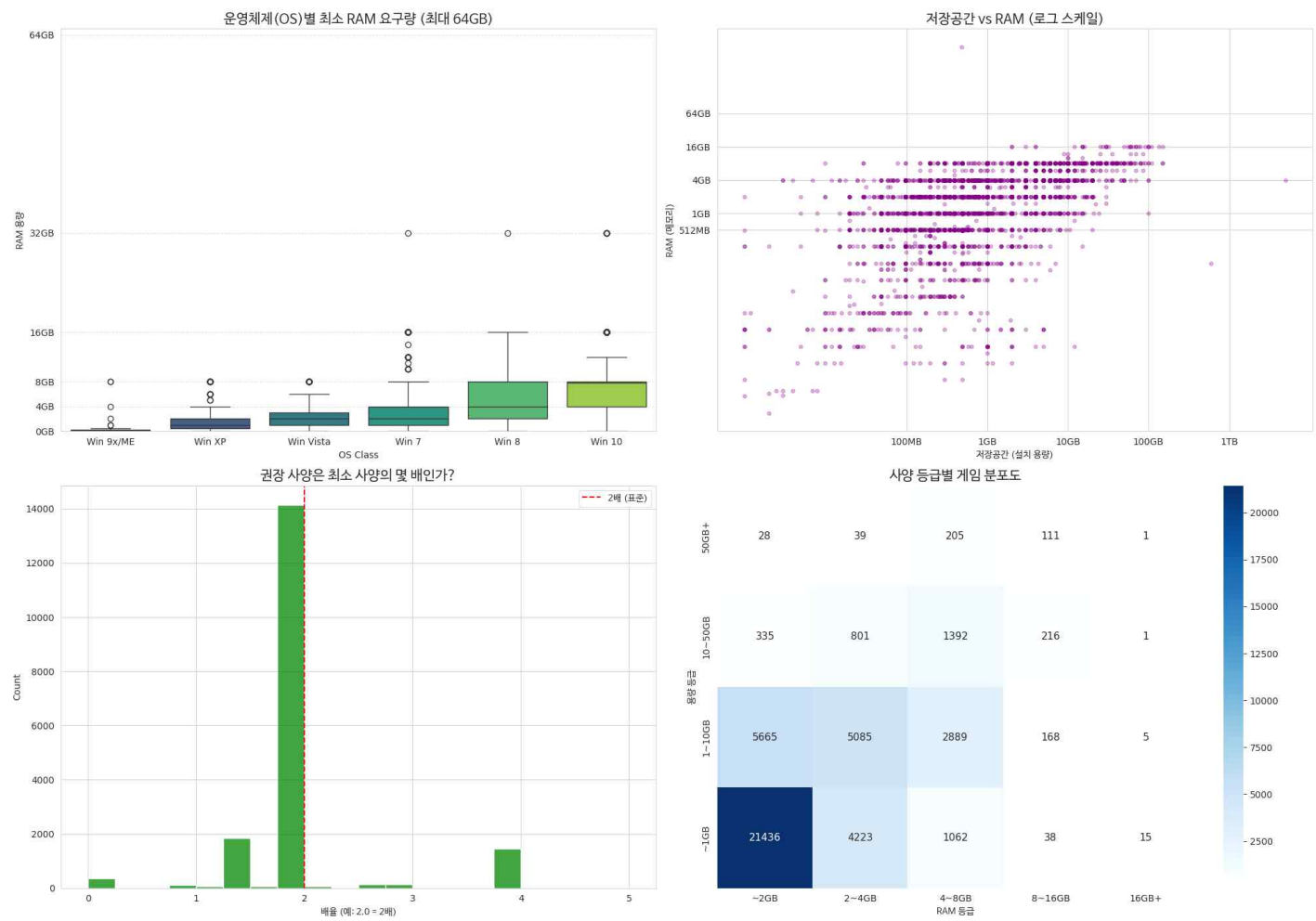


NVIDIA는 GTX 500 → 900 시리즈까지는 완만한 상승 곡선을 보이다가, RTX 20 시리즈(튜링)에서 80% 수준의 큰 폭의 점프가 발생한다. 이후 RTX 30 시리즈에서도 40%대 향상이 이어지지만, RTX 40 → 50 시리즈에 이르러 향상률이 20% → 1.7% 수준으로 급격히 둔화된다.

AMD RX 5000 → 6000 사이에서 약 ****88.8%****의 매우 큰 향상률을 기록하며 RTX 30과 경쟁했으나 RX 6000 → 7000 → 9000으로 갈수록 전 세대 대비 증가폭이 각각 19.8%, 12.3%로 점차 줄어드는 모습을 보인다.

결론적으로, 양사 모두 최근 세대에서는 성능 증가율이 점차 둔화되는 “성숙기”에 접어든 상태임을 알 수 있다.

4.4 게임 요구 사양 분포 분석 (대규모 PCGamingWiki 데이터)



주요 관찰 결과 OS 세대가 올라갈수록 최소 RAM 요구량이 꾸준히 증가하고 있으며, Win 9x/ME 시대 게임은 대부분 512MB~1GB 이하였다. Win 7 시점부터 4GB 이상 요구가 본격적으로 등장, Win 10 기반 게임들은 8GB를 요구하는 비중이 매우 높다.

저장공간과 RAM의 상관관계 산점도(로그 스케일)를 보면, 설치 용량이 수십 GB를 넘는大作일수록 RAM 요구량도 대체로 큰 편이다. 다만 인디 게임 등 예외도 존재해 상관계수는 완전히 높진 않다.

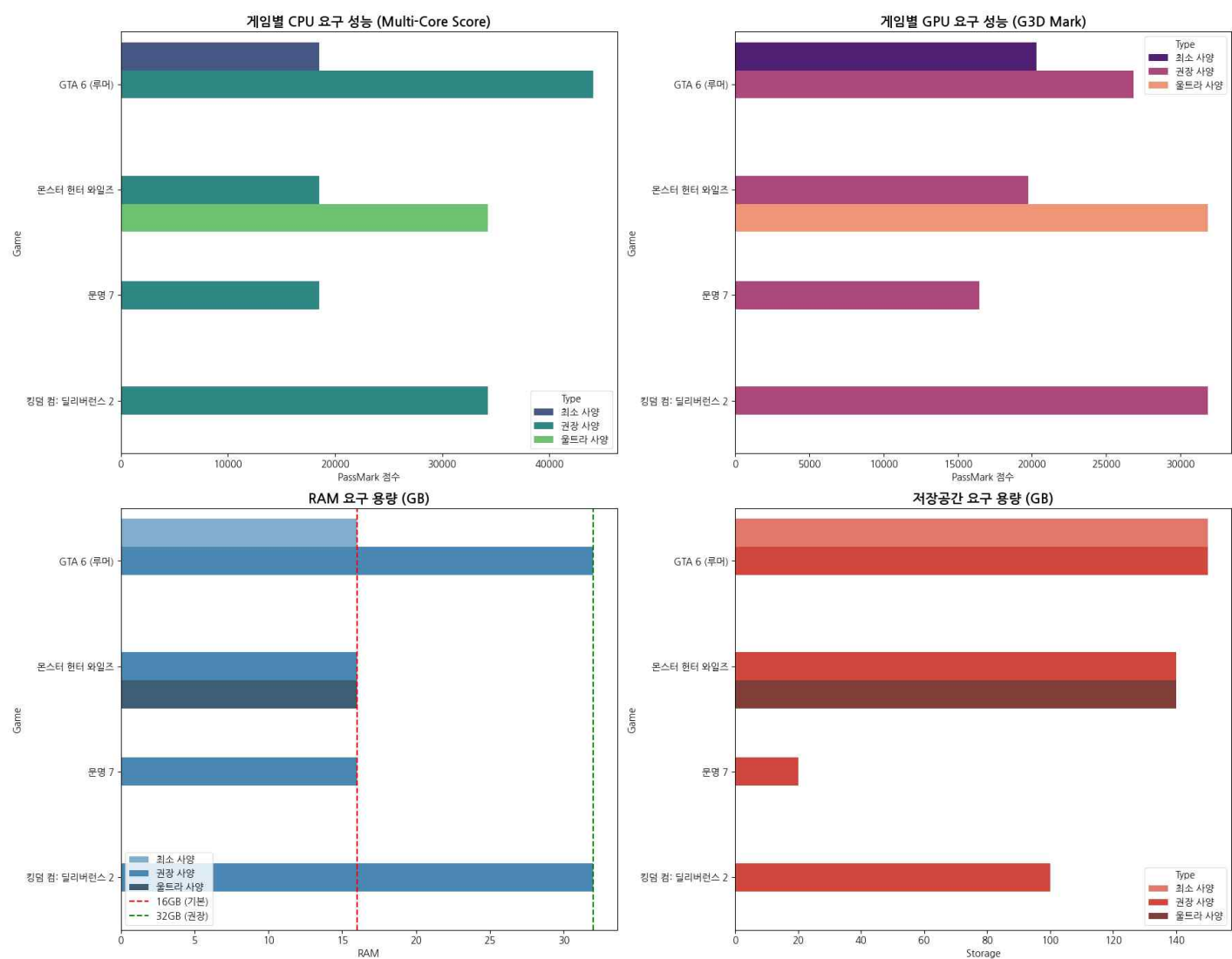
RAM 기준 권장/최소 비율 Ratio를 보면, 대부분의 게임의 권장 사양은 최소 사양이 2.0 배 근처에 많이 몰려 있다. (예: 최소 8GB, 권장 16GB) 3배 이상 요구하는 게임은 드물며 특수 케이스에 가깝다.

사양 등급별 게임 분포 히트맵에서 가장 짙게 색이 칠해진 영역은

- RAM: ~2GB, 2~4GB, 4~8GB 구간
- 저장공간: ~1GB, 1~10GB, 10~50GB 구간.

최근 대작이 늘어났음에도, 전체 게임 풀에서는 여전히 “중간 사양” 게임이 대부분이라는 점을 확인할 수 있다.

4.5 2025년 기대작 기준 권장 PC 사양 도출



선정 타이틀은 GTA 6 (루머), 몬스터 헌터 와일즈, 문명 7, 킹덤 컴: 딜리버런스 2이며, 분석 방법은 각 게임의 최소·권장·울트라 사양에 명시된 CPU/GPU 이름을 CPU_Data.csv, GPU_Data.csv에서 검색하여 PassMark 점수로 치환 후, 권장/울트라 사양의 최댓값을, “당분간 대부분의 고사양 게임을 무난히 즐길 수 있는 목표치”로 정의한다.

결과 요약:

- CPU: 분석 대상 중 최고 요구치는 킹덤 컴: 딜리버런스 2 울트라 사양 수준 기준 Multi-Core Score 기준 약 40,000점대 이상이면 상위 게임들을 커버 가능.
- GPU: RTX 4070 / RX 7800 XT 급에서 G3D Score가 상위권에 위치함으로 약 25,000~30,000점 전후의 G3D 성능이면 울트라 옵션에도 대응 가능.
- RAM: 최소 16GB, 일부 타이틀은 권장 32GB까지 요구하며, 추세상 16GB는 이제 “최소 기준”, 32GB가 사실상 권장으로 이동하는 중.
- 저장공간: GTA 6 루머 기준 150GB, 다른 대작도 100GB 이상이 흔해짐. 향후를 고려하면 1TB급 NVMe SSD가 현실적인 선택으로 보인다.

이를 종합하여, 2025년 기준 “고사양 게임까지 넉넉히 커버 가능한 권장 사양”은 다음과 같이 정리할 수 있다.

목표 CPU 점수 (Multi): 44,106 점 이상

목표 GPU 점수 (G3D): 31,829 점 이상

권장 RAM 용량: 32 GB

필요 저장공간: 150 GB (NVMe SSD)

[추천 CPU 모델 (목표 성능 충족)]

Model	Multi_Score	MSRP
Intel Core i7 13700KF	45730	\$384
Intel Core i7 13700K	45786	\$409
Intel Core i7 14700K	52216	\$409

[추천 GPU 모델 (목표 성능 충족)]

Model	G3D_Score
NVIDIA GeForce RTX 4070 Ti Super	31829
NVIDIA GeForce RTX 5070 Ti	32545
NVIDIA GeForce RTX 4080 Super	34221

5. 결론

1. CPU·GPU 성능은 여전히 상승 중이지만, 최근 세대에서는 전 세대 대비 향상률이 확연히 둔화되는 추세를 보인다. 이는 제조 공정의 한계와 아키텍처 성숙기를 반영하는 결과로 보이며, 당분간 “세대 교체 = 두 배 성능”과 같은 극적인 변화는 기대하기 어려워 보인다.
2. 대규모 게임 사양 데이터 분석 결과, 권장 RAM은 최소 RAM의 약 2배가 되는 경우가 가장 많았고, Win 10 시대 기준으로는 8GB가 최소, 16GB 이상이 권장 수준으로 자리잡았다.
3. 2025년 기대작들의 요구 사양을 PassMark 점수로 환산해 본 결과, CPU 4만점 / GPU 2.5만점 / RAM 32GB / 1TB SSD 정도면, 향후 몇 년간 출시될 대부분의 AAA 타이틀을 높은 옵션으로 즐기는 데 충분한 수준으로 판단된다.