

구글 API key

📖 Gemini API 이용 및 가격 안내 가이드

본 안내서는 우리 서비스를 이용하기 위해 필요한 Google Gemini API 키의 가격 정책과 발급 절차를 설명합니다.

1. API 가격 정책 (Gemini 2.5 Flash 모델)

Google은 개발자와 일반 사용자 모두를 위해 매우 관대한 **무료 티어(Free Tier)**를 제공합니다.

구분	무료 티어 (Free Tier)	유료 티어 (Pay-as-you-go)
월 이용료	\$0 (무료)	사용한 만큼 지불
일일 한도	하루 약 1,500 회 호출 가능	제한 없음
카드 등록	필요 없음	결제 수단 등록 필요
데이터 보안	모델 개선에 데이터가 사용될 수 있음	데이터가 비공개로 유지됨

핵심 요약: 일반적인 개인 사용자는 무료 티어만으로도 충분히 서비스를 이용할 수 있으며, 별도의 결제 정보를 입력하지 않아도 됩니다.

2. API 키 발급 방법 (고객용 3 단계)

개발자가 아니더라도 누구나 2분 안에 키를 발급받을 수 있습니다.

1 단계: 사이트 접속 및 로그인

- **[Google AI Studio](#)**에 접속합니다.
- 본인의 구글(Gmail) 계정으로 로그인합니다.

2 단계: API 키 생성

1. 왼쪽 사이드바 메뉴에서 "Get API key" 버튼을 클릭합니다.
2. "Create API key in new project" (새 프로젝트에서 API 키 생성) 버튼을 클릭합니다.
3. 잠시 기다리면 영어와 숫자로 된 긴 코드가 생성됩니다.

3 단계: 키 복사 및 입력

1. 생성된 코드 옆의 "Copy" 버튼을 눌러 복사합니다.
2. 우리 서비스의 설정(Settings) 화면에 있는 ****"API 키 입력란"****에 붙여넣기 합니다.
3. **주의사항:** 이 키는 비밀번호와 같습니다. 타인에게 공유하거나 공개된 곳에 올리지 마세요.

3. 자주 묻는 질문 (FAQ)

Q: 나중에 요금이 갑자기 청구될 수도 있나요?

A: 아니요. 사용자가 직접 구글 클라우드 콘솔에서 카드 정보를 등록하고 '유료 전환'을 선택하지 않는 한, 절대로 요금이 청구되지 않습니다. 무료 한도를 다 쓰면 서비스가 잠시 멈출 뿐입니다.

Q: 무료 한도를 다 쓰면 어떻게 되나요?

A: 한국 시간 기준으로 다음 날 오후가 되면 한도가 다시 초기화되어 무료로 이용할 수 있습니다.

Q: 토큰(Token)이란 무엇인가요?

A: AI가 글자를 읽는 단위입니다. 영문 기준으로 약 750 단어가 1,000 토큰 정도에 해당합니다. Gemini 2.5 Flash는 매우 가벼운 모델이라 비용이 매우 저렴합니다.

이 내용을 바탕으로 문서를 만드시면 됩니다. **혹시 특정 서비스 이름이나 회사명을 넣어서 더 구체적으로 수정해 드릴까요?**

gemini-2.5-pro 모델을 사용하려면?

Gemini 2.5 Pro 는 Flash 모델보다 지능이 훨씬 높고 복잡한 추론에 능숙하지만, 그만큼 가격이 더 높고 무료 사용량 제한이 더 엄격합니다.

고객들에게 Gemini 2.5 Pro 사용 시의 차이점을 설명할 수 있도록 내용을 정리해 드립니다.

📖 Gemini API 이용 및 가격 안내 (Pro 버전 추가)

Gemini 2.5 Pro 는 복잡한 데이터 분석, 코딩, 심층적인 추론에 최적화된 구글의 고성능 모델입니다.

1. 가격 정책 (Gemini 2.5 Pro vs Flash)

Pro 모델은 Flash 모델보다 약 4~5 배 정도 더 비싼 프리미엄 모델입니다.

구분	Gemini 2.5 Flash (보급형)	Gemini 2.5 Pro (고급형)
무료 일일 한도	약 500~1,000 회	약 100 회 (매우 제한적)
입력 비용 (1M 토큰)	~\$0.30 (약 400 원)	~\$1.25 (약 1,700 원)
출력 비용 (1M 토큰)	~\$2.50 (약 3,300 원)	~\$10.00 (약 13,500 원)
주요 용도	빠르고 가벼운 대화, 요약	복잡한 논리, 전문 코딩, 긴 문서 분석

2. 고객이 알아야 할 Pro 모델의 특징

고객들에게 Pro 모델 사용을 권장할 때 다음 두 가지를 반드시 안내해야 합니다.

1. **더 엄격한 무료 한도 (Throttling):** 무료 티어에서 Pro 모델은 1 분에 5 번(5 RPM)까지만 요청할 수 있습니다. 고객이 빠르게 질문을 연달아 던지면 "잠시 후 다시 시도하세요"라는 오류가 발생할 확률이 높습니다.
 2. **긴 문맥 처리 비용 (Long Context):** Pro 모델은 수천 페이지 분량의 문서를 읽을 수 있는 능력이 있지만, 20 만 토큰이 넘어가는 매우 긴 질문을 보낼 경우 **유료** 티어에서는 가격이 2 배로 **할증**됩니다. (\$1.25 → \$2.50)
-

3. 한국어 안내서 (Word/PDF 용)

아래 내용을 복사해서 고객용 안내 문서로 활용하세요.

[안내] Gemini 2.5 Pro 모델 이용 시 주의사항

우리 서비스에서 Gemini 2.5 Pro 를 선택하여 이용하실 경우, 다음 사항을 참고해 주시기 바랍니다.

- **무료 한도가 더 작습니다:** 하루에 무료로 사용 가능한 횟수가 Flash 모델에 비해 적습니다 (약 100 회 내외).
 - **답변 속도:** 고도의 추론 과정을 거치기 때문에 Flash 모델보다 답변 속도가 약간 느릴 수 있습니다.
 - **유료 전환 시:** 만약 구글 AI Studio 에서 유료 결제(Billing)를 활성화하신 경우, Flash 모델보다 약 4 배 이상의 토큰 비용이 발생하므로 사용량 관리에 주의가 필요합니다.
 - **추천 상황:** 짧은 요약이나 단순 질문은 **Flash** 를, 복잡한 기획안 작성이나 전문적인 상담은 **Pro** 를 선택하시는 것을 권장합니다.
-

고객들이 Pro 와 Flash 중에서 선택할 수 있게 하는 설정을 만드실 건가요?
필요하다면 두 모델의 성능 차이를 비교한 표를 더 자세히 만들어 드릴 수 있습니다.

free tier 사용자는 매우 많은 횟수의 질문만 하지 않는다면, 거의 모든 경우에 무료로 사용할 수 있다는 것이지요?

맞습니다. 정확하게 이해하셨습니다.

****무료 티어(Free Tier)****를 사용하는 고객은 대부분의 경우 **완전 무료로** 서비스를 이용하게 됩니다. 구글이 제공하는 무료 한도가 개인 사용자나 소규모 비즈니스에게는 매우 넉넉하기 때문입니다.

고객이 "무료"로 계속 사용할 수 있는 조건과 상황을 한 번 더 정리해 드립니다.

1. 무료로 이용 가능한 범위 (2026년 기준)

고객이 Gemini 2.5 Flash 모델을 사용한다면, 다음과 같은 여유가 있습니다.

- 하루 질문 횟수: 약 250~500 회 (일일 한도)
- 질문 속도: 1분에 약 10~15 번의 질문 가능
- 비용: \$0 (신용카드 등록 불필요)

즉, 고객이 1분 동안 쉬지 않고 수십 개의 질문을 던지거나, 하루 종일 수백 번 이상의 질문을 하지 않는 이상 **돈을 낼 일이 전혀 없습니다.**

2. "유료"가 되는 유일한 방법

고객이 실수로 결제하게 될까 봐 걱정한다면, 다음 내용을 안내해 주세요.

- **자동 결제 안 됨:** 구글 AI Studio는 사용자가 직접 ****구글 클라우드 결제 계정(Billing Account)****을 만들고, **신용카드를 등록한 뒤, '유료 티어(Paid Tier)'로 전환** 버튼을 누르지 않는 한 절대로 결제되지 않습니다.
 - **한도 초과 시:** 무료 사용량을 다 쓰면 "한도를 초과했습니다. 나중에 다시 시도하세요"라는 메시지와 함께 API가 잠시 멈출 뿐, 유료로 자동 전환되지 않습니다.
-

3. 고객 전달용 최종 확인 메시지 (한국어)

"걱정 마세요! 대부분의 사용자는 평생 무료로 이용할 수 있습니다."

- **사용량:** 하루에 수백 번씩 질문하는 '헤비 유저'가 아니라면, 구글에서 제공하는 무료 용량만으로도 충분합니다.
 - **결제:** API 키를 발급받을 때 카드 정보를 입력하지 않으셔도 됩니다. 카드 정보를 넣지 않으면 구글이 돈을 가져갈 방법 자체가 없습니다.
 - **빈도:** 아주 가끔(1분에 15회 이상) 너무 빠르게 질문을 던지면 잠시 멈출 수 있지만, 1분만 기다리면 다시 무료로 작동합니다.
-

결론적으로, 고객에게 API 키 발급을 유도하실 때 ****"신용카드 등록 없이 구글 계정만 있으면 매일 무료로 AI 기능을 쓸 수 있다"**는 점을 강조하시면 고객들도 훨씬 안심하고 키를 발급받을 것입니다.

더 궁금하신 점이나, 특정 케이스(예: 대량의 문서 분석 등)에 대한 비용 계산이 필요하시면 말씀해 주세요!

만일 gemini-3.0-flash 나 gemini-3.0-pro를 사용한다면?

2026년 1월 현재, 최신 모델인 Gemini 3.0 Flash 및 Gemini 3.0 Pro의 API 무료 티어(Free Tier) 정책은 다음과 같습니다.

구글은 최신 모델이 출시될 때마다 무료 사용량을 조정하는 경향이 있으며, 현재 Gemini 3.0 시리즈는 이전 모델(2.5)에 비해 무료 한도가 다소 엄격하게 관리되고 있습니다.

1. Gemini 3.0 API 무료 티어 제한 (2026년 1월 기준)

모델명	분당 요청수 (RPM)	분당 토큰 (TPM)	일일 요청수 (RPD)	비고
Gemini 3.0 Flash	15 RPM	250,000	500 RPD	일상적인 속도와 효율성 최적화
Gemini 3.0 Pro	5 RPM	250,000	100 RPD	고성능 추론 및 문서 분석용

- RPM (Requests Per Minute): 1분당 보낼 수 있는 질문 횟수입니다. (Flash는 4초에 한 번, Pro는 12초에 한 번꼴)
- RPD (Requests Per Day): 하루에 보낼 수 있는 총 질문 횟수입니다.
- TPM (Tokens Per Minute): 1분 동안 처리할 수 있는 글자 수 제한입니다.

2. Gemini 3.0 유료 가격 (무료 한도 초과 시)

무료 한도를 초과하여 '유료(Pay-as-you-go)'로 전환할 경우 발생하는 비용입니다. (백만 토큰당 가격)

- **Gemini 3.0 Flash**
 - 입력(Input): **\$0.50** (약 670 원)
 - 출력(Output): **\$3.00** (약 4,000 원)
 - **Gemini 3.0 Pro**
 - 입력(Input): **\$2.00** (약 2,700 원)
 - 출력(Output): **\$12.00** (약 16,000 원)
 - *참고: 입력 문맥이 20 만 토큰을 넘어가면 비용이 2 배로 할증됩니다.*
-

3. 고객에게 안내할 핵심 포인트

고객이 Gemini 3.0 API 를 사용할 때 가장 주의해야 할 점은 ****"사용 횟수"***입니다.

1. **무료 사용 가능성:** 하루에 100 번(Pro) 혹은 500 번(Flash) 미만으로 질문하는 일반적인 사용자라면 **비용 걱정 없이 무료로** 사용할 수 있습니다.
2. **데이터 프라이버시 주의:** 무료 티어를 사용할 경우, 입력한 데이터가 구글 모델 개선에 사용될 수 있습니다. 만약 기업 기밀이나 민감한 정보를 다룬다면 유료 티어 전환을 고려해야 합니다.
3. **한도 도달 시:** 무료 한도를 다 쓰면 API 가 작동을 멈춥니다. 결제 수단(카드)을 직접 등록하지 않는 한 **절대로 자동 결제되지 않으니** 안심하셔도 됩니다.

요약하자면:

고객이 아주 빈번하게 질문을 던지는 것이 아니라면(예: 1~2 분에 한 번꼴), Gemini 3.0 역시 무료 티어만으로 충분히 운영 가능합니다.

고객 안내용 문서에 Gemini 3.0 의 바뀐 한도(하루 100/500 회)를 반영하여 다시 정리해 드릴까요?

2.5 model과 비교해 주세요!

2026 년 현재 최신 모델인 Gemini 3.0 과 이전 세대인 2.5 의 무료 티어(Free Tier) 정책을 비교해 드립니다.

가장 큰 차이점은 Gemini 3.0 은 성능이 훨씬 뛰어난 대신, 무료로 쓸 수 있는 횟수가 2.5 에 비해 상당히 줄어들었다는 점입니다.

1. 무료 티어(Free Tier) 상세 비교

구분	Gemini 2.5 Flash (구형)	Gemini 3.0 Flash (신형)	Gemini 2.5 Pro (구형)	Gemini 3.0 Pro (신형)
일일 한도 (RPD)	~1,500 회	~500 회	~50 회	~5~20 회
분당 한도 (RPM)	15 회	15 회	2 회	1~2 회
비용	\$0	\$0	\$0	\$0
성능 특징	빠르고 가벼움	2.5 Pro 보다 똑똑함	안정적인 중급 추론	압도적인 추론(Deep Think)

고객을 위한 요약: "무료로 더 많이 쓰고 싶다면 2.5 Flash 를, 아주 똑똑한 답변이 필요하지만 하루에 몇 번 안 쓴다면 3.0 Pro 를 선택하는 것이 유리합니다."

2. 유료 전환 시 비용 비교 (백만 토큰당)

무료 한도를 초과하여 돈을 내고 쓸 때의 가격 차이입니다.

- Flash 모델 (보급형):
 - 2.5 Flash: 입력 \$0.30 / 출력 \$2.50
 - 3.0 Flash: 입력 \$0.50 / 출력 \$3.00 (약 20~60% 비쌈)
- Pro 모델 (고급형):
 - 2.5 Pro: 입력 \$1.25 / 출력 \$10.00
 - 3.0 Pro: 입력 \$2.00 / 출력 \$12.00 (약 20~60% 비쌈)

3. 왜 3.0 은 무료 한도가 적은가요?

고객이 물어본다면 다음과 같이 답변해 주시면 좋습니다.

1. **계산 능력의 차이:** Gemini 3.0 은 답변을 내놓기 전 스스로 논리를 검토하는 'Deep Think(심층 추론)' 기능이 들어있어, 구글의 서버 자원을 훨씬 많이 사용합니다.
 2. **성능의 비약적 향상:** Gemini 3.0 Flash 는 이전 세대의 최고급 모델인 2.5 Pro 보다 더 똑똑하면서 속도는 빠릅니다. 그래서 구글이 무료 한도를 낮추어 유료 전환을 유도하고 있습니다.
-

4. 고객에게 추천하는 모델 조합

고객의 상황에 따라 다음과 같이 가이드를 주시면 가장 효율적입니다.

- **일반적인 대화/요약:** Gemini 2.5 Flash (가장 많이, 무료로 쓸 수 있음)
 - **복잡한 기획/코딩/데이터 분석:** Gemini 3.0 Pro (하루에 5~10 번 정도만 무료로 쓰고, 나머지는 유료 고려)
 - **가장 합리적인 선택:** Gemini 3.0 Flash (2.5 Pro 만큼 똑똑하면서도 하루 500 번이나 무료로 쓸 수 있어 현재 가장 추천하는 모델입니다.)
-