

BLOCK ELEMENTS

Syntax	Result	Notes
# Heading 1	H1 — largest heading	Space after # required (Co
## Heading 2	H2	Up to ##### for H6
Heading\n===	H1 setext style	=== for H1, --- for H2
- item or * item	Unordered list	- * + all work
1. item	Ordered list	Number value doesn't matte
> text	Blockquote	Nestable with >>
--- or ***	Thematic break <hr>	3+ hyphens/asterisks
indented	Code block (4 spaces)	No language tag
```lang	Fenced code block	Preferred; supports lang l

## GFM EXTENSIONS (4 OFFICIAL)

 gfm-extensions.md
# Tables
Col1   Col2
----- :----:  ← :center: / right:
data   data
# Task Lists
- [x] Done task
- [ ] Todo task
# Strikethrough
~~deprecated text~~
# Autolinks (bare URLs)
<a href="https://jsonl.ai">https://jsonl.ai</a> <-- auto-linked

## INLINE ELEMENTS

Syntax	Result	Notes
<b>**bold**</b> or <code>__bold__</code>	Bold text	** preferred
<i>*italic*</i> or <code>_italic_</code>	Italic text	* preferred in words
<b><i>***bold italic***</i></b>	Bold + italic	Triple delimiter
<code>`inline code`</code>	Code span	Use `` for backticks
<code>[text](url)</code>	Inline link	<code>[text](url 'title')</code>
<code>[text][ref]</code>	Reference link	<code>[ref]: url elsewhere</code>
<code>![alt](url)</code>	Image	Same as link + !
<code>*escaped*</code>	Literal *	Backslash any punct.
<code>&lt;https://url&gt;</code>	Autolink	CommonMark style

## STRING ESCAPE REFERENCE

<code>\\</code> Backslash	<code>*</code> Asterisk
<code>_</code> Underscore	<code>\`</code> Backtick
<code>\#</code> Hash / pound	<code>\[</code> Open bracket
<code>\]</code> Close bracket	<code>\(</code> Open paren
<code>\)</code> Close paren	<code>\!</code> Exclamation
<code>\+</code> Plus	<code>\-</code> Hyphen/dash
<code>\.</code> Period	<code>\{</code> Open brace

## MARKDOWN FLAVORS

**Original** Gruber · 2004  
Intentionally underspecified

**CommonMark** 0.31.2 · 2024  
Unambiguous, 500+ tests

**GFM** GitHub · 2017  
CommonMark + 4 extensions

**MDX** Community · 2018  
Markdown + JSX components

## WHY EVERY LLM OUTPUTS MARKDOWN

### Token Efficiency

~70% fewer tokens than HTML for same structure

### Training Data

GitHub, StackOverflow, Reddit = mostly Markdown

### Structural Clarity

# = topic, ## = subtopic, - = list item

### Human Readable

Readable as plain text without rendering

### Universal Support

Every major platform renders it correctly

## PYTHON: PARSE & PROCESS MARKDOWN

```

 md-utilities.py
pip install markdown-it-py
from markdown_it import MarkdownIt

md = MarkdownIt()
html = md.render(markdown_string)

Extract code blocks from LLM output
import re
pattern = r'```(\w*)\n(.*?)```'
blocks = re.findall(pattern, md,
 flags=re.DOTALL)

Strip Markdown → plain text
text = re.sub(r'\*[1,3](.+?)\*[1,3]',
 r'\1', text)

```

## COMMON MISTAKES

**No space after #**  
# Title ✓ vs #Title X (CommonMark)

**Underscore in words**  
Use *asterisks* not _underscores_ in words

**No blank lines**  
Block elements need blank lines before/after

**Mixing list markers**  
Use only - or only * within one list

**Tabs vs spaces**  
CommonMark defines tabs specially — use spaces

**No comments**  
HTML <!-- --> works but not portable