

Basics

One two *three* four **five** six.

Seven eight

Nine

ten eleven

$\sqrt{\text{don't mind me}}$

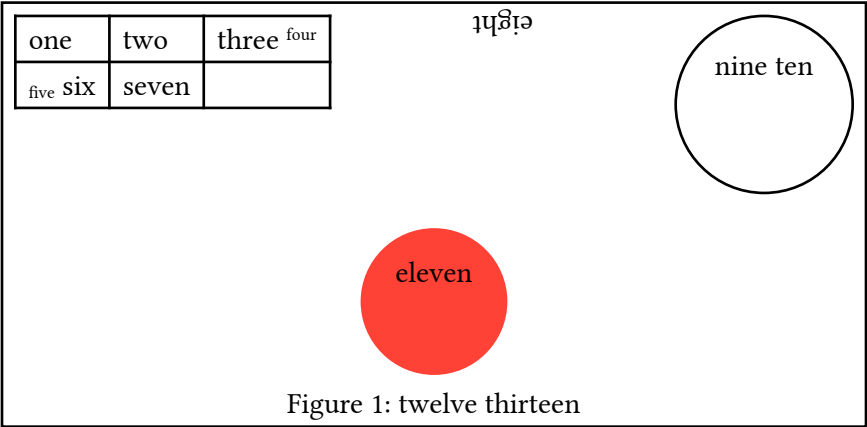
twelve.

Thirteen **fourteen**

- **fifteen**
- sixteen seventeen
- eighteen!

(characters: 101, words: 18, sentences: 3)

More basics



```
(characters: 59, words: 13, sentences: 0) (  
  (  
    "one",  
    "two",  
    ("three ", "four"),  
    ("five", " six"),  
    "seven",  
  ),  
  "eight",  
  "nine ten",  
) ,  
("eleven", "twelve thirteen"),  
)
```

Punctuation

“One **two**, three!” **Four**, five.

Six, **seven**, eight.

Raw tree: (

```
" \"One ",  
  "two",  
  ", three!\" ",  
  "Four",  
  ", five. ",  
  ("Six, ", "seven", ", eight."),  
)
```

Stats: (characters: 41, words: 8, sentences: 3)

Word edge cases

One two three [Words: 3] Acronyms count as O.N.E. word! [Words: 5]

Hyphen-words [Words: 1] are one, but En-Dash [Words: 2] are two.

Punctuation doesn't count !? [Words: 3] Qu'est-ce que c'est ? [Words: 3]

The amount is \$4,599.99! [Words: 4] One (or so). [Words: 3]

Scoped counts

Guess what, this box contains 15 words!

Full statistics: (characters: 81, words: 15, sentences: 1)

One two three four. There are 10 total words and 48 characters.

Master function

Hello, stats are in! (characters: 51, words: 10, sentences: 1)

One two three four. There are 10 total words and 48 characters.

Sentences

Pour quoi ? Qu'est-ce que c'est !?

“I don't know anything.”

(characters: 49, words: 9, sentences: 3)

Excluding elements by type

Not me.

One, two, three. ~~Not me, either.~~ Four.

~~Words: 4~~

Excluding elements by label

One two

Three, four.

Not me!

Five, six.

(characters: 26, words: 6, sentences: 2)

One, two, three, four. That was 4, not counting this sentence!

Where-selectors

One

Not me!

Two three four five

(characters: 19, words: 5, sentences: 0))

Custom counters

Hello there are vowels here.

(vowels: 10)

Raw text extraction

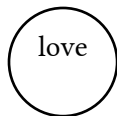
Hello, *this* is a bunch of content.

New line.

What we have

- This and that
- etcetera

what we **need**



Hello, this is a bunch of content .

New line.

What we have This and that etcetera what we need love

Issue #1

Figures might have:

- no caption field, or
- a caption field with the value none.

Hello from the figure body.

Ciao.

(characters: 28, words: 6, sentences: 2)