import re

input\_file = "preproinsulin-seq.txt"

output\_file = "preproinsulin-seq-clean.txt"

with open(input\_file, "r") as f:

text = f.read()

print(text)

count\_letters = 0

for i in text:

count\_letters += 1

print (count\_letters, "\n")

# Remove all characters, except small letters;

text = re.sub(r'[^a-z]', '', text)

print(text, "\n")

count\_letters = 0

for i in text:

count\_letters += 1

print (count\_letters)

with open(output\_file, "w") as f:

f.write(text)

Ctrl+k+c to make a comment

Ctrl+k+u to unmake a comment