## 8 consejos de procesamiento de archivos en Python

|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | | |  | | --- | | 1. Al [escribir archivos](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vcHltLmRldi9jcmVhdGluZy1hbmQtd3JpdGluZy1maWxlLXB5dGhvbi8_X19zPWJpaTA1N21najA0Y3p4Z3JrYnVsJnV0bV9zb3VyY2U9ZHJpcCZ1dG1fbWVkaXVtPWVtYWlsJnV0bV9jYW1wYWlnbj13ZWVrbHkrZW1haWwrLSsyMDIzLTA3LTI2JnV0bV9jb250ZW50PVdlZWtseStQeXRob24rdGlwJTNBKzgrZmlsZStwcm9jZXNzaW5nK3RpcHMifQ.5P-Fl0AYpXZRUPEWpftPPwmvhf7inY2s21iQ9iW2efA) (e idealmente al leerlos también) [use un bloque with](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9..2ExQn6jQSF_h5iV6eupv1_x5esnKlMs-ElaA2iqQmLQ) para cerrar automáticamente su archivo cuando haya terminado de trabajar con ellos. 2. Cuando trabaje con archivos de texto muy grandes, [procese el archivo línea por línea](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vcHltLmRldi9yZWFkaW5nLWZpbGVzLWxpbmUtbGluZS8_X19zPWJpaTA1N21najA0Y3p4Z3JrYnVsJnV0bV9zb3VyY2U9ZHJpcCZ1dG1fbWVkaXVtPWVtYWlsJnV0bV9jYW1wYWlnbj13ZWVrbHkrZW1haWwrLSsyMDIzLTA3LTI2JnV0bV9jb250ZW50PVdlZWtseStQeXRob24rdGlwJTNBKzgrZmlsZStwcm9jZXNzaW5nK3RpcHMifQ.L3LEZSmAUsBP7Kdja-qsFk1bFwFNMbyBhA615VsSsZ8) haciendo un bucle sobre él (esto solo almacenará 8 KB del archivo en la memoria a la vez gracias a la forma en que funciona el almacenamiento en búfer de archivos). 3. Puede procesar archivos binarios grandes [fragmento por fragmento](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9..z5PcGDsmsUyeXFP4jcuvsjwys5mdmIGuwDBN89w2d2c) para evitar leerlos en la memoria todos a la vez. 4. Si es posible que sus archivos de texto no estén en UTF-8, asegúrese de [especificar la codificación de sus archivos](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vcHltLmRldi91bmljb2RlLWNoYXJhY3Rlci1lbmNvZGluZ3MtaW4tcHl0aG9uLz9fX3M9YmlpMDU3bWdqMDRjenhncmtidWwmdXRtX3NvdXJjZT1kcmlwJnV0bV9tZWRpdW09ZW1haWwmdXRtX2NhbXBhaWduPXdlZWtseStlbWFpbCstKzIwMjMtMDctMjYmdXRtX2NvbnRlbnQ9V2Vla2x5K1B5dGhvbit0aXAlM0ErOCtmaWxlK3Byb2Nlc3NpbmcrdGlwcyJ9.AWreL0GTRYrURJMiN2io3tcONtTYvyVTc0qxxrjAtdw) al abrirlos. 5. Cuando trabaje con archivos que no son de confianza que pueden tener líneas extremadamente largas, en lugar de hacer un bucle línea por línea, llame al método readline con un [tamaño máximo](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vdHdpdHRlci5jb20vdHJleWh1bm5lci9zdGF0dXMvMTQ0ODA0MDM5NTc0NTU5OTQ4OT9fX3M9YmlpMDU3bWdqMDRjenhncmtidWwmdXRtX3NvdXJjZT1kcmlwJnV0bV9tZWRpdW09ZW1haWwmdXRtX2NhbXBhaWduPXdlZWtseStlbWFpbCstKzIwMjMtMDctMjYmdXRtX2NvbnRlbnQ9V2Vla2x5K1B5dGhvbit0aXAlM0ErOCtmaWxlK3Byb2Nlc3NpbmcrdGlwcyJ9.BfRWj0YJkITkIfc2wxyJU6ZNPqMIbrVIug47ko6RCu4) . Ignore este consejo si sabe que el archivo que no es de confianza es pequeño (debido a los límites de carga de archivos, por ejemplo) 6. Al manipular rutas de archivos, use objetos pathlib.Path . De hecho, tiendo a preferir [pathlib](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vdHJleWh1bm5lci5jb20vMjAxOS8wMS9uby1yZWFsbHktcGF0aGxpYi1pcy1ncmVhdC8_X19zPWJpaTA1N21najA0Y3p4Z3JrYnVsJnV0bV9zb3VyY2U9ZHJpcCZ1dG1fbWVkaXVtPWVtYWlsJnV0bV9jYW1wYWlnbj13ZWVrbHkrZW1haWwrLSsyMDIzLTA3LTI2JnV0bV9jb250ZW50PVdlZWtseStQeXRob24rdGlwJTNBKzgrZmlsZStwcm9jZXNzaW5nK3RpcHMifQ.XTMmlTb-G7s5HsX0ujFku4vQclzSwB7hQxLZKMTojP8" \t "_blank) casi siempre que trabajo con archivos en Python. 7. Si alguna vez necesita leer un archivo dos veces, es posible que desee utilizar el método [de búsqueda](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vd3d3LnB5dGhvbm1vcnNlbHMuY29tL3NlZWtpbmctaW4tZmlsZXMvP19fcz1iaWkwNTdtZ2owNGN6eGdya2J1bCZ1dG1fc291cmNlPWRyaXAmdXRtX21lZGl1bT1lbWFpbCZ1dG1fY2FtcGFpZ249d2Vla2x5K2VtYWlsKy0rMjAyMy0wNy0yNiZ1dG1fY29udGVudD1XZWVrbHkrUHl0aG9uK3RpcCUzQSs4K2ZpbGUrcHJvY2Vzc2luZyt0aXBzIn0.OmnHPQld4SOwshdRY3XZjVL2ZKER91Hb_jSokoZOvfA) . 8. Si necesita asegurarse de no sobrescribir un archivo o desea agregarlo al final de un archivo, eche un vistazo a [los modos de archivo de Python](https://t.dripemail2.com/c/eyJhbGciOiJIUzI1NiJ9.eyJhdWQiOiJkZXRvdXIiLCJpc3MiOiJtb25vbGl0aCIsInN1YiI6ImRldG91cl9saW5rIiwiaWF0IjoxNjkwMzcyOTE4LCJuYmYiOjE2OTAzNzI5MTgsImFjY291bnRfaWQiOiIzOTQ1ODYwIiwiZGVsaXZlcnlfaWQiOiIzc2t3dGR3YTBxOWN6dzg4a2J2biIsInVybCI6Imh0dHBzOi8vd3d3LnB5dGhvbm1vcnNlbHMuY29tL2ZpbGUtbW9kZXMtaW4tcHl0aG9uLz9fX3M9YmlpMDU3bWdqMDRjenhncmtidWwmdXRtX3NvdXJjZT1kcmlwJnV0bV9tZWRpdW09ZW1haWwmdXRtX2NhbXBhaWduPXdlZWtseStlbWFpbCstKzIwMjMtMDctMjYmdXRtX2NvbnRlbnQ9V2Vla2x5K1B5dGhvbit0aXAlM0ErOCtmaWxlK3Byb2Nlc3NpbmcrdGlwcyJ9.CeOJCn0N0QlTBS1B4rmg_lmJRCA_goRJXKS-WdDKqBI). | | |