

Microsoft vodi forum za povratne informacije o Excelu na UserVoiceu (<https://oreil.ly/y1XwU>) gde svako može da pošalje novu ideju o kojoj će drugi glasati. Zahtev za funkcije sa najvećim brojem glasova je „Python as an Excel scripting language” (Python kao skript jezik za Excel) i ima otprilike dvostruko više glasova od drugog kandidata za najviše glasova. Iako se ništa nije dogodilo od kada je ideja dodata 2015. godine, korisnici programa Excel dobili su novu nadu krajem 2020. godine kada je Guido van Rossum, tvorac Pythona, poslao tvit (https://oreil.ly/N1_7N) da mu je „penzionisanje bilo do sadno” i da će se pridružiti Microsoftu. Ne znam da li njegov potez ima uticaja na integraciju programa Excel i Python. Znam, međutim, šta ovu kombinaciju čini tako ubedljivom i kako možete početi da koristite Excel i Python zajedno – danas. Ovo je ukratko ono o čemu ova knjiga govori.

Glavna pokretačka snaga priče o *Pythonu za Excel* je činjenica da živimo u svetu podataka. Danas su ogromni skupovi podataka dostupni svima i o svemu. Često su ti skupovi podataka toliko veliki da više ne staju u tabelu. Pre nekoliko godina ovo se možda nazivalo *big data* (veliki podaci), ali u današnje vreme skup podataka od nekoliko miliona redova zaista nije ništa posebno. Excel je evoluirao da bi se nosio sa tim trendom: uveo je Power Query za učitavanje i čišćenje skupova podataka koji se ne uklapaju u tabelu i Power Pivot, dodatak za analizu podataka na ovim skupovima podataka i predstavljanje rezultata. Power Query je zasnovan na Power Query M jeziku formule (M), dok Power Pivot definiše formule pomoću izraza za analizu podataka (Data Analysis Expressions, DAX). Ako želite da automatizujete nekoliko stvari u Excel datoteci, onda biste koristili Excelov ugrađeni jezik za automatizaciju, Visual Basic for Applications (VBA). To jest, za nešto prilično jednostavno, na kraju možete koristiti VBA, M i DAX. Jedan problem sa ovim je što svi ovi jezici rade samo u Microsoftovom svetu, najistaknutiji u Excelu i Power BI-u (Power BI ću ukratko predstaviti u poglavlju 1).

Python, je programski jezik opšte namene koji je postao jedan od najpopularnijih izbora među analitičarima i stručnjacima za podatke. Ako koristite Python sa programom Excel, moći ćete da koristite programski jezik koji je dobar u svim aspektima priče, bilo da se radi o automatizaciji programa Excel, pristupu i pripremi skupova podataka ili obavljanju zadataka analize i vizualizacije podataka. Ono što je najvažnije, možete koristiti svoje veštine rada sa Pythonom izvan programa Excel: ako treba da povećate svoju računarsku moć, lako biste mogli da premestite svoj kvantitativni model, simulaciju ili aplikaciju za mašinsko učenje u oblak, gde vas čekaju praktično neograničeni računarski resursi.

Zašto sam napisao ovu knjigu

Tokom svog rada na xlwingsu, Excelovom paketu za automatizaciju koji ćemo sresti u delu IV ove knjige, u bliskom sam kontaktu sa mnogim korisnicima koji koriste Python za Excel – bilo da je to putem programa za praćenje problema (<https://oreil.ly/ZJQkB>) na GitHubu, pitanje na StackOverflowu (<https://stackoverflow.com>) ili na fizičkom događaju poput sastanka ili konferencije.

Od mene se redovno traži da preporučim resurse za početak rada sa Pythonom. Iako ne nedostaje Python uvoda, oni su često ili previše opšti (ništa o analizi podataka) ili previše specifični (potpuni naučni uvod). Međutim, korisnici programa Excel imaju tendenciju da budu negde u sredini: oni rade sa podacima, ali potpuni naučni uvod može biti previše tehnički. Oni često imaju posebne zahteve i pitanja na koja nema odgovora ni u jednom postojećem materijalu. Neka od ovih pitanja su:

- Koji paket Python-Excel mi je potreban za koji zadatak?
- Kako da premestim svoju Power Query vezu sa bazom podataka na Python?
- Šta je ekvivalent Excelovog AutoFiltera ili zaokretne tabele u Pythonu?

Napisao sam ovu knjigu za vas koji nemate dovoljno znanja o Pythonu da biste mogli automatizovati svoje Excel zadatke i iskoristiti Pythonove analize podataka i alate za naučno računarstvo u Excelu i to bez stranputica.

Kome je namenjena ova knjiga

Ako ste napredni korisnik programa Excel koji želi da prevaziđe granice programa Excel savremenim programskim jezikom, ova knjiga je za vas. Najčešće to znači da svakog meseca provodite sate preuzimajući, čisteći i kopirajući/umećući velike količine podataka u kritične tabele. Iako postoje različiti načini za prevazilaženje ograničenja programa Excel, ova knjiga će se fokusirati na to kako koristiti Python za ovaj zadatak.

Trebalo bi da imate osnovno razumevanje programiranja: pomaže ako ste već napisali funkciju ili for petlju (bez obzira u kom programskom jeziku) i imate ideju o tome šta je ceo broj ili niz znakova. Možda ćete čak moći da savladate ovu knjigu ako ste navikli da pišete složene ćelijske formule ili imate iskustva sa podešavanjem snimljenih VBA makroa. Ne očekuje se da imati iskustvo vezano za Python, budući da postoje uvodi u sve alate koje ćemo koristiti, uključujući i uvod u sam Python.

Ako ste iskusni VBA programer, naći ćete redovna poređenja između Pythona i VBA koja će vam omogućiti da rešavate uobičajene probleme.

Ova knjiga takođe može biti od pomoći ako ste programer na Pythonu i morate da naučite o različitim načinima na koje Python može da se nosi sa Excel aplikacijom i Excel datotekama da bi mogao da izabere pravi paket s obzirom na zahteve vaših poslovnih korisnika.

Kako je ova knjiga organizovana

U ovoj knjizi ću vam pokazati sve aspekte priče *Python za Excel* podeljeno na četiri dela:

Deo I: Uvod u Python

Ovaj deo počinje razmatranjem razloga zašto je Python tako prijatan saputnik za Excel pre nego što predstavimo alate koje ćemo koristiti u ovoj knjizi: distribuciju Anaconda Python, Visual Studio Code i Jupyter beležnice. Ovaj deo će vas naučiti dovoljno i Pythona da savladate ostatak ove knjige.

Deo II: Uvod u pandas

pandas je Pythonova biblioteka za analizu podataka. Naučićemo kako da zamenite Excel radne sveske kombinacijom Jupyter beležnica i pandas. Obično je pandas kôd lakši za održavanje i efikasniji od Excel radne sveske, a možete raditi i sa skupovima podataka koji ne staju u tabelu. Za razliku od Excela, pandas vam omogućava da pokrenete svoj kôd gde god želite, uključujući oblak.

Deo III: Čitanje i pisanje Excel datoteka bez Excela

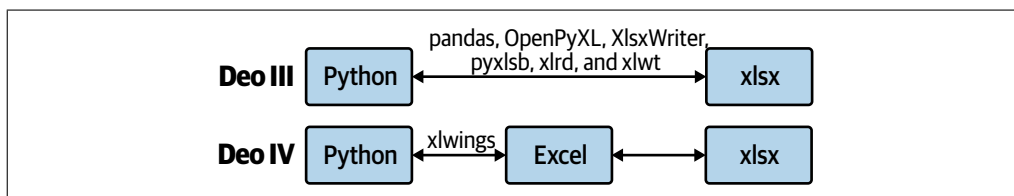
Ovaj deo se bavi manipulacijom Excel datotekama pomoću jednog od sledećih Python paketa: pandas, OpenPyXL, XlsxWriter, pyxlsb, xlrd i xlwt. Ovi paketi mogu da čitaju i pišu Excel radne sveske direktno na disku i kao takvi zamenjuju Excel aplikaciju: pošto vam nije potrebna instalacija programa Excel, oni rade na bilo kojoj platformi koju Python podržava, uključujući Windows, macOS i Linux. Tipičan slučaj upotrebe paketa za čitanje je čitanje podataka iz Excel datoteka koje svakog jutra primete od spoljne kompanije ili sistema i skladištenje njihovog sadržaja u bazu podataka. Tipičan slučaj upotrebe paketa za pisanje je obezbeđivanje funkcionalnosti koje se nalazi iza čuvenog dugmeta „Export to Excel” (Izvezi u Excel) koje možete pronaći u gotovo svakoj aplikaciji.

Deo IV: Programiranje Excel aplikacije sa xlwings

U ovom delu ćemo videti kako možemo koristiti Python sa paketom xlwings za automatizaciju Excel aplikacije umesto čitanja i pisanja Excel datoteka na disk. Zbog toga ovaj deo zahteva da imate lokalnu instalaciju programa Excel. Naučićemo kako da otvorimo Excel radne sveske i da njima rukujemo dok su ispred nas. Osim čitanja i pisanja datoteka putem programa Excel, izgrađićemo i interaktivne Excel alate: oni nam omogućavaju da pritisnemo dugme da bi Python obavio nešto što ste možda prethodno radili sa VBA makroima, poput računarski zahtevnih izračunavanja. Takođe ćemo naučiti kako pisati korisnički definisane funkcije¹ (UDF) u Pythonu umesto VBA.

Važno je razumeti fundamentalnu razliku između čitanja i pisanja Excel *datoteka* (deo III) i programiranje Excel *aplikacije* (deo IV) kao što je prikazano na slici P-1.

¹ Microsoft je počeo da koristi izraz *prilagođene funkcije* (eng. *custom functions*) umesto UDF. U ovoj knjizi ću ih i dalje nazivati UDF.



Slika P-1. Čitanje i pisanje Excel datoteka (deo III) u odnosu na programiranje Excela (deo IV)

Pošto deo III ne zahteva instalaciju programa Excel, sve radi na svim platformama koje Python podržava: Windows, macOS i Linux, međutim, radiće samo na onim platformama koje podržava Microsoft Excel, tj. Windows i macOS, pošto se kôd oslanja na lokalnu instalaciju Microsoft Excela.

Python i Excel verzije

Ova knjiga zasnovana je na Pythonu 3.8, verziji Pythona koja dolazi sa najnovijom verzijom distribucije Anaconda Python u vreme pisanja ovog teksta. Ako želite da koristite noviju verziju Pythona, sledite uputstva na veb stranici ove knjige (<https://xlwings.org/book>), ali pazite da ne koristite stariju verziju. Povremeno ću ostaviti komentar ako se nešto promeni sa Pythonom 3.9.

Ova knjiga takođe očekuje da koristite savremenu verziju programa Excel, što znači najmanje Excel 2007 u sistemu Windows i Excel 2016 u sistemu macOS. Lokalno instalirana verzija programa Excel koja dolazi sa pretplatom na Microsoft 365 takođe će savršeno funkcionisati – u stvari, čak je i preporučujem, jer ima najnovije funkcije koje nećete pronaći u drugim verzijama programa Excel. To je bila i verzija koju sam koristio za pisanje ove knjige, pa ako koristite drugu verziju programa Excel, ponekad ćete možda videti malu razliku u nazivu ili lokaciji stavke menija.

Konvencije korišćene u ovoj knjizi

U ovoj knjizi se koriste sledeće tipografske konvencije:

Italic

Označava nove termine, URL-ove, adrese e-pošte, nazive datoteka i oznake tipa datoteka.

Konstantna širina

Koristi se za listing programa, kao i u pasusima za upućivanje na programske elemente kao što su imena promenljivih ili funkcija, baze podataka, tipovi podataka, promenljive okruženja, izrazi i ključne reči.

Konstantna širina podebljano

Prikazuje komande ili drugi tekst koji korisnik treba doslovno da otkuca.

Kurziv konstantne širine

Prikazuje tekst koji treba zameniti vrednostima koje je uneo korisnik ili vrednostima određenim kontekstom.



Ovaj element označava savet ili sugestiju.



Ovaj element označava opštu napomenu.



Ovaj element ukazuje na upozorenje ili oprez.

Korišćenje primera koda

Održavam veb stranicu (<https://xlwings.org/book>) sa dodatnim informacijama koje će vam pomoći oko ove knjige. Posetite je, posebno ako naiđete na problem.

Dodatni materijal (primeri koda, vežbe itd.) dostupan je za preuzimanje na <https://github.com/fzumstein/python-for-excel>. Da biste preuzeli ovo prateće spremište, kliknite na zeleno dugme Code, a zatim odaberite Download ZIP. Kada je preuzeto, kliknite desnim tasterom miša na datoteku u Windowsu i izaberite Extract All da biste raspakovali sadržaj datoteke u folder. Na macOS-u jednostavno dvaput kliknite na datoteku da biste je raspakovali. Ako znate kako da radite sa Gitom, takođe biste mogli da koristite Git za kloniranje spremišta na vaš lokalni čvrsti disk. Folder možete staviti gde želite, ali ja ću ga povremeno pominjati na sledeći način u ovoj knjizi:

```
C:\Users\imekorisnika\python-for-excel
```

Jednostavnim preuzimanjem i raspakovanjem ZIP datoteke u Windowsu dobićete strukturu foldera sličnu ovoj (obratite pažnju na ponavljajuća imena foldera):

```
C:\...\Downloads\python-for-excel-1st-edition\python-for-excel-1st-edition
```

Kopiranje sadržaja ovog foldera u jedan koju pravite pod `C:\Users\<imekorisnika>\python-for-excel` moglo bi vam olakšati praćenje. Iste napomene važe i za macOS, tj. kopirajte datoteke u `/Users/<imekorisnika>/python-for-excel`.

Ako imate tehničko pitanje ili problem sa primerima koda, pošaljite e-poruku na bookkuestions@oreilli.com.

Ova knjiga je tu da vam pomogne da obavite svoj posao. Uopšteno, ako se uz ovu knjigu nudi primer koda, možete ga koristiti u svojim programima i dokumentaciji. Ne morate da nas kontaktirate za dozvolu ako ne reprodukujete značajan deo koda. Na primer, za pisanje programa koji koristi nekoliko delova koda iz ove knjige nije potrebna dozvola. Za prodaju ili distribuciju primera iz ove knjige potrebna je dozvola. Za odgovaranje na pitanje citiranjem ove knjige i citiranjem primera koda nije potrebna dozvola. Za uključivanje značajne količine primera koda iz ove knjige u dokumentaciju vašeg proizvoda potrebna je dozvola.

Cenimo, ali generalno ne zahtevamo, navođenje. Navođenje obično uključuje naslov, autora, izdavača i ISBN broj. Na primer: „*Python za Excel* Felix Zumstein (O’Reilly). Autorska prava 2021 Zoomer Analytics LLC, 978-1-492-08100-5.”

Ako smatrate da upotreba primera koda ne spada u poštnu upotrebu ili gore navedenu dozvolu, slobodno nas kontaktirajte na permissions@oreilli.com.

Mikro knjiga

Preko 37 godina Mikro knjiga donosi odabrane IT naslove i naslove opšteobrazovnih tema u ediciji *Za neupućene...* Sa preko 460 objavljenih naslova pomogla je u obrazovanju desetina hiljada IT stručnjaka. Za više informacija o našim izdanjima pogledajte na www.mikroknjiga.rs

Kako da nas kontaktirate

Komentare i pitanja u vezi ove knjige uputite izdavaču:

MIKRO KNJIGA, DOO
Kneza Višeslava 34
11030 Beograd, Srbija
+381 11 3058-044

e-mail: redakcija@mikroknjiga.rs

Facebook: <http://facebook.com/mikroknjiga>

Twitter: <http://twitter.com/mikroknjiga>

Na ovoj strnaici (<https://oreil.ly/py4excel>) naći ćete greške, primere i sve dodatne informacije o ovoj knjizi..

Web prikaz ove knjige nalazi se na adresi
<https://www.mikroknjiga.rs/store/prikaz.php?ref=PYEX>

Sve vaše sugestije, primedbe i prijave eventualnih grešaka u knjizi su dobrodošle.

Zahvalnice

Kao autor po prvi put, veoma sam zahvalan na pomoći koju sam dobio od toliko ljudi na ovom putu – koji su mi znatno olakšali ovo putovanje!

U O'Reillyju želim da se zahvalim svojoj urednici Melissi Potter, koja je uradila sjajan posao održavajući me motivisanim i u tempu i koja mi je pomogla da ovu knjigu dovedem u čitljiv oblik. Takođe bih želeo da se zahvalim Michelle Smith, koja je sa mnom radila na početnoj verziji knjige, i Danielu Elfanbaumu, koji se nije umorio odgovorajući na moja tehnička pitanja.

Veliko hvala svim mojim kolegama, prijateljima i klijentima koji su uložili mnogo sati u čitanje najranijih oblika mojih nacрта. Njihove povratne informacije bile su ključne za lakše razumevanje knjige, a neke studije slučaja inspirisane su stvarnim problemima u Excelu koje su podelili sa mnom. Zahvaljujem se Adamu Rodrigezu, Manu Beeslaru, Simonu Schieggju, Rui Da Costi, Jurgu Nageru i Christopheu de Montrichardu.

Dobio sam korisne povratne informacije i od čitalaca rane verzije knjige koja je objavljena na O'Reilly platformi za učenje na mreži. Hvala vam Felipe Maion, Ray Doue, Kolyu Minevski, Scott Drummond, Volker Roth i David Ruggles!

Imao sam veliku sreću što su knjigu recenzirali visoko kvalifikovani recenzenti i zaista cenim njihov trud koji su uložili pod velikim vremenskim pritiskom. Hvala vam na vašoj pomoći, Jordan Goldmeier, George Mount, Andreas Clenow, Werner Brönnimann i Eric Moreira!

Posebna zahvalnost ide Björnu Stielu, koji nije bio samo tehnički recenzent, već je neko od koga sam naučio i mnoge stvari o kojima pišem u ovoj knjizi. Uživao sam radeći sa vama ovih nekoliko godina!

Poslednje, ali ne i najmanje važno, želeo bih da se zahvalim Ericu Reinoldsu, koji je 2016. spojio svoj ExcelPython projekat u bazu koda xlwings. Redizajnirao ceo paket od nule, čineći moj užasan API iz ranih dana stvar prošlosti. Mnogo vam hvala!