

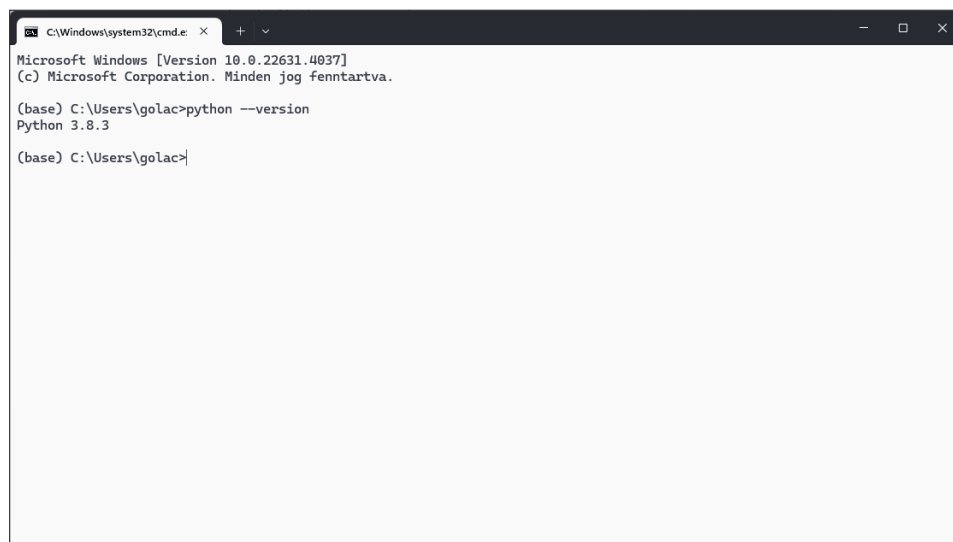
Tantusz Könyvek – Gépi tanulás

Az R beszerzése és használata

A könyvben az Anacondát használjuk az adatelemzési kódok bemutatásához. Először telepítened kell az Anacondát a könyv 4. fejezetében leírtak szerint. Miután telepítetted az Anacondát, a következő szakaszokban található útmutatás alapján létrehozatsz egy R-környezetet, amelyben a könyv R-kódjait futtathatod.

Az R telepítése az Anacondához

Az R telepítéséhez meg kell nyitnod az Anaconda-parancssort. Ha Windows-felhasználó vagy, ehhez a Start ⇨ Minden program (Mind) ⇨ Anaconda3 ⇨ Anaconda Prompt lehetőséget kell választanod. Ekkor megjelenik az Anaconda-parancssor, amelyet az 1. ábrán is láthatsz. Figyeld meg a (base) jelölést a prompt elején. Ez az aktuális környezet – az alapkörnyezet.



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.22631.4037]
(c) Microsoft Corporation. Minden jog fenntartva.

(base) C:\Users\golac>python --version
Python 3.8.3

(base) C:\Users\golac>
```

1. ábra.

Nyiss meg egy Anaconda-parancssort az R telepítéséhez.

Az Anaconda-parancssorban írd be a `conda create -n R_kornyezet r-essentials r-base` parancsot, majd nyomd meg az Entert. Ezzel a paranccsal létrehozol egy új Anaconda-környezetet `R_kornyezet` néven. Amikor R-kóddal szeretnél dolgozni, mindig az `R_kornyezet` környezetet használd. A `conda`, az Anaconda parancssori segédprogramja ebben a környezetbe telepíti az `essentials` és a `base` R-csomagot. Ezek a csomagok elegendőek ahhoz, hogy hozzáláss a munkához. Később telepíthetsz más csomagokat is, ha szükségesek az elemzések adott fajtájához. A telepítési folyamat során egy sor üzenet jelenik majd meg, melyek végén azoknak a csomagoknak a listája fog szerepelni, amelyeket a `conda` telepíteni fog, ahogy a 2. ábrán is láthatod.

A telepítési folyamat elindításához írd be az `y` betűt, majd nyomd meg az Entert. Ekkor a `conda` megkezdí a csomagok letöltését és kicsomagolását. Amikor a letöltések befejeződnek, a `conda` különböző tranzakciós folyamatokat hajt végre, köztük esetlegesen egyes letöltött csomagok végrehajtható (vagy más) formára való lefordítását is. Ez a folyamat eltarthat egy darabig, ezért érdemes lehet a kezéd ügyében tartani egy jó könyvet, vagy egy kávé, amit elszüröcsölhetsz.

2. ábra.
A conda segédprogram
megmutatja, hogy mely
csomagokat fogja
telepíteni.

```

C:\Windows\system32\cmd.e
+ v
The following NEW packages will be INSTALLED:

_r-mutex      pkgs/r/win-64::_r-mutex-1.0.0-anaconda1_1
anyio         pkgs/main/win-64::anyio-4.6.2-py312h95532_0
argon2-cffi   pkgs/main/noarch::argon2-cffi-21.3.0-pyhd3eb1b0_0
argon2-cffi-bindings pkgs/main/win-64::argon2-cffi-bindings-21.2.0-py312h2bbff1b_0
asttokens     pkgs/main/noarch::asttokens-2.0.5-pyhd3eb1b0_0
async-lru     pkgs/main/win-64::async-lru-2.0.4-py312h95532_0
attrs         pkgs/main/win-64::attrs-24.2.0-py312h95532_0
babel         pkgs/main/win-64::babel-2.11.0-py312h95532_0
beautifulsoup4 pkgs/main/win-64::beautifulsoup4-4.12.3-py312h95532_0
bleach        pkgs/main/noarch::bleach-4.1.0-pyhd3eb1b0_0
brotli-python pkgs/main/win-64::brotli-python-1.0.9-py312hd77b12b_8
bzip2         pkgs/main/win-64::bzip2-1.0.8-h2bbff1b_6
ca-certificates pkgs/main/win-64::ca-certificates-2024.9.24-h95532_0
certifi       pkgs/main/win-64::certifi-2024.8.30-py312h95532_0
cffi          pkgs/main/win-64::cffi-1.17.1-py312h827c3e9_0
charset-normalizer pkgs/main/noarch::charset-normalizer-3.3.2-pyhd3eb1b0_0
colorama      pkgs/main/win-64::colorama-0.4.6-py312h95532_0
comm          pkgs/main/win-64::comm-0.2.1-py312h95532_0
debugpy       pkgs/main/win-64::debugpy-1.6.7-py312hd77b12b_0
decorator     pkgs/main/noarch::decorator-5.1.1-pyhd3eb1b0_0
defusedxml    pkgs/main/noarch::defusedxml-0.7.1-pyhd3eb1b0_0
executing     pkgs/main/noarch::executing-0.8.3-pyhd3eb1b0_0
expat         pkgs/main/win-64::expat-2.6.3-h5da7b33_0
h11           pkgs/main/win-64::h11-0.14.0-py312h95532_0
httpcore      pkgs/main/win-64::httpcore-1.0.2-py312h95532_0
httpx         pkgs/main/win-64::httpx-0.27.0-py312h95532_0
idna          pkgs/main/win-64::idna-3.7-py312h95532_0

```



VIGYÁZAT

Sajnos előfordulhat, hogy nem sikerül létrehoznod a környezetet, mert a parancs végrehajtása elakad. Ezt úgy oldhatod meg a legegyszerűbben, ha frissíted az Anaconda csomagfeloldóját, melyhez a következő címen találhatsz útmutatást: <https://www.anaconda.com/blog/a-faster-conda-for-a-growing-community>.



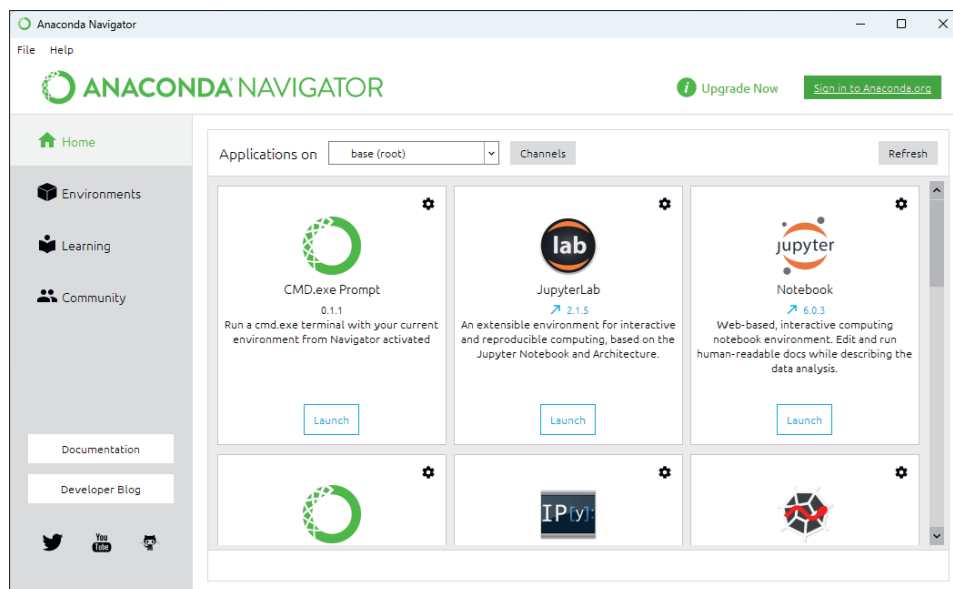
FONTOS

Miután a telepítési folyamat befejeződött, még mindig a (base) környezetben vagy. Ahhoz, hogy az (R_környezet) környezetben dolgozhass, be kell írnod a **conda activate R_környezet** parancsot az Anaconda-parancssorba. A platformon támogatott bármely könyvtárparancsot használhatod, hogy megnézd a conda által telepített fájlok listáját. Ha szeretnél kilépni az (R_környezet) környezetből, és visszatérni a (base) környezetbe, írd be a **conda deactivate** parancsot, majd nyomd meg az Entert.

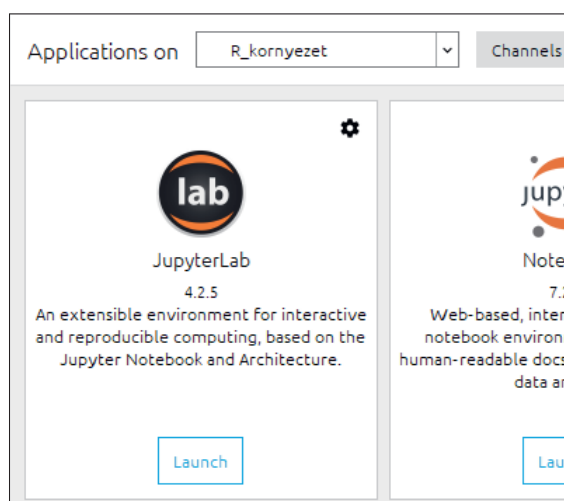
Az R-környezet elindítása

Szerencsére nem az Anaconda-parancssorban fogsz dolgozni, amikor R-kódot szeretnél kezelni. Azonban a Notebookot sem szabad közvetlenül elindítanod, mert ezzel a (base) környezetet indítanád el, neked pedig az (R_környezet) környezetre van szükséged. Ehelyett indítsd el az Anaconda Navigátort, melyet a 3. ábrán is láthatsz. Figyeld meg, hogy ebben a segédprogramban az Anaconda által támogatott összes GUI-s eszközt elérheted (és ha nem találsz, amire szükséged van, bármikor telepíthetsz továbbiakat).

Az R elindításához először ki kell választanod az R_környezet környezetet az Applications On (Alkalmazások futtatása ezen) legördülő listában, ahogy a 4. ábrán is láthatod. Ekkor megváltoznak a megjelenített eszközök, annak megfelelően, hogy mit használhatsz ebben a környezetben. A könyv szempontjából fontos eszköz a Notebook. Figyeld meg a négyzetének jobb felső sarkában látható fogaskerék ikont. Ha módosítanod kell a Notebook verzióját, kattints az ikonra, és válaszd az Install Specific Version (Adott verzió telepítése) lehetőséget a listában.



3. ábra.
Az Anaconda Navigatorban
számos hasznos eszköz
érhető el.



4. ábra.
Ha környezetet váltasz,
azzal gyakran megváltozik
az elérhető eszközök
listája is.

Az alkalmazás elindításához kattints a Launch (Indítás) gombra. Ekkor láthatod, hogy a Notebook elindul a böngészőben, ugyanúgy, mint amikor a Pythonnal dolgozol, de most R-t fogsz használni.

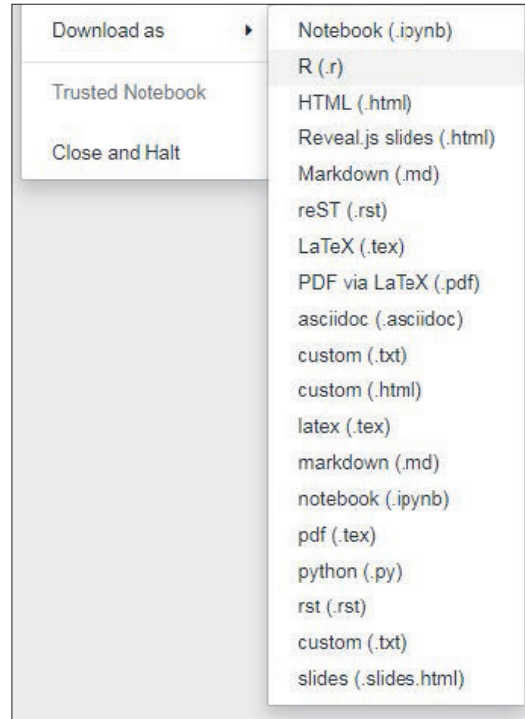
Az R-kódtár meghatározása

Mivel a Notebookot használod, minden pontosan ugyanúgy fog működni az R-rel is, mint a Pythonnal, ahogyan arról a 4. fejezet „Határozzuk meg a kódtárat” című részében olvashatsz. A könyvben a TKGTR nevű R-kódtárat használjuk. A „Határozzuk meg a kódtárat” című szakaszban található utasítások alapján hozhatsz létre egy R-adattárat, majd kísérletezhetsz egy R-fájlal. A New (Új) legördülő listából azonban a Python 3 helyett az R-t kell választanod. Ezt leszámítva minden ugyanúgy működik, mint a Python esetében. Az ebben a szakaszban használt tesztfájl a TKGTR_04_Sample.ipynb. Az R-fájlok esetén a fájlneveket, a megjegyzéseket és a változóneveket is angolul olvashatod.



TIPP

Fontos megjegyezni, hogy az R használatakor letöltheted a kódot .r fájlként, ahogy azt az 5. ábrán is láthatod. Ami azt illeti, nemcsak a Notebook formátumai közül választhatsz, hanem számos R-specifikus formátum közül is. Az egyszerűség kedvéért a könyvhöz tartozó letölthető forrásban minden forráskód esetében .ipynb fájlokat használunk, hogy biztosan elérd a fájlhoz szükséges megjegyzéseket.



5. ábra.

Az R-kódot .r fájlba mentheted el, de a .r fájlokban nem szerepelnek a Notebook megjegyzései.

A következő kóddal egy gyors tesztet végezhet az R-környezeteden:

```
print(R.version.string)
eredmény <- system2('conda', args='list -n base anaconda$', stdout=TRUE, wait=TRUE)
print(,Anaconda-verzió: ,)
print(eredmény)
```

A legjobb, ha a `system2()` függvényt használod a conda futtatásához, mivel az itt megadott módon be kell állítanod az `stdout` és a `wait` argumentumot. Figyeld meg, hogy a conda a parancs, és az `args` argumentumban kell megadni azt, hogy mit szeretnél csinálni a conda paranccsal.

```
Miután rákattintottál a Run (Futtatás) gombra, a következő kimenet fog megjelenni:
[1] „R version 4.0.3 (2020-10-10)”
[1] „Anaconda-verzió: „
[1] „# packages in environment at C:\\Users\\golac\\ANACON~1:”
[2] „#”
[3] „# Name Version Build Channel”
[4] „anaconda 2020.07 py38_0 „
```

Ebben a kimenetben a könyvhöz használt eszközverziókat, köztük az R 4.0.3-as és az Anaconda 2020.07-es verzióját láthatod. A legjobb eredmények elérése érdekében győződj meg arról, hogy a verzióadataid megegyeznek az itt láthatóakkal.