# Csoporttagok:

* Petrik Ádám (GVERV7)
* Balog Barnabás (YRVN14)
* Szilveszter Milán (CZZYLG)
* Gergely Jácint (NPCAPG)

GitHub Repository: https://github.com/gerjacint23/MSc-ML-GroupWork

# Regressziós feladat:

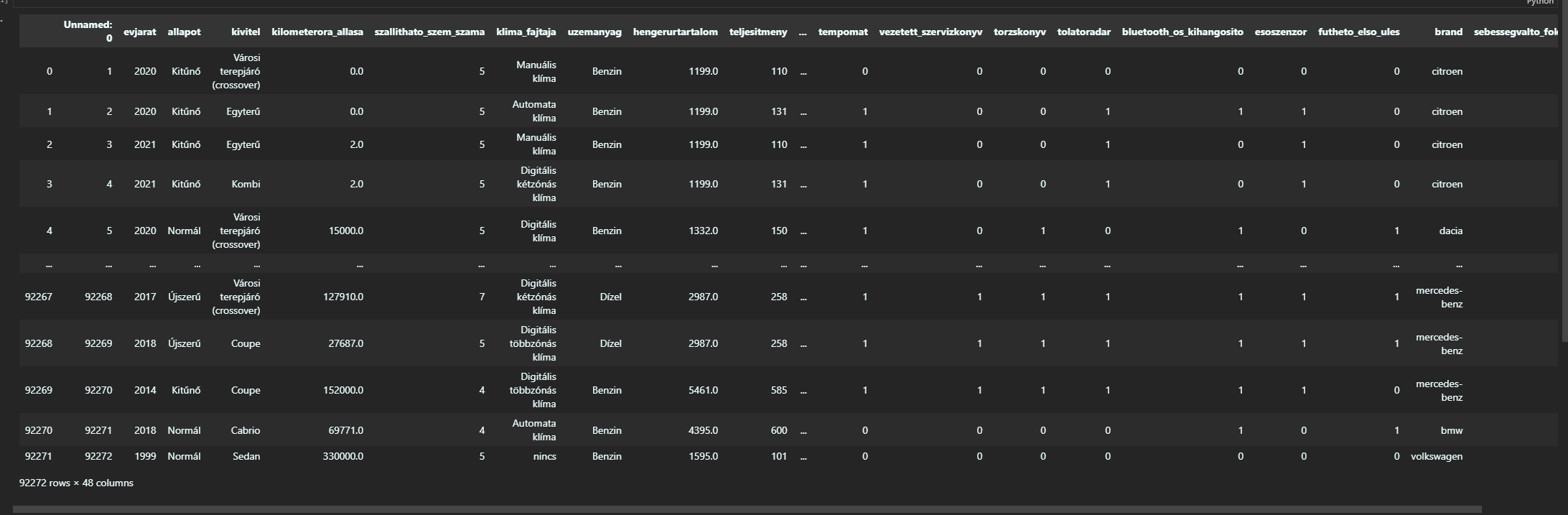
### „Nagy” adathalmaz:

Az adathalmaz a használtautó.hu autókereskedő oldalra feltöltött hirdetések adataiból áll. Az adathalmaz kb. 92 ezer rekordot és 48 attribútumot tartalmaz. A modell célja, hogy az elérhető paraméterekből megbecsüljük az egyes gépjárművek vételárát.

Adattisztítási lehetőségek:

* Karakterkódolási hibák javítása
* Üres értékek helyettesítése
* Szöveges értékek számokká alakítása

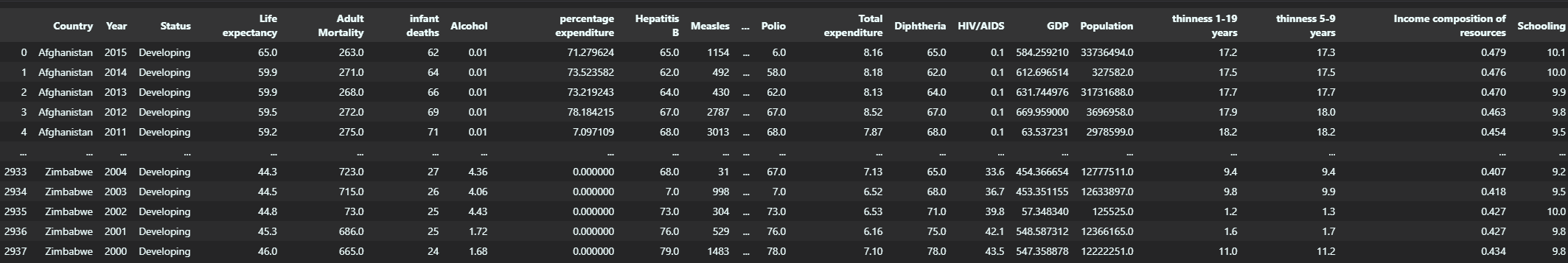
**Adatforrás:** <https://www.kaggle.com/datasets/nndorszakmry/hasznltautk-adatai?resource=download>

****

### „Kis” adathalmaz:

Annak az elemzésére szolgál az adathalmaz, hogy melyek azok a változók, amelyek kulcsfontosságú szerepet játszanak a várható élettartam alakításában. Az adathalmaz közel 3000 rekordot tartalmaz 22 oszloppal.

**Adatforrás**: <https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who/data>

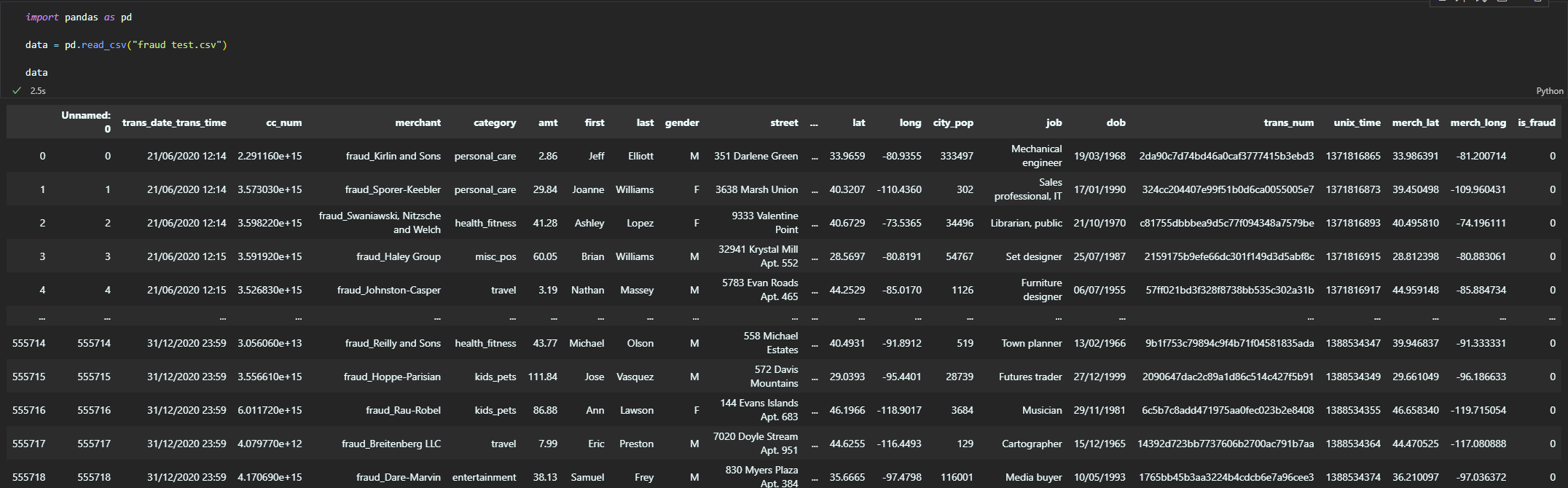


# Osztályozási feladat:

### „Nagy” adathalmaz:

A modell célja a hamis hitelkártya-tranzakciók felderítése, annak érdekében, hogy az ügyfeleket ne terheljék olyan tételekkel, amelyeket nem ők vásároltak. Az adathalmaz több mint 550 ezer rekordot és 23 attribútumot tartalmaz.

**Adathalmaz:** <https://www.kaggle.com/datasets/kelvinkelue/credit-card-fraud-prediction/data>



### „Kis Adathalmaz”:

Ez az adatkészlet az Agaricus és Lepiota családba tartozó 23 kopoltyús gombafajnak megfelelő hipotetikus minták leírását tartalmazza. Minden egyes faj azonosítva van, mint biztosan ehető, biztosan mérgező, vagy ismeretlen ehetőségű és nem ajánlott.

A feltanított modell célja a különböző gombákról eldönteni, hogy mérgezők vagy ehetőek-e.

Az adathalmaz 8123 rekordot és 23 oszlopot tartalmaz.

A képen képernyőkép látható

Automatikusan generált leírás

**Adatforrás**: <https://archive.ics.uci.edu/dataset/73/mushroom>