### **INFORME**

# 1. Context de projecte

La història del *e-commerce* comença amb la primera venta online realitzada el 1994. Aquest és el primer exemple d'un consumidor adquirint un producte venut per un negoci a través d'Internet. Avui dia podem trobar infinitat de productes a Internet i de fet s'estima que les ventes realitzades per Internet adquiriran un valor de 27 trilions de dòlars aquest any 2020.

En el context e-commerce s'ha seleccionat una pàgina web on es venen productes hardware electrònics, en concret, d'entre tots els productes s'han seleccionat ordinadors.

### 2. Informació sobre el dataset

tonat named lan namentamistianung dala
ataset conté les característiques dels dors que es venen al lloc web així com les erístiques de venta com ara el preu actual, a anterior, etc.
A GeForce GTX 1650 TF - 8 GB Memory 256 GB SSD - Win 10", "1,200",1650.00,24Kg[14], Free Shipping test SSD Win 10 Pro High-End 15.6" FHD 144 Hz Garning Laptop", "2,179",2499.00,13% (35), Free Shipping CASF CREE MIZSD 15.6" Windows 10 Home Shelt (Only & Newege)", 549,599.99,21% (41), Free Shipping CASF CREE MIZSD 15.6" Windows 10 Home, FUES SN, NAN, NA, Free Shipping CASF CREE MIZSD 15.6" Windows 10 Home, TUPSD ST. AVT 23 Rotebook PC Computer "799, NAN, NA, Free Shipping CASF CREE MIZSD 16 Graphics, Windows 10 Home, S322FL-E871, Transparent Silver", 799, NAN, NA, 52.99 Shipping CASF CREE MIZSD CASF CASF CASF CASF CASF CASF CASF CASF

#### 2.4.Contingut

Cada ordinador es correspon amb un registre del conjunt de dades, on s'hi recullen les següents característiques:

- Marca: Recull el nom de la marca del ordinador.
- **Producte:** Recull el nom oficial del producte que es ven i característiques hardware.
- Preu Actual: Recull el preu de venta actual del ordinador
- **Preu Anterior**: Recull el preu de venta anterior del ordinador.
- **Rebaixa**: Recull el percentatge de descompte comparant el preu actual i anterior.
- **Rating:** Recull el nombre de comentaris que ha rebut del producte.
- Cost Envio: Recull el cost que té enviar el producte.

# 3. Agraïments

Newegg Inc. és líder en ventes de hardware electrònic a Nord Amèrica, amb abast a més de 80 països arreu d'Europa, Asia Pacífic, Sud Amèrica i Orient Mitjà. Fundada el 2001, la companyia publica una selecció dels últims productes electrònics.

Agraïm Newegg ja que gràcies a la seva pàgina web ha estat possible el procés de web scraping a través de l'algoritme dissenyat en un Jupyter Notebook.

## 4. Inspiració

L'objectiu de la descàrrega d'aquest conjunt de dades és la comparació d'ordinadors. Es pretén respondre a preguntes com ara:

- Quants ordinadors podem trobar dins d'un determinat rang de preu?
- Quants ordinadors podem trobar amb cost envio gratis?
- Quants ordinadors estan rebaixats?
- D'acord amb les seves característiques, quin és el millor ordinador que podem obtenir un preu més assequible?
- Quin és l'ordinador que ha rebut més comentaris?

Com a usuari, l'anàlisi d'aquestes dades ens permet obtenir el millor ordinador d'acord amb un conjunt de característiques seleccionades.

#### 5. Llicència

La llicència escollida per la publicació de les nostres dades seria una **CC0: Public Domain License.** El motiu és que volem donar drets d'autor a qualsevol persona i que sense cap tipus de condició pugui fer ús de les dades. L'objectiu és que aquestes puguin ser usades per un bé comú amb total llibertat.

#### 6. Fitxers

- **webscraper.ipynb**. Conté l'algoritme encarregat de dur a terme el procés de webscraping i també d'escriure el fitxer de sortida
- ordinadors.csv. Conté les dades extretes del lloc web a través del algoritme contingut al fitxer webscraper.ipynb