

# CAT-PAC4-enun

June 3, 2020

## 1 Programació per a la ciència de dades - PAC4

En aquest Notebook trobareu un exercici que suposa la quarta activitat d'avaluació continuada (PAC) de l'assignatura. Aquesta PAC constarà d'un únic exercici a resoldre, que englobarà molts dels conceptes coberts durant l'assignatura.

L'objectiu d'aquest exercici serà desenvolupar un **paquet de Python**, fora de l'entorn de Notebooks, que ens permeti resoldre el problema donat. Aquest haurà d'incloure el corresponent codi organitzat lògicament (separat en mòduls, organitzats per funcionalitat), la documentació de codi (docstrings) i tests. A més, s'hauran d'incloure els corresponents arxius de documentació d'alt nivell (README), així com els arxius de llicència i dependències (`requirements.txt`).

## 2 Exercici

Se'ns demana que implementem un paquet de Python que sigui capaç de realitzar una anàlisi senzilla de dades de Bitcoin. Bitcoin és una criptomonedra descentralitzada, que opera enviant *transaccions* a través d'una xarxa P2P. Aquestes transaccions s'agrupen en una estructura de dades coneguda com a *bloc*.

Haureu de generar funcions que us permetin trobar: 1. El número de transaccions per cada bloc. 2. El valor de cada transacció de cada bloc. 3. El temps entre blocs. 4. La mitjana de la mida del bloc per hora. 5. El número de transaccions per hora.

A més, haureu de generar codi que permeti representar tots els resultats gràficament. Per a cada mètrica, és important que penseu quin tipus de gràfica és la més adient per a representar-la.

El codi haurà d'estar correctament comentat, incloent documentació de funcions, i correctament testejat.

### 2.1 Les dades

Les dades a analitzar ens són proporcionades en dos arxius separats: `blocks.json` i `txs.json`.

`blocks.json` conté, per cada línia del fitxer, la informació d'un bloc en format JSON. De forma anàloga, `txs.json` conté, per cada línia, la informació d'una transacció també en format JSON.

Els arxius contenen informació sobre 144 blocs el que equival a, aproximadament, un dia en dades. Les transaccions de l'arxiu `txs.json` corresponen a els blocs de l'arxiu `blocks.json`.

Fent una ullada als arxius proporcionats, podreu veure que tant les transaccions com els blocs contenen força informació. Per a resoldre l'exercici proposat, només serà necessari fer-ne servir una quantitat molt reduïda d'aquesta.

### 2.1.1 Bloc

Pel que fa a un bloc, ens interessarà saber quines transaccions conté (camp tx), quan va ser generat (camp time) i quina és la seva mida (camp size).

### 2.1.2 Transacció

Pel que fa a una transacció, ens interessarà saber a quin bloc pertany i quin és el valor d'aquesta transacció.

Pel que fa a quin bloc pertany una transacció, l'identificador que apareix a la llista tx dins del bloc correspon amb el camp txid dins la transacció. A més, la transacció inclou el camp blockhash que correspon amb el camp hash del bloc.

Per altra banda, el valor de cada transacció no el trobareu directament a cap camp, sinó que l'haureu de calcular. Per fer-ho, haureu de sumar els valors dels camps `value` dins del camp `vout`. El camp `vout` és una llista de diccionaris.

**Exemple** Donat el bloc següent:

```
{
  "hash": "0f84abb78891a4b9e8bc9637ec5fb8b4962c7fe46092fae99e9d69373bf7812a",
  "confirmations": 1,
  "strippedsize": 216,
  "size": 252,"weight": 900,
  "height": 6126,
  "version": 536870912,
  "versionHex": "20000000",
  "merkleroot": "254d5cc8d2b1889a2cb45f7e3dca8ed53a3fcfa32e8b9eac5f68c4f09e7af7bd",
  "tx": [
    "254d5cc8d2b1889a2cb45f7e3dca8ed53a3fcfa32e8b9eac5f68c4f09e7af7bd"
  ],
  "time": 1590830080,
  "mediantime": 1590170439,
  "nonce": 1,
  "bits": "207fffff",
  "difficulty": 4.656542373906925e-10,
  "chainwork": "00000000000000000000000000000000000000000000000000000000000000002fde",
  "nTx": 1,
  "previousblockhash":
    "553ace55253369ed716ae9864ad307f1368285974762576e54441b255842f462"
}
```

Podem veure que la seva mida és 252, va ser creat a 1590830080 (*unix timestamp*) i conté una única transacció que te com a identificador 254d5cc8d2b1889a2cb45f7e3dca8ed53a3fcfa32e8b9eac5f68c4f09e7af7bd.

Si busquem aquesta mateixa transacció al fitxer de transaccions:

```
{
  "txid": "254d5cc8d2b1889a2cb45f7e3dca8ed53a3fcfa32e8b9eac5f68c4f09e7af7bd",
  "hash": "a8e125eb6d7ab883177d8ab228a3d09c1733d1ca49b7b2dff4b057eeb80ff9be",

```



```
"blocktime": 1590830080
}
```

Podem veure que el seu identificador (txid) correspon amb el que havíem vist al bloc, aquesta té un valor de 17.00000000 BTC, que correspon a la suma dels camps value de la llista vout.

## 2.2 Criteris de correcció

Aquesta PAC es valorarà seguint els criteris següents:

- **Funcionalitat** (5 punts): Es valorarà que el codi implementi correctament el que demana l'enunciat.
- **Documentació** (0.5 punts): Totes les funcions del exercici d'aquesta PAC hauran de ser correctament documentades utilitzant docstrings (en el format que preferiu).
- **Modularitat** (1 punt): Es valorarà la modularitat del codi (tant l'organització del codi en fitxers com la creació de funcions).
- **Estil** (0.5 punts): El codi ha de seguir la guia d'estil de Python (PEP8), exceptuant els casos on fer-ho compliqui la llegibilitat del codi.
- **Tests** (2 punts): El codi ha de contenir una o diverses *suites* de testos que permetin comprovar el bon funcionament de les funcions implementades.
- **Requeriments** (0.5 punts): Hi haurà d'haver un fitxer de requeriments que llisti (només) les llibreries necessàries per a executar el codi.
- **README i llicència** (0.5 punts): Es valorarà la creació d'un fitxer de README, que presenti el projecte i expliqui com executar-lo, així com la inclusió de la llicència sota la que es distribueix el codi (podeu triar la que vulgueu).

### Important:

**Nota 1:** De la mateixa manera que en la PAC2, els criteris transversals es valoraran de manera proporcional a la part de la funcionalitat implementada.

Per exemple, si el codi només implementa la meitat de la funcionalitat demanada, i la documentació d'aquesta part és perfecta, aleshores la puntuació corresponent a la part de documentació seria de 0.25.

**Nota 2:** És imprescindible que el paquet que lliureu s'executi correctament a la màquina virtual, i que el fitxer de README que inclogueu expliqui clarament com s'ha d'executar el vostre codi per tal de generar les gràfiques resultants de l'anàlisi.

**Nota 3:** Lliureu el paquet com a un únic arxiu .zip al Registre d'Avaluació Continua.