

1. feladat

Egy baráti társaság 2 csapatra felosztva játszottak egymás ellen. Írj AWK szkriptet, amely a paraméterként kapott fájlban kiszámítja az **A** csapat pontszámait! Az **A** csapat tagjait a nevük alapján a következő szerint ismerheted fel:

- Egy nagybetűvel kezdődik.
- Utána minimum 5 és maximum 17 kis- vagy nagybetű szerepel.
- Ezt egy '.' karakter követi.
- Végül pedig pontosan 4 darab számjegy, melyek egyike sem lehet 3 vagy 7.
- Csak az angol ábécé betűire kell számítanod.

Mindenki, akire illeszkedik ez a minta, az **A** csapatban játszott. A bemeneti fájlban ';' karakterrel vannak egymástól elválasztva az adatok, melyek sorrendje: név, pontszám.

Elvárt kimenet a `game.csv` fájlra: 228

2. feladat

Írj AWK szkriptet, amely a paraméterben érkező dalszövegben reguláris kifejezéssel megszámolja, hogy hányszor szerepel benne a „mama” részszó! Ügyelj rá, hogy kezdődhet kis- és nagybetűvel is, továbbá az utolsó 'a' betű tetszőleges mennyiségben, de legalább egyszer szerepel! Tehát például az alábbiakat is bele kell számolni: „Mama”, „mama”, „Mamaaa” ...

Elvárt kimenet a `rhapsody.txt` fájlra: 7

3. feladat

A kapott inputban információk találhatóak a napról. Írj AWK szkriptet, amely a paraméterben kapott fájlból kiírja azokat a szavakat egymás után új sorba, amelyek tartalmazznak legalább egy számot és közvetlen utána egy % jelet. Tehát például a „70%”, „0,1%-ot”, „2%-ban” szavak megfelelnek az előbbi feltételnek. Egy szónak tekinthetsz minden szóközzel elválasztott szót. A program a legvégén még írja ki, hogy összesen hány ilyen illeszkedést talált!

Elvárt kimenet a `nap.txt` fájlra:

```
99,87%-a
0,1%-ot
70%
28%
2%-ban
35%
6
```

4. feladat

Tamás tudni szeretné, hogy 2021-ben mely napokon tudna elég időt rászánni arra, hogy az egyik kedvenc játékával játszani tudjon. Az input fájl első sora tartalmazza a fejléceket. Az első oszlopban található a dátum az alábbi két formátumban:

- EEEE.HH.NN
- EEEE-HH-NN

Mind a kettő megfelelő formátum. A második oszlopban található, hogy ez éppen melyik napra esik, de ezeknek a napoknak egy rövidített kódolása található a sorokban, mégpedig: H:hétfő, K:kedd, SZE:szerda, CS:csütörtök, P:péntek, SZO:szombat, V:vasárnap.

A harmadik oszlopban található az, hogy előre láthatóan aznap a 24 órából hány órát kell eltöltenie egyéb elfoglaltságokkal. Ezek az oszlopok ';' karakterrel vannak elválasztva egymástól. A feladatod megtalálni azokat a dátumokat, amelyek:

1. 2021-es dátumok, a hónapra nem kell figyelned, hogy esetleg érvénytelen számú hónapot tartalmazhat, mert biztosan megfelelő lesz, illetve Tamás minden 3-ra, 5-re és 7-re végződő dátumokon egészen biztosan nem szeretne játszani.
2. Továbbá csak kedden (K), szerdán (SZE) vagy pénteken (P) játszik, egyéb napokon biztosan nem.
3. Tamás csak akkor látja értelmét játszani, ha egy nap legalább 7 óra hosszát el tud ezzel tölteni, egyéb esetben inkább el sem kezdi (ezt számold ki a 3. oszlopban található érték alapján). Továbbá felteheted, hogy Tamás nem szokott aludni, a teste megfelelően regenerálódik a játékok okozta élvezet hatására.

Minden ezeknek megfelelő dátumot (de csak azt, egyéb információ nem kell a sorba) írd ki a konzolra soronként, majd az utolsó sorba írd ki, hogy ez összesen hány óra játékot jelent erre az évre.

Elvárt kimenet utolsó 10 sora a `wow.csv` fájlra:

```
2021-11-12
2021.11.16
2021.11.19
2021-11-24
2021.11.26
2021.11.30
2021-12-10
2021-12-14
2021.12.31
806
```

5. feladat

Egy hosszú szorgalmi időszak után a legtöbb diáknak nem sok ereje marad a vizsgák felkészülésére. Egyesek csak azt szeretnék, hogy végre érjen véget a félév és megjelenjen a leckekönyvben az összes kurzus mellett az a csodálatos pipa. Kuba szeretné minél előbb letudni a vizsgaidőszakot, hogy megpihenjen ezen keserves félév után.

Segíts hát neki, és írd szkriptet, ami a lehetséges tantárgyakból összeállít egy sorrendet, ami alapján a leggyorsabban végezhet! A `csv` fájl 2 oszlopot tartalmaz. Az első oszlopban van a tantárgy neve, míg a második oszlopban a dátum, melyek ';' karakterrel vannak egymástól elválasztva. Egészen biztosan nincs egy napra két különböző vizsga is hirdetve, továbbá felteheted, hogy minden vizsga egy hónapban van, így neked csak a napot kell figyelembe vened (a dátum formátuma `év.hónap.nap`). A sorrendiséget úgy kell meghatároznod, hogy minden egyes tárgyból kiválasztod a legkorábbi időpontot és egy **tömbben** tárolod el (amelynek indexe a tárgy neve, értéke pedig a legkorábbi időpont legyen). Ezután ugyanebben a formátumban írd ki a konzolra az eredményt, tehát 1-1 sorban 1-1 kurzus található ';' -vel elválasztva azzal a dátummal, amelyik a legkorábbi időpont!

Elvárt kimenet a vizsga.csv fájlra:

```
Szkriptnyelvek;2020.12.19
Kozelito_es_szimbolikus_szamitasok;2020.12.20
Adatbazisok;2020.12.14
Rendszerfejlesztés_I;2020.12.18
Programozás_II;2020.12.15
```

6. feladat

Egy kísérlet vége felé jártok, és egy nagyon titkos számítás lehetséges eredményét kell előállítanod. Mivel egy nagyon komplex feladatról van szó, ezért kémcsőben DNS számítás során bizonyos eljárások használatával végeztétek el a számításokat. Sajnos túl előre elkezdtetek örülni a sikernek, így az esti buli miatt Tamás által hozott vinjak és egyéb szennyeződés keveredett a kémcsőbe. A feladatod, hogy megtaláld azokat a DNS szekvenciákat, amelyek tartalmazhatják még így is a megfelelő eredményt. Egy DNS szekvencia 4 lehetséges bázis sorozataként tevődik össze. Ezek a bázisok az adenin (A), guanin (G), citozin (C) és a timin (T). Ezek a bázisok egymással csak bizonyos módon tudnak kapcsolódni. Az adenin a timinnel 2 hidrogén kötés során, míg a citozin a guaninnal 3 hidrogénkötéssel (ezek egymás komplementerei).

Amit biztosan tudni, hogy a lehetséges eredményt adó DNS szekvencia hossza pontosan 20 karakter hosszú. Illetve tudni, hogy az eredmény DNS szekvencia kezdő részszekvenciájának komplementere az alábbi sorozat (tehát neked a fent említett kapcsolódások alapján ennek a sorozatnak a komplementerére kell szűrnöd): „AATCGCCT”. Az utolsó 4 karakter pedig biztosan ugyanazon bázisból tevődik össze (tehát lehetséges értékek: AAAA vagy TTTT vagy CCCC vagy GGGG). A köztük lévő karakterekről mindösszesen annyit tudni, hogy minden bázis duplán szerepel (tehát az első 8 karakter után, ha van egy 'A' betű akkor a következő karakternek is biztosan 'A'-nak kell lennie, a harmadik viszont lehet bármi, például 'T' de ekkor az őt követő bázisnak szintén 'T'-nek kell legyen). Egyéb tulajdonság nem ismert, az összes erre illő DNS szekvenciát írd ki a konzolra egymás után új sorban!

Elvárt kimenet a dns.txt fájlra:

```
TTAGCGGAAACCGGCCTTTT
TTAGCGGAAAGGTTCCCCC
TTAGCGGACCTTCCAATTTT
TTAGCGGATTAATTAATTTT
TTAGCGGATTAATTAACCCC
TTAGCGGAAACCTTCCAAA
TTAGCGGATTAATTCCTTTT
TTAGCGGATTAACCAACCCC
```