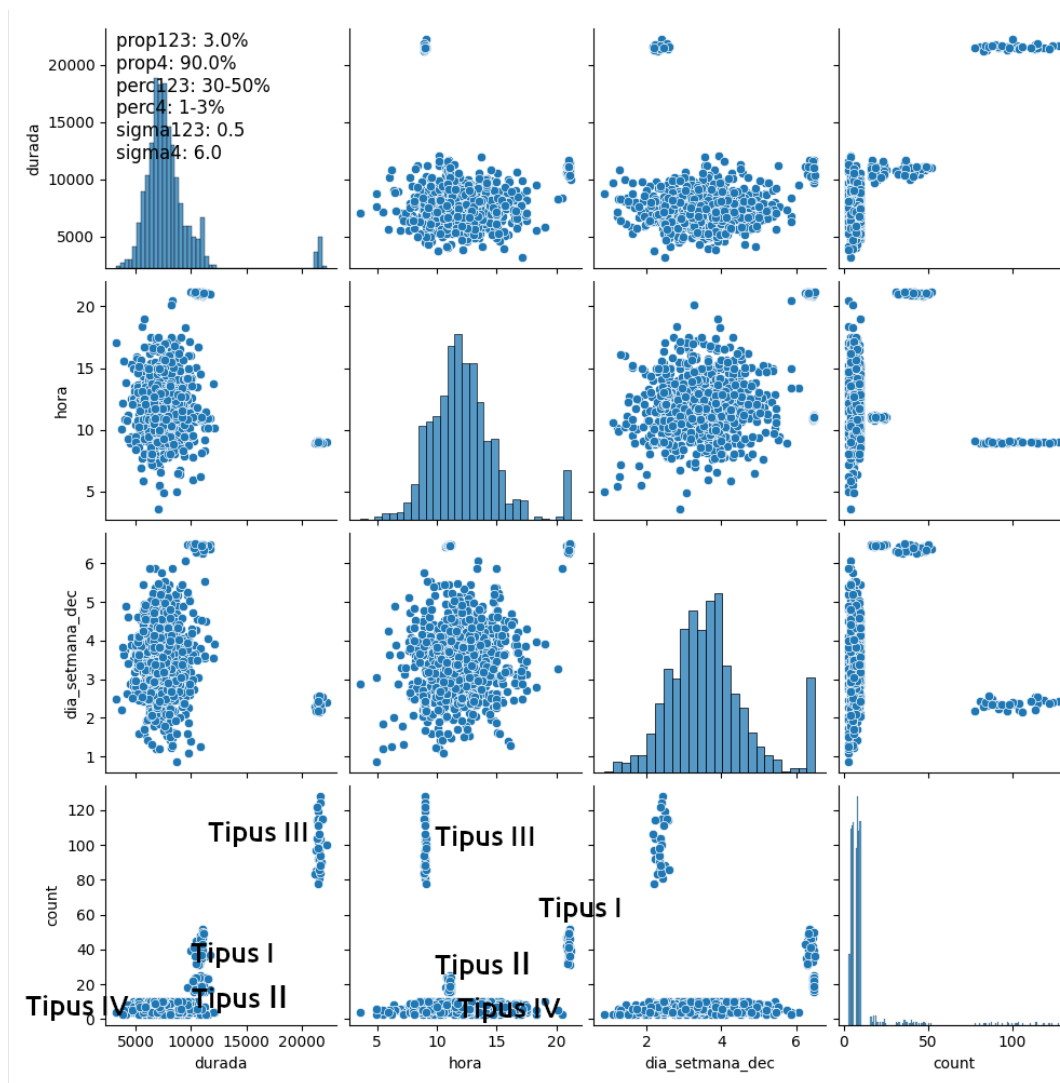


# Informe de les simulacions de Pàrking

Després d'haver generat 27 imatges de simulacions que hem aconseguit variant els paràmetres, podem ara saber sota quines condicions es podran distingir els 4 clústers amb una mica de garanties.

Analitzarem 4 imatges diferents, corresponents a les imatges 8.png, 15.png, 21.png i 0.png. Les analitzem en una progressió de manera que amb la primera podem distingir els 4 clústers, i cada vegada serà més difícil.

## 8.png



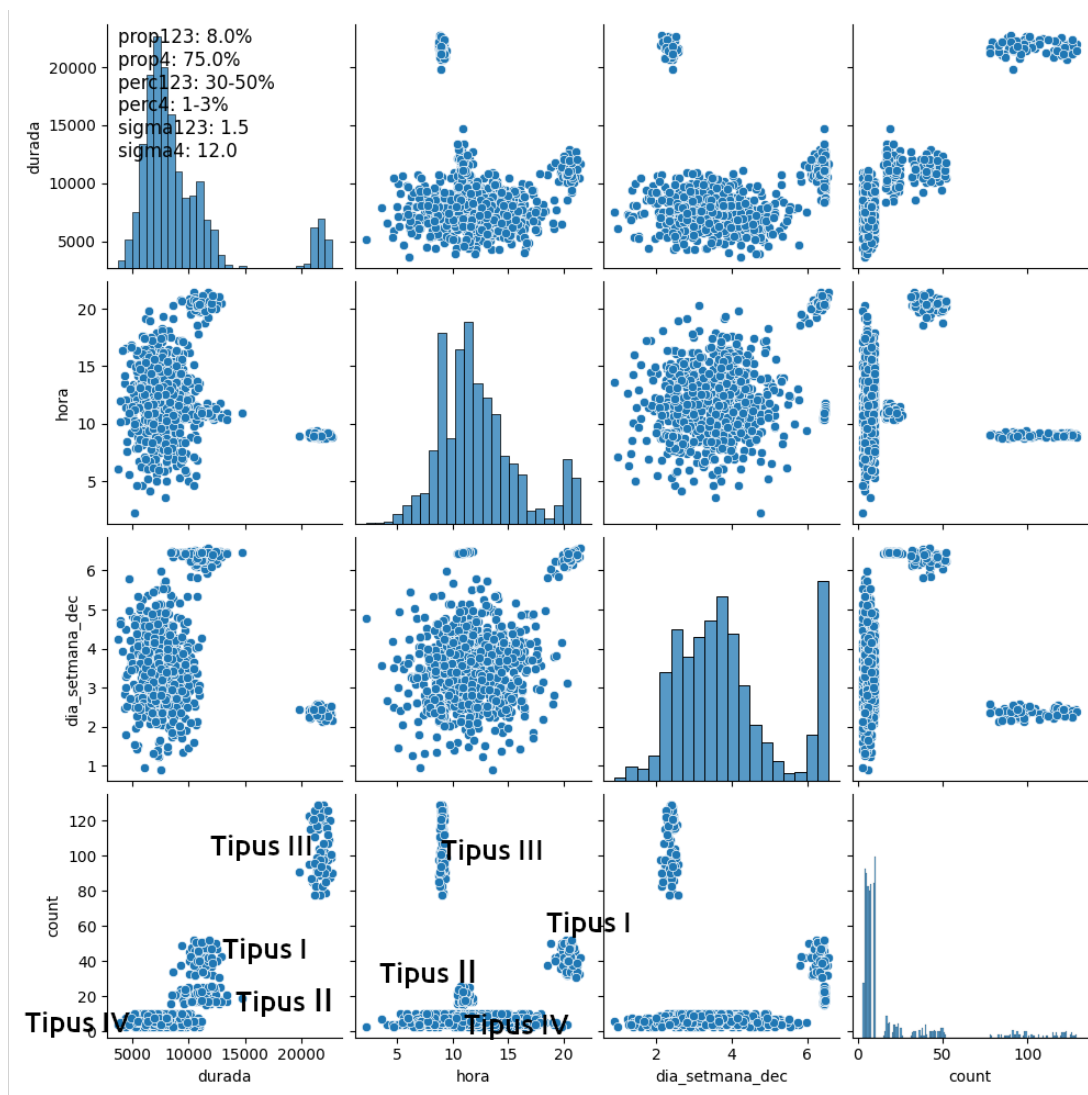
Aquesta és la gràfica més fàcil d'interpretar, doncs es distingeixen clarament els 4 clústers. Les variables clau per distingir els clústers són la durada, hora, i sobretot el count. I menys clau és el dia\_setmana\_dec (dia setmana decimal).

A la imatge 8.png distingim bé els clústers perquè la desviació típica dels tipus I, II i III és 0.5 (mitja hora), que significa que els valors estan molt concentrats i això facilita localitzar els clústers. Si ens fixem en hora/count veiem que:

- Tipus I (diumenges vespre). Efectivament veiem com les mitjanes d'entrada al pàrking són una mica més tard de les 20h.
- Tipus II (dissabte al matí). És difícil de distingir del tipus IV, és com un mugró que li surt al tipus IV. Aquest mugró és degut a què el count és més elevat, són assidus del dissabte al matí. Si ens fixem en el dia\_setmana\_dec ara sí que podem distingir el Tipus II del Tipus IV, però compte que en d'altres gràfiques tendeixen a fusionar-se el tipus I i el tipus II si ens fixem en el dia\_setmana\_dec.
- Tipus III (gent que va a treballar els dies feiners). És fàcil de localitzar perquè entren a treballar cap a les 9, i el count és un valor alt (són assidus al pàrking).
- Tipus IV (cotxes normals, soroll, representen el 90% i tenen una desviació gran). El tipus IV és el núvol de punts gran que dificulta la separació dels clústers. Té com a tret principal que el count és baix, i per tant poden aflorar els altres clústers.

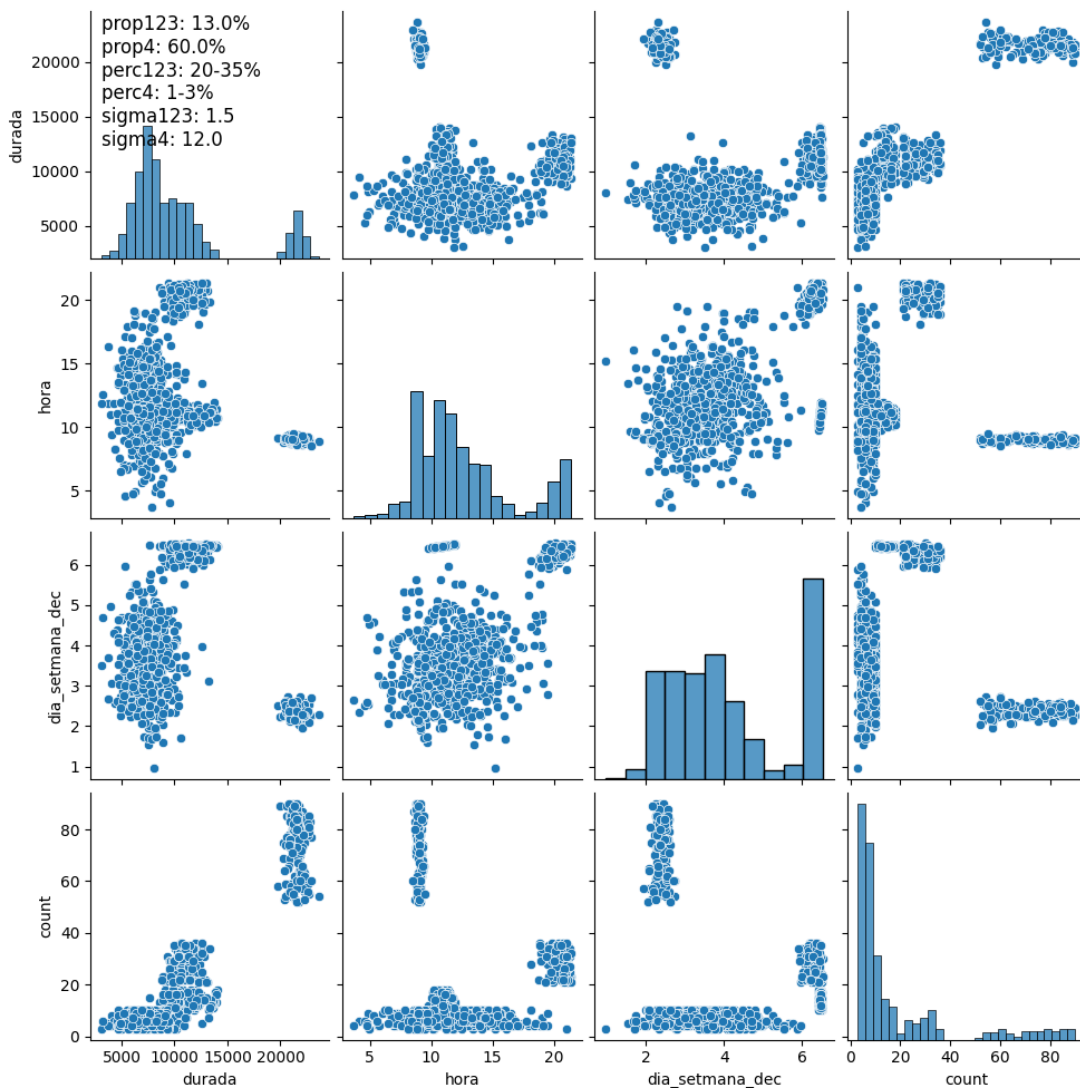
Veiem, doncs, que el **count** és un paràmetre important per poder ubicar un cotxe en un dels clústers. Això significa (com és lògic) que quan un cotxe visita el pàrking per primera vegada no se li pot assignar un clúster. Només quan passa un període llarg de temps (per ex un any), i veiem el comportament dels cotxes, podrem treure conclusions i assignar-lo a un clúster. El paràmetre count té molt a veure amb la proporció i el percentatge. En aquest cas encara que la proporció és petita, el percentatge és gran.

## 15.png



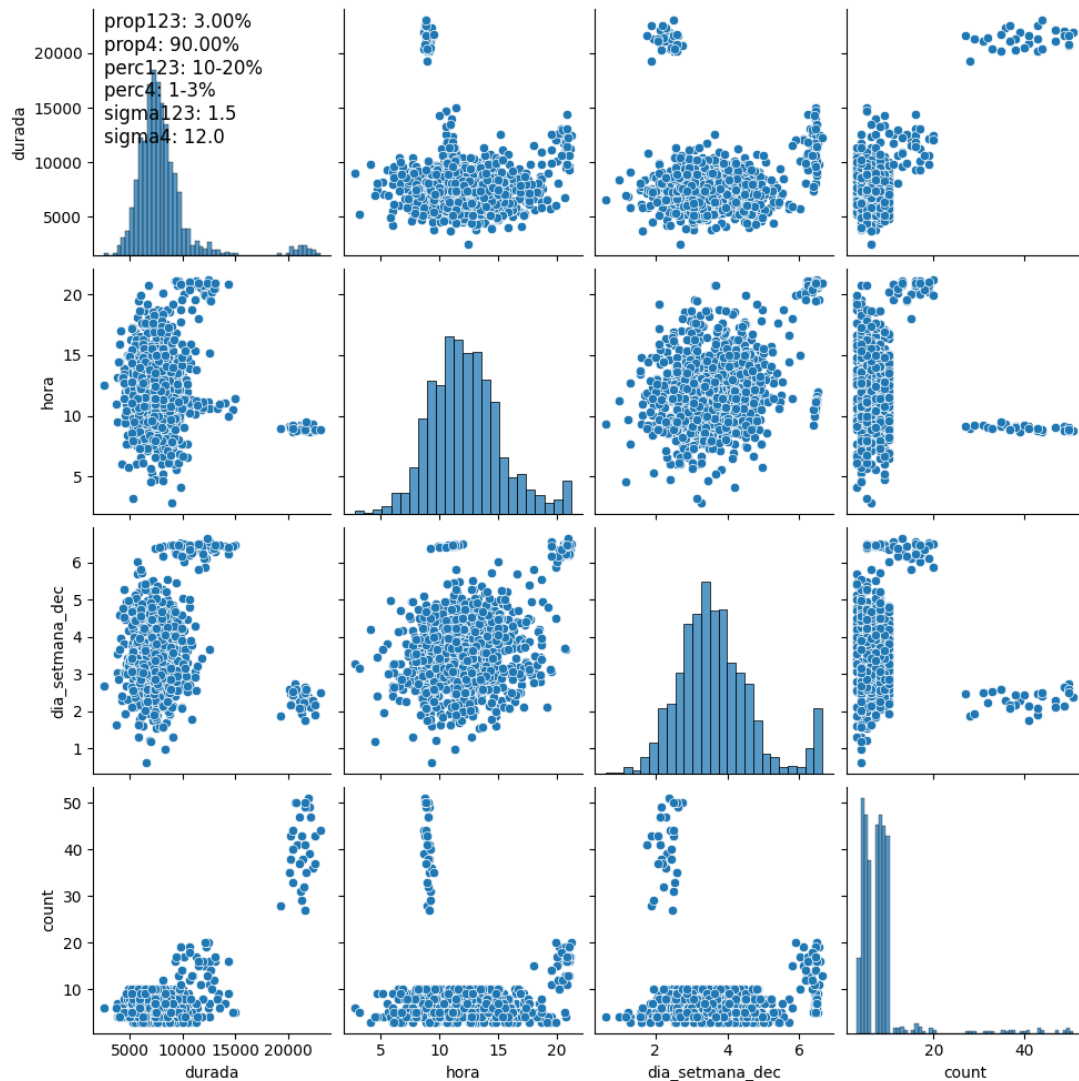
En l'arxiu 15.png tenim els clústers que encara es poden distingir, però ara és un cas diferent. Com que sigma val 1.5 (marge de  $\pm 1.5$  hores en l'hora d'entrada i en la durada) fa que els núvol de punt són més dispersos. Si ens fixem en *dia\_setmana\_dec*/*count*, veiem com el tipus I i tipus II (a la dreta de la imatge ja no es distingeixen (a 8.png es distingien clarament)). Com que les proporcions són del 8% pels tres primers tipus (i el 75% pel soroll), ara fa que tinguem més quantitat de punts en els tipus I, II i III, i això ajuda també a distingir els clústers (com es veu a *durada*/*count*). Si ens fixem en *hora*/*count* encara podem distingir el tipus II i tipus IV.

## 21.png



Aquí ja es fa difícil distingir els clústers, si ens fixem en durada/count, els tipus I, II i IV s'han fusionat. Si ens fixem en hora/count, el mugró del tipus II ja no es distingeix tant respecte el núvol de punts del tipus IV. Si ens fixem en dia\_setmana\_dec/count, els tipus I, II s'han fusionat. Aquestes fusions és degut a què sigma és alt (1.5h) i el número de punts I, II i III ha augmentat (el soroll és el 60% ara).

0.png



Per tot el que hem vist aquest és el cas més difícil (o impossible) de separar els clústers, doncs tenim una dispersió alta (1.5h), i el número de punts dels tipus I, II i III és baix (proporció 3%, percentatge 10-20%). El tipus III es pot distingir clarament, però els tipus I i tipus II es fusionen. I mirant hora/count ara serà impossible fer aflorar el tipus II.