Csapattagok: Angeli Imre, Nagy Zsuszanna, Papp Mercédesz

Felépítés:

Az elkészült dossziénk 3 fejezetből áll. Az első fejezet egy általános áttekintéssel és érdekességekkel kezdődik. Ezután mélyebben belementünk a filmek különböző tulajdonságaiba, mint például a filmek játékideje vagy a besorolás. Megnéztük és kigyűjtöttük, hogy kapott-e Oscar díjat az adott film. A következő fejezetben pénzügyi oldalról nézegettük meg a filmeket. Az utolsó fejezetben általánosabb adatok elemzéséről szólt, olyanoké, amiket a filmet kedvelők tudhatnak, utána szoktak nézni, akár ez alapján kiválasztva a következő megnézendő filmet.

Vizualizációk:

A lejátszási idő esetén azt láthatjuk, hogy az átlagos, 120-140 perc közötti hosszúágú filmek a legjellemzőbbek. A szignifikánsan magas, illetve alacsony értékek viszonylag kis számban fordulnak elő. Ennek megfelelően a hisztogramon a normális eloszlásra jellemző haraggörbéhez hasonló alakzat rajzolódik ki, ugyanakkor attól enyhén jobbra eltér. Ennek legvalószínűbb oka, hogy míg a leghosszabb filmek lejátszási ideje korlátlan hosszú lehet (akár több mint 5 óra), addig a legalacsonyabb értékeknek van egy alsó korlátjuk (0 perc).

A halmozott dobozábrán azt figyelhetjük meg, hogy a lejátszási idő eloszlása erősen eltér a műfajok szerinti bontásban. A medián szempontjából a történelmi témájú filmek a leghosszabbak (140 perc). Ugyanakkor a legmagasabb kiugró értékeket az akciófilmek, vígjátékok, illetve krimik esetén tapasztalhatjuk. A viszonylag rövid terjedelem leginkább a vígjátékokra, illetve a családi filmekre jellemző.

A cenzúra besorolás esetén a perecdiagram alapján azt láthatjuk, hogy az összes film több mint 40%-a a korhatáros filmek közé sorolható. Az egyes arányok ugyanakkor közel sem egyeznek meg egymással országok szerinti bontásban, ahogy azt a 100%-ig halmozott sávdiagramm is mutatja. Japánban, Indiában, illetve Franciaországban a nem értékelt filmek vannak többségben, Németországban, az Egyesült Államokban, illetve az Egyesült királyságban pedig a korhatáros filmek a leggyakoribbak. Láthatjuk tehát, hogy jelentősek a különbségek a besorolásokban az egyes országok filmkínálatát tekintve, tehát a két változó nem tekinthető függetlennek egymástól.

Erősen eltérnek egymástól az összbevételek is a besorolások, illetve a műfajok függvényében. A hőtérképen azt láthatjuk, hogy a legtöbb besorolás esetén a dráma filmekhez kapcsolódik a legnagyobb bevétel, azonban az általános besorolású, valamint a szülői felügyelettel megtekinthető filmek körében az akciófilmekhez kapcsolódik a legmagasabb befolyt összeg.

Összefüggések:

A filmek esetén nem különíthető el tisztán egy műfaj, az azonban egyértelmű, hogy a dráma a legnézettebb mind közül ( összesen 184 besorolás) . A műfaj dominanciája megmutatkozik az Oscar díjak esetén is, ugyanis közel 44%-uk nyert.

Első ránézésre azt gondolnánk, hogy ha egy film magas IMDB értékeléssel rendelkezik, akkor biztosan nem lehetett veszteséges. Vegyük példának a 12 Angry Men című filmet, amely egyidőben ötödik a toplistán és második a legnagyobb veszteséggel járó filmek listáján is. A toplistán első helyet foglaló The Shawshank Redemption című film szintén a veszteséges filmek táborát erősíti.

A dosszié összeállításakor tüzetesen megvizsgáltunk minden mennyiségi változót és kerestük a köztük lévő kapcsolatokat. A legerősebb egyirányú kapcsolatot (R2=70%) a költségvetés és a filmek bruttó összértéke között találtuk. A korreláció erőssége azt mutatja, hogy a nagyobb költésgvetésű filmek várhatóan sikeresebbek. A költségevetés eloszlását tekintve megfigyelhető, hogy a filmek közel 50%-ka az első osztályközbe koncentrálódik. Ha tovább vizsgálódunk, akkor az látható, hogy ezeknek a filmeknek közel 35%-ka nem térült meg, a második osztályköz esetén pedig ez a filmek 22%-ára jellemző. Leginkább az első két osztályköz esetén figyelhető meg ez a tendencia, a többi osztály esetén jóval kisebb arányban vannak jelen a veszteséges filmek és a nagyobb költségvetési kategóriákban ez az arány 0-ra csökken. Ennek okára a jelen helyzetben rendelkezésre álló adatok alapján nem lehet pontos választ adni, ugyanis nem lehet teljes mértékben elkülöníteni, hogy ez a tendencia az egyenlőtlen gyakorisági kategóriák kialakításának a következménye vagy pedig a fentebb bemutatott korrelációs kapcsolat szüleménye.