

**OKOSTELEFONOK VILÁGSZERTE**

Készítette: Raduly Izabela Andreea

1. **Bevezető**

Az utóbbi évtizedekben a mobiltelefonok nagymértékű fejlődésének lehetünk szemtanúi. A több 40 évvel ezelőtti első hívás óta már napjaink egyik legfontosabb eszközévé váltak, köszönhetően a fejlett technológiának. Napjainkban szinte nincs olyan ember akinek ne lenne okostelefonja, amlely az alapfunkciókon kívül sok olyan funkcióval rendelkezik, ami megkönnyíti a kommunikációt és elénk tárja a világot néhány négyzetcentiméteren.

Projektem témájaként azért választottam az „Okostelefonok világszerte” témát, mert mindig is vonzott ez a terület és szerettem volna megtudni többet az okostelefonok piacáról.

Az adatvizualizációhoz a Statista oldalt használtam, ugyanis ez az egyik legmegbízhatóbb adatforrás és rengeteg adatot lehet rajta találni, így én is megtaláltam ott mindent, amire szükségem volt.

1. **Adatfelfedezés, normalizálás**

Mivel az adataimat nagyrészt a Statisztáról gyűjtöttem ezért formailag nagyon hasonlítottak, de volt némi eltérés is. Első sorban, letöltöttem őket .xlsx formátumba. A hasonlóságokat kihasználva egy sablon segítségével normalizáltam az adatokat, amit persze egyes esetben bővíteni kellett.

A sablon tartalma a következő alaplépésekre osztható:

1. A kívánt xlsx fájl beolvasása Driveról.   
   xls = pd.ExcelFile(".xlsx")
2. Mivel 2 oldalból áll az excel fájl, ezért a df2-be mentettem azt amelyik az adatokat tartalmazza (Data)  
   df2 = pd.read\_excel(xls, 'Data')
3. Oszlopok elnevezése  
   df2.columns=['dumb','…','….', '….'] #oszlopok átnevezese
4. Mivel a fájl tartalmazott egy üres oszlopot, ami nekem nem volt hasznos, törölni kellett az alábbi sor segítségével  
   df2.pop("dumb")
5. Első 4,egyébként üres, sor törlése, mert nincs szükségem rájuk  
   df2=df2.iloc[4:]
6. Mentés .csv formátumban indexelés nélkül  
   df.to\_csv(".csv",index=False)

Ezt a sablont még néhány helyen bővítettem bizonyos oszlopok értékeinek 100-al való osztásával, ami későbbiekben a százalékos megjelenítésnél volt hasznos.  
 Pl. df2["Male"] = df2["Male"] /100

Előfordult, hogy évszámot kellett átalakítanom a következőképp:

df2['year']=pd.to\_datetime(df2['year'].astype('str'))

Volt olyan is, hogy vesszőt kellett “törölni”

df2=df2.replace(',','')

A térképekhez szükség volt némi geokódolásra is ahol létrehoztam egy úgynevezett “koordináta szótárat”,amit hozzáfűztem a már formázott adathalmazhoz. Így a tabla bővült a az ország hosszúsági és szélességi koordinátáival.

1. **Vizualizáció menete**

Vizualizációs eszközként a Power BI –t választottam, mert az tűnt számomra a legegyszerűbbnek. (Itt megjegyzem azt is, hogy mégsem volt olyan egyszerű számomra, mint ahogy elképzeltem). Az eszközzel való megismerkedés után, már nem volt baj a munka menetével.

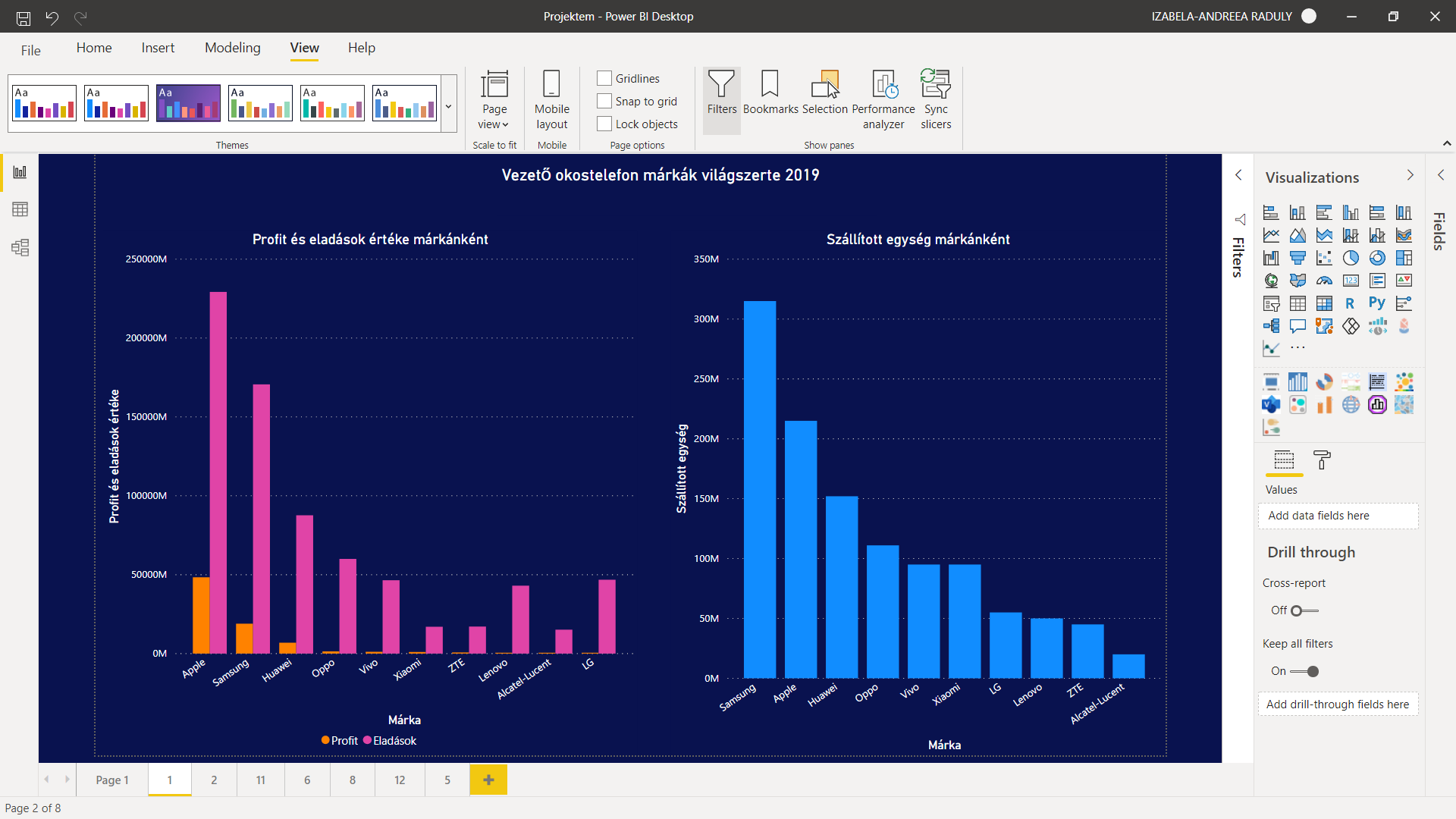
Nulladik lépésként lejegyeztem, hogy mindegyik adathalmazhoz milyen típusú vizualizációt készítek. A konkrét munkafolyamatot azzal kezdtem, hogy letöltöttem a Power BI Desktop verzióját, mert sokkal praktikusabbnak találtam a webesnél. Második lépésként létrehoztam egy új „Report”-ot, és ezek után feltöltöttem minden .csv fájlt amire szükségem volt. Miután mindegyik vizualizációhoz kiválasztottam a megfelelő diagramot, beállítottam a megfelelő mezőket, az évszámok és márkák általában az X tengelyen találhatók, a számbeli értékek pedig az Y-on. A térképek megjelenítéséhez, Latitude,Longitude-ot is beállítottam. Amiután kirajzolódott a megfelelő vizualizáció elkezdtem a design-al foglalkozni. Beállítottam a hátteret, az elemek és a betűk színeit és méretét valamint elhelyeztem úgy, hogy jól nézzen ki.

Továbbiakban részletezem mindegyik vizualizációt külön-külön.

1. **Vezető okostelefon márkák világszerte 2019 –ben**

Ez a vizualizáció 3 paramétert vesz figyelembe: Profit, eladások értéke, és szállított egységek. Az Axis-hoz beállítottam a Márkát , a Value-hoz pedig egyik esteben a Profit és eladások értéke, valamint a Szállított egységet

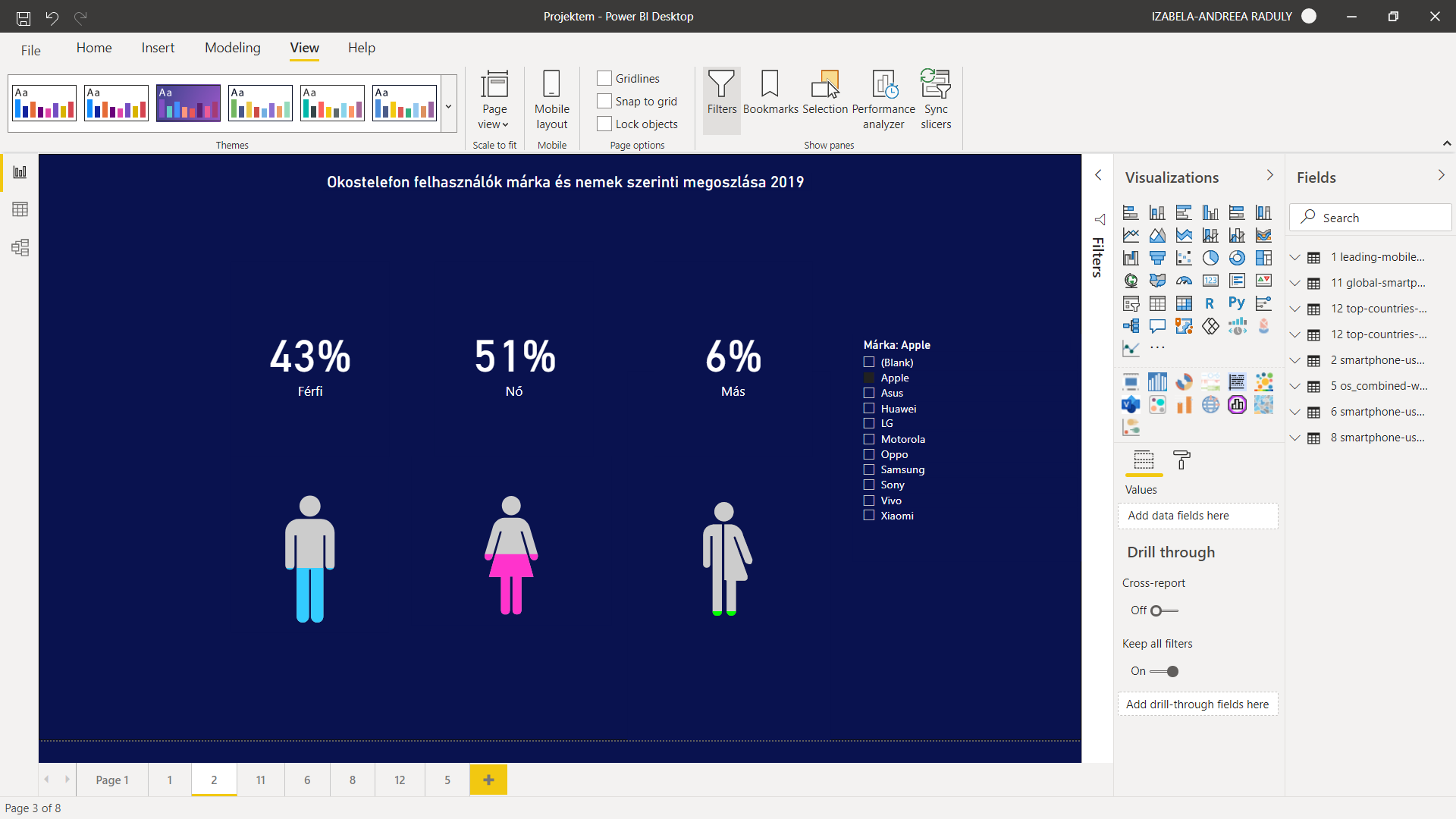
Legmegfelelőbbnek találtam a „Clustered column chart”-al való megjelnítést, mert könnyen összehasonlíthatóak a márkák. Profit és eladások értékében az APPLE áll az első helyen, szállított egységeknél pedig a Samsung



1. **Okostelefon felhasználók márka és nemek szerinti megoszlása 2019**

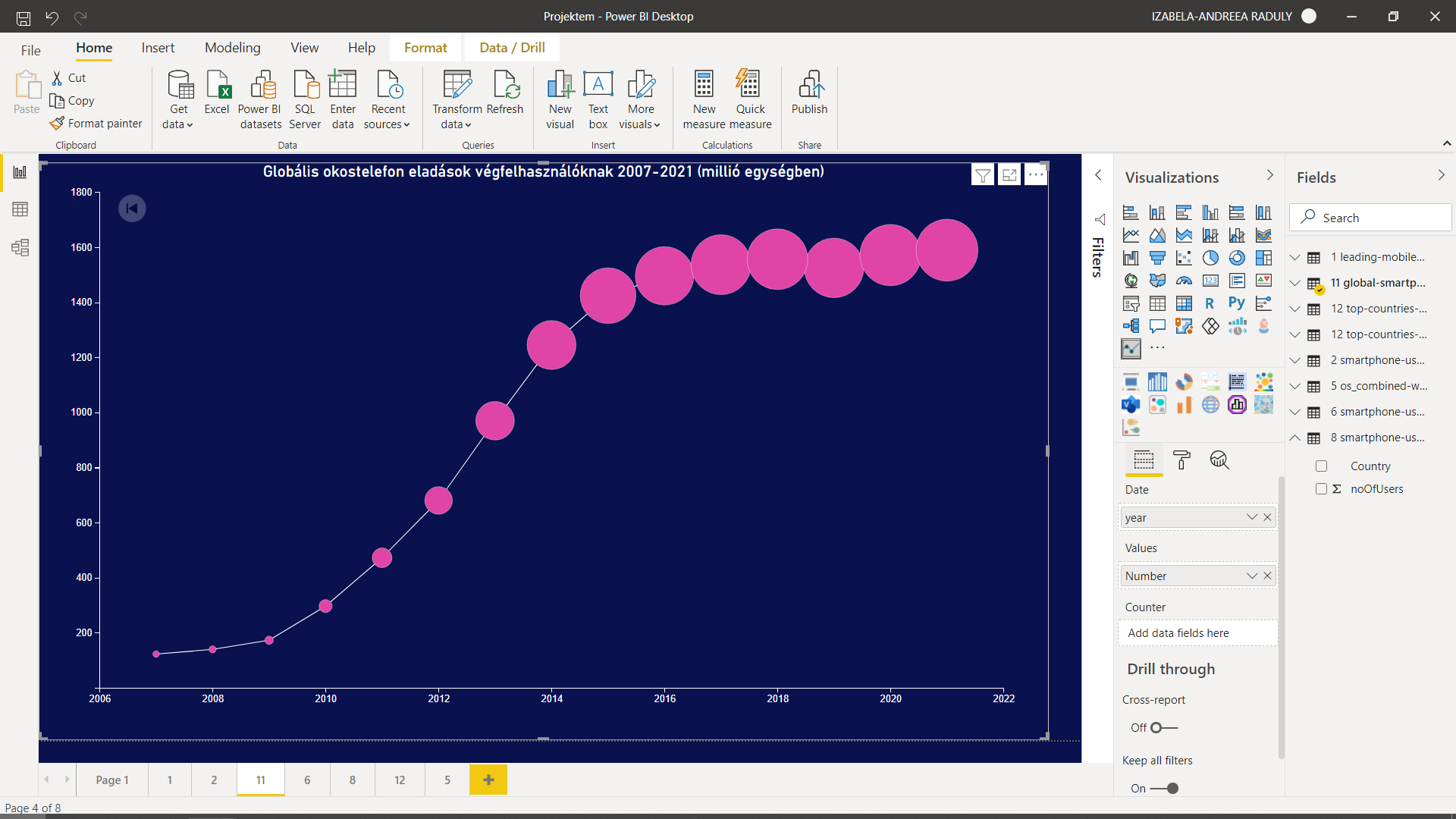
Ez a vizualizáció nemek szerint mutatja be egyes márkák felhasználóit.

Ehez Infographic Designer-t használtam, mert itt beszúrhatók különböző ikonok, amivel látványosabbá varázsolhatjuk a vizualizációt. Erre pedig 3 emberformájú alakzatot használtam, ami a nemeket jelképezi. Az általuk képviselt százalékok által megszineződnek, addig a szintig amit a 100%-hoz képest képviselnek. Továbbá használtam ehez még Cardokat a számok syemléltetésére és Slicert a márkák szerinti szűrésre.



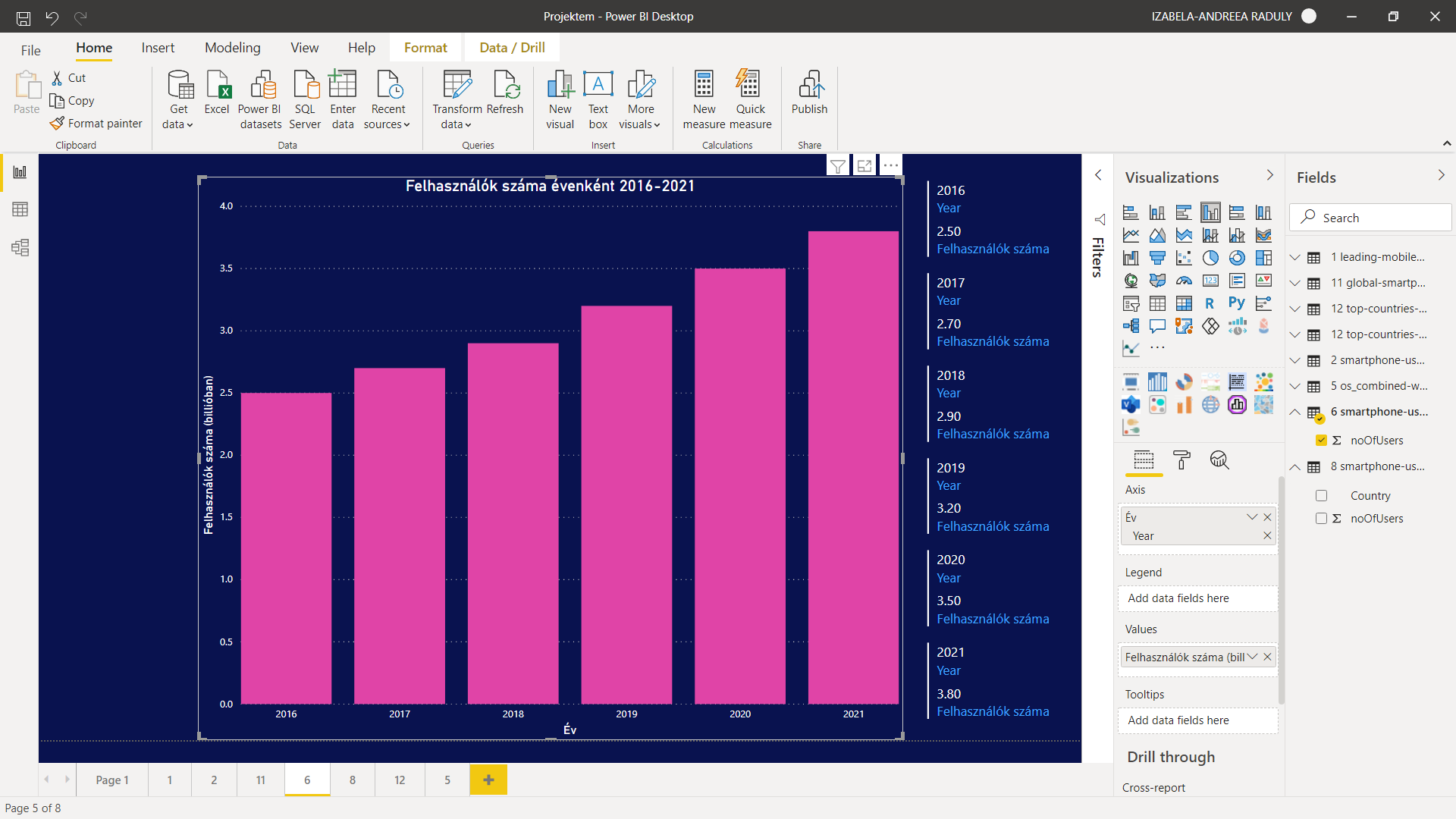
1. **Globális okostelefon eladások végfelhasználóknak 2007-2021 (millió egységben)**

Ez a vizualizáció az eladások növekedését ábrázolja évenként globális szinten. Ugyanakkor egy előrejelzést is kapunk a 2020,2021 –es évekre. Ehez Line dot chart-ot használtam ami animációval mutatja a növekedést.



1. **Felhasználók száma évenként 2016-2021**

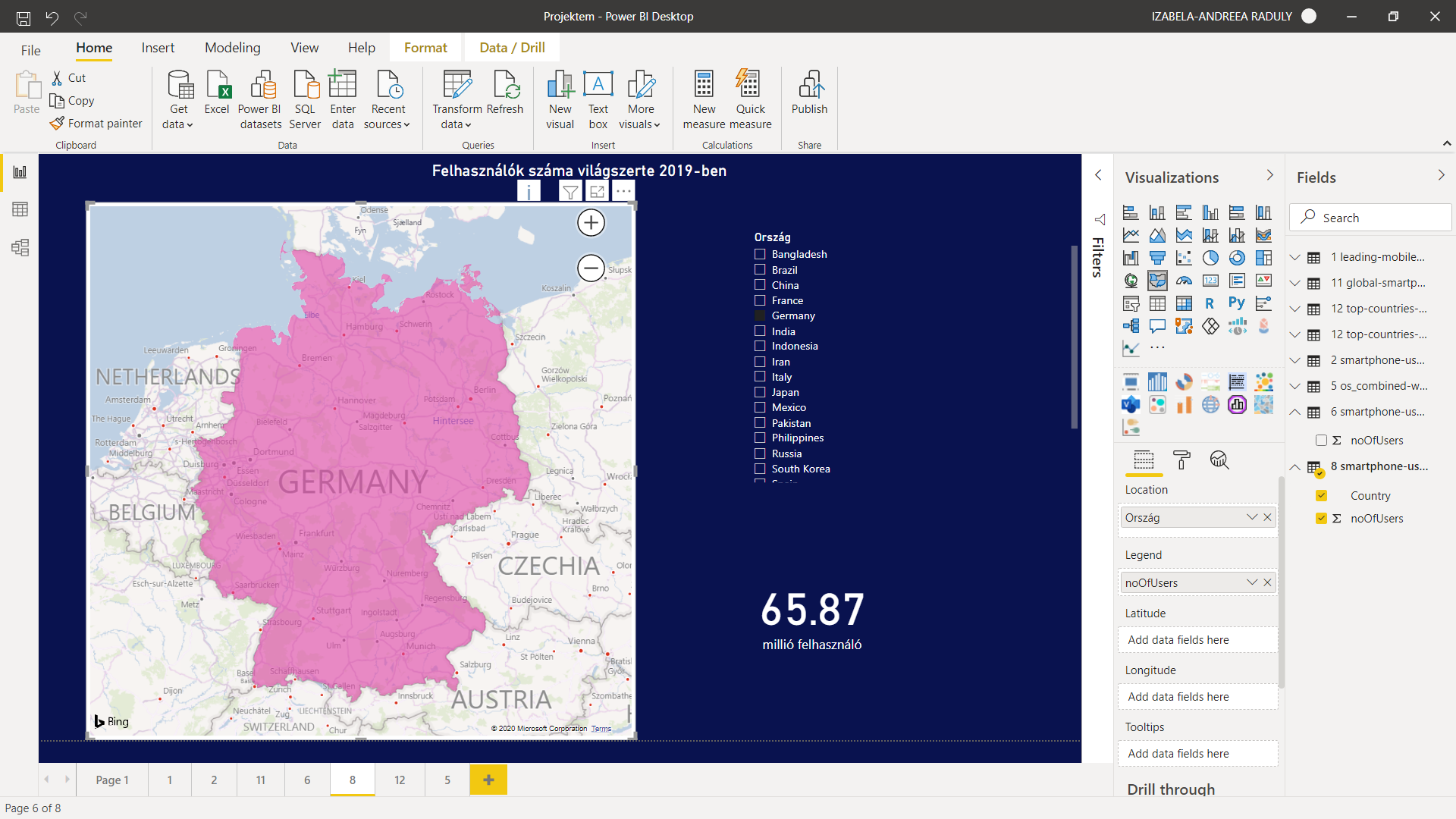
Ahogyan a fenti vizualizációban, itt is kapunk előrejelzést a 2020-as és 2021 es évekre. Ehez a vizualizációhoz a még említett „Clustered colmn chart”-ot használtam , ez által látszanak legjobban a különbségek. Hogy még átláthatóbbak legyenek az adatok felhasználtam még „Multi-row card”-ot is.



1. **Felhasználók száma világszerte 2019-ben**

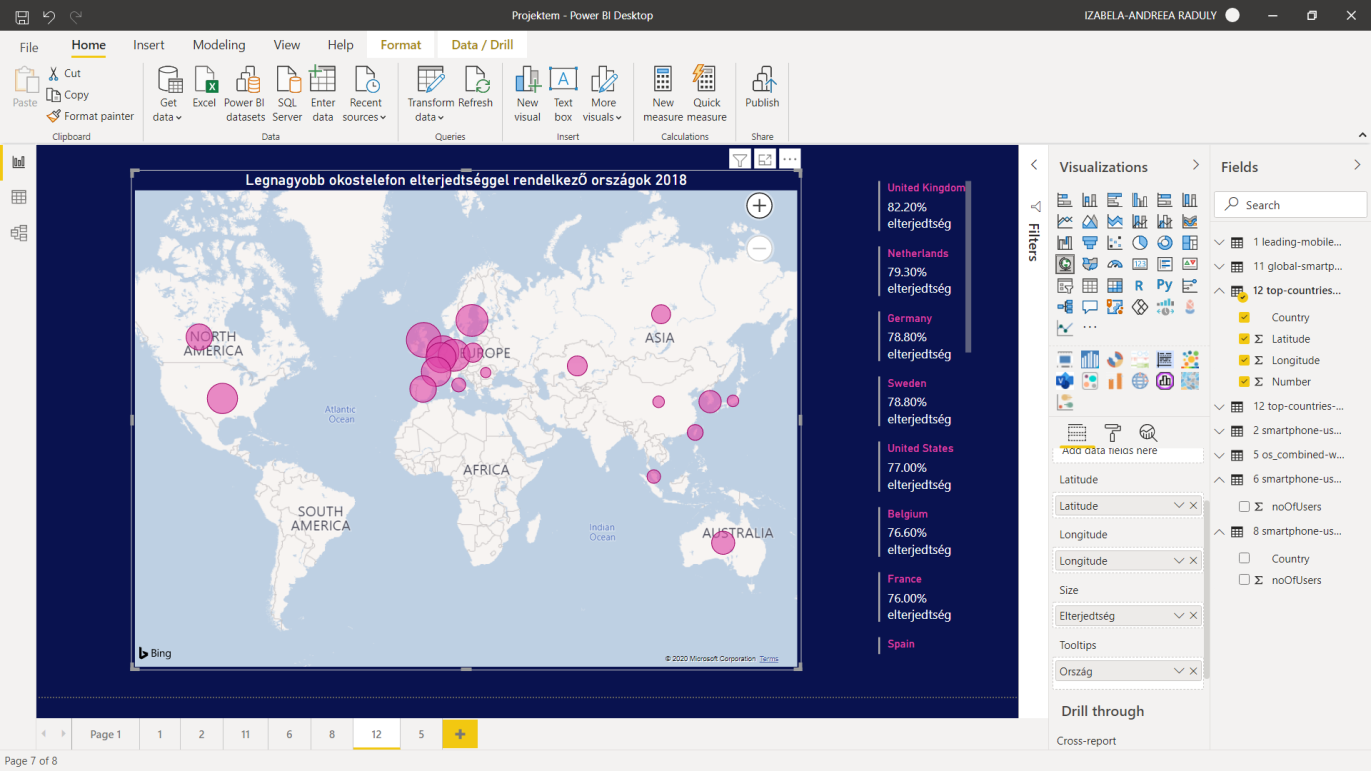
Ez a vizualizáció országonként mutatja be a felhasználók számát a 2019-es évre.

Mivel országokról van szó, így „Filled map”-et választottam a vizualizációhoz. A többi vizualizációhoz hasonlóan itt is lehet szűrni az Országokat, valamint külön megjelenik a felhasználók száma a kiválasztott országban.



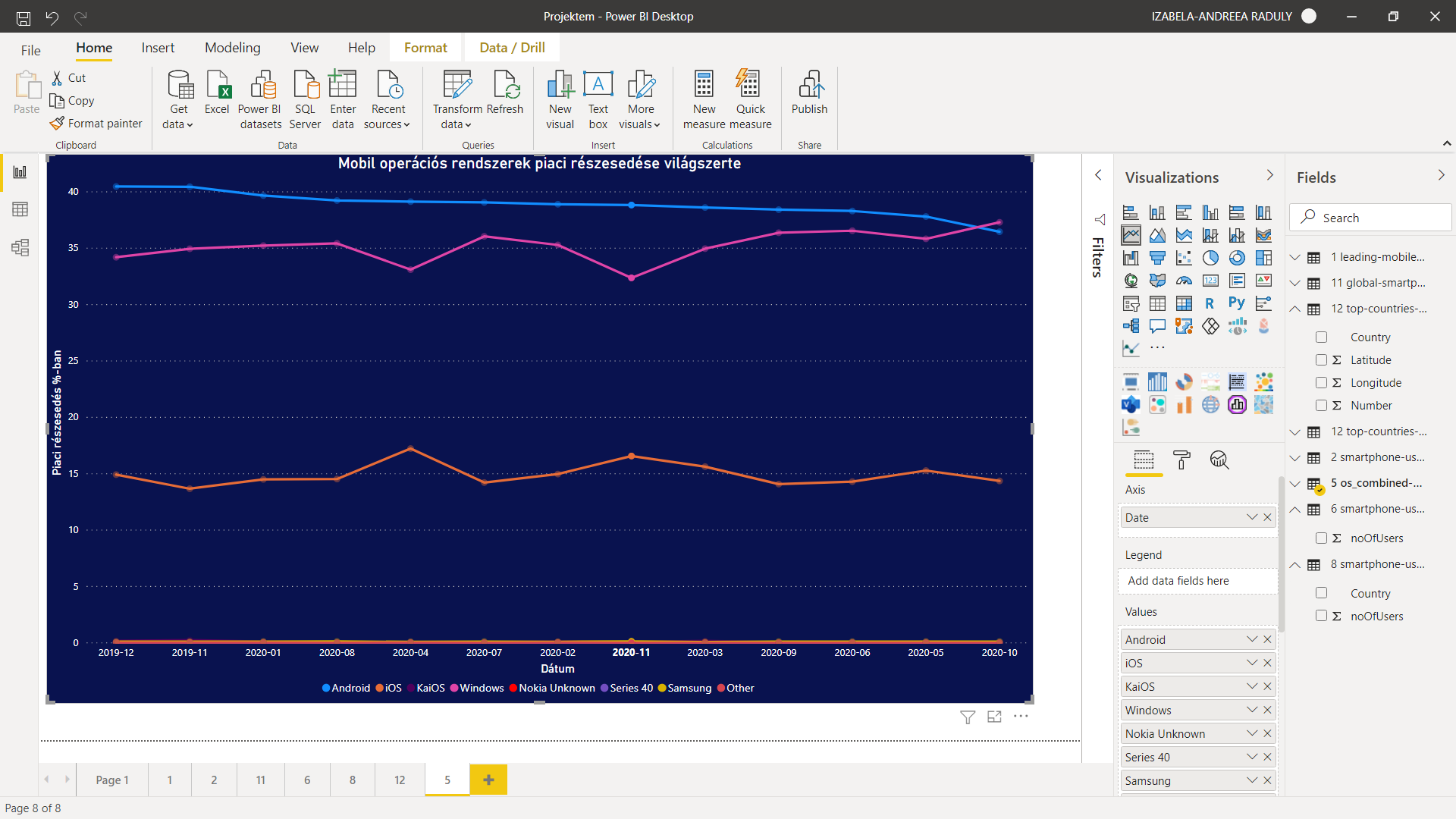
1. **Legnagyobb okostelefon elterjedtséggel rendelkező országok 2018**

Ehez a vizualizációhoz „Map”-et használtam, mert a térképen megjelenő körök remekül szemléltetik az elterjedtség nagyságát. A térképet a geokódolás által nyert hosszusági és szélességi koordináták segítségével jelenítettem meg, valamint a körök nagyságát az elterjedtség nagysága befolyásolja. A könnyebb átláthatóság kedvéért felkerült ismét egy „Multi-row card” az adatokkal.



1. **Mobil operációs rendszerek piaci részesedése világszerte**

Ez a vizualizáció bemutatja a többé-kevésbé ismert operációs rendszerek piaci részesedését az utóbbi egy évből. Ennél a vizualizációnál fontos volt, hogy a operációs renszerek párhuzamosan jelenjenek meg és úgyanakkor érzékelhető legyen az idő is, valamint a piaci részesedés. Ennek következtében a legmegfelelőbb választás a „Line chart” lett. Az X tengelyen megtalálhatók az utóbbi 1 év hónapjai, az Y tengelyen pedig a piaci részesedések %-ban.



1. **Megjelenítési forma előnyei és hátárnyai**

A megjelenítési forma előnyei talán azok lehetnek, hogy rendkívül átlátható, nincs zsúfolva. Ugyanakkor kellőképpen részletes és praktikus.

Hátránya talán annyi lenne, hogy a grafikonról, illetve magáról a térképekről vagy a sávokról hiányolom az adatokat. A számok persze megtalálhatók a tengelyeken, vagy mellette, viszont így tévesztőbb.

1. **Alternatívák**

A fennebb említett hátrányok miatt alternatívaként olyan vizualizációkat használnék, ahol az adatok megjelennek a grafikon vagy a diagram belsejében, pl. az adott sávon, vagy térképek esetén mondjuk a számok az oszágokon.

1. **Mit tanulunk ebből?**

Az adatvizualizáció elészítése során megtanultam, hogyan kell az adatokat úgy megjeleníteni, hogy helyes legyen, átlátható és érthető a célközönség számára. Ugyanakkor rájöttem, hogy milyen adathalmazokhoz milyen vizualizáció passzol és, hogy egyáltalán nem mindegy miket használunk fel.

A vizualizáció végeredményéből pedig rájöttem, hogy a mobiltelefonok világában nem is úgy vannak a dolgok, ahogy én azt hittem, szóval magán a folyamaton kívül, nagyon sok olyan információt tudtam meg, ami nem teljesen ehez a tantárgyhoz köthető.