

Metody analýzy dat

Analýza videoherního průmyslu

ONDŘEJ ŘEHÁČEK (REH0063)

December 11, 2016

1 Dataset

Dataset obsahuje statistické informace o prodeji více než 16 tisíců nejprodávanějších videoherních titulů v rozdílných regionech. Je volně dostupný na [kaggle.com](https://www.kaggle.com/gregorydsmith) a byl publikován Gregory Smithem 10. Října 2016. Obsahuje tedy poměrně aktuální data čerpána z ověřeného zdroje ve videoherním světě - [vgchartz.com](https://www.vgchartz.com), pomocí Python skriptu ¹. pro sběr dat z této webové stránky

1.1 Motivace

Videoherní průmysl je již stejně dominantní jako filmový či hudební. Postupem času s vývojem techniky se vyvíjí i tato forma zábavy a v dnešní době mohou videohry vyprávět stejně komplexní příběhy jako celovečerní filmy. Není žádným tajemstvím, že jednotlivé kontinenty mají podstatně jinak rozložený prodej a zájem o videohry. S pomocí tohoto datasetu a jeho analýzy chci tyto rozdíly zdokumentovat a detailně popsat.

1.2 Atributy

Rank	Pořadí titulu v počtu prodejů.
Name	Název herního titulu
Platform	Platforma na které titul běží.
Year	Rok vydání hry.
Genre	Žánr herního titulu.
Publisher	Vydavatel / Vývojář titulu.
NA_Sales	Prodeje na území státu serevní ameriky. (V miliónech)
JP_Sales	Prodeje na území Japonska. (V miliónech)
EU_Sales	Prodej na území státu Evropy. (V miliónech)
Other_Sales	Prodej na území zbylých států. (V miliónech)
Global_Sales	Celkový počet prodaných kusů.

¹<https://github.com/GregorUT/vgchartzScrape>

1.3 Úpravy a chybějící data

Přibližně 27 titulů bylo opraveno, protože obsahovaly nesmyslné hodnoty. Původně je v datasetu asi 329 položek, které nemají kompletní data - jedná se převážně o chybějící rok vydání. Některé z položek byly manuálně doplněny o zjištěná data.

1.4 Otázky, pro které hledáme odpovědi.

Kromě obecné analýzy dat, hledáme odpovědi na následující specifitější otázky:

- Jak se zvětšoval videoherní trh s postupem času ?
- Jaké jsou oblíbené žánry titulů v rozdílných regionech ?
- Jaké platformy byly nejspěšnější ve svých generacích ?
- Které žánry byly nejdominantnější na určitých platformách ?
- Které žánry jsou nejdominantnější v určitých regionech ?
- Preferují určité regiony rozdílné vývojáře ?

1.5 Analýza

Analýza datasetu je provedena v R (IDE RStudio 1.0.44) a byly využity následující balíky:

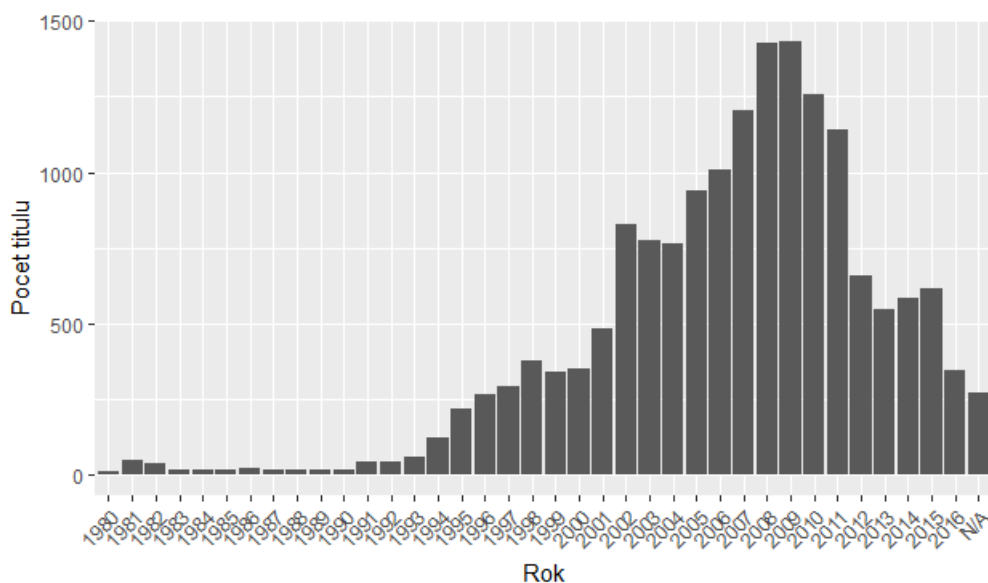
- dplyr
- ggplot2
- dplyr
- scales
- reshape2

2 Obecná analýza

2.1 Analýza atributů

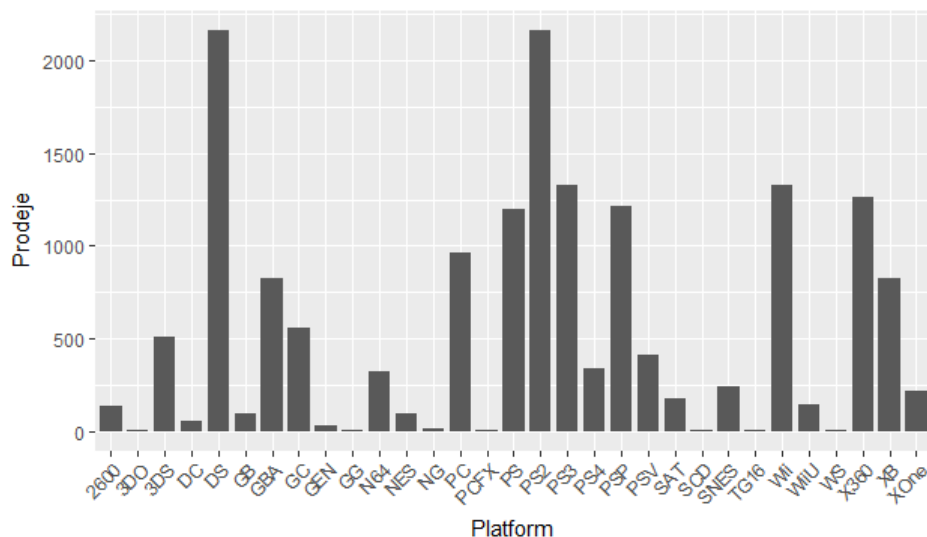
	Length	Min	Mean	Max	Type
Rank	16600	1	8300	16600	Integer
Name	X	X	X	X	String
Platform	X	X	X	X	String
Year	16600	1980	1993	2016	Mixed
Genre	X	X	X	X	String
Publisher	X	X	X	X	String
NA_Sales	16600	0.0000	0.2647	41.4900	Float
EU_Sales	16600	0.0000	0.1467	29.0200	Float
JP_Sales	16600	0.0000	0.0778	10.2200	Float
Other_Sales	16600	0.0000	0.0480	10.5700	Float
Global_Sales	16600	0.0000	0.5374	82.7400	Float

2.2 Počet titulů na rok.

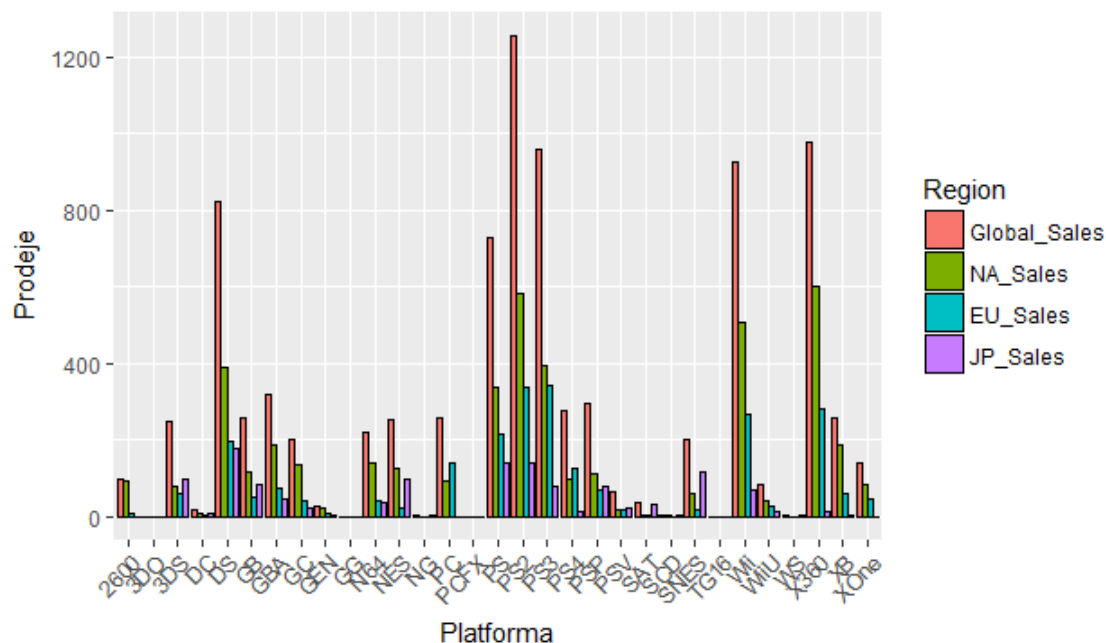


- Nárůst v roce 1994 po dostupnosti **Sony Playstation (102 mil.)**
- Po roce 2001, kdy se na trh dostává historicky nejprodávanější konzole **Sony Playstation 2 (> 155 mil.)**.
- V letech 2005-2011 najdeme hodnoty nejvyšší, vzhledem k tomu, že trh ovládá **Nintendo DS (154 mil.)** , **Nintendo Wii (101 mil.)**, **Xbox 360 (85 mil.)**, **Playstation 3 (84 mil.)** najednou - všechny astronomicky úspěšné produkty.

2.3 Rozložení platforem

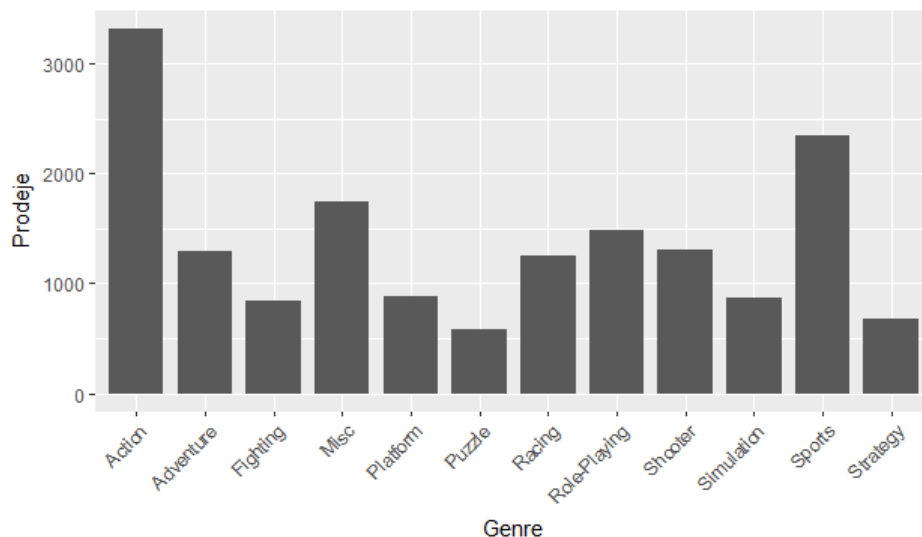


Zde lze potvrdit tvrzení uvedené u grafu 2.1. Jasně dominují právě ty nejlépe prodávané platformy, které mohou za nárůsty umístěných titulů v době kdy byly uvedeny a prodávány.

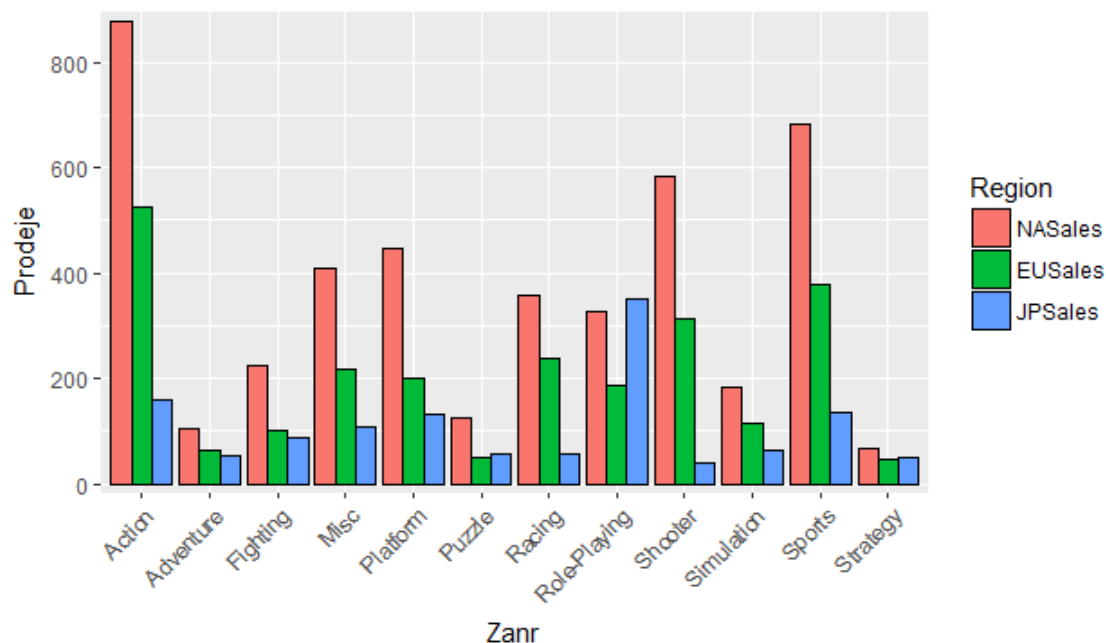


Také se projevuje preference jednotlivých regionů - Sony ze svou řadou Playstation jasně ovládá domovský japonský trh společně s Nintendem, naopak značka Xbox od Microsoftu je v Japonsku naprostým propadákem. Na druhé straně v USA - domácí Xbox převládá nad ostatními konzolami. Evropa je regionem "neutrálním".

2.4 Oblíbené žánry



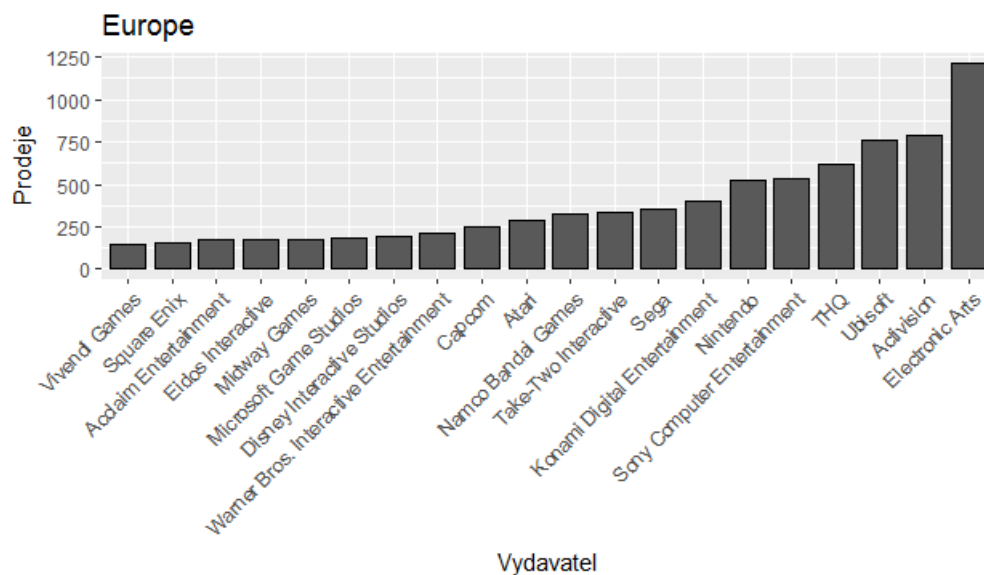
Nejoblíbenějším žánrem jsou akční hry. Na druhém místě doprovázeny sportovními hrami, které se uvádí na trh pravidelně každý rok a proto představují velkou část datasetu. Nejméně se vyskytují logické a strategické hry.



Při rozdělení prodejů jednotlivých žánru do regionu se opět projevují již "zaseté" předpoklady. Japonskému trhu vládou RPG hry, jejíž převaha je na grafu velmi výrazná. Evropa a Severní amerika naopak preferuje hry akční a sportovní.

2.5 Vydavatelé a tituly podle regionu

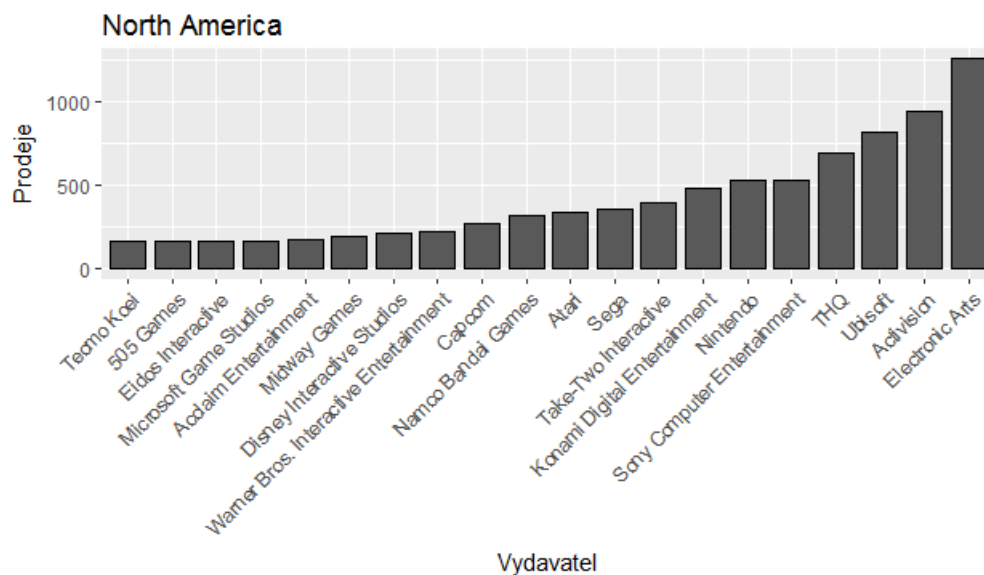
2.5.1 Evropa



Název titulu	Prodeje v regionu (mil.)
Wii Sports	29.02
Grand Theft Auto V	23.04
Mario Kart Wii	12.88
FIFA 15	12.40
Call of Duty: Modern Warfare 3	11.29

Evropu ovládá Electronic Arts díky sportovním seriím jako je FIFA, NHL, které jsou na území EU velmi oblíbené a patří opakovaně k nejprodávanějším titulům. První místo v prodeji videoher patří Nintendu s Wii Sports, vzhledem k tomu, že byl po dlouhou dobu dodáván s každou konzolí v základním balení. Historický úspěch zaznamenal taky Rockstar Games se svým Grand Theft Auto V.

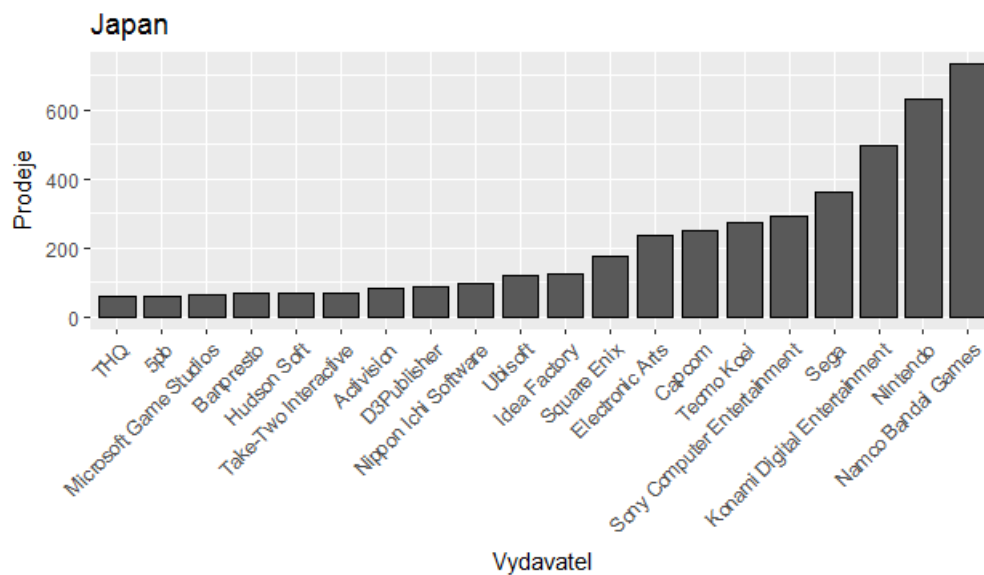
2.5.2 Severní Amerika



Název titulu	Prodeje v regionu (mil.)
Nintendo Wii Sports	41.49
Super Mario Bros.	32.48
Duck Hunt	26.93
Tetris	26.17
Grand Theft Auto V	23.46

Stejně jako v EU, nejlépe se zde daří domácímu Electronic Arts a Activision díky herním značkám jako je akční Battlefield, Call of Duty či sportovní tituly série NFL. Historicky nejprodávanější tituly však patří společnosti Nintendo, vzhledem k expanzi do spojených států a dominanci jejich systému NES, SNES a Game Boy v průběhu 80 a 90 tých let na uzemí USA.

2.5.3 Japonsko

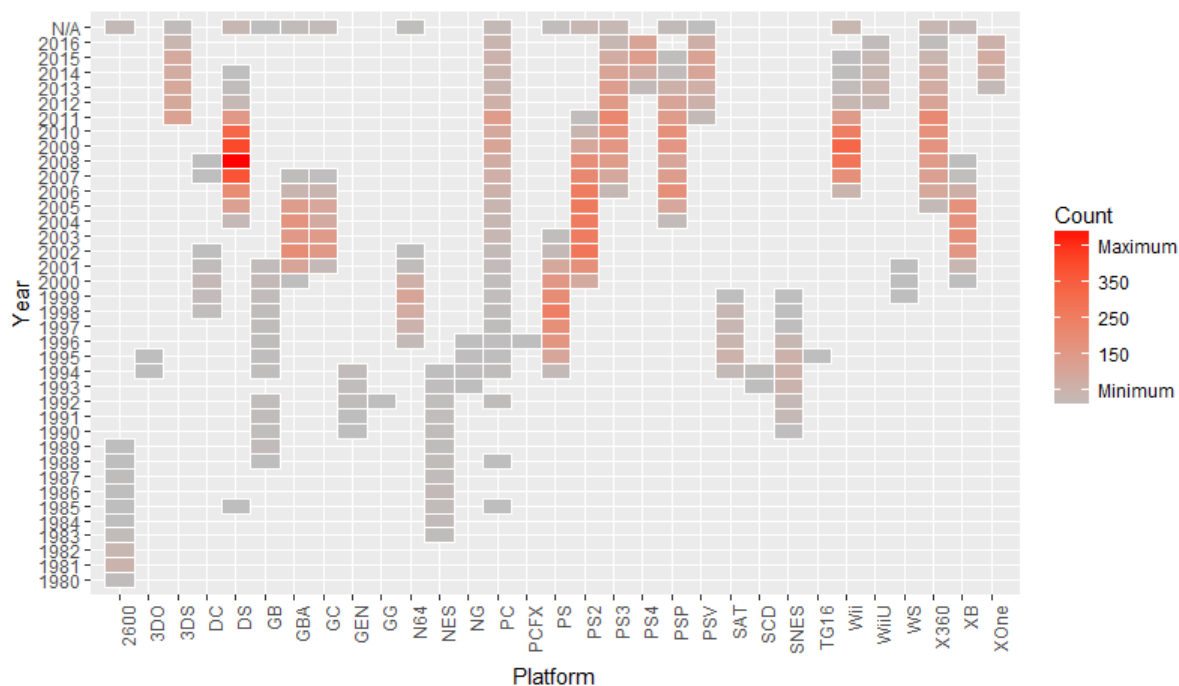


Název titulu	Prodeje v regionu (mil.)
Pokemon Red/Pokemon Blue	10.22
Pokemon Gold/Pokemon Silver	7.20
Super Mario Bros.	6.96
New Super Mario Bros.	6.50
Pokemon Diamond/Pokemon Pearl	6.04

Klasicky v Japonském regionu je situace úplně jiná. Je zde jasně vidět preferování lokálních značek jako je Namco, Nintendo, Konami, Square. Stejně tak je rozdílná situace u nejprodávanejích titulech které kompletně patří domácímu Nintendo a značkám jako Pokémon a Super Mario Brothers.

3 Specifická analýza

3.1 Životnost platform



Na tomto grafu můžeme nejen sledovat životnost jednotlivých konzolí, ale také úspěch jednotlivých generací. Zde je opět naprosto zjevné že předchozí generace byla ta nejúspěšnější. Zjistili jsme také, že:

- Největší počet nejúspěšnějších titulů vzniká vždy uprostřed životního cyklu platformy.
- V datasetu jde skvěle vidět, jak se generace konzolí prolínají (Atari 2600 -> Nintendo NES) a známé značky se vyvíjí (Sony Playstation -> Sony Playstation 4).
- Životnost herní platformy (Nepočítaje osobní počítače.) je průměrně 10 let.