



TECHNICKÁ UNIVERZITA V LIBERCI  
Fakulta mechatroniky, informatiky  
a mezioborových studií ■

# Případová studie - Analýza celosvětových prodejů počítačových her

## Zápočtová práce

*Studijní program:* N2610 – Elektrotechnika a informatika

*Studijní obor:* 1802T007 – Informační technologie

*Autor práce:* **Bc. Tomáš Moravec**

*Akademický rok:* 2016/2017



# 1 Zadání

Pomocí více než 16 500 záznamů, z největší celosvětové databáze prodeje počítačových her VGChartz se pokusím analyzovat, jaký je klíč k vydání úspěšné hry pro jednotlivé kontinenty, případně jak se trendy měnily v průběhu posledních let.

## 2 Zdrojová data

### 2.1 Získání dat

Data ke zpracování vybrané úlohy, tedy analýza celosvětových prodejů počítačových her, jsem se rozhodl hledat na největších data-miningových portálech, zejména jsem procházel UCI, Kaggle a KDNNuggets. Konečná zdrojová data jsem si vybral z portálu Kaggle (jedná se o projekt společnosti Google, nyní již Alphabet Inc.), přesněji na adrese <https://www.kaggle.com/gregorut/videogamesales>, kde jsem našel největší množství dat ke zpracování. Zdrojem dat je portál VGChartz a získána byla pomocí scrape bota, dostupného na adrese <https://github.com/GregorUT/vgchartzScrape>.

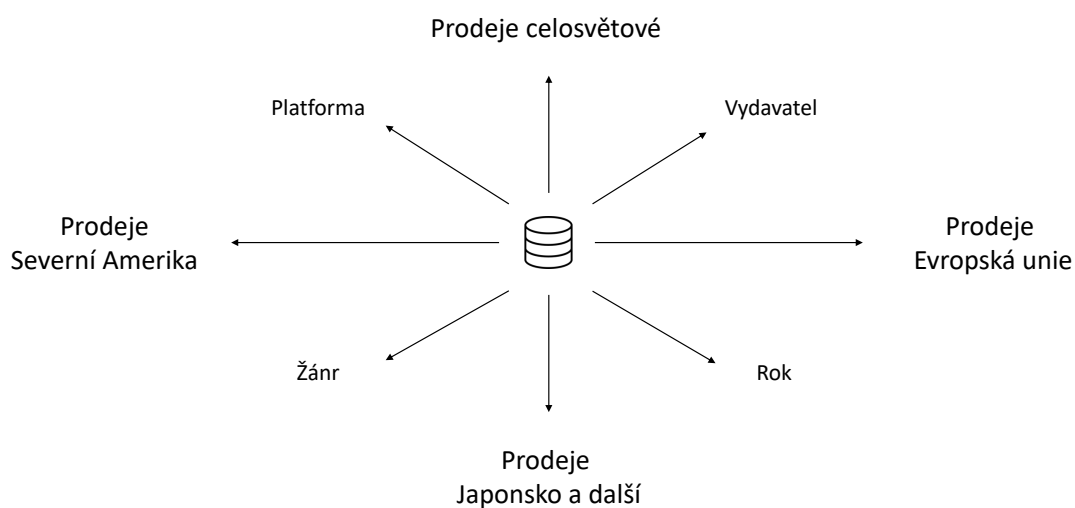
### 2.2 Struktura dat

Data obsahují jak globální prodeje, tak prodeje na jednotlivých trzích. Do tabulky se dostali pouze hry, které měly více než 100 000 prodaných kopií.

**Přesněji jsou data strukturovaná takto:**

- Rank - Pořadí v tabulce
- Name - Celý název hry
- Platform - Platforma (zařízení), pro které hra vyšla (PC,PS4, atd.)
- Year - Rok vydání hry

- Genre - Herní žánr
- Publisher - Vydavatelství hry
- NA\_Sales - Prodeje v Severní Americe (v milionech)
- EU\_Sales - Prodeje v Evropě (v milionech)
- JP\_Sales - Prodeje v Japonsku (v milionech)
- Other\_Sales - Prodeje ve zbytku světa (v milionech)
- Global\_Sales - Prodeje celosvětově (v milionech), jedná se o součet ostatních



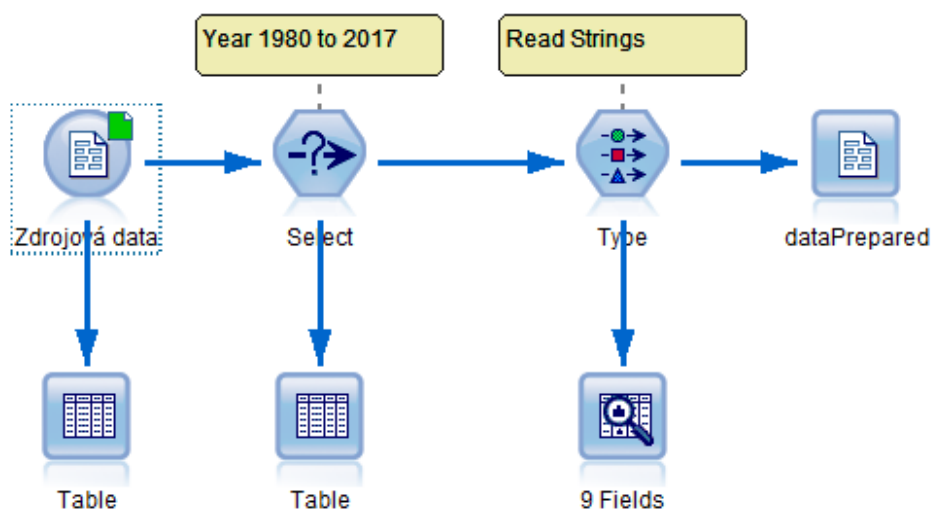
Obrázek 2.1: Grafické znázornění dat

## 3 Příprava dat

V tomto příkladu budeme pracovat s dodanou datovou maticí. Je však třeba ověřit kvalitu dat a analyzovat rozdělení jednotlivých proměnných.

### 3.1 Zlepšení kvality dat

Po bližším prozkoumání dat jsem zjistil, že valná většina dat je velice nekvalitních. Pro prozkoumání jednotlivých nekvalitních záznamů a několika méně úspěšných pokusech o zlepšení kvality jsem přišel na to, že pokud omezím rok vydání, na období, ve kterém jsou vydávány všechny moderní hry, tedy od roku 1980 (Wikipedia) do dnes, tedy do roku 2017, data získají 100% kvalitu. Po prozkoumání již kvalitních dat jsem došel k pouhým 2 121 záznamům. Postup je znázorněn na obrázku níže.

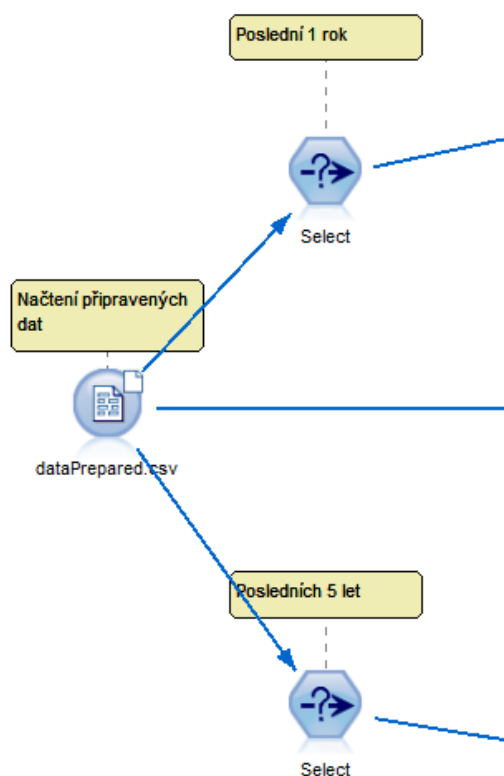


Obrázek 3.1: Postup výběru kvalitních dat

## 4 Vyhodnocení dat

### 4.1 Rozdělení na období prodejů

Prvním úkolem bylo, zvolit si správné období tak, aby bylo možné vidět nejenom aktuální dění na herním trhu, ale také proměnu trhu v posledních letech. Zvolil jsem si období prodejů za 1 poslední rok (2016 a začátek 2017) a dále za 5 posledních let (2011 až začátek 2017). Data obsahují i několik her z ledna 2017, která jsem započítal do roku 2016.



Obrázek 4.1: Rozdělení dat na zvolená období

## 4.2 Zpracování jednotlivých období

Jednotlivá období jsem zpracovával identicky, zaměřím se tedy pouze na poslední 1 rok. Rozhodl jsem se zjistit, jak zisky ovlivňují platformy, vydavatelství a žánr hry.

### 4.2.1 Celosvětové prodeje v závislosti na platformě

Nejdříve jsem se zaměřil na získání informací o globálních zisků z jednotlivých platform, tedy zařízení, pro které probíhal vývoj hry a na které si jí lze zahrát.

	Platform	Global_Sales_Sum
1	PS4	15.110
2	XOne	2.410
3	PS3	0.510
4	3DS	0.360
5	PSV	0.140
6	PC	0.010

Obrázek 4.2: Tabulka celosvětových prodejů pro jednotlivé platformy za poslední rok

Z tabulky je patrné, že dnes nejpopulárnější platformy jsou ty nejvýdělečnější.

### 4.2.2 Celosvětové prodeje v závislosti na vydavatelství

Dále jsem se tedy zaměřil na samotná vydavatelství, která za vývojářské studio vydávají hru, stejně jako knižní vydavatelství vydává hru za spisovatele. Zajímalo mě tedy, zda jsou zisky ovlivněny výběrem vydavatelství.

	Publisher	Global_Sales_Sum
1	Ubisoft	7.810
2	Electronic Arts	4.770
3	Sony Computer Entertainment	4.360
4	Namco Bandai Games	0.430
5	Warner Bros. Interactive Entertainment	0.320
6	Konami Digital Entertainment	0.170
7	Sega	0.160
8	Aksys Games	0.120
9	Nippon Columbia	0.060
10	Square Enix	0.060

Obrázek 4.3: Tabulka celosvětových prodejů pro deset největších vydavatelství za poslední rok

Není překvapivé, že největší vydavatelství, které mají největší distribuční síť, dokáží hře vygenerovat značně vyšší zisky. Jedná se tedy o ekvivalent ke knižnímu vydavatelství.

### 4.2.3 Celosvětové prodeje v závislosti na žánru hry

Poslední zkoumanou kategorií je žánr hry, tedy zda se jedná o strategickou, sportovní, nebo jinou hru.

	Genre	Global_Sales_Sum
1	Shooter	9.880
2	Sports	4.940
3	Action	2.940
4	Role-Playing	0.290
5	Adventure	0.280
6	Misc	0.120
7	Racing	0.080
8	Strategy	0.010

Obrázek 4.4: Tabulka celosvětových prodejů pro jednotlivé žánry her za poslední rok

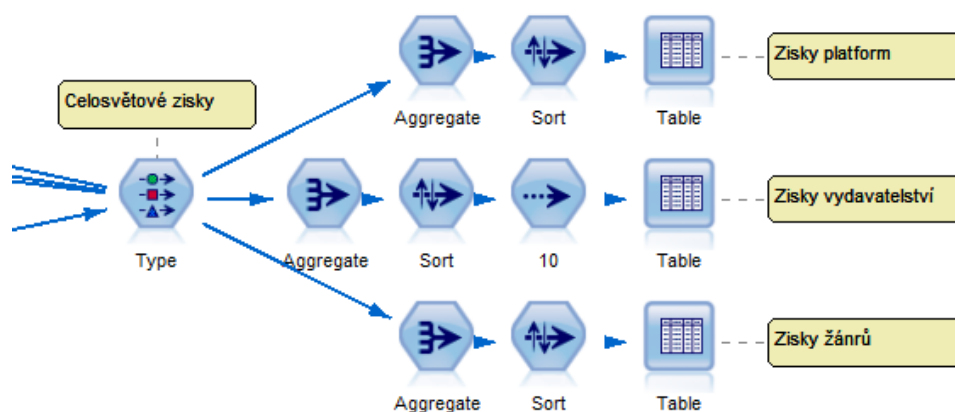
Ve výsledku jsem byl velice překvapen, že sportovní hry mají takovéto prodeje. Hlubším průzkumem jsem zjistil, že je to díky prodejem rodinných her, například



pro Kinect, které ovládáte pohybem vlastního těla. Tyto hry jsou v posledních letech velice populární.

#### 4.2.4 Postup získávání jednotlivých výsledků

Na obrázku níže je ukázáno, jak probíhalo získávání zobrazených dat.



Obrázek 4.5: Získávání výsledků pro jednotlivé parametry

Na konci jsem ještě dělal průzkum vždy pro 3 největší vydavatelství (Sony, EA, Ubisoft), pro které platformy a žánry tvoří hry. Zda se drží nejvýdělečnějších, nebo se vymykají. Zjištění bylo takové, že největší vydavatelství striktně vydávají hry pouze na nejvýdělečnější platformy, stejně tak hry nejvýdělečnějších žánrů.

### 4.3 Změny trhu

Vhledem k tomu, že prozkoumávána byla data jak za poslední rok, tak za posledních 5 let. Došel jsem k následujícím výsledkům.

#### 4.3.1 Zisky z platform, vydavatelství a žánrů

U platforem se v průběhu let nezměnilo v podstatě nic. Jednou za několik let vždy vychází modernější verze pro každou platformu, je tedy jasné že zisky za posledních

5 let jsou primárně ze starších verzí těchto platforem. Nicméně jiné rozdíly nejsou.

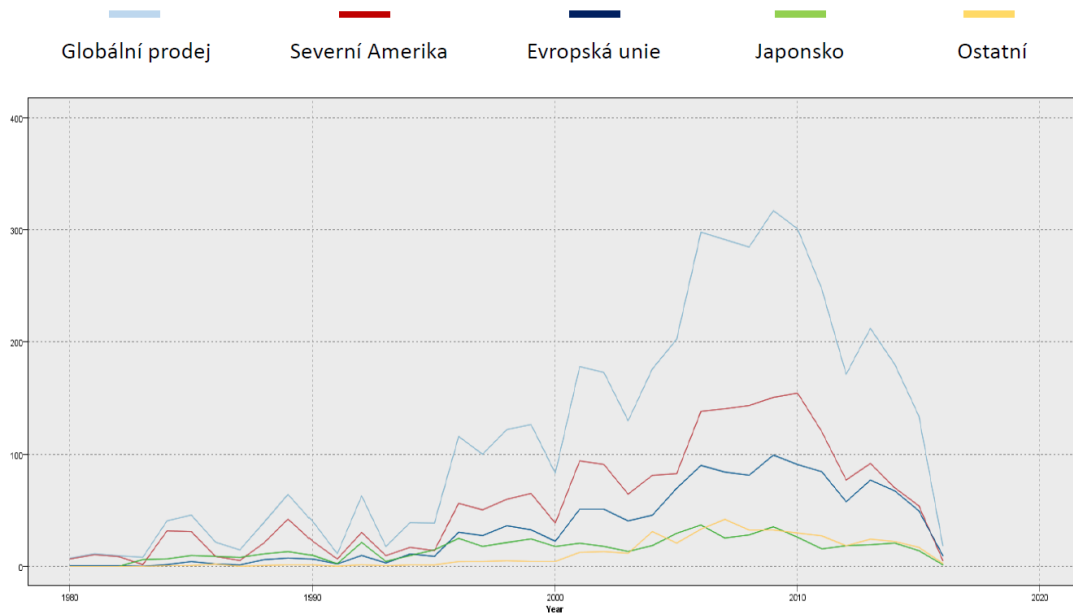
	Platform	Global_Sales_Sum
1	PS3	227.730
2	X360	213.210
3	PS4	179.990
4	3DS	139.660
5	XOne	71.290
6	Wii	49.770

Obrázek 4.6: Získávání výsledků pro jednotlivé parametry za posledních 5 let

U ostatních parametrů (vydavatelství a žánr) se jedná o stejné výsledky. Tedy trendy se za poslední roky příliš nezměnily. Stejně tak největší vydavatelství vydávají pořád na největší platformy a vydávají hry primárně s nejprodávanějšími žánry.

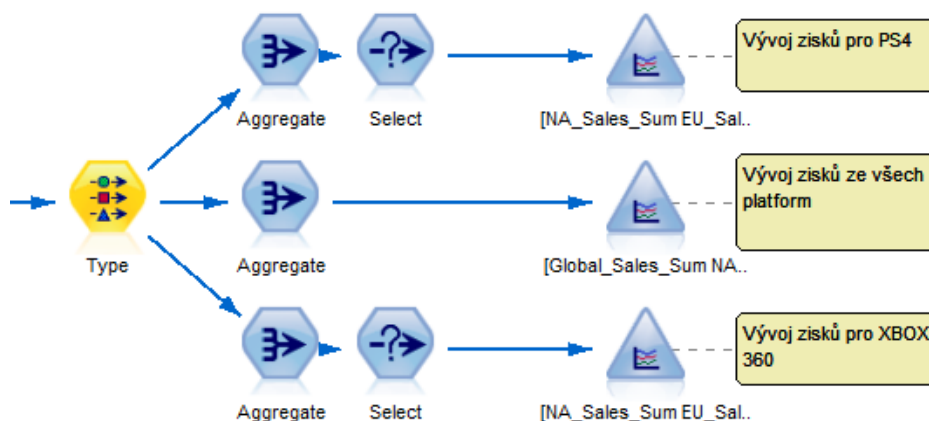
## 4.4 Vývoj zisků za všechna období

Pro zajímavost jsem ke konci své práce prozkoumával historické vývoje zisků skrze všechna zaznamenaná období. Má očekávání byla, že prodeje budou každoročně růst, nebo stagnovat, nicméně, k mému překvapení, získal jsem naprostý opak.



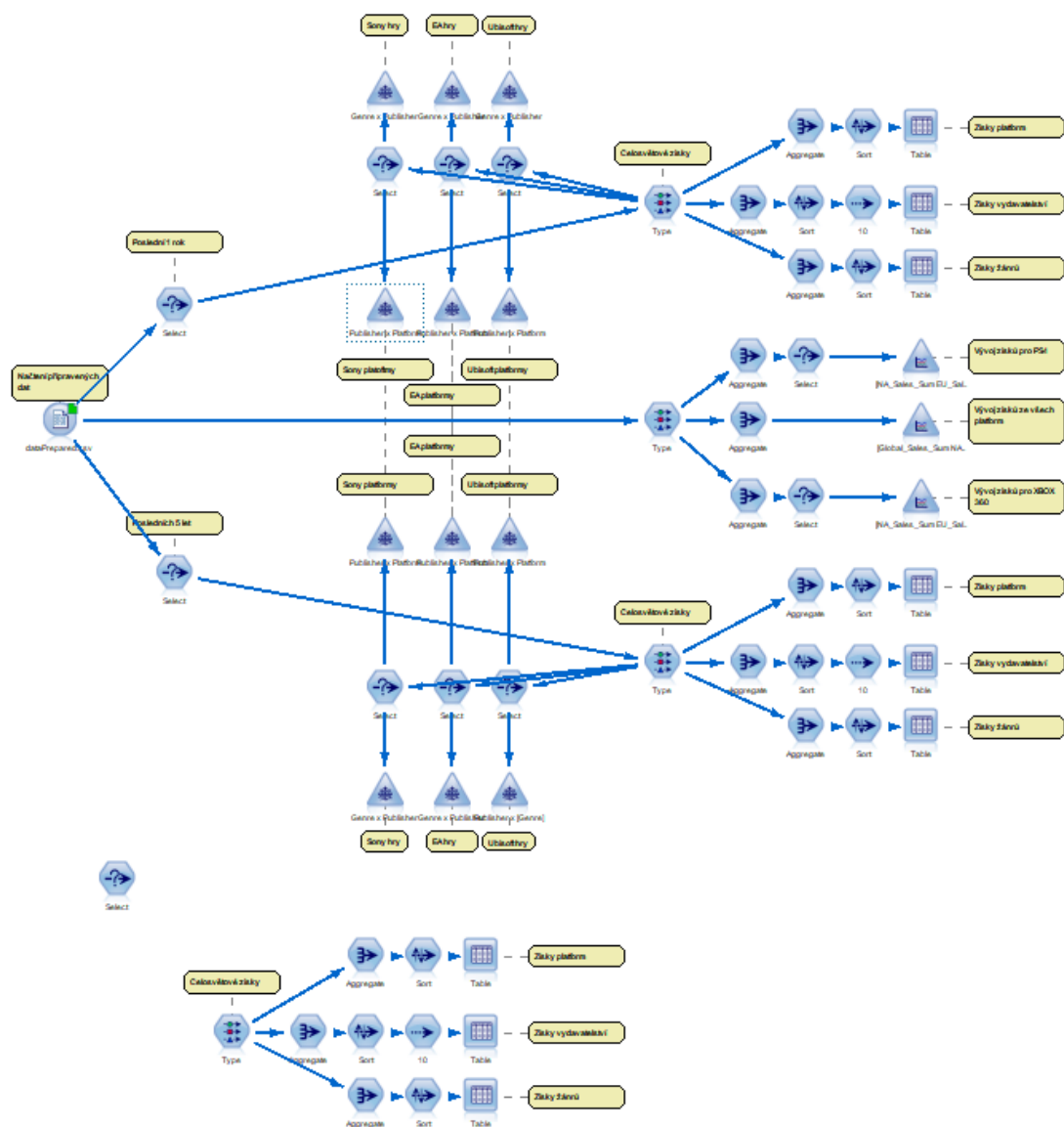
Obrázek 4.7: Vývoj zisků za všechna období

Ze získaných dat jsem byl zmatený, udělal jsem tedy analýzu závislosti prodeje na dvou nejvýdělečnějších platformách a na všech platformách, výsledky byli ale totožné. Níže je možné vidět postup při získávání grafů v závislosti na parametrech.



Obrázek 4.8: Zjištění vývoje zisků v závislosti na jednotlivých parametrech

Prohledal jsem, zda nemám chybu v postupu, kterou jsem neodhalil a vzhledem k totožným výsledkům jsem tedy závěrem považoval výsledky za správné, i když zvláštní. Vysvětlení přišlo až po prezentaci mé práce, viz závěr (nové skutečnosti), od jednoho ze studentů. Pro zajímavost přidávám vzhled kompletního proudu.



Obrázek 4.9: Kompletní datový proud v modeleru

## 5 Závěr

### 5.1 Shrnutí práce

Výsledkem práce je, že pokud chcete maximalizovat možné zisky ze hry, je nutné využít jednoho ze tří největších distributorů her (Sony, EA, Ubisoft), dále je nutné se zaměřit na nejpoblárnější platformy (Playstation, Xbox) a zvolit poslední nejoblíbenější žánr (Akční, Sportovní). Myslím, že tyto skutečnosti byli pro člověka se zájmem o herní průmysl jasné již od začátku. Nicméně jsem byl velice překvapen z celosvětových prodejů, které jasné ukazovaly extrémní pád prodejů her. Důvod k takovému závěru je vysvětlen v následujícím odstavci.

### 5.2 Nové skutečnosti

Po prezentaci mé práce na cvičení, mi bylo kolegou Karlem Šírem sděleno, že se o portál, ze kterého data čepají, zajímá a že prodeje jsou pouze z kamenných obchodů a chybí tedy největší distribuční síť současnosti, jako je Steam, Origina a další. Bylo to pro mě na jednu stranu nečekané zjištění, na druhou stranu tato skutečnost objasnila, proč data ukazují extrémní úpadek prodeje her v posledních letech, trh se jednoduše adaptuje na prodeje online a kamenné obchody ztrácí na oblibě. Navazující práce na toto téma by mohla zahrnovat data ze všech prodejních sítí, díky čemuž bychom získali přesnější obraz o aktuálním dění na herním trhu.