

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Dominik Posavec**

**Analiza statistike prodaje videoigara**

**SEMINARSKI RAD**

**Varaždin, 2021.**

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Dominik Posavec**

**Matični broj: 45046/16–R**

**Studij: Organizacija poslovnih sustava**

**Analiza statistike prodaje videoigara**

**SEMINARSKI RAD**

**Mentor/Mentorica:**

Prof. dr. sc. Kornelije Rabuzin

**Varaždin, lipanj 2021.**

## **Sažetak**

Svrha ovog seminarskog rada je prikazivanje rada skladišta podataka, njihova svrha, namjena i uloga, te prikaz same analize podataka koja se odvija na temelju statistike prodaje videoigara. Kako bi što bolje precizirali svrhu ovog rada, potrebno je dakle izraditi skladište podataka, provesti određene ETL procese i prikazati postupak prijenosa tih podataka, odnosno prikaz u obliku izvještaja. Kako bi se rad mogao što kvalitetnije i efikasnije izraditi, potrebno je koristiti sljedeće alate: MySQL Workbench i PowerBI Desktop verziju od strane Microsoft-a.

**Ključne riječi:** skladišta podataka, ETL, alati, statistika prodaje, videoigre, PowerBI, MySQL

# Sadržaj

Sadržaj .....	iv
1. Uvod .....	1
2. Osnove skladišta podataka.....	2
2.1. Karakteristike kreiranog skladišta podataka.....	2
2.2. Tehnologije korištene za rad .....	3
3. Izrađivanje skladišta podataka.....	4
3.1. Odabrani skup podataka .....	4
3.2. Korištenje ETL-a nad podacima.....	5
3.3. Model zvijezde .....	7
3.3.1. Dimenzijska tablica rankname .....	8
3.3.2. Dimenzijska tablica platform .....	8
3.3.3. Dimenzijska tablica publisher .....	9
3.3.4. Dimenzijska tablica genre .....	10
3.3.5. Dimenzijska tablica year .....	10
3.3.6. Dimenzijska tablica sales .....	11
3.4. UPDATE nad tablicom vgsales.....	11
3.5. ALTER tablice vgsales.....	12
3.6. Činjenična tablica fact_table .....	12
4. Analiza uz pomoć Power BI alata .....	13
4.1. Generirani izvještaji.....	16
5. Zaključak .....	19
Popis literature.....	20
Popis slika .....	21

# 1. Uvod

U današnje novije vrijeme tehnologije primorani smo upotrebljavati sve više i više kvalitetnija skladišta podataka, naravno uz uvjet da su ona zaštićena od neovlaštenog upada od strane zlonamjernih osoba. Isto tako s razvojem tehnologije a samim time i rastom različitih skupova podataka dolazi do potrebe da se ti podaci spremaju na određenim mjestima kako bi se mogli što učinkovitije kasnije pripremati za analizu istih. Kako bi neko poduzeće razvilo svoje sposobnosti upravljanja podacima na što kvalitetnijim metodama, potrebno je da se sama skladišta podataka informatiziraju i da budu što funkcionalnija kod njihovog rada i pristupa. Isto tako, živimo u dobu gdje se globalne promjene događaju svakodnevno i vrlo brzo, te nam je stoga izuzetno važno i pravovremeno reagirati na svaku novonastalu situaciju neovisno da je ona izazvana unutarnjim ili vanjskim čimbenicima. Osobe koje su odgovorne moraju svakodnevno donositi vrlo važne odluke koje za neke mogu rezultirati vrlo dobro, dok neke odluke mogu čak biti i katastrofalne za daljnje poslovanje. Kako bi se odluke mogle ispravno donositi potrebno je imati pristup što relevantnijim podacima, to jest, da nam informacije budu pregledne, lako razumljive i da se temelje na analizi provjerenih i integriranih podataka.

Ovaj se seminarski rad upravo bazira na izradi skladišta podataka gdje ćemo uz pomoć istog moći prikazati kvalitetno spremljene podatke u obliku izvještaja koji su netom prije kreiranja izvještaja ispravno analizirani unutar kreiranog skladišta podataka. Pomoć kod izrade praktičnog dijela seminara uveliko se odražava i na korištene tehnologije kao što su MySQL Workbench i PowerBi o kojima ćemo nešto više reći u narednim poglavljima. Spomenimo da se u radu nalazi i postupak prikupljanja podataka koji su korišteni u svrhu ovog rada, strukturiranje istih, obavljanje ETL procesa nad podacima i slično. Također, u ovome radu prikazat ćemo izradu i opis dimenzijskih i činjeničnih tablica, te prikaz istih u obliku izvještaja.

## 2. Osnove skladišta podataka

Prema nekim definicijama skladišta podataka možemo opisati kao temelj za dobivanje izvještaja koji su ključni za donošenje kvalitetnih poslovnih odluka. Isto tako, skladišta podataka objedinjuju podatke koji su prikupljeni tokom dugi niz godina poslovanja poduzeća i koja su dospjela iz različitih izvora. Također ona nam predstavljaju i novu generaciju računarskog sustava za dodatne potpore u odlučivanju. Obuhvaćaju podatke koje su subjektivnog područja i koji su potrebni za odlučivanje u nekom određenom području s nekim ciljem, svrhom i namjenom. Samim time, skladišta sadrže i logički integrirane podatke iz raznih aplikacija dok su ti isti podaci prema nekim pravilima nepromjenjivi u smislu obrade. Kako bi se omogućilo predviđanje i uspoređivanje podataka, skladišta su također vremenski usmjerena jer sadrže podatke koji opisuju pojave karakterizirane za dulje vremensko razdoblje. [1]

Dakle, skladišta podataka predstavljaju subjektivno orijentirane, integrirane, nepromjenjive i vremenski dinamične skupove podataka koji se koriste kao potpora kod odlučivanja u nekim bitnim fazama. Ono je i proces gdje podaci prikupljaju i upravljaju iz mnogih drugih izvora koji su agregirani, ekstrahirani, detaljni i obogaćeni ciljem potpore analize i donošenju procesu odluka. Po sadržaju i po nekim tehničkim zahtjevima, skladište podataka je i drugačije u odnosu na transakcijski sustav iako se transakcijski sustav ponaša kao njegova pretpostavka, no međutim, skladište podataka možemo shvatiti i kao sučelje koje dijeli operativne obrade od aplikacija za potporu u odlučivanju. [1]

### 2.1. Karakteristike kreiranog skladišta podataka

Izrađivanje skladišta podataka u ovome radu koriste se podaci koji su preuzeti sa stranice kaggle.com i odnose se na analiziranje statistike prodaje videoigara. Kaggle.com je web stranica koja sadrži velik i širok spektar skupova podataka iz različitih područja. Uz mogućnost korištenja skupova podataka, navedena web stranica nudi i ostale mogućnosti kao što su učenja kroz različitih lekcija, kako bi osoba mogla proširiti svoje znanje u ostalim područjima IT-a. Uz pomoć preuzetih podataka, moći ćemo dobiti uvid u karakteristike videoigara, primjerice zastupljenost poduzeća koje najviše proizvodi i prodaje videoigre, zatim kojeg su žanra, odnosno namjena, platforma za koju je videoigra namijenjena i slično.

Dataset koji je preuzeti sadrži sljedeća obilježja i podatke:

- Rank – broj pod kojim se nalazi određena videoigra
- Name – naziv videoigre
- Platform – platforma za koju je namijenjena

- Year – godina izdanja videoigre
- Genre – vrsta odnosno žanr videoigre
- Publisher – naziv izdavača koji je plasirao videoigru na tržište
- NA\_Sales – sveukupni broj NA prodaje
- EU\_Sales – sveukupno broj prodaju u EU
- JP\_Sales – sveukupni broj JP prodaje
- Other\_Sales – sveukupni broj ostale prodaje
- Global\_Sales – globalni sveukupni broj prodaje

## 2.2. Tehnologije korištene za rad

Za ovaj projektni zadatak korištene su sljedeće tehnologije: MySQL Workbench, PowerBI i MySQL. Izrada skladišta podataka obavlja se u alatu MySQL Workbench koji nam ujedno predstavlja i grafički prikaz dizajniranih baza podataka koji također integrira razvoj SQL-a, administraciju, dizajn baze podataka, stvaranje i održavanje u jedinstveno razvojno okruženje. Alat je besplatan i dostupan svim uzrastima te je namijenjen radu na bilo kojem operacijskom sustavu. Također, omogućava smanjenje rizika, troškova i složenosti u razvoju, uvođenju i upravljanju kritičnim poslovnim MySQL aplikacijama na kojima rade administratori baze podataka. Uz prethodna obilježja također omogućuje i modeliranje, generiranje, upravljanje bazom podataka, pogodan je za sinkronizaciju i usporedbu shema. Administratorima omogućuje da izvrše reverzni inženjering postojećih baza podataka ili aplikacija kako bi arhitektima koji rade na dizajniraju baza podataka pružili bolji uvid u dizajn same baze podataka. Pomoću migracija baze podataka osigurava i korištenje čarobnjaka za vizualnu migraciju podataka kao što je migracija baza podataka koja uključuju migracije s Microsoft SQL poslužiteljem, MS Access i slično. Pruža mogućnosti korisnicima da definiraju i ciljaju na određeni izvor podataka netom prije samo procesa migracije te ostale brojne prednosti. [2]

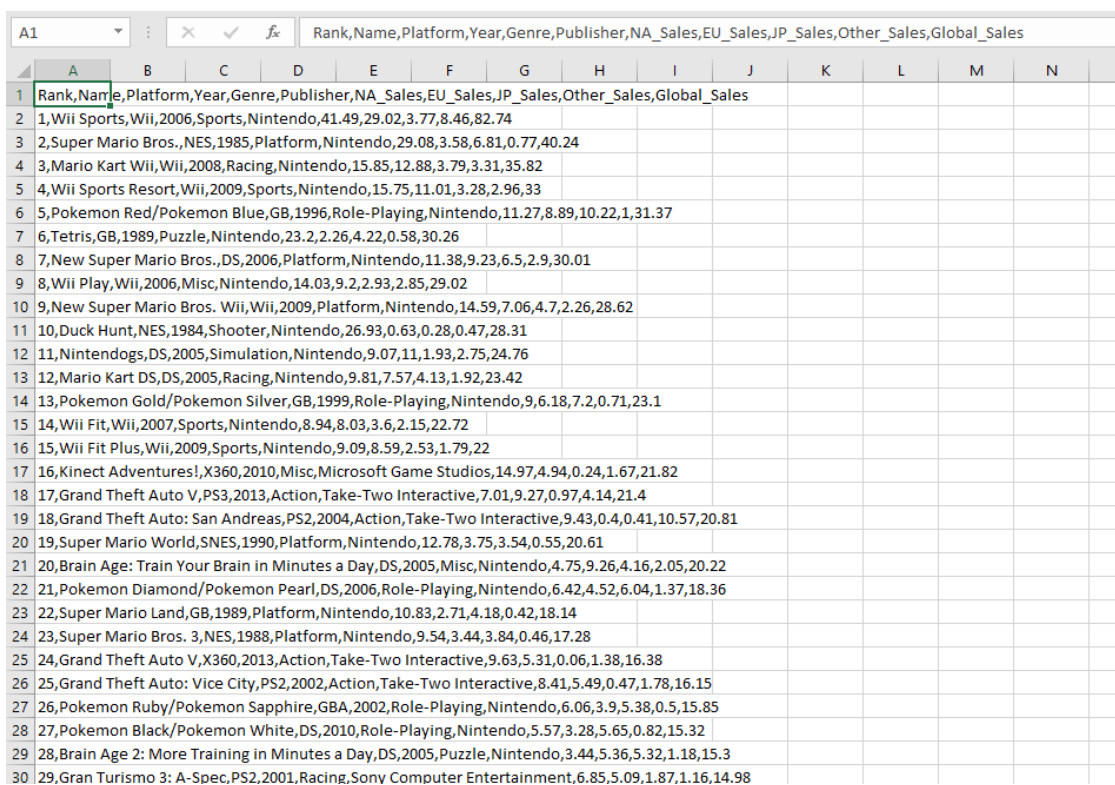
Alat koji služi za prikazivanje izvještaja je PowerBI i on se definira kao zbirka softverskih usluga i aplikacija koji zajedno rade i kako bi naše nepovezane izvore podataka pretvorili u koherentne, interaktivne i vizualno predočene. Alat se može koristiti na računalu, pod nazivom Power BI Desktop, zatim kao internetska usluga, Power BI usluga i kao verzija na mobilnim operacijskim sustavima koji je namijenjen za Android, Windows ili iOS uređaje. Potrebno je spomenuti kako alat sadrži i dva bitna elementa a to su Power BI Report Builder koji služi za stvaranje izvješća i PowerBI Report Server koji se definira kao lokalni poslužitelj izvješća na kojem možemo objavljivati vlastita izvješća. [3]

### 3. Izrađivanje skladišta podataka

U ovome poglavlju naglasak je na pronalaženju i prikazivanju izvora podataka koji će se dalje koristiti za rad s podacima i u konačnici prikazivanje izvještaja istih kao i djelovanje ETL (eng. Extract, Transform, Load) nad njime.

#### 3.1. Odabrani skup podataka

Odabrani skup podataka koji je pronađen i preuzet odnosi se na prodaju videoigara i navedemo skupu podataka (eng. Dataset) možemo pristupiti na kaggle.com web adresi kao što smo to prethodno naveli u radu. Prilikom preuzimanja datasea, podaci su pohranjeni u format naziva .csv (eng. Commaseparated values). Broj redaka koji se nalazi u datasetu je točno 16599 i svaki je redak različit u odnosu na ostale retke.



A1	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
2	2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
3	3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33
5	5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1	31.37
6	6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26
7	7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.5	2.9	30.01
8	8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.2	2.93	2.85	29.02
9	9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.7	2.26	28.62
10	10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
11	11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11	1.93	2.75	24.76
12	12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92	23.42
13	13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9	6.18	7.2	0.71	23.1
14	14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.6	2.15	22.72
15	15	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.09	8.59	2.53	1.79	22
16	16	Kinect Adventures!	X360	2010	Misc	Microsoft Game Studios	14.97	4.94	0.24	1.67	21.82
17	17	Grand Theft Auto V	PS3	2013	Action	Take-Two Interactive	7.01	9.27	0.97	4.14	21.4
18	18	Grand Theft Auto: San Andreas	PS2	2004	Action	Take-Two Interactive	9.43	0.4	0.41	10.57	20.81
19	19	Super Mario World	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61
20	20	Brain Age: Train Your Brain in Minutes a Day	DS	2005	Misc	Nintendo	4.75	9.26	4.16	2.05	20.22
21	21	Pokemon Diamond/Pokemon Pearl	DS	2006	Role-Playing	Nintendo	6.42	4.52	6.04	1.37	18.36
22	22	Super Mario Land	GB	1989	Platform	Nintendo	10.83	2.71	4.18	0.42	18.14
23	23	Super Mario Bros. 3	NES	1988	Platform	Nintendo	9.54	3.44	3.84	0.46	17.28
24	24	Grand Theft Auto V	X360	2013	Action	Take-Two Interactive	9.63	5.31	0.06	1.38	16.38
25	25	Grand Theft Auto: Vice City	PS2	2002	Action	Take-Two Interactive	8.41	5.49	0.47	1.78	16.15
26	26	Pokemon Ruby/Pokemon Sapphire	GBA	2002	Role-Playing	Nintendo	6.06	3.9	5.38	0.5	15.85
27	27	Pokemon Black/Pokemon White	DS	2010	Role-Playing	Nintendo	5.57	3.28	5.65	0.82	15.32
28	28	Brain Age 2: More Training in Minutes a Day	DS	2005	Puzzle	Nintendo	3.44	5.36	5.32	1.18	15.3
29	29	Gran Turismo 3: A-Spec	PS2	2001	Racing	Sony Computer Entertainment	6.85	5.09	1.87	1.16	14.98

Slika 1. Prikaz dijela preuzetog skupa podataka, [autorski rad]

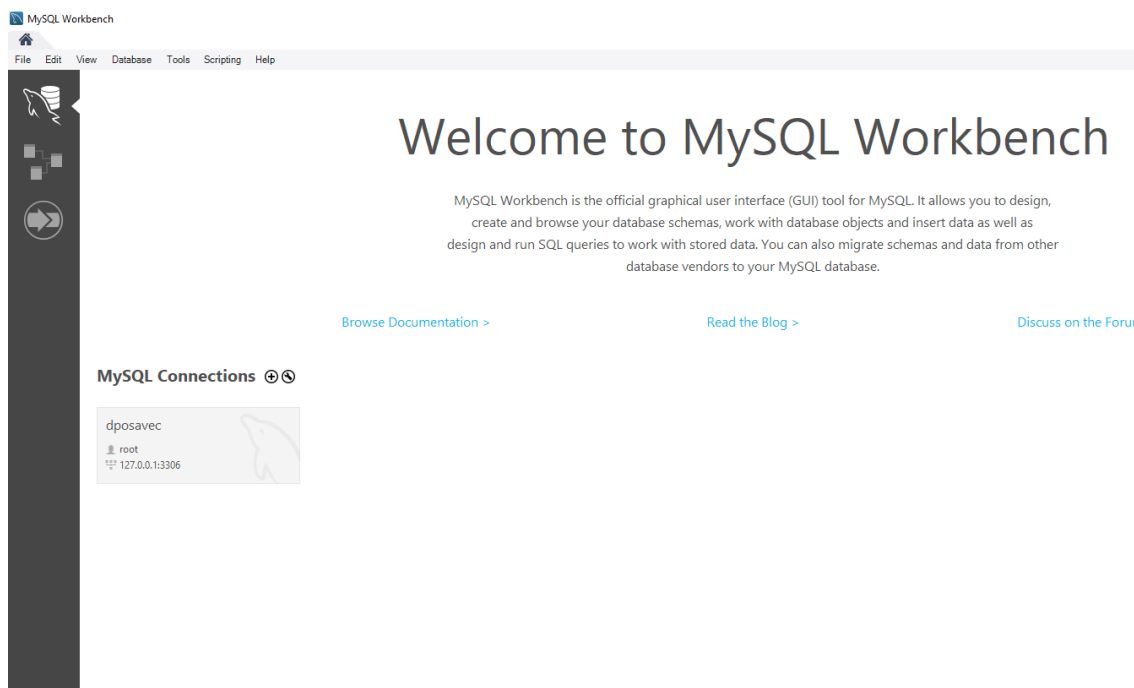
Na prethodnoj fotografiji moguće je vidjeti dio skupa podataka koji je preuzet i koji će se koristiti u svrhu ovog rada. Za prikaz prethodne fotografije koristio se alat MS Excel. Svaki redak, odnosno svaki zapis sadrži sljedeće attribute redom navedene: Rank, Name, Platform, Year, Genre, Publisher, NA\_Sales, EU\_Sales, JP\_Sales, Other\_Sales i Global\_Sales.



## 3.2. Korištenje ETL-a nad podacima

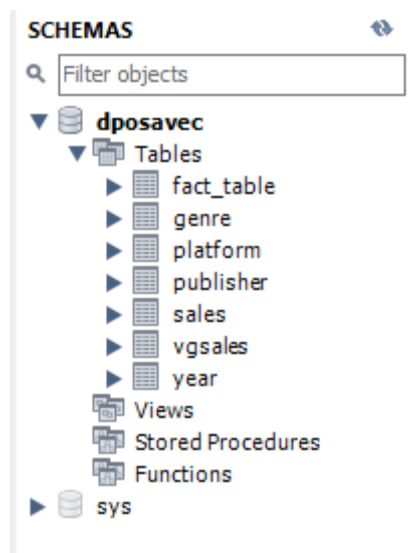
ETL možemo definirati kao postupak koji ekstrahira podatke iz različitih izvornih sustava, te iste podatke transformira i učitava ih u skladište podataka (eng. Data Warehouse). Da bi se zadržala vrijednost koje nam alat nudi za donošenje odluka, sustav skladišta podataka mora se mijenjati s poslovnim promjenama. ETL je ponavljajuća aktivnost koji se može raditi na dnevnoj, tjednoj ili mjesečnoj razini i mora biti vrlo dobro dokumentirana i automatizirana. ETL pomaže tvrtkama u analiziranju svojih poslovnih podataka za donošenje kritičnih poslovnih odluka. Transakcijske baze podataka ne mogu odgovoriti na složena poslovna pitanja koja se odnose na ETL, stoga sam ETL pruža metodu premještanja podataka iz različitih izvora u skladište podataka. Omogućuje provedbu uzoraka podataka između izvora i ciljanog sustava. Isto tako njegov postupak može izvoditi složene transformacije i zahtijeva dodatno područje koje će se koristiti za pohranu podataka. Isto tako ETL nam pomaže u migraciji podataka u skladište podataka i pomaže u poboljšavanju produktivnosti jer se kodificira i ponovo upotrebljava bez potrebe za tehničkim vještinama, te se u osnovi dijeli na tri osnovna koraka, a to su ekstrakcija, transformacija i učitavanje. [4]

Netom prije same provedbe ETL principa i metoda, potrebno je omogućiti spajanje MySQL Workbench alata sa serverom koji je u ovom primjeru nazvan „dposavec“ uz pomoć kojeg je dalje potrebno izraditi sve potrebne tablice unutar iste baze.



Slika 2. Prikaz servera, [autorski rad]

Prethodna slika prikazuje početni zaslon MySQL Workbench alata i naziv servera sa kojim radimo.



Slika 3. Popis tablica, [autorski rad]

Prethodna fotografija prikazuje popis tablica koje se nalaze u shemu i u nastavku ćemo prikazati rezultate koji se nalaze u dotičnim tablicama.

Na sljedećoj fotografiji možemo vidjeti učitani dataset, te numeričke i tekstualne vrijednosti koje se nalaze u istoj, a dok fotografija nakon nje prikazuje atribute navedenog dataseta.

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1	31.37
6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26
7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.5	2.9	30.01
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.2	2.93	2.85	29.02
9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.7	2.26	28.62
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31
11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11	1.93	2.75	24.76
12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92	23.42
13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9	6.18	7.2	0.71	23.1
14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.6	2.15	22.72
15	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.09	8.59	2.53	1.79	22
16	Kinect Adventures!	X360	2010	Misc	Microsof...	14.97	4.94	0.24	1.67	21.82
17	Grand Theft Auto V	PS3	2013	Action	Take-T...	7.01	9.27	0.97	4.14	21.4
18	Grand Theft Auto: San And...	PS2	2004	Action	Take-T...	9.43	0.4	0.41	10.57	20.81
19	Super Mario World	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61
20	Brain Age: Train Your Brain ...	DS	2005	Misc	Nintendo	4.75	9.26	4.16	2.05	20.22
21	Pokemon Diamond/Pokemo...	DS	2006	Role-Playing	Nintendo	6.42	4.52	6.04	1.37	18.36
22	Super Mario Land	GB	1989	Platform	Nintendo	10.83	2.71	4.18	0.42	18.14
23	Super Mario Bros. 3	NES	1988	Platform	Nintendo	9.54	3.44	3.84	0.46	17.28
24	Grand Theft Auto V	X360	2013	Action	Take-T...	9.63	5.31	0.06	1.38	16.38

Slika 4. Prikaz dataseta, [autorski rad]




Table Name:

Schema: **dposavec**

Charset/Collation:

Engine:

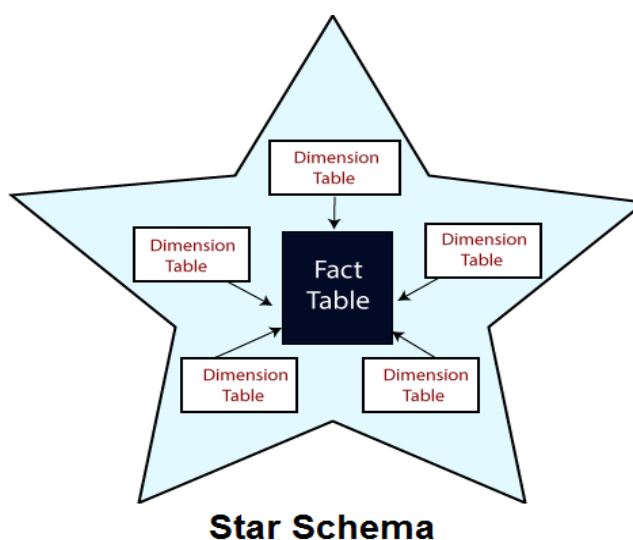
Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
Ranked	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Name	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Platform	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Year	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Genre	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Publisher	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
NA_Sales	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
EU_Sales	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
JP_Sales	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Other_Sales	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Global_Sales	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Slika 5. Prikaz atributa tablice vgsales, [autorski rad]

### 3.3. Model zvijezde

Zvjezdasti oblik sheme možemo definirati kao osnovni oblik dimenzijskog modela gdje se podaci organiziraju u dimenzije i činjenice.



Slika 6. Model zvijezde, [ 5]

Činjenica se definira kao događaj koji se broji ili mjeri poput prijave ili prodaje a dok dimenzija može uključivati referentne podatke o činjenici koji se odnose na podatke o datumima, nekim predmetima ili samim kupcima. Sama shema modela zvijezda nam predstavlja relacijsku shemu čiji dizajn predstavlja višedimenzionalni model podataka, još se opisuje i kao eksplicitna shema skladišta podataka iz razloga što se entiteti u shemi simuliraju

kao zvijezde s točkama koje se razilaze od središnje tablice. Tablice činjenice može uključivati ili činjenicu na razini detalja ili činjenice koje uključuju agregirane činjenice koje se također mogu nazivati i sažetim tablicama. Dimenzijske tablice obično se sastoje od jedne ili više hijerarhija koje kategoriziraju podatke. Primarni ključevi svake tablice dimenzija dio su složenih primarnih ključeva tablice činjenica. [5]

### 3.3.1. Dimenzijska tablica rankname

Dimenzijska tablica rankname sastoji se od dva atributa. Prvi atribut je naziva *idname*, a drugi atribut je naziva *nameRank*. Kod atributa *idname*, koji je ujedno i primarni ključ postavljen je na *auto increment*, a dok je atribut *nameRank* dodijeljen tip podatka VARCHAR (255). Da bi dobili rezultat tablice koji se prikazan na sljedećoj fotografiji potrebno je napisati sljedeći upit:

```
INSERT INTO rankname (nameRank) SELECT DISTINCT Name FROM vgsales;
```



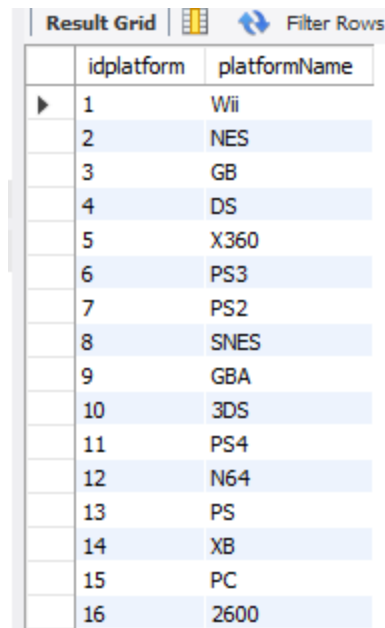
idname	nameRank
716	Spider-Man: The Movie
717	Dragon Quest VII: Warriors...
718	Uncharted: The Nathan Dr...
719	God of War
720	Nintendo Land
721	Michael Jackson: The Exper...
722	Tony Hawk's Pro Skater 3
723	Winning Eleven: Pro Evoluti...
724	The Elder Scrolls IV: Oblivion
725	Zelda II: The Adventure of ...
726	Need for Speed: Most Wan...
727	Resistance: Fall of Man
728	Sonic the Hedgehog
729	Kingdom Hearts II
730	Asteroids
731	Fable II

Slika 7. Dimenzijska tablica rankname, [autorski rad]

### 3.3.2. Dimenzijska tablica platform

Sljedeća dimanezijska tablica platform sastoji se od dva atributa, *idplatform* koji je ujedno i primarni ključ te je postavljen na *auti increment* i *platformName* koji sadrži tip podatka VARCHAR (255). Rezultat tablice prikazan je na sljedećoj fotografiji i sadrži sljedeći upit:

`INSERT INTO platform(platformname) SELECT DISTINCT Platform FROM vgsales;`



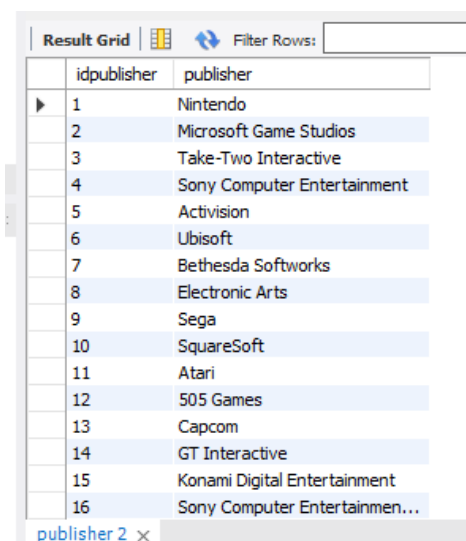
	idplatform	platformName
▶	1	Wii
	2	NES
	3	GB
	4	DS
	5	X360
	6	PS3
	7	PS2
	8	SNES
	9	GBA
	10	3DS
	11	PS4
	12	N64
	13	PS
	14	XB
	15	PC
	16	2600

Slika 8. Dimenzijska tablica platform, [autorski rad]

### 3.3.3. Dimenzijska tablica publisher

U nastavku se nalazi prikaz rezultata dimenzijske tablice publisher koja se sastoji od atributa primarnog ključa naziva *idpublisher*, koji također je postavljen na *auto increment*, te drugog atributa *publisher* koji se odnosi na samo ime izdavača videoigre i sadrži tip podatka VARCHAR (255). Upit koji je korišten za dobivanje rezultata glasi:

`INSERT INTO publisher(publisher) SELECT DISTINCT Publisher FROM vgsales;`



	idpublisher	publisher
▶	1	Nintendo
	2	Microsoft Game Studios
	3	Take-Two Interactive
	4	Sony Computer Entertainment
	5	Activision
	6	Ubisoft
	7	Bethesda Softworks
	8	Electronic Arts
	9	Sega
	10	SquareSoft
	11	Atari
	12	505 Games
	13	Capcom
	14	GT Interactive
	15	Konami Digital Entertainment
	16	Sony Computer Entertainmen...

Slika 9. Dimenzijska tablica publisher, [autorski rad]

### 3.3.4. Dimenzijska tablica genre

Dimenzijska tablica genre sastoji se također od dva atributa, a to su atribut *idgenre* i *genreName* i on se odnosi na naziv žanra dotične videoigre. Kao i kod prethodnih primarnih ključeva, tako je i ovdje *idgenre* postavljen *auto increment*, *not null*, a dok je *genreName* dodijeljen tip podatka VARCHAR(255). Za prikaz rezultata tablice kao što je to prikazano sljedećoj fotografiji korišten je sljedeći upit:

```
INSERT INTO genre(genreName) SELECT DISTINCT Genre FROM vgsales;
```



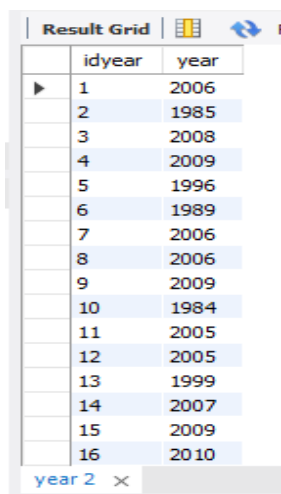
	idgenre	genreName
▶	1	Sports
	2	Platform
	3	Racing
	4	Role-Playing
	5	Puzzle
	6	Misc
	7	Shooter
	8	Simulation
	9	Action
	10	Fighting
	11	Adventure
	12	Strategy
✱	NULL	NULL

Slika 10. Dimenzijska tablica genre, [autorski rad]

### 3.3.5. Dimenzijska tablica year

Sljedeća dimenzijska tablica sastoji se od atributa *idyear* i *year*. Prvom atributu je dodijeljen *auto increment* i *not null*, a drugom atributu je dodijeljen tip podatka INT. Upit glasi:

```
INSERT INTO year(year) SELECT DISTINCT Year FROM vgsales;
```



	idyear	year
▶	1	2006
	2	1985
	3	2008
	4	2009
	5	1996
	6	1989
	7	2006
	8	2006
	9	2009
	10	1984
	11	2005
	12	2005
	13	1999
	14	2007
	15	2009
	16	2010

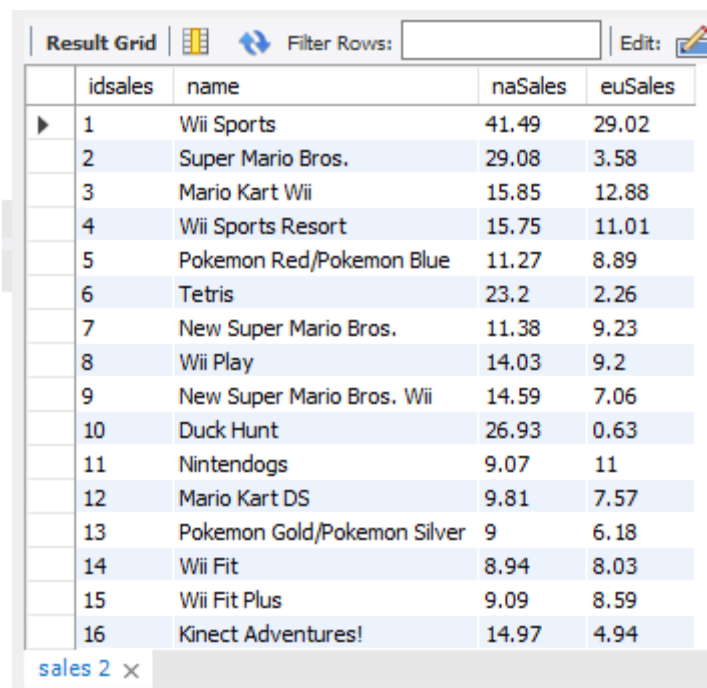
year 2 ✕

Slika 11. Dimenzijska tablica year, [autorski rad]

### 3.3.6. Dimenzijska tablica sales

Sljedeća dimenzijska tablica naziva sales sastoji se od ukupno četiri atributa, a to su redom navedeni: *idsales*, *name*, *naSales* i *euSales*. U ovoj tablici atributu *idsales* koji je ujedno i primarni ključ ove tablice, dodijeljen *auto increment* i *not null*, atributu *name* koji se odnosi na naziv videoigre, dodijeljen je tip podatka VARCHAR (255), a dok su ostalim atributima *naSales* i *euSales* dodijeljeni tipovi podataka DOUBLE. Kako bi dobili rezultate tablice koji su prikazani na sljedećoj fotografiji korišten je sljedeći upit:

```
INSERT INTO sales(name,naSales,euSales) SELECT Name, NA_Sales,EU_Sales
FROM vgsales;
```



	idsales	name	naSales	euSales
▶	1	Wii Sports	41.49	29.02
	2	Super Mario Bros.	29.08	3.58
	3	Mario Kart Wii	15.85	12.88
	4	Wii Sports Resort	15.75	11.01
	5	Pokemon Red/Pokemon Blue	11.27	8.89
	6	Tetris	23.2	2.26
	7	New Super Mario Bros.	11.38	9.23
	8	Wii Play	14.03	9.2
	9	New Super Mario Bros. Wii	14.59	7.06
	10	Duck Hunt	26.93	0.63
	11	Nintendogs	9.07	11
	12	Mario Kart DS	9.81	7.57
	13	Pokemon Gold/Pokemon Silver	9	6.18
	14	Wii Fit	8.94	8.03
	15	Wii Fit Plus	9.09	8.59
	16	Kinect Adventures!	14.97	4.94

Slika 12. Dimenzijska tablica sales, [autorski rad]

### 3.4. UPDATE nad tablicom vgsales

U nastavku su prikazani upiti koji se odnose na ažuriranje tablice vgsales:

```
UPDATE vgsales v INNER JOIN genre g on v.Genre = g.genreName
SET v.Genre = g.idgenre;
```

```
UPDATE vgsales v INNER JOIN platform p on v.Platform = p.platformName
SET v.Platform = p.idplatform;
```

```
UPDATE vgsales v INNER JOIN publisher p on v.Publisher = p.publisher
```

```
SET v.Publisher = p.idpublisher;
```

```
UPDATE vgsales v INNER JOIN name n on v.Name = n.idname
```

```
SET v.Name = n.nameRank;
```

```
UPDATE vgsales v INNER JOIN sales s on v.Name = s.name
```

```
SET v.Name = s.idsales;
```

### 3.5. ALTER tablice vgsales

U nastavku se nalaze upiti koji se odnose na prethodno izrađene tablice:

```
ALTER TABLE vgsales MODIFY Ranked int;
```

```
ALTER TABLE vgsales MODIFY Genre int;
```

```
ALTER TABLE vgsales MODIFY Platform int;
```

```
ALTER TABLE vgsales MODIFY Publisher int;
```

```
ALTER TABLE vgsales MODIFY Name int;
```

### 3.6. Činjenična tablica fact\_table

U ovom primjeru prikazat ćemo činjeničnu tablicu koja se sastoji od ukupno šest atributa, a to su:

- idfact\_table
- eu\_sales
- idplatform
- idpublisher
- idsales
- idgenre

Atribut koji su označeni sa id prefiksom, predstavljaju vanjske ključeve odgovarajuće dimenzijske tablice koje su povezane sa prethodnim činjeničnim tablicama. Kako bi se dohvatile sve potrebne vrijednosti koristi se sljedeći upit:

```
INSERT INTO fact_table(eu_sales, idplatform, idpublisher, idsales,  
idgenre) SELECT EU_Sales, Platform, Publisher, Name,  
Genre FROM dposavec.vgsales;
```



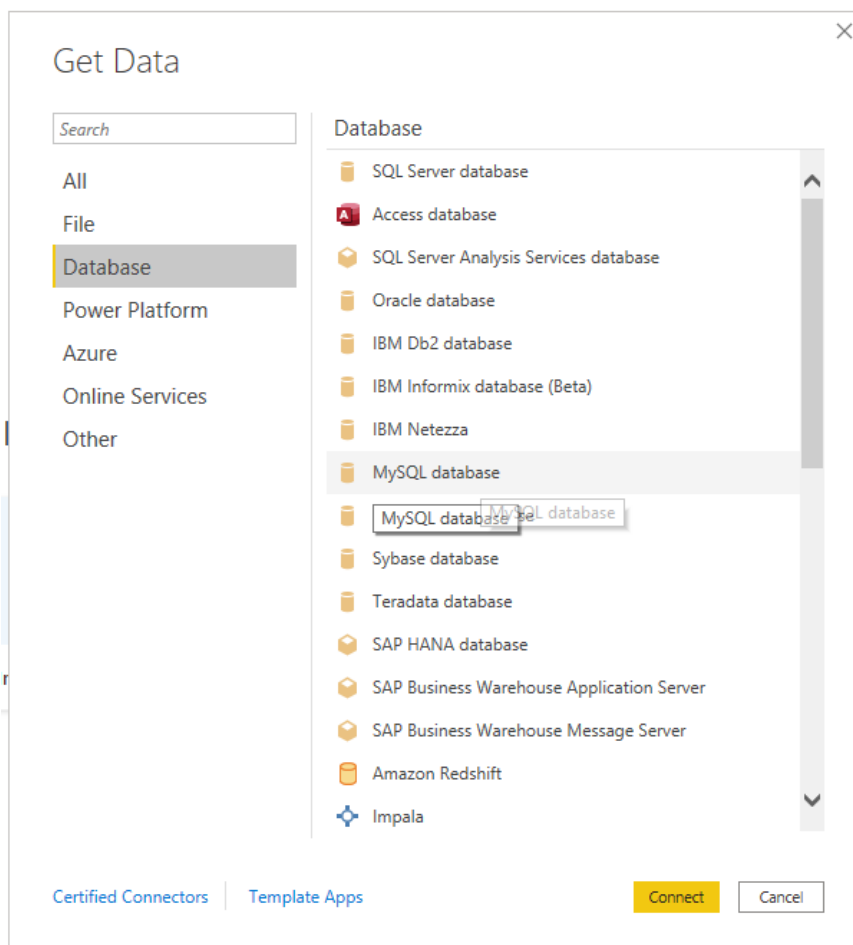
	idfact_table	eu_sales	idplatform	idpublisher	idsales	idgenre
1	1	29.02	1	1	1	1
2	2	3.58	2	1	2	2
3	3	12.88	1	1	3	3
4	4	11.01	1	1	4	1
5	5	8.89	3	1	5	4
6	6	2.26	3	1	6	5
7	7	9.23	4	1	7	2
8	8	9.2	1	1	8	6
9	9	7.06	1	1	9	2
10	10	0.63	2	1	10	7
11	11	11	4	1	11	8
12	12	7.57	4	1	12	3
13	13	6.18	3	1	13	4
14	14	8.03	1	1	14	1
15	15	8.59	1	1	15	1
16	16	4.94	5	2	16	6

Slika 13. Prikaz činjenične tablice fact\_table, [autorski rad]

## 4. Analiza uz pomoć Power BI alata

Alat Power BI paket je alat za poslovnu analitiku koji pružaju uvidu u sva područja koja se odnose na organizacije nekog područja. Sadrži i mogućnost stvaranja personaliziranih nadzornih ploča s jedinstvenim pogledom na tvrtku, isto tako možemo koristiti i relevantne podatke. Bez obzira da li se korisnikovi podaci nalaze u excel proračunskim tablicama, lokalni izvori podataka, streaming podaci ili usluge u oblaku, alat Power BI omogućuje korisniku integraciju podataka s različitih izvora. [6]

Prilikom otvaranja desktop verzije alata Power BI, potrebno je dohvatiti podatke koje želimo iskoristiti u radu s izvještajima. Na sljedećoj fotografiji nalazi se putanja koju je potrebno izvršiti i spojiti na izvor s kojega želimo dohvatiti podatke koje se nalazi na adresi 127.0.0.1 i port 3306. Radi se o MySQL-u a baza podataka je naziva dposavec.

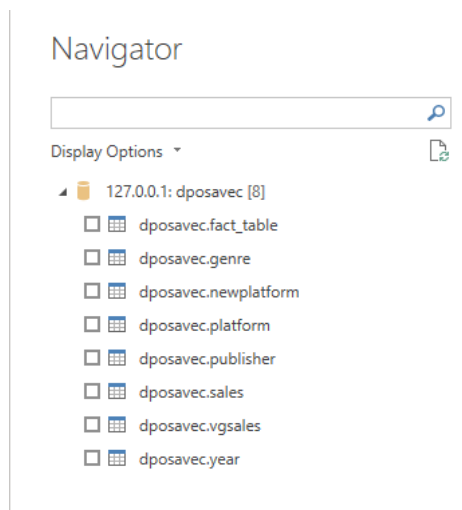


Slika 14. Postupak dodavanja podataka u PowerBI, [autorski rad]

Na sljedećoj fotografiji nalazi se mjesto na kojemu je potrebno popuniti polja kako bi se dohvatili podaci sa korisnikovog vlastitog servera i baze podataka.

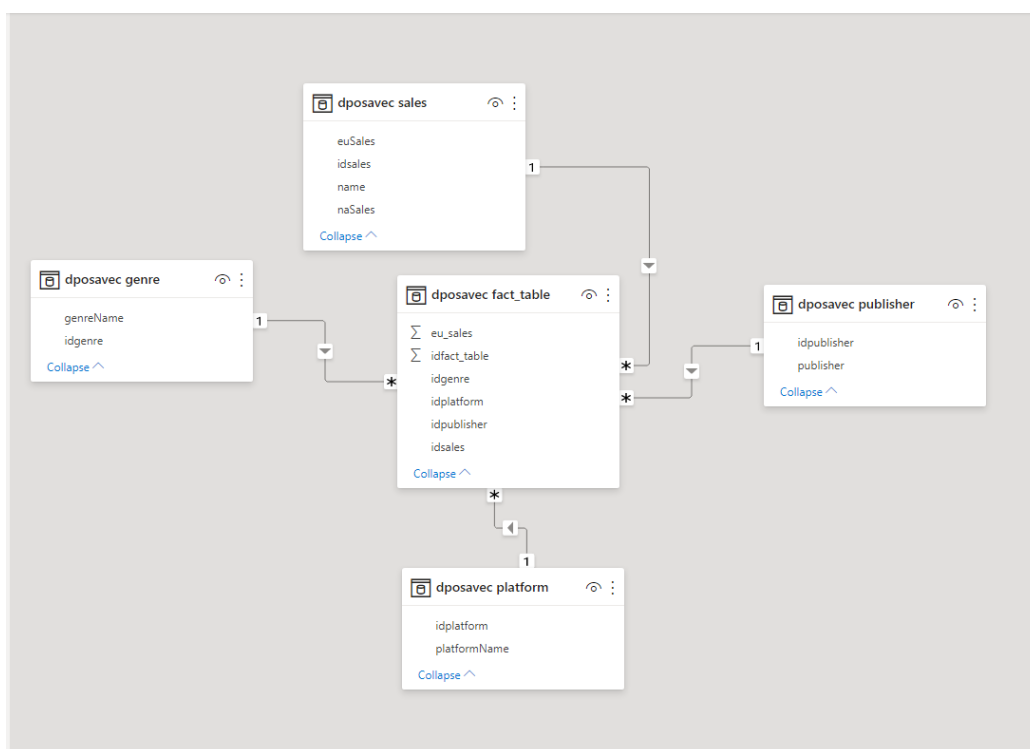


Slika 15. Popunjavanje korisnikovih podataka, [autorski rad]



Slika 16. Prikaz tablica u PowerBI, [autorski rad]

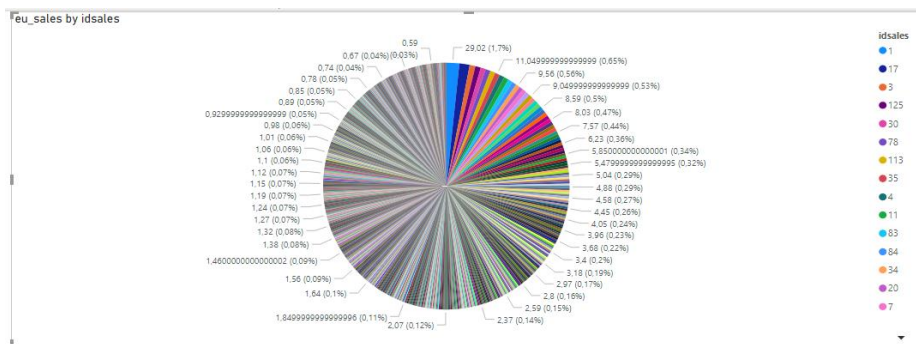
Nakon prethodnog koraka dobivamo dijaloški okvir koji je prikazan kao na prethodnoj fotografiji i tada možemo odabrati sve tablice sa kojima želimo izrađivati izvještaje. Na sljedećoj fotografiji nalazi se dobiveni generirani model zvijezda kojega dobijemo nakon odabira u navigatoru svih tablica.



Slika 17. Generirani model zvijezde, [autorski rad]

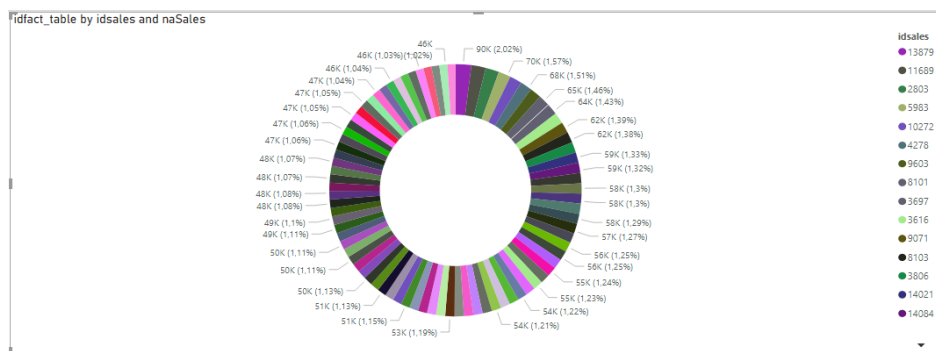
## 4.1. Generirani izvještaji

Pomoću alata Power BI u nastavku će biti prikazani primjeri izvještaja. Na sljedećoj fotografiji nalazi se prikaz prodaje videoigara na području Europe.

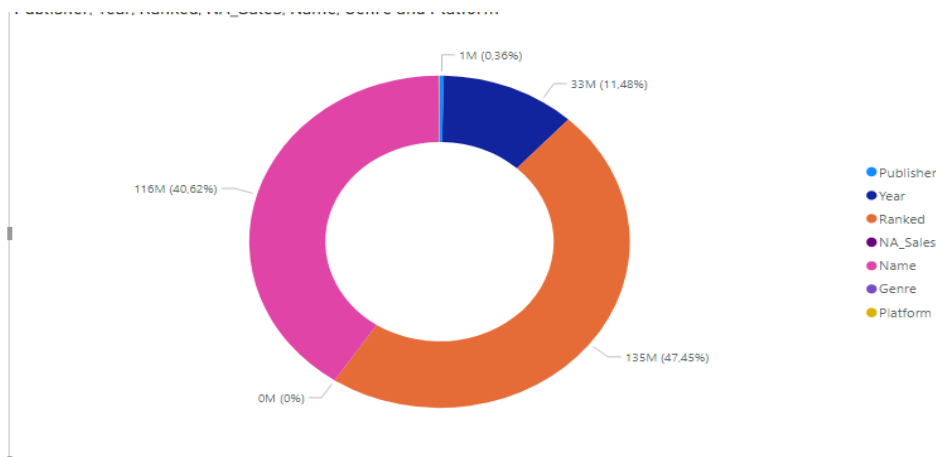


Slika 18. Prikaz prodaje videoigara u Europi, [autorski rad]

Na sljedećoj fotografiji nalazi se izvještaj prodaje koji se odnose na prodaju u određenim područjima.



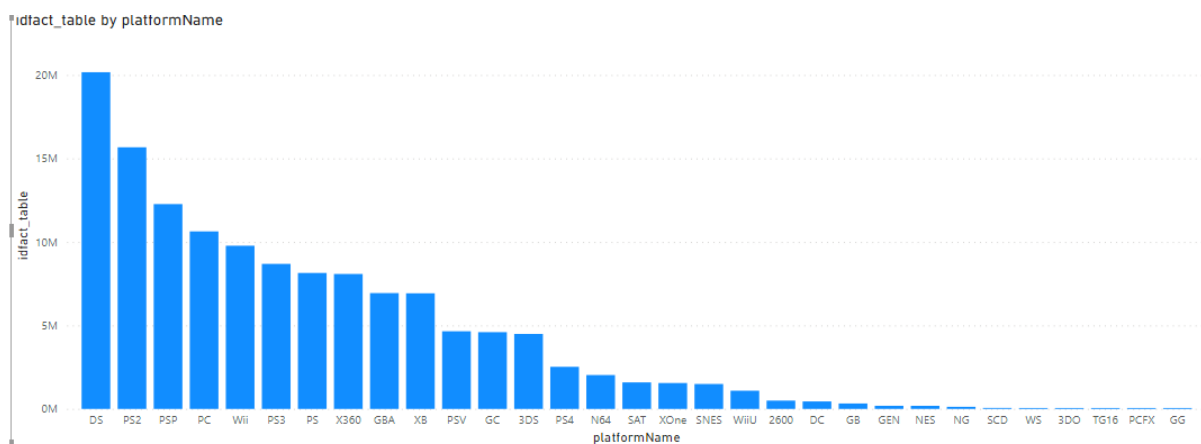
Slika 19. Prikaz prodaje u određenim područjima, [autorski rad]



Slika 20. Prikaz zastupljenosti atributa u bazi podataka, [autorski rad]

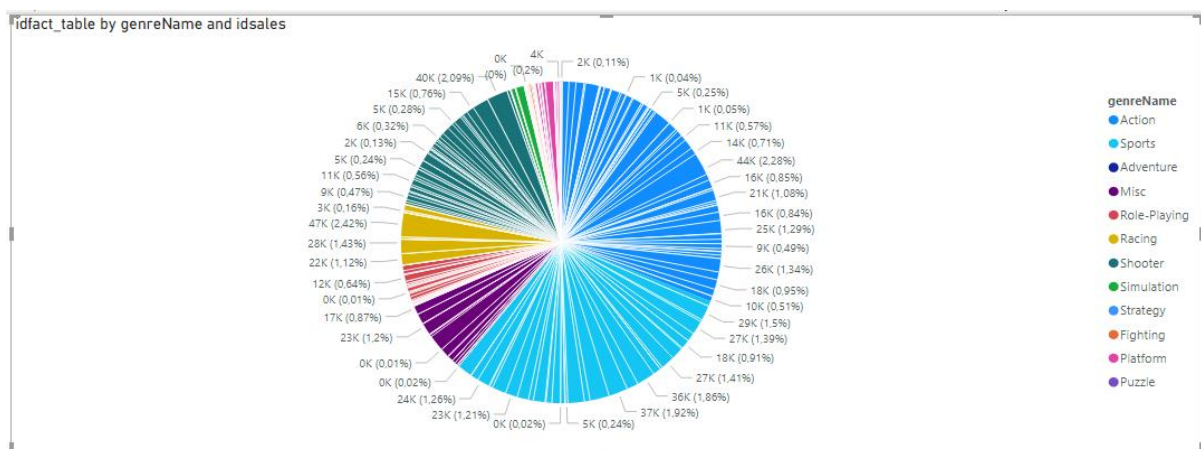
Prethodna fotografija prikazuje zastupljenost svakog atributa u koji se odnosi na izdavače, naziv videoigre, žanr i ostalo.

Sljedeća fotografija zastupljenost određene videoigre po njezinoj platformi, pa tako možemo vidjeti tri najzastupljenije platforme koje su u vodstvu, a to su: DS, PS2 i PSP.



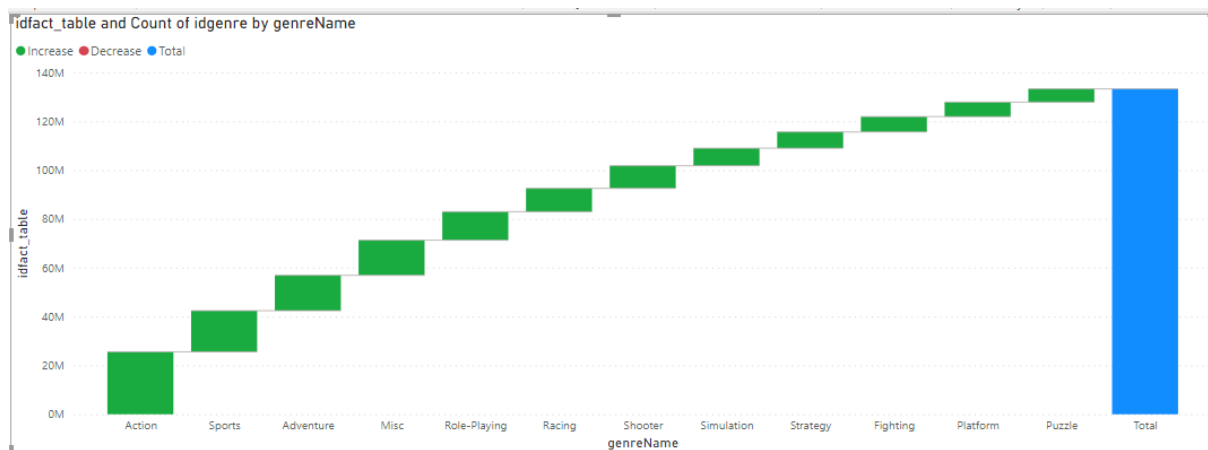
Slika 21. Prikaz najzastupljenijih platformi, [autorski rad]

Pomoću sljedeće fotografije možemo vidjeti rezultate prodaje videoigara prema žanru na koje se odnose, a to mogu biti akcijske videoigre, avanturističke, sportske, zabavne i slično.

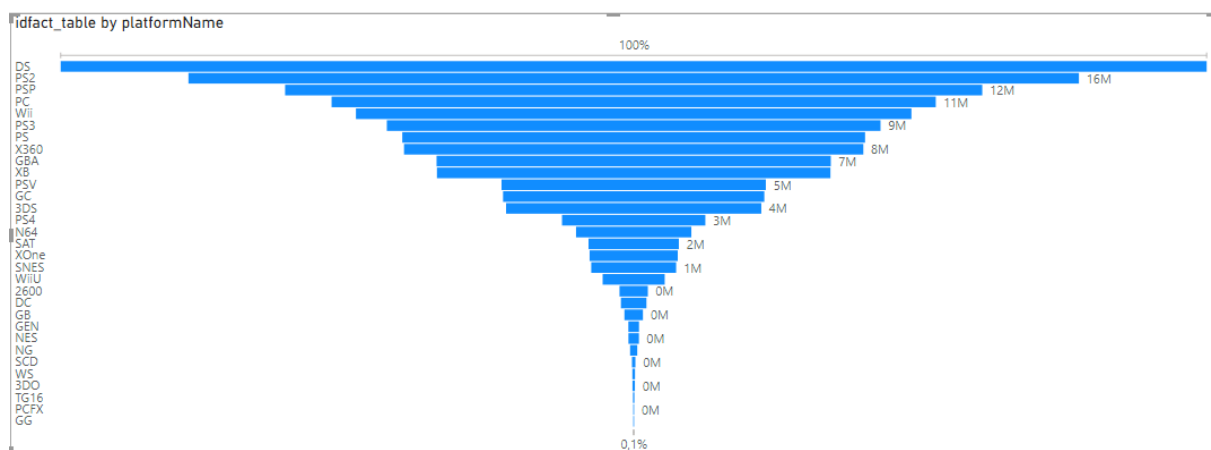


Slika 22. Prikaz prodaje po žanru, [autorski rad]

Sljedeća fotografija prikazuje zbroj pojedinačnih žanrova po njihovim i rasponu prodaje u dotičnim kategorijama.



Slika 23. Prikaz pojedinačnih žanrova, [autorski rad]



Slika 24. Prikaz prodaje po platformama, [autorski rad]

Prethodna fotografija prikazuje zastupljenost prodaje videoigara po platformama.

## 5. Zaključak

Ovaj projekt izrađen je s ciljem upoznavanja novih tehnologija, poput Power BI alat kao i ponavljanje znanja prethodnih godina koje se odnose na baze podataka. Samim time, uz pomoć novih tehnologija i daljnjim istraživanjima uspješno je riješena analiza statistike prodaje videoigara koja je ujedno i glavna tema ovog seminarskog rada. Pomoći određenih upita dobili smo uvid u rješenja potrebnih dimenzijskih tablica, činjenične tablice, te u konačnici i vizualizirani izvještaj uz pomoć kojega smo na kraju jasno predočili rješenja. Uz prethodno spomenute alate koji su bili korišteni u ovome radu Power BI i MySQL Workbench, koristio se i besplatan dataset koji je bio preuzet s web stranice Kaggle.com. Uz velikog broj zapisa i bazi podataka sama rješenja u konačnici su bila vrlo kvalitetna zahvaljujući vrlo kvalitetnim i preciznim podacima u tablici. Na samome kraju možemo reći kako se skladišta podataka u današnje vrijeme tehnologije vrlo brzo razvijaju u smislu širenja podataka, zahtjevnosti što većeg kapaciteta za skladištenje podataka. Za upravljanje skladištem podataka i rad s njihovim podacima, potrebni su stručnjaci koji su specijalizirani za rad u takvom području podataka jer s rastom tehnologije dolazi i njihovog proboja, odnosno napada zlonamjernih osoba s ciljem otuđivanja podataka. Isto tako možemo reći kako nam skladište podataka predstavlja i jedinstvenu sliku poslovne realnosti kao i definiranje cjelokupne poslovne strategije. Samo skladište podataka nam predstavlja informacijsku tehnologiju koja se sa svojim aspektima i činjenicama opisuje kao kooperativan i timski rad koji se usko vezan uz stručnjake na informatičkom ali i u bilo kojem području, u globalu ono nam predstavlja poslovni i tehnički informacijski svijet.

## Popis literature

- [1] K.Ćurko (2001.): *Skladište podataka – sustav za potporu odlučivanju*, dostupno na: <https://hrcak.srce.hr/file/45126> , pristupano 25.5.2021.
- [2] Ottomatik.io (2021.): *Guide To MySQL Workbench – Understand How It Works*, dostupno na: <https://ottomatik.io/blog/mysql-workbench/> , pristupano 25.5.2021.
- [3] Docs.microsoft.com (2021.): *What is Power BI?*, dostupno na: <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview> , pristupano: 25.5.2021.
- [4] Guru99.com (2021.): *ETL (Extract, Transform, and Load) Process in Data Warehouse*, dostupno na: <https://www.guru99.com/etl-extract-load-process.html> , pristupano 25.5.2021.
- [5] Javatpoint.com (2021.): *What is Star Schema?*, dostupno na: <https://www.javatpoint.com/data-warehouse-what-is-star-schema>, pristupano 25.5.2021.
- [6] Ambit-group.com (2021.): *Microsoft Power Platform*, dostupno na: <https://ambit-group.com/en/products-solutions/microsoft-power-platform>, pristupano 25.5.2021.



## Popis slika

Slika 1. Prikaz dijela preuzetog skupa podataka, [autorski rad] .....	4
Slika 2. Prikaz servera, [autorski rad] .....	5
Slika 3. Popis tablica, [autorski rad] .....	6
Slika 4. Prikaz dataseta, [autorski rad] .....	6
Slika 5. Prikaz atributa tablice vgsales, [autorski rad] .....	7
Slika 6. Model zvijezde, [ 5] .....	7
Slika 7. Dimenzijska tablica rankname, [autorski rad] .....	8
Slika 8. Dimenzijska tablica platform, [autorski rad] .....	9
Slika 9. Dimenzijska tablica publisher, [autorski rad] .....	9
Slika 10. Dimenzijska tablica genre, [autorski rad] .....	10
Slika 11. Dimenzijska tablica year, [autorski rad] .....	10
Slika 12. Dimenzijska tablica sales, [autorski rad] .....	11
Slika 13. Prikaz činjenične tablice fact_table, [autorski rad] .....	13
Slika 14. Postupak dodavanja podataka u PowerBI, [autorski rad] .....	14
Slika 15. Popunjavanje korisnikovih podataka, [autorski rad].....	14
Slika 16. Prikaz tablica u PowerBI, [autorski rad] .....	15
Slika 17. Generirani model zvijezde, [autorski rad].....	15
Slika 18. Prikaz prodaje videoigara u Europi, [autorski rad] .....	16
Slika 19. Prikaz prodaje u određenim područjima, [autorski rad].....	16
Slika 20. Prikaz zastupljenosti atributa u bazi podataka, [autorski rad] .....	16
Slika 21. Prikaz najzastupljenijih platformi, [autorski rad] .....	17
Slika 22. Prikaz prodaje po žanru, [autorski rad] .....	17
Slika 23. Prikaz pojedinačnih žanrova, [autorski rad].....	18
Slika 24. Prikaz prodaje po platformama, [autorski rad] .....	18