

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

**Kristijan Maoduš**

**Analiza rabljenih automobila s Ebay-  
Kleinanzeigen**

**SEMINARSKI RAD**

**Varaždin, 2020.**

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

**Kristijan Maoduš**

**JMBAG: 0016116726**

**Studij: Baze podataka i baze znanja**

**Analiza rabljenih automobila s Ebay-Kleinanzeigen**

**SEMINARSKI RAD**

**Mentor:**

Prof. dr. sc. Kornelije Rabuzin

**Varaždin, lipanj 2020.**

*Kristijan Maoduš*

### **Izjava o izvornosti**

Izjavljujem da je moj završni/diplomski rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

*Autor potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi*

---

## **Sažetak**

Seminarski rad za potrebe kolegija *Skladišta podataka i poslovna inteligencija*. Kroz rad je prikazana izrada i implementacija modela skladišta korištenjem Microsoft SQL Server-a i kreiranje izvještaja na temelju podataka iz istog skladišta korištenjem alata Microsoft Power BI.

**Ključne riječi:** skladišta podataka; poslovna inteligencija; Power BI; SQL Server; SSMS; data warehouse

# Sadržaj

1. Uvod .....	1
2. Opis korištenih alata .....	1
3. Opis domene prikazane modelom skladišta .....	1
4. Opis korištenog skupa podataka .....	2
5. Opis provedenog ETL procesa .....	5
6. Opis modela izrađenog skladišta podataka .....	9
6.1. Dimenzijska tablica 'DimGearbox' .....	9
6.2. Dimenzijska tablica DimFuelType .....	10
6.3. Dimenzijska tablica 'DimBrand' .....	11
6.4. Dimenzijska tablica 'DimModel' .....	12
6.5. Dimenzijska tablica 'DimDate' .....	14
6.6. Dimenzijska tablica 'DimAdvertisement' .....	15
6.7. Činjenična tablica 'FactCars' .....	17
6.8. Model zvijezde .....	19
7. Prikaz izvještaja napravljenih u Microsoft Power BI alatu .....	20
8. Zaključak .....	24
Popis literature .....	25
Popis slika .....	26
Popis tablica .....	28

# 1. Uvod

Kroz ovaj seminarski rad prikazat ćemo izradu i implementaciju skladišta podataka u svrhu analiziranja rabljenih automobila preko oglasnika Ebay-Kleinanzeigen. Za potrebe kreiranja skladišta korišten je Microsoft SQL Server, a za izvještaje također Microsoftov alat Power BI. Podaci koje sam koristio nalazili su se u csv datoteci. Više o alatima koji su korišteni i podatkovnom skupu reći ću u narednim poglavljima. Također, opisat ću domenu prikazanu modelom skladišta, model izrađenog skladišta, opis provedenih ETL procesa, te naposljetku prikaz izvještaja kreiranih uz pomoć Power BI alata.

## 2. Opis korištenih alata

Alati koji su korišteni za potrebe ovog seminarskog rada su:

- Microsoft Excel
- Microsoft SQL Server
- Microsoft SQL Server Management Studio
- Microsoft Power BI [1]

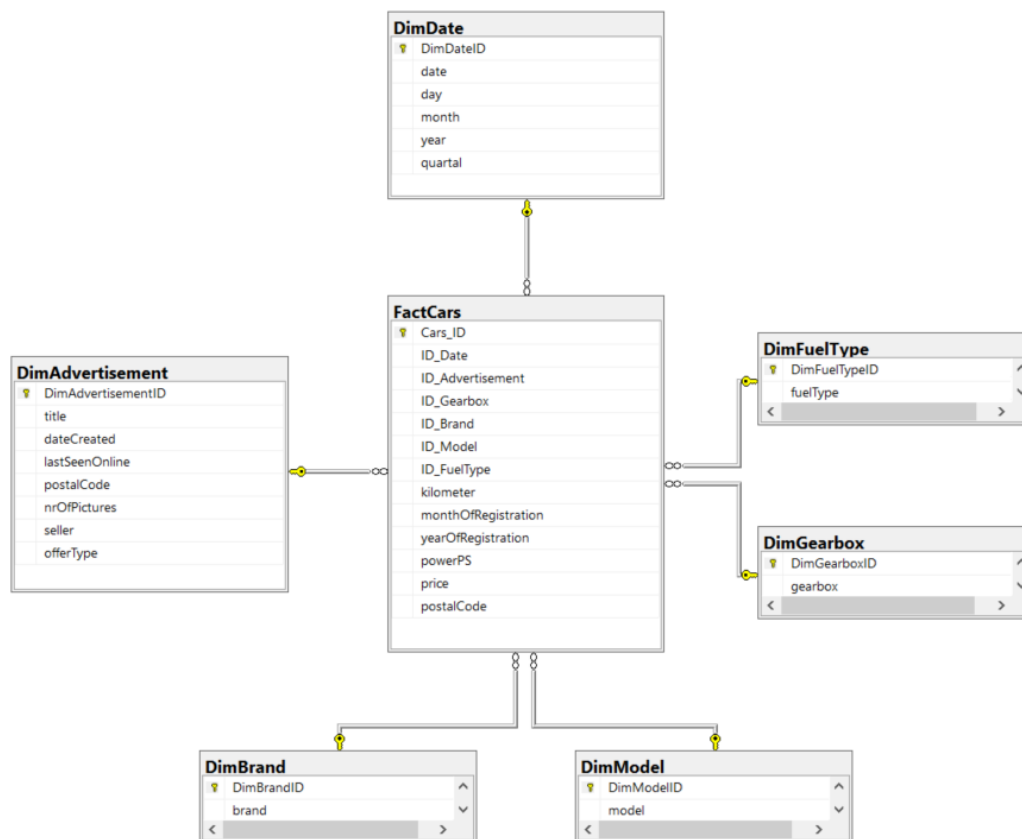
Kao što možemo vidjeti svi navedeni alati su od tvrtke Microsoft i savršeno su kompatibilni. Excel sam koristio za potrebe modificiranja početne .csv datoteke, koju sam zatim učitao (eng. *import*) u SQL Server Management Studio (kraće SSMS) verzije 18.5. Nakon što sam izradio model skladišta i proveo ETL (eng. *Extract Transform Load*) proces transformirane podatke učitao sam u Power BI Desktop program pomoću kojega su podaci vizualizirani tj. pomoću kojega su izrađeni izvještaji.

Power BI je zapravo zbirka softverskih usluga i aplikacija koji zajedno rade na pretvaranju nepovezanih izvora podataka u koherentne, vizualno imerzivne i interaktivne uvide. Podaci mogu biti tablice iz Excela ili kolekcija hibridnih skladišta podataka bila ona spremljena lokalno ili u oblaku [2].

## 3. Opis domene prikazane modelom skladišta

Svrha izrađenog skladišta je prikaz činjenica o rabljenim automobilima s određenog oglasa, te prikaz svih mjera auta dostupnih na oglasu kako bi otkrili povezanosti između specifikacija auta, cijene oglasa, najpopularnijih proizvođača, modela i slično.

Na slici 1. prikazan je model skladišta kako bi dobili dojam samog skladišta podataka, a detaljnije će biti opisan u poglavlju 6.



Slika 1. Model skladišta podataka

## 4. Opis korištenog skupa podataka

Izvor podataka je csv datoteka koju sam preuzeo s <https://www.kaggle.com/orgesleka/used-cars-database>. Podaci su dohvaćeni preko skripte [3] sa njemačkog oglasnika Ebay-Kleinanzeigen. Datoteka sadrži više od 370 000 zapisa i 20 stupaca, odnosno atributa. Izgled početne datoteke vidimo na slici ispod.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	dateCrawled	name	seller	offerType	price	abtest	vehicleType	yearOfRegistration	gearbox	powerPS	model	kilometer	monthOfRegistration	fuelType	brand	notRepairedDamage	dateCreated	nrOfPictures	postalCode	lastSeen			
2	2016-03-24 11:52:17	Golf_3_1,6	privat	Angebot	480	test	1993	manuell	0	golf	150000,0	benzin	volkswagen	2016-03-24 00:00:00	0,70435	2016-04-07 03:16:57							
3	2016-03-24 10:58:45	As_Sportback_2_7	TDI	privat	Angebot	18300	test	coupe	2011	manuell	190_125000,5	diesel	audi	ja	2016-03-24 00:00:00	0,66954	2016-04-07 01:46:50						
4	2016-03-14 12:52:21	Jeep_Grand_Cherokee_Overland	privat	Angebot	9800	test	2004	automatik	163	grand	125000,0	diesel	jeep	2016-03-14 00:00:00	0,90480	2016-04-05 12:47:46							
5	2016-03-17 16:54:04	GOLF_4_1_4	3TUREP	privat	Angebot	1500	test	kleinwagen	2001	manuell	75	golf	150000,6	benzin	volkswagen	nein	2016-03-17 00:00:00	0,91074	2016-03-17 17:40:17				
6	2016-03-31 17:25:20	Skoda_Fabia_1,4	TDI	PO	Classic	privat	Angebot	3600	test	kleinwagen	2008	manuell	69	fabia	90000,7	diesel	skoda	nein	2016-03-31 00:00:00	0,60437	2016-04-06 10:17:21		
7	2016-04-04 17:36:29	BMW_316i_e36	Limousine	Bastlerfahrzeug	Export	privat	Angebot	650	test	limousine	1995	manuell	102	3er	150000,10	benzin	bmw	ja	2016-04-04 00:00:00	0,33775	2016-04-06 19:17:07		
8	2016-04-01 20:48:51	Peugeot_206_CC_110	Platinum	privat	Angebot	2200	test	cabrio	2004	manuell	109_2	reihe	150000,8	benzin	peugeot	nein	2016-04-01 00:00:00	0,67112	2016-04-05 18:18:39				
9	2016-03-21 18:54:58	VW_Derby_BJ_80	Scheunenfund	privat	Angebot	0	test	limousine	1980	manuell	50	andere	40000,7	benzin	volkswagen	nein	2016-03-21 00:00:00	0,19948	2016-03-25 16:47:58				
10	2016-04-04 23:42:13	Ford_C_Max_Titanium_1_0_L	EcoBoost	privat	Angebot	14500	control	bus	2014	manuell	125_c_max	30000,8	benzin	ford	2016-04-04 00:00:00	0,94505	2016-04-04 23:42:13						
11	2016-03-17 10:53:50	VW_Golf_4_5	tuerig_zu_verkaufen_mit_Anhaengerkupplung	privat	Angebot	999	test	kleinwagen	1998	manuell	101	golf	150000,0	volkswagen	2016-03-17 00:00:00	0,27472	2016-03-31 17:17:06						
12	2016-03-26 19:54:18	Mazda_3_1,6	Max	Titanium_1_0	privat	Angebot	2000	control	limousine	2004	manuell	125_c_max	30000,8	benzin	ford	2016-03-26 00:00:00	0,96224	2016-04-06 10:45:34					
13	2016-04-07 10:06:22	Volkswagen_Passat_Variant_2_0	TDI	Comfortline	privat	Angebot	2799	control	kombi	2005	manuell	140	passat	150000,12	diesel	volkswagen	ja	2016-04-07 00:00:00	0,57290	2016-04-07 10:25:17			
14	2016-03-15 22:49:09	VW_Passat_Facelift_35l	"Sitzer"	privat	Angebot	999	control	kombi	1995	manuell	115	passat	150000,11	benzin	volkswagen	2016-03-15 00:00:00	0,37269	2016-04-01 13:16:16					
15	2016-03-21 21:37:40	VW_PassAT_1,9	TDI	131_PS_LEDER	privat	Angebot	2500	control	kombi	2004	manuell	131	passat	150000,2	volkswagen	nein	2016-03-21 00:00:00	0,90762	2016-03-23 02:50:54				
16	2016-03-21 12:57:01	Nissan_Navara_2,5D	DPF	SE4x4_Klima_Sitzheizg	Bluetooth	Doppelkabine	privat	Angebot	17999	control	suv	2011	manuell	190	navara	70000,3	diesel	nissan	nein	2016-03-21 00:00:00	0,04177	2016-04-06 07:45:42	
17	2016-04-11 21:39:15	KA_Luftthansa_Edition_450E_VB	privat	Angebot	450	test	kleinwagen	1910_0	ka	5000,0	benzin	ford	2016-03-11 00:00:00	0,24148	2016-03-19 08:46:47								
18	2016-04-01 12:46:46	Polo_6n_1_4	privat	Angebot	300	test	2016_60	polo	150000,0	benzin	volkswagen	2016-04-01 00:00:00	0,38871	2016-04-01 12:46:46									
19	2016-03-20 10:25:19	Renault_Twingo_1,2_16V	Aut	privat	Angebot	1750	control	kleinwagen	2004	automatik	75	twingo	150000,2	benzin	renault	nein	2016-03-20 00:00:00	0,65599	2016-04-06 13:16:07				
20	2016-03-23 15:48:05	Ford_C_Max_2_0	TDi	DPF	Titanium	AMG	privat	Angebot	7550	test	bus	2007	manuell	136_c_max	150000,6	diesel	ford	nein	2016-03-23 00:00:00	0,88361	2016-04-05 18:45:11		
21	2016-04-01 22:58:47	Mercedes_Benz_A_160	Classic_Klima	privat	Angebot	1850	test	bus	2004	manuell	102_a_klasse	150000,1	benzin	mercedes_benz	nein	2016-04-01 00:00:00	0,49565	2016-04-05 22:46:05					
22	2016-04-01 19:56:48	Volkswagen_Scirocco_1,4	TSI	Sport	privat	Angebot	10400	control	coupe	2009	manuell	160	scirocco	100000,4	benzin	volkswagen	nein	2016-04-01 00:00:00	0,75365	2016-04-05 16:45:49			
23	2016-03-27 11:38:00	BMW_530i_TUV_717	Scheckheftgepflegt	sehr_guter_Zustand	privat	Angebot	3699	test	limousine	2002	automatik	231	5er	150000,7	benzin	bmw	nein	2016-03-27 00:00:00	0,68309	2016-04-07 06:44:26			
24	2016-03-13 24:52:51	Opel_Meriva_1,3	privat	Angebot	2900	test	2018	manuell	90	meriva	150000,5	benzin	opel	nein	2016-03-23 00:00:00	0,49716	2016-03-31 01:16:33						
25	2016-03-12 19:43:07	Stadtfilzter	privat	Angebot	450	test	kleinwagen	1997	manuell	50	arosa	150000,5	benzin	seat	nein	2016-03-12 00:00:00	0,09526	2016-03-21 01:46:11					
26	2016-03-13 20:40:49	MERCEDES_200E_TUV	04/2016	privat	Angebot	500	test	limousine	1990	manuell	118	andere	150000,10	benzin	mercedes_benz	ja	2016-03-13 00:00:00	0,35390	2016-03-13 20:40:49				
27	2016-03-18 21:44:09	BMW_530d_touring	Vollausstattung_NAVI	privat	Angebot	2500	control	kombi	2002	automatik	193	5er	150000,5	diesel	bmw	ja	2016-03-18 00:00:00	0,73765	2016-03-18 21:14:09				
28	2016-03-10 19:38:18	Citroen_C4_Grand_Picasso	privat	Angebot	5555	control	2017	manuell	125_c4	125000,4	citroen	nein	2016-03-10 00:00:00	0,31199	2016-03-16 09:16:46								
29	2016-03-07 12:51:23	Honda_Civic_1,4_1	VTEC	Comfort	privat	Angebot	6900	test	limousine	2008	manuell	95	civic	60000,11	benzin	honda	nein	2016-03-07 00:00:00	0,12621	2016-03-26 09:44:53			
30	2016-03-09 11:56:38	Volkswagen_T3	andere	privat	Angebot	1990	test	bus	1981	manuell	50	transporter	5000,1	benzin	volkswagen	nein	2016-03-09 00:00:00	0,87471	2016-03-10 07:44:33				
31	2016-03-08 19:55:19	Fiat_Punto_1,2	privat	Angebot	600	test	kleinwagen	2003	manuell	60	punto	150000,3	benzin	fiat	nein	2016-03-08 00:00:00	0,86199	2016-03-09 11:45:28					
32	2016-04-03 15:48:11	Mercedes_Benz_E_250_D_Original	Zustand_II	privat	Angebot	3300	test	limousine	1995	automatik	113_e_klasse	150000,1	diesel	mercedes_benz	nein	2016-04-03 00:00:00	0,53879	2016-04-05 15:16:05					
33	2016-03-29 16:57:02	Renault_clio_1,2	TUV	2016	privat	Angebot	899	control	2016	manuell	60	clio	150000,6	benzin	renault	2016-03-29 00:00:00	0,37075	2016-03-29 17:43:07					
34	2016-03-15 20:59:01	Golf_3_1	1,4	privat	Angebot	245	test	limousine	1994_0	golf	150000,2	benzin	volkswagen	nein	2016-03-15 00:00:00	0,44145	2016-03-17 18:17:43						
35	2016-03-25 21:48:47	BMW_325i_Aut	privat	Angebot	18000	test	limousine	2007	automatik	218	3er	20000,5	benzin	bmw	nein	2016-03-25 00:00:00	0,39179	2016-04-07 04:45:21					
36	2016-03-17 18:55:12	Mercedes_Benz_E_200_CDI	Automatik	Classic	privat	Angebot	3500	control	limousine	2004	automatik	122_e_klasse	150000,11	diesel	mercedes_benz	nein	2016-03-17 00:00:00	0,67071	2016-03-30 15:46:10				
37	2016-03-08 07:54:46	VW_Golf_3	privat	Angebot	350	test	2016	manuell	75	golf	150000,4	benzin	volkswagen	nein	2016-03-08 00:00:00	0,19386	2016-03-08 09:44:50						
38	2016-03-11 11:50:07	Opel_Kadett_E_C	privat	Angebot	1600	control	andere	1991	manuell	75	kadett	70000,0	opel	2016-03-11 00:00:00	0,02943	2016-04-07 03:46:09							
39	2016-03-28 17:50:15	Renault_Kangoo_1,9	Diesel	privat	Angebot	1500	test	2016_0	kangoo	150000,1	diesel	renault	nein	2016-03-28 00:00:00	0,46483	2016-03-30 09:18:02							
40	2016-04-01 17:45:07	Abschleppwagen_Vw_UT_135	000	gruene_Plakette_TUV_8/2017	privat	Angebot	11900	test	andere	2002	manuell	129	andere	150000,11	diesel	volkswagen	nein	2016-04-01 00:00:00	0,10551	2016-04-05 12:47:30			
41	autos																						

Slika 2. Prikaz originalne csv datoteke

Pošto su podaci na njemačkom, preveo sam ih na engleski jezik kojeg sam se držao i prilikom izrade samog skladišta podataka. Unutar Microsoft Excela razdvojio sam podatke po stupcima i dodao sam stupac *csvID* s jedinstvenim identifikatorom pošto izvorna datoteka nije imala jedan. Također obrisao sam neke redove koji su bili ekstremi tj. previše su odstupali od prosjeka i nisu imali smisla. Isto tako obrisao sam redove koji nisu bili potpuni. Modificirana csv datoteka prikazana je na slici 2.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
csvID	dateCrawled	name	seller	offerType	price	vehicleType	yearOfRegistration	gearbox	powerPS	model	kilometer	monthOfRegistration	fuelType	brand	notRepairedDamage	dateCreated	nrOfPictures	postalCode	lastSeen		
1	16.3.2016	1148 Alfa Romeo_147_GTA	private	offer	8999	2005	0	147	125000	0	gasoline	alfa_romeo	0	10.3.2014 00:00	9	79669	7.4.2016 4				
2	23.3.2016	1442 Der_VAN_ist_us_OSTERN_reisefertig	private	offer	6000	van	1995	automatic	200	andere	150000	0	gasoline	chevrolet	0	20.3.2015 00:00	8	24558	7.4.2016 3		
3	23.3.2016	148 Mercedes_Benz_C220_BT_70_Tronic_AMG	private	offer	47900	limousine	2014	automatic	0_c_klasse	20000	2	diesel	mercedes	0	11.6.2015 00:00	2	46145	6.4.2016 9			
4	5.3.2016	2144 Fiat_Ducato	private	offer	10400	van	1993	0_ducato	3000	1	diesel	fiat	0	18.6.2015 00:00	1	76275	5.4.2016 17				
5	2.4.2016	1947 Mercedes_S1X_200_AMG_Performance_19	private	offer	14250	cabrio	2005	manual	163_sik	150000	6	gasoline	mercedes	0	7.8.2015 00:00	12	76228	7.4.2016 9			
6	8.3.2016	1757 Cabrio_Peugeot_206_CC_Platinum_suv_Wrivate	private	offer	5200	cabrio	2006	manual	109_2_reihe	80000	11	gasoline	peugeot	0	10.8.2015 00:00	13	26382	5.4.2016 20			
7	14.3.2016	843 VW_Polo_1,2_Cricket_2005_mit_Tauschmc	private	offer	600	compact	2005	manual	54_polo	150000	1	gasoline	volkswagen	1	4.9.2015 00:00	14	72669	15.3.2016 14			
8	15.3.2016	156 VW_Polo_1,2_Cricket_2005_mit_Tauschmc	private	offer	600	compact	2005	manual	54_polo	150000	1	gasoline	volkswagen	1	4.9.2015 00:00	7	72669	15.3.2016 15			
9	5.3.2016	2151 OPEL_CORSA_Apfelgruen_dringend_zu_ver	private	offer	9000	compact	2013	manual	90_corsa	50000	4	diesel	opel	0	9.9.2015 00:00	5	98739	5.4.2016 17			
10	5.3.2016	1635 Audi_Q7_3_0_TDI_7_Sitzer_Vollausstattung	private	offer	27900	suv	2009	automatic	239_q7	125000	10	diesel	audi	0	14.10.2015 00:00	3	44651	14.3.2016 15			
11	5.3.2016	1753 RENAULT_CLIO_1,6_EDITION_DYNAMIQUE	private	offer	3490	limousine	2006	manual	86_clio	150000	3	gasoline	renault	0	2.11.2015 00:00	12	40880	17.3.2016 15			
12	5.3.2016	1431 Audi_A4_Avant_1,8_TSI_Ambition_neuer	private	offer	15500	station wagon	2009	manual	160	4	gasoline	audi	0	8.11.2015 00:00	0	59410	6.4.2016 23				
13	23.3.2016	2373 Geefleugel_Peugeot_207_Sport_V17_Sport	private	offer	1589	station wagon	2009	manual	52_reihe	70000	3	gasoline	peugeot	0	10.11.2015 00:00	1	22641	7.4.2016 9			
14	14.3.2016	1842 Lancia_Typhon_1_2_8_V_ane_Zwei_Tov	private	offer	5400	compact	2009	manual	144_typhon	40000	4	gasoline	lancia	0	12.11.2015 00:00	7	31119	6.4.2016 16			
15	14.3.2016	745 X4_3.0_Drive_Asch_Navi_Xenon_Leder	private	offer	5400	compact	2015	automatic	258_reihe	100000	3	diesel	bmw	0	13.11.2015 00:00	8	85366	5.4.2016 20			
16	5.3.2016	2042 W202_C880_v6_schmidt_fangm_Av	private	offer	5999	limousine	1997	automatic	157_c_klasse	150000	9	gasoline	mercedes	0	17.01.2015 00:00	4	28816	6.4.2016 15			
17	5.3.2016	2419 Mazda_Cabrio_Mk1_Bs_90_Sp_Silver	private	offer	5900	cabrio	1995	manual	90_ms_reihe	125000	3	gasoline	mazda	0	23.11.2015 00:00	0	63741	23.3.2016 7			
18	19.3.2016	1709 TUV_Neu_Vivaro_Combo_1200_3_GC	private	offer	11999	van	2008	manual	1000	50000	0	gasoline	mercedes	0	23.11.2015 00:00	0	72417	23.3.2016 17			
19	14.3.2016	827 Sest_M10_Nightblue_nit_Winterreifen	private	offer	7999	limousine	2013	manual	75_m1	30000	6	gasoline	sest	0	24.11.2015 00:00	14	60437	5.4.2016 22			
20	3.4.2016	2150 Seat_M10_Nightblue_20_Summer_Ubernat	private	offer	7999	limousine	2013	manual	75_m1	30000	6	gasoline	sest	0	24.11.2015 00:00	5	60437	7.4.2016 11			
21	14.3.2016	141 Subaru_Impreza_GD_500_Super_Optim_Klim	private	offer	900	limousine	1997	automatic	115_impreza	150000	4	gasoline	subaru	0	5.12.2015 00:00	10	2943	23.3.2016 15			
22	7.3.2016	141 Subaru_Impreza_GD_500_Super_Optim_Klim	private	offer	900	limousine	1997	automatic	115_impreza	150000	4	gasoline	subaru	0	5.12.2015 00:00	7	2943	23.3.2016 20			
23	18.3.2016	1041 Subaru_Impreza_GD_8_2_0_444_Super_Privat	private	offer	900	limousine	1997	automatic	115_impreza	150000	4	gasoline	subaru	0	5.12.2015 00:00	8	2943	23.3.2016 18			
24	28.3.2016	1331 Audi_A3_Ambition	private	offer	6950	limousine	2004	automatic	150	43	100000	1	gasoline	audi	0	6.12.2015 00:00	2	61389	6.4.2016 14		
25	11.3.2016	1247 Audi_A4_2_0_TDI_Multitronic_S_Line/MOTOR_S	private	offer	9999	station wagon	2006	automatic	140	44	150000	3	diesel	audi	0	17.12.2015 00:00	15	88975	5.4.2016 12		
26	11.3.2016	1481 BMW_650_N51_Gas_190_mw_TV	private	offer	7500	limousine	2005	automatic	120	10000	0	gasoline	bmw	0	27.12.2015 00:00	11	16317	27.12.2015 12			
27	14.3.2016	1548 Neuvagen_Mercedes_Benz_180_Style	private	offer	2500	van	2015	manual	122_klasse	5000	11	gasoline	mercedes	0	28.11.2015 00:00	1	76287	27.3.2016 16			
28	24.3.2016	1419 Mercedes_S180_Farbe_127_Antikgitter	private	offer	17200	cabrio	1985	automatic	155	11	90000	2	gasoline	mercedes	0	10.12.2015 00:00	15	75177	5.4.2016 12		
29	15.3.2016	2157 Smart_Fortwo_coupé_pulsant_48kW_16V	private	offer	2400	limousine	2004	automatic	61_fortwo	90000	6	gasoline	smart	0	10.12.2015 00:00	6	10155	6.4.2016 15			
30	14.3.2016	19 Volkswagen_Golf_1_2_TDI_100kW_16V	private	offer	2015	compact	2015	manual	101_golf	60000	1	gasoline	volkswagen	0	10.12.2015 00:00	1	2291	14.3.2016 16			
31	3.4.2016	257 Audi_80_1_90_TDI_90_PS_Avant_Euro_2_1	private	offer	750	station wagon	1993	manual	90	80	150000	11	diesel	audi	0	2.11.2016 00:00	3	91437	5.4.2016 19		
32	14.3.2016	155 Audi_80_1_90_TDI_90_PS_Avant_Euro_2_1	private	offer	750	station wagon	1993	manual	90	80	150000	11	diesel	audi	0	2.11.2016 00:00	7	92648	7.4.2016 18		
33	6.3.2016	1748 Opel_Mokassa_Altan_city_Vier_Fuehrersitze	private	offer	5500	sunder	2009	automatic	0	40000	9	diesel	somitag	0	2.11.2016 00:00	6	31071	6.4.2016 13			
34	14.3.2016	181 Audi_80_1_90_TDI_90_PS_Avant_Euro_2_1	private	offer	1400	station wagon	1993	manual	90	80	150000	12	diesel	audi	0	2.11.2016 00:00	1	91479	5.4.2016 19		
35	5.3.2016	1753 Opelgolddot_V8_R32_zu_Hand	private	offer	14499	2006	2006	manual	250	90000	9	diesel	volkswagen	0	3.11.2016 00:00	0	74955	7.4.2016 14			
36	5.3.2016	1844 VW_Polo_1_2_TSI_Style_Model_Panorama	private	offer	8000	compact	2012	manual	105_polo	100000	2	gasoline	volkswagen	0	6.11.2016 00:00	3	40432	19.3.2016 9			
37	5.3.2016	1645 Opel_Frontera_A_Sport_2_0	private	offer	1099	suvs	1997	manual	116_andere	150000	6	gasoline	opel	1	7.1.2016 00:00	13	52441	15.3.2016 2			
38	5.3.2016	1441 Opel_Insignia_1_0_cdti_Sport_ecof	private	offer	1999	station wagon	2012	automatic	700	15000	1	gasoline	opel	1	7.1.2016 00:00	11	46211	17.3.2016 18			
39	16.3.2016	1544 Fiat_500_C_1_2_Pop	private	offer	7950	cabrio	2010	manual	69	500	80000	3	gasoline	fiat	0	7.21.2016 00:00	9	38154	7.4.2016 14		
	cars4g																				



U tablici ispod opisani su sadržani atributi unutar csv datoteke.

Tablica 1. Atributi korištenog skupa

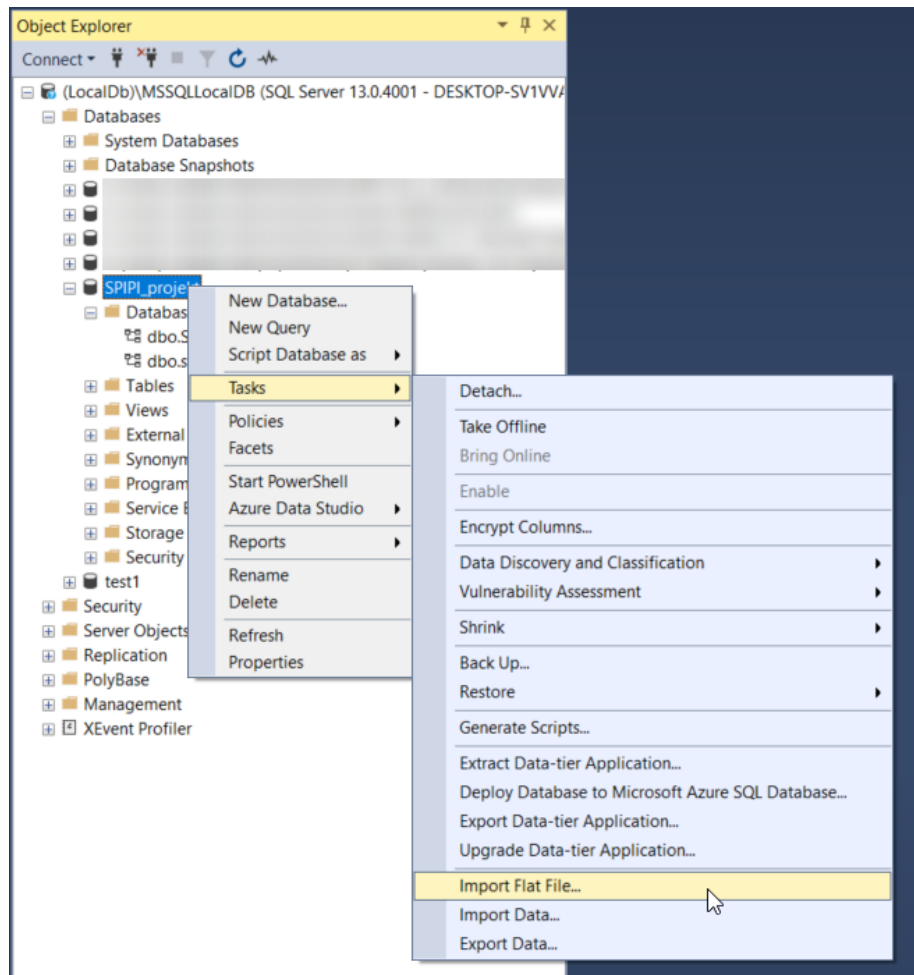
Atribut	Opis
csvID	jedinstveni identifikator
dateCrawled	datum dohvata oglasa skriptom
name	naziv oglasa
seller	tip prodavača – svi su privatni, tako da nam je ovaj atribut praktički beskoristan
offerType	tip ponude – sve su za prodaju, tako da nam je ovaj atribut praktički beskoristan
price	cijena auta u eurima
vehicleType	tip karoserije vozila
yearOfRegistration	godina prve registracije
gearbox	vrsta mjenjača
powerPS	snaga auta u konjskim snagama <sup>1</sup> (njem. <i>PferdeStärke</i> )
kilometer	kilometraža auta
monthOfRegistration	mjesec prve registracije
fuelType	vrsta goriva
brand	proizvođač
notRepairedDamage	oštećenja na autu koja nisu popravljena (0 znači da nema takvih oštećenja, odnosno sve je popravljeno, a 1 označava suprotno)
dateCreated	datum kreiranja oglasa
nrOfPictures	broj slika dostupnih u oglasu
postalCode	njemački poštanski brojevi
lastSeenOnline	datum kada je oglas zadnji put viđen online putem skripte

---

<sup>1</sup> 1 HP (eng. *horsepower*) = 1.0142777265087 PS (njem. *PferdeStärke*)

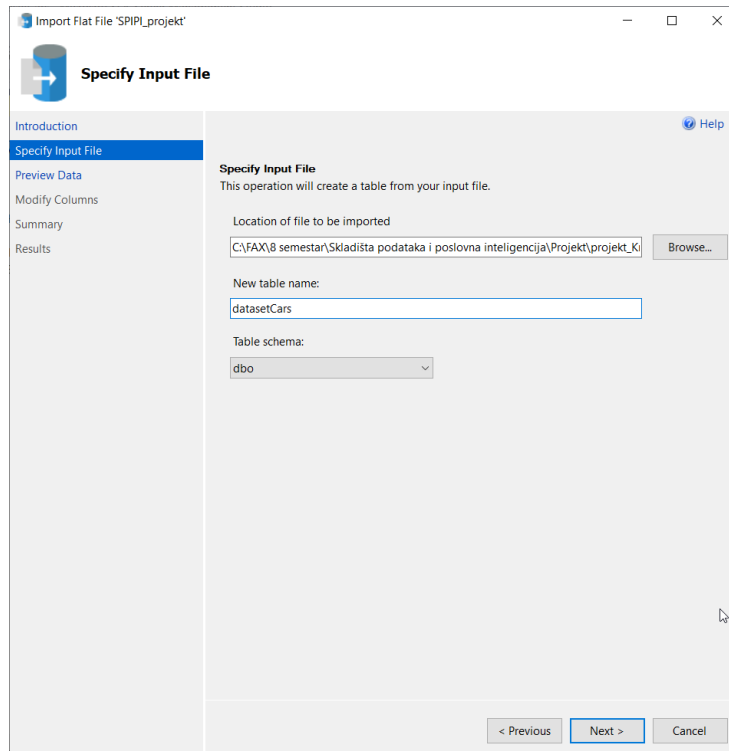
## 5. Opis provedenog ETL procesa

Nakon modifikiranja csv datoteke spremni smo za učitavanje u SQL Server. Kreirana je baza **SPIPI\_projekt**. Pritiskom desnog klika na bazu i odabira *Tasks->Import Flat File* učitavamo željenu datoteku.



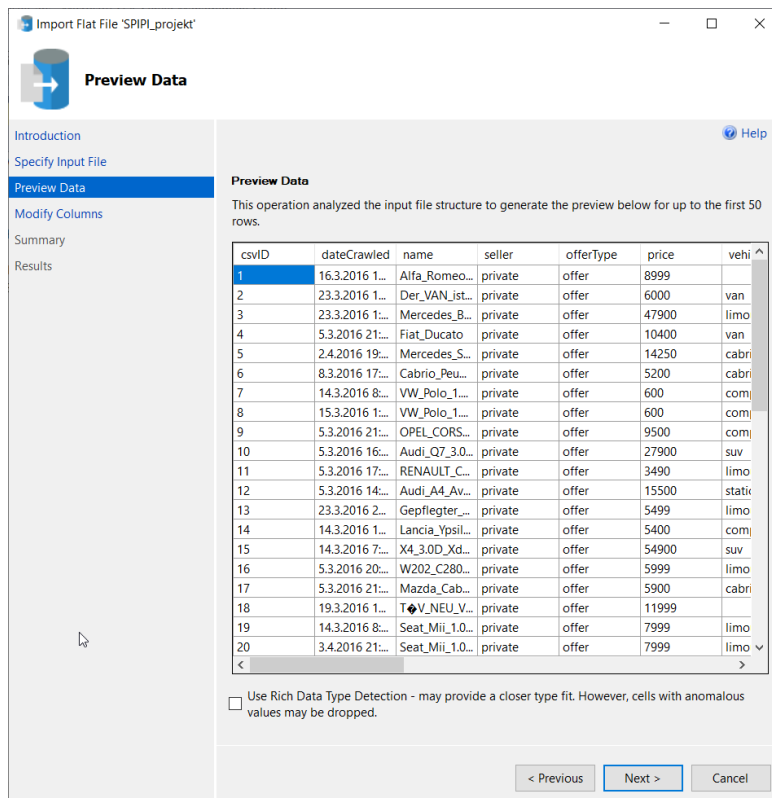
Slika 4. Postupak odabira opcije učitavanja datoteke

Zatim nam se otvara prozor kao na slici 5. gdje odabiremo putanju datoteke i imenujemo tablicu koja je u ovome slučaju *datasetCars*.



Slika 5. Odabir putanje do datoteke

SSMS nudi pregled (eng. *preview*) podataka u idućem prozoru koji se otvara. Vidljivo na slici ispod.



Slika 6. 'Preview' učitanih podataka

DESKTOP-SV1VVAB\...- dbo.datasetCars		DESKTOP-SV1VVAB\...- dbo.DimGearbox	
	Column Name	Data Type	Allow Nulls
▼	csvID	int	<input type="checkbox"/>
	dateCrawled	datetime2(7)	<input checked="" type="checkbox"/>
	name	nvarchar(100)	<input checked="" type="checkbox"/>
	seller	nvarchar(50)	<input checked="" type="checkbox"/>
	offerType	nvarchar(50)	<input checked="" type="checkbox"/>
	price	int	<input checked="" type="checkbox"/>
	vehicleType	nvarchar(50)	<input checked="" type="checkbox"/>
	yearOfRegistration	int	<input checked="" type="checkbox"/>
	gearbox	nvarchar(50)	<input checked="" type="checkbox"/>
	powerPS	int	<input checked="" type="checkbox"/>
	model	nvarchar(50)	<input checked="" type="checkbox"/>
	kilometer	int	<input checked="" type="checkbox"/>
	monthOfRegistration	int	<input checked="" type="checkbox"/>
	fuelType	nvarchar(50)	<input checked="" type="checkbox"/>
	brand	nvarchar(50)	<input checked="" type="checkbox"/>
	notRepairedDamage	char(1)	<input checked="" type="checkbox"/>
	dateCreated	datetime2(7)	<input checked="" type="checkbox"/>
	nrOfPictures	int	<input checked="" type="checkbox"/>
	postalCode	int	<input checked="" type="checkbox"/>
	lastSeen	datetime2(7)	<input checked="" type="checkbox"/>

Slika 7. Dizajn tablice datasetCars

Podaci su sada učitani i kreirana je tablica *datasetCars*. Izgled prikazan na slici 7.

csvID	dateCrawled	name	seller	offerType	price	vehicleType	yearOfRegistration	gearbox	powerPS	model	kilometer	monthOfRegistration	fuelType	brand	notRepairedDamage	dateCreated
1	2016-03-16 11:48:00.0000000	Alfa_Romeo_147_GTA	private	offer	8999	NULL	2005	NULL	0	147	125000	0	gasoline	alfa_romeo	0	2014-03-1
2	2016-03-23 14:42:00.0000000	Der_VAN_ist_zu_OSTER...	private	offer	6000	van	1995	automatic	200	andere	150000	0	gasoline	chevrolet	0	2015-03-2
3	2016-03-23 01:48:00.0000000	Mercedes_Benz_C220_B...	private	offer	47900	limousine	2014	automatic	0	c_klasse	20000	2	diesel	mercedes_b...	0	2015-06-1
4	2016-03-05 21:44:00.0000000	Fiat_Ducato	private	offer	10400	van	1993	NULL	0	ducatto	30000	1	diesel	fiat	0	2015-06-1
5	2016-04-02 19:47:00.0000000	Mercedes_SLK_200_AM...	private	offer	14250	cabrio	2005	manual	163	slk	150000	6	gasoline	mercedes_b...	0	2015-08-4
6	2016-03-08 17:57:00.0000000	Cabrio_Peugeot_206_CC...	private	offer	5200	cabrio	2006	manual	109	2_reihe	80000	11	gasoline	peugeot	0	2015-08-1
7	2016-03-14 08:43:00.0000000	VW_Polo_1_2_Cricket_2...	private	offer	600	compact	2005	manual	54	polo	150000	1	gasoline	volkswagen	1	2015-09-4
8	2016-03-15 01:56:00.0000000	VW_Polo_1_2_Cricket_2...	private	offer	600	compact	2005	manual	54	polo	150000	1	gasoline	volkswagen	1	2015-09-4
9	2016-03-05 21:51:00.0000000	OPEL_CORSA_Apfelgru...	private	offer	9500	compact	2013	manual	90	corsa	50000	4	diesel	opel	0	2015-09-4
10	2016-03-05 16:55:00.0000000	Audi_Q7_3.0_TDI_7_Sit...	private	offer	27900	suv	2009	automatic	239	q7	125000	10	diesel	audi	0	2015-10-1
11	2016-03-05 17:53:00.0000000	RENAULT_CLIO_1.6_EDL...	private	offer	3490	limousine	2006	manual	88	clio	150000	3	gasoline	renault	0	2015-11-4
12	2016-03-05 14:31:00.0000000	Audi_A4_Avant_1.8_TFSI...	private	offer	15500	station wag...	2009	manual	160	a4	100000	11	gasoline	audi	0	2015-11-4
13	2016-03-23 23:37:00.0000000	Gepfliegter_Peugeot_20...	private	offer	5499	limousine	2009	manual	95	2_reihe	70000	3	gasoline	peugeot	0	2015-11-1
14	2016-03-14 18:44:00.0000000	Lancia_Ypsilon_1.2_8_V...	private	offer	5400	compact	2009	manual	44	ypsilon	40000	9	gasoline	lancia	0	2015-11-1
15	2016-03-14 07:45:00.0000000	X4_3.0D_xdrive_AHK...	private	offer	54900	suv	2015	automatic	258	x_reihe	10000	3	diesel	bmw	0	2015-11-1
16	2016-03-05 20:42:00.0000000	W202_C280_V6_schmi...	private	offer	5999	limousine	1997	automatic	197	c_klasse	150000	9	gasoline	mercedes_b...	0	2015-11-1
17	2016-03-05 21:49:00.0000000	Mazda_Cabrio_MX5_NA...	private	offer	5900	cabrio	1995	manual	90	mx_reihe	125000	5	gasoline	mazda	0	2015-11-2

Slika 8. Kreirana tablica *datasetCars*

U tablicu sam naknadno dodao stupce s jedinstvenim identifikatorima za svaku dimenziju, koje sam zatim popunio sa upitima na slici 9. Na slici 29. vidimo sam izgled tablice činjenične tablice koja je zapravo tablica početnog skupa podataka bez obrisanih stupaca.

```
SQLQuery1.sql - (Lo...B\KristijanPC (53))*  DESKTOP-SV1VVAB\L...t1 - dbo.FactCars  DESKTOI

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_Advertisement = b.DimAdvertisementID
FROM dbo.datasetCars a
INNER JOIN dbo.DimAdvertisement b ON a.name = b.title;

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_Brand = b.DimBrandID
FROM dbo.datasetCars a
INNER JOIN dbo.DimBrand b ON a.brand = b.brand;

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_Date = b.DimDateID
FROM dbo.datasetCars a
INNER JOIN dbo.DimDate b ON a.dateCreated = b.date;

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_FuelType = b.DimFuelTypeID
FROM dbo.datasetCars a
INNER JOIN dbo.DimFuelType b ON a.fuelType = b.fuelType;

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_Gearbox = b.DimGearboxID
FROM dbo.datasetCars a
INNER JOIN dbo.DimGearbox b ON a.gearbox = b.gearbox;

UPDATE dbo.datasetCars
SET dbo.datasetCars.ID_Model = b.DimModelID
FROM dbo.datasetCars a
INNER JOIN dbo.DimModel b ON a.model = b.model;
```

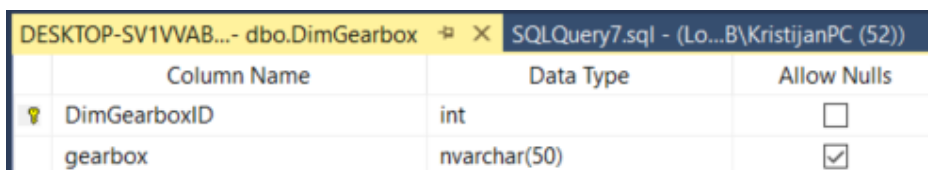
Slika 9. Upiti za popunjavanje datasetCars tablice

## 6. Opis modela izrađenog skladišta podataka

U ovome poglavlju reći ćemo nešto više o samom modeliranju skladišta. Za potrebe projekta korišten je takozvani model zvijezde (eng. *star model*) pošto je ovo relacijsko skladište podataka. Ovaj često korišten model zahtijeva klasifikaciju tablica na dimenzijske i činjenične tablice. Dimenzijske tablice opisuju poslovne subjekte i odgovaraju na pitanje prema čemu mjerimo. Najčešće korištena dimenzijska tablica je datum koju sam i ja koristio, a čije kreiranje će biti opisano malo kasnije. Dimenzijske tablice sadrže atribut jedinstvenog identifikatora, čijom kombinacijom obično kreiramo složeni primarni ključ činjenične tablice, no sve je stvar dizajna [4].

### 6.1. Dimenzijska tablica 'DimGearbox'

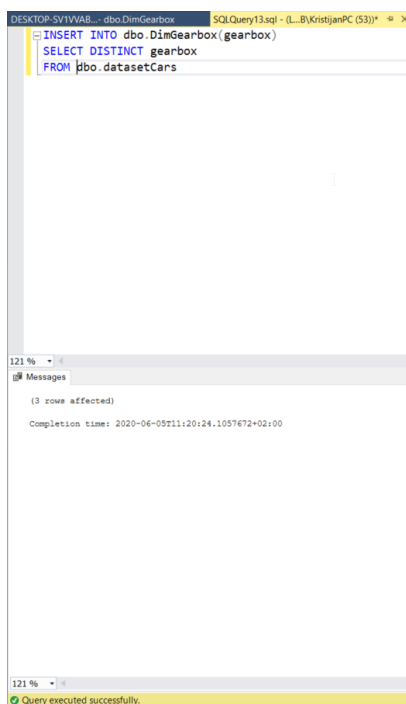
Tablica sadrži samo dva atributa, a to su DimGearBoxID tipa int koji je ujedno i primarni ključ tablice, te gearbox tipa nvarchar koji sadrži tip mjenjača brzina.



Column Name	Data Type	Allow Nulls
DimGearboxID	int	<input type="checkbox"/>
gearbox	nvarchar(50)	<input checked="" type="checkbox"/>

Slika 10. Struktura tablice DimGearbox

Upit s kojim je popunjena tablica prikazan je na slici ispod.



Slika 11. Upit za unos podataka u tablicu DimGearbox

DESKTOP-SV1VVAB...- dbo.DimGearbox		
	DimGearboxID	gearbox
▶	1	automatic
	2	NULL
	3	manual

Slika 12. Tablica DimGearbox

## 6.2. Dimenzijska tablica DimFuelType

Ova dimenzijska tablica također sadrži samo dva atributa, a to su DimFuelTypeID tipa int koji je ujedno i primarni ključ tablice, te fuelType tipa nvarchar koji označava vrstu goriva.

DESKTOP-SV1VVAB\...dbo.DimFuelType			
	Column Name	Data Type	Allow Nulls
▶	DimFuelTypeID	int	<input type="checkbox"/>
	fuelType	nvarchar(50)	<input checked="" type="checkbox"/>

Slika 13. Struktura tablice DimFuelType

Upit s kojim je popunjena tablica prikazan je na slici ispod.

```

INSERT INTO dbo.DimFuelType(fuelType)
SELECT DISTINCT fuelType
FROM dbo.datasetCars;

```

121 %

Messages

(8 rows affected)

Completion time: 2020-06-05T11:32:42.2570926+02:00

121 %

Query executed successfully.

Slika 14. Upit za unos podataka u tablicu DimFuelType

DESKTOP-SV1VVAB\...dbo.DimFuelType		
	DimFuelTy...	fuelType
▶	1	hybrid
	2	cng
	3	other
	4	NULL
	5	diesel
	6	electric
	7	gasoline
	8	lpg

Slika 15. Tablica DimFuelType

### 6.3. Dimenzijska tablica 'DimBrand'

Tablica sadrži tri atributa, a to su DimCarsID tipa int koji je ujedno i primarni ključ tablice, brand tipa nvarchar koji predstavlja proizvođača, te model tipa nvarchar koji predstavlja model automobila.

DESKTOP-SV1VVAB\...1 - dbo.DimBrand		SQLQuery25.sql - (L...B\KristijanPC (57))	
	Column Name	Data Type	Allow Nulls
▶	DimBrandID	int	<input type="checkbox"/>
	brand	nvarchar(50)	<input checked="" type="checkbox"/>

Slika 16. Struktura tablice DimBrand

Upit s kojim je popunjena tablica prikazan je na slici ispod.

```

SQLQuery1.sql - (Lo...B\KristijanPC (53))*
INSERT INTO dbo.DimBrand
SELECT DISTINCT brand
FROM dbo.datasetCars;

```

121 %

Messages

(371822 rows affected)

Slika 17. Upit za unos podataka u tablicu DimBrand



DESKTOP-SV1VVAB\...1 - dbo.DimBrand		
	DimBrandID	brand
▶	339	rover
	340	honda
	341	dacia
	342	jeep
	343	volvo
	344	subaru
	345	alfa_romeo
	346	suzuki
	347	saab
	348	renault
	349	seat
	350	1guar
	351	mitsubishi
	352	volkswagen
	353	trabant
	354	land_rover
	355	bmw
	356	kia
	357	nissan
	358	mercedes_b...
	359	chevrolet
	360	smart
	361	sonstige_au...
	362	citroen
	363	mini
	364	porsche
	365	toyota
	366	mazda
	367	fiat
	368	daihatsu
	369	opel
	370	skoda
	371	hyundai
	372	chrysler
	373	daewoo
	374	lada
	375	ford
	376	lancia
	377	audi
	378	peugeot

Slika 18. Tablica DimBrand

## 6.4. Dimenzijska tablica 'DimModel'

Tablica sadrži tri atributa, a to su DimCarsID tipa int koji je ujedno i primarni ključ tablice, brand tipa nvarchar koji predstavlja proizvođača, te model tipa nvarchar koji predstavlja model automobila.

DESKTOP-SV1VVAB\...1 - dbo.DimBrand		DESKTOP-SV1VVAB\...1 - dbo.DimModel	
	Column Name	Data Type	Allow Nulls
	DimModelID	int	<input type="checkbox"/>
	model	nvarchar(50)	<input checked="" type="checkbox"/>

Slika 19. Struktura tablice DimModel

Upit s kojim je popunjena tablica prikazan je na slici ispod.

```
SQLQuery1.sql - (Lo...B\KristijanPC (53))* X DESKTOP-SV1VVAB
INSERT INTO dbo.DimModel (model)
SELECT DISTINCT model
FROM dbo.datasetCars;
```

121 %

Messages

(371822 rows affected)

Completion time: 2020-06-05T15:13:47.4648889+02:00

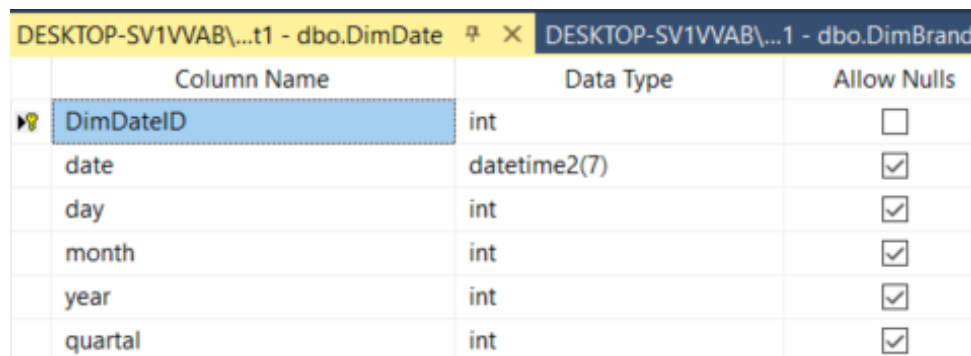
Slika 20. Upit za unos podataka u tablicu DimModel

DESKTOP-SV1VVAB\...1 - dbo.DimModel		
	DimModelID	model
	1	fabia
	2	kappa
	3	145
	4	ptcruiser
	5	discovery
	6	tigra
	7	kuga
	8	zafira
	9	freelander
	10	lodgy
	11	serie_3
	12	a3
	13	calibra
	14	legacy
	15	159
	16	touran
	17	v40
	18	clio
	19	almera
	20	cordoba
	21	c5
	22	sportage
	23	cl

Slika 21. Tablica DimModel

## 6.5. Dimenzijska tablica 'DimDate'

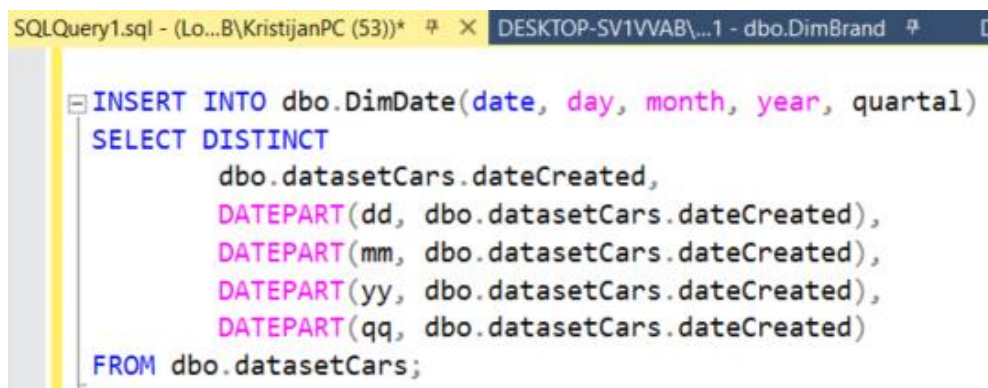
Tablica sadrži šest atributa, a to su DimDateID tipa int koji je ujedno i primarni ključ tablice, brand tipa nvarchar koji predstavlja proizvođača, te model tipa nvarchar koji predstavlja model automobila.



Column Name	Data Type	Allow Nulls
DimDateID	int	<input type="checkbox"/>
date	datetime2(7)	<input checked="" type="checkbox"/>
day	int	<input checked="" type="checkbox"/>
month	int	<input checked="" type="checkbox"/>
year	int	<input checked="" type="checkbox"/>
quartal	int	<input checked="" type="checkbox"/>

Slika 22. Struktura tablice DimDate

Upit s kojim je popunjena tablica prikazan je na slici ispod.



```
INSERT INTO dbo.DimDate(date, day, month, year, Quartal)
SELECT DISTINCT
    dbo.datasetCars.dateCreated,
    DATEPART(dd, dbo.datasetCars.dateCreated),
    DATEPART(mm, dbo.datasetCars.dateCreated),
    DATEPART(yy, dbo.datasetCars.dateCreated),
    DATEPART(qq, dbo.datasetCars.dateCreated)
FROM dbo.datasetCars;
```

Slika 23. Upit za unos podataka u tablicu DimDate

DESKTOP-SV1VVAB\...t1 - dbo.DimDate						
SQLQuery1.sql - (Lo...B\KristijanPC (53))*						
	DimDateID	date	day	month	year	quartal
1		2016-04-02 ...	2	4	2016	2
2		2016-03-27 ...	27	3	2016	1
3		2016-03-05 ...	5	3	2016	1
4		2016-02-27 ...	27	2	2016	1
5		2016-01-07 ...	7	1	2016	1
6		2015-03-20 ...	20	3	2015	1
7		2016-01-10 ...	10	1	2016	1
8		2016-03-21 ...	21	3	2016	1
9		2015-11-17 ...	17	11	2015	4
10		2016-03-12 ...	12	3	2016	1
11		2016-01-27 ...	27	1	2016	1
12		2016-03-22 ...	22	3	2016	1
13		2016-04-07 ...	7	4	2016	2
14		2016-03-01 ...	1	3	2016	1
15		2016-01-31 ...	31	1	2016	1
16		2016-03-26 ...	26	3	2016	1
17		2016-04-06 ...	6	4	2016	2
18		2016-02-10 ...	10	2	2016	1
19		2016-01-29 ...	29	1	2016	1
20		2016-03-19 ...	19	3	2016	1
21		2015-11-23 ...	23	11	2015	4
22		2016-02-28 ...	28	2	2016	1
23		2016-02-29 ...	29	2	2016	1
24		2016-02-21 ...	21	2	2016	1
25		2016-04-04 ...	4	4	2016	2
26		2016-02-16 ...	16	2	2016	1
27		2016-03-17 ...	17	3	2016	1
28		2016-03-20 ...	20	3	2016	1
29		2016-01-06 ...	6	1	2016	1
30		2016-02-23 ...	23	2	2016	1
31		2015-11-24 ...	24	11	2015	4
32		2016-02-05 ...	5	2	2016	1
33		2015-11-13 ...	13	11	2015	4
34		2016-01-08 ...	8	1	2016	1
35		2016-02-24 ...	24	2	2016	1
36		2015-12-30 ...	30	12	2015	4
37		2016-01-22 ...	22	1	2016	1
38		2016-03-15 ...	15	3	2016	1
39		2016-02-06 ...	6	2	2016	1
40		2016-03-24 ...	24	3	2016	1

Slika 24. Tablica DimDate

## 6.6. Dimenzijska tablica 'DimAdvertisement'

Tablica sadrži pet atributa, a to su DimDateID tipa int koji je ujedno i primarni ključ tablice, brand tipa nvarchar koji predstavlja proizvođača, te model tipa nvarchar koji predstavlja model automobila.

SQLQuery13.sql - (L...B\KristijanPC (53))*				DESKTOP-SV1VVAB...DimAdvertisement	
	Column Name	Data Type	Allow Nulls		
1	DimAdvertisementID	int	<input type="checkbox"/>		
2	title	nvarchar(100)	<input checked="" type="checkbox"/>		
3	dateCreated	datetime2(7)	<input checked="" type="checkbox"/>		
4	lastSeenOnline	datetime2(7)	<input checked="" type="checkbox"/>		
5	postalCode	int	<input checked="" type="checkbox"/>		
6	nrOfPictures	int	<input checked="" type="checkbox"/>		
7	seller	nvarchar(50)	<input checked="" type="checkbox"/>		
8	offerType	nvarchar(50)	<input checked="" type="checkbox"/>		

Slika 25. Struktura tablice DimAdvertisement

Upit s kojim je popunjena tablica prikazan je na slici ispod. Tablica sadrži preko 370000 zapisa.

SQLQuery25.sql - (L...B\KristijanPC (55))\*

```

INSERT INTO dbo.DimAdvertisement(title, dateCreated, lastSeenOnline, postalCode, nrOfPictures, seller, offerType)
SELECT name, dateCreated, lastSeen, postalCode, nrOfPictures, seller, offerType
FROM dbo.datasetCars;

```

100 %

Messages

(371822 rows affected)

Completion time: 2020-06-08T22:04:05.4185214+02:00

100 %

Query executed successfully. (LocalDb)\MSSQLLocalDB (13... | DESKTOP-SV1V...

Slika 26. Upit za unos podataka u tablicu DimAdvertisement

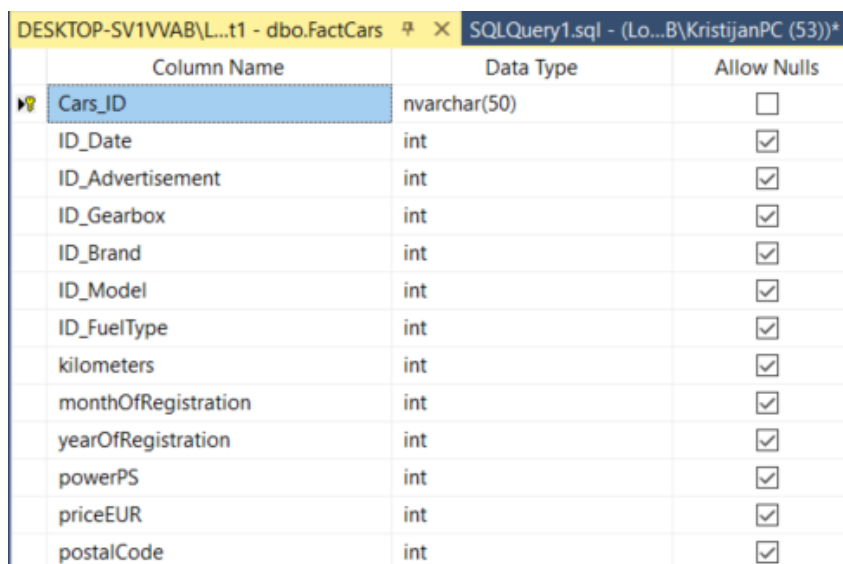
DimAdvertisementID	title	dateCreated	lastSeenOnline	postalCode	nrOfPictures	seller	offerType
1	Alfa_Romeo_147_GTA	2014-03-10 00:00:00.0000000	2016-04-07 04:46:00.0000000	78669	9	private	offer
2	Der_VAN_at_au_OSTERN_reisefertig	2015-03-20 00:00:00.0000000	2016-04-07 03:45:00.0000000	24558	8	private	offer
3	Mercedes_Benz_C220_BT_7Q_Tronic_AMG_Modell1hr...	2015-06-11 00:00:00.0000000	2016-04-06 09:47:00.0000000	46145	2	private	offer
4	Fiat_Ducato	2015-06-18 00:00:00.0000000	2016-04-05 17:46:00.0000000	76275	3	private	offer
5	Mercedes_SLK_200_AMG_Performance_19_Zoll__absolut...	2015-08-07 00:00:00.0000000	2016-04-07 09:15:00.0000000	76228	12	private	offer
6	Cabrio_Peugeot_206_CC_Platinum_zum_Winterpreis!	2015-08-10 00:00:00.0000000	2016-04-05 20:46:00.0000000	26382	13	private	offer
7	VW_Polo_1.2_Cricket_2005_mit_Tauschmotor_54TKM_8...	2015-09-04 00:00:00.0000000	2016-03-15 14:16:00.0000000	72669	14	private	offer
8	VW_Polo_1.2_Cricket_2005_mit_Tauschmotor_54TKM_Ki...	2015-09-04 00:00:00.0000000	2016-03-15 15:49:00.0000000	72669	7	private	offer
9	OPEL_CORSA_Apfelgruen_dringend_zu_verkaufen!	2015-09-09 00:00:00.0000000	2016-04-05 17:46:00.0000000	98739	5	private	offer
10	Audi_07_3.0_TDI_7_Sitzer_Vollausstattung_Standheizung	2015-10-14 00:00:00.0000000	2016-03-14 15:17:00.0000000	44651	9	private	offer
11	RENAULT_OJIO_1.6_EDITION_DYNAMIQUE	2015-11-02 00:00:00.0000000	2016-03-17 15:46:00.0000000	40890	12	private	offer
12	Audi_A4_Avant_1.8_TFSI_Ambition__neuer_Motor_bei_81.1...	2015-11-08 00:00:00.0000000	2016-04-06 23:15:00.0000000	56410	0	private	offer
13	Gefegter_Peugeot_207_Sport_VTI_Sport	2015-11-10 00:00:00.0000000	2016-04-07 09:15:00.0000000	22941	1	private	offer
14	Lancia_Ypsilon_1.2_8_V_vie_neu_TW_09/2016	2015-11-12 00:00:00.0000000	2016-04-06 03:16:00.0000000	53119	7	private	offer
15	X4_3.0D_Xdrive_AHK_Navi_Xenon_Leder_PDC_u_Kamera...	2015-11-13 00:00:00.0000000	2016-04-05 20:16:00.0000000	85368	8	private	offer
16	VW202_C280_V6_schmidt_felgen__amg_ATP_sport_H&R...	2015-11-17 00:00:00.0000000	2016-04-06 05:46:00.0000000	28816	4	private	offer
17	Mazda_Cabrio_MX5_NA_BJ_95_90_PS_Schiffslaugen...	2015-11-23 00:00:00.0000000	2016-03-23 07:18:00.0000000	63741	0	private	offer

Query executed successfully. (LocalDb)\MSSQLLocalDB (13... | DESKTOP-SV1VVAB\Kristi... test1 00:00:03 371.822 rows

Slika 27. Tablica DimAdvertisement

## 6.7. Činjenična tablica 'FactCars'

Tablica sadrži nekoliko atributa, a to su Cars\_ID tipa nvarchar koji je ujedno i primarni ključ tablice, ID\_Date tipa int koji je vanjski ključ od dimenzijske tablice DimDate, ID\_Advertisement tipa int koji je vanjski ključ od dimenzijske tablice DimAdvertisement, ID\_Gearbox tipa int koji je vanjski ključ od dimenzijske tablice DimGearbox, ID\_Date tipa int koji je vanjski ključ od dimenzijske tablice DimDate, ID\_Brand tipa int koji je vanjski ključ od dimenzijske tablice DimBrand, ID\_Model tipa int koji je vanjski ključ od dimenzijske tablice DomModel, ID\_FuelType tipa int koji je vanjski ključ od dimenzijske tablice DimFuelType, kilometers tipa int koja predstavlja kilometražu auta, monthOfRegistration tipa int koji predstavlja mjesec prve registracije auta, yearOfRegistration tipa int koji predstavlja godinu prve registracije auta, powerPS tipa int koji predstavlja snagu auta u konjskim snagama (njem. *PferdeStärke*), priceEUR tipa int koji je cijena automobila izražena u eurima i postalCode koji predstavlja njemački poštanski broj mjesta u kojemu se automobil s oglasa nalazi.



Column Name	Data Type	Allow Nulls
Cars_ID	nvarchar(50)	<input type="checkbox"/>
ID_Date	int	<input checked="" type="checkbox"/>
ID_Advertisement	int	<input checked="" type="checkbox"/>
ID_Gearbox	int	<input checked="" type="checkbox"/>
ID_Brand	int	<input checked="" type="checkbox"/>
ID_Model	int	<input checked="" type="checkbox"/>
ID_FuelType	int	<input checked="" type="checkbox"/>
kilometers	int	<input checked="" type="checkbox"/>
monthOfRegistration	int	<input checked="" type="checkbox"/>
yearOfRegistration	int	<input checked="" type="checkbox"/>
powerPS	int	<input checked="" type="checkbox"/>
priceEUR	int	<input checked="" type="checkbox"/>
postalCode	int	<input checked="" type="checkbox"/>

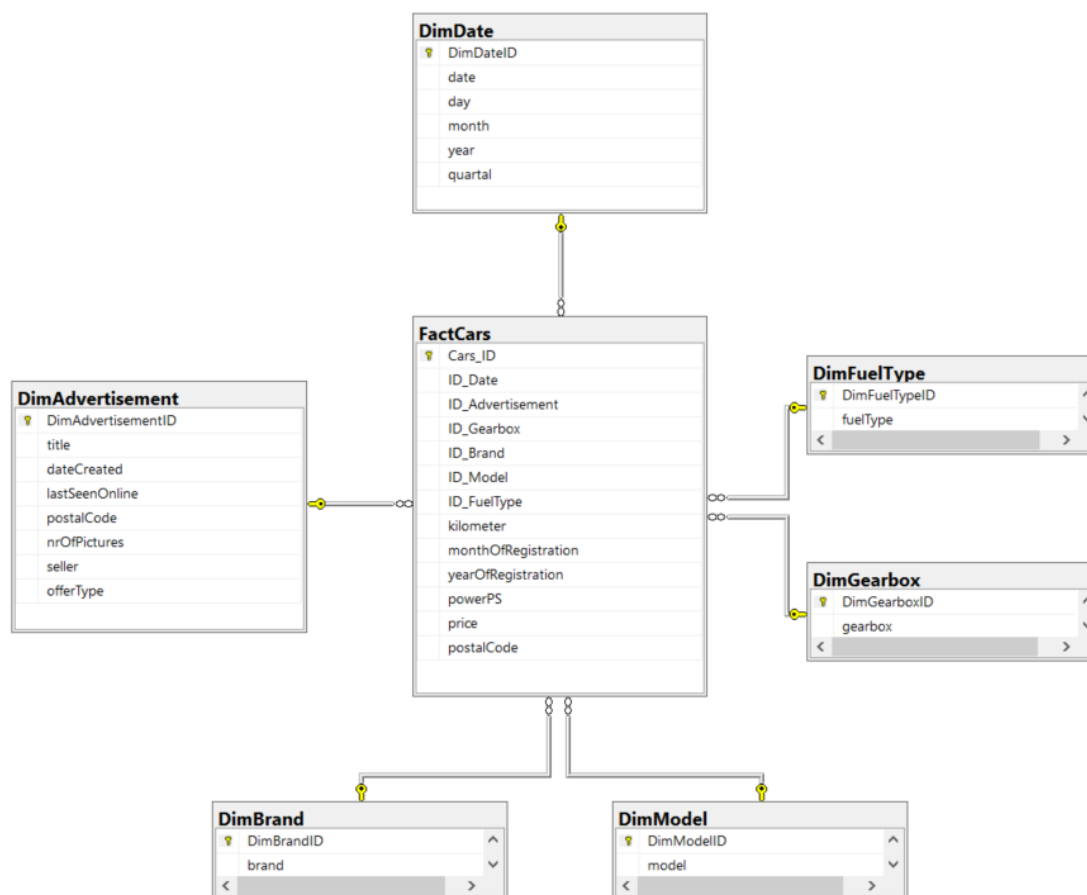
Slika 28. Struktura tablice FactCars

DESKTOP-SV1VVAB\L...t1 - dbo.FactCars												
DESKTOP-SV1VVAB\L...t1 - dbo.FactCars												
SQLQuery1.sql - (L...B\KristijanPC (53))*												
SQLQuery28.sql - (L...B\KristijanPC (56))*												
Cars_ID	ID_Date	ID_Advertis...	ID_Gearbox	ID_Brand	ID_Model	ID_FuelType	kilometers	monthOfRe...	yearOfRegi...	powerPS	priceEUR	postalCode
10	94	1	NULL	345	151	7	125000	0	2005	0	8999	79669
100	15	100	1	352	177	NULL	150000	2	2008	140	6950	37081
1000	64	1000	3	348	18	7	90000	12	2006	111	4150	66802
10000	3	10000	3	352	130	7	150000	4	2004	115	3000	42855
100000	114	100000	3	357	181	7	125000	1	2003	65	2200	45657
100001	114	100001	3	377	NULL	7	150000	2	1964	0	1800	88145
100002	114	100002	NULL	361	NULL	NULL	150000	0	2005	0	1000	42275
100003	114	3637	NULL	369	NULL	NULL	150000	0	2000	0	400	26810
100004	114	100004	3	375	85	5	150000	2	2005	136	2000	33829
100005	114	100005	3	356	38	NULL	150000	8	2004	75	600	63067
100006	114	100006	3	369	87	7	40000	7	1968	145	8500	18233
100007	114	2165	3	357	181	7	100000	10	2000	0	850	41836
100008	114	100008	1	377	61	5	150000	9	2008	190	10400	40468
100009	114	100009	3	369	167	7	150000	12	1999	75	1000	26215
10001	3	6108	3	357	19	7	150000	3	2001	90	1600	17099
100010	114	100010	3	377	133	7	150000	3	1991	130	1	94474
100011	114	15552	NULL	352	240	NULL	150000	0	2000	75	0	14476
100012	114	100012	3	356	NULL	5	150000	9	2006	75	1999	80937
100013	114	100013	3	352	130	7	150000	11	1995	75	350	26340
100014	114	100014	3	352	130	7	150000	3	2017	70	0	12524
100015	114	100015	3	352	177	5	125000	1	2000	100	2222	45134
100016	114	81658	3	355	182	5	150000	0	2002	150	800	66482
100017	114	100017	NULL	375	104	NULL	150000	0	2016	0	100	24211
100018	114	100018	NULL	355	182	NULL	5000	7	1997	0	250	94036
100019	114	100019	3	352	177	5	150000	5	2001	131	2800	59755
10002	3	10002	3	346	224	NULL	80000	2	2009	92	4900	50181
100020	114	79479	3	352	177	5	150000	12	2007	0	4000	90419
100021	114	56096	3	369	208	7	125000	10	2002	58	1600	13357
100022	114	100022	NULL	349	123	NULL	150000	0	2006	0	4750	36124
100023	114	954	3	369	208	7	150000	3	2017	60	600	42857
100024	114	100024	3	375	163	7	150000	6	1999	100	1200	58706
100025	114	100025	3	352	130	7	150000	9	2016	80	7600	41748
100026	114	100026	3	353	180	5	150000	0	1995	70	1200	73461

Slika 29. Tablica FactCars

## 6.8. Model zvijezde

Model zvijezde sastoji se od jedne činjenične tablice i šest dimenzijskih tablica koje su opisane u prethodnim poglavljima.

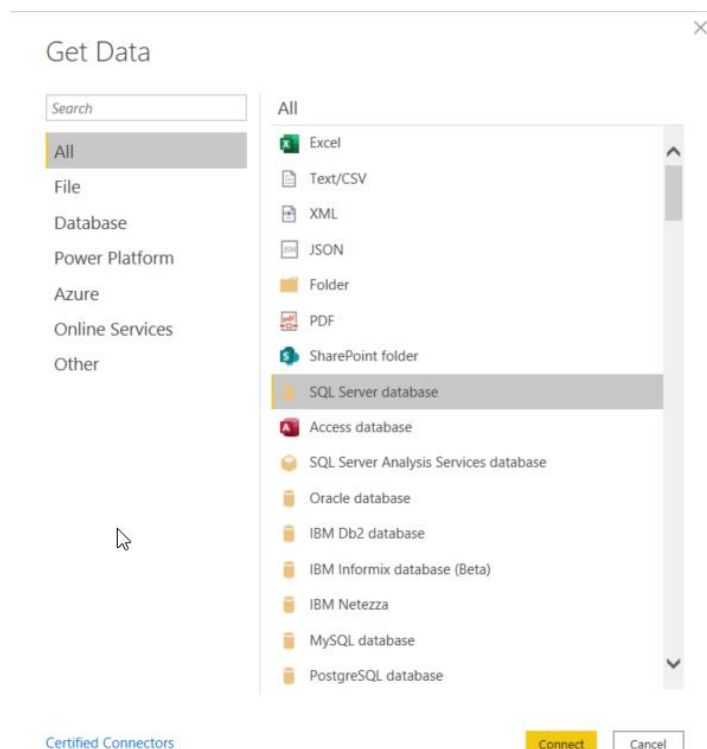


Slika 30. Model zvijezde

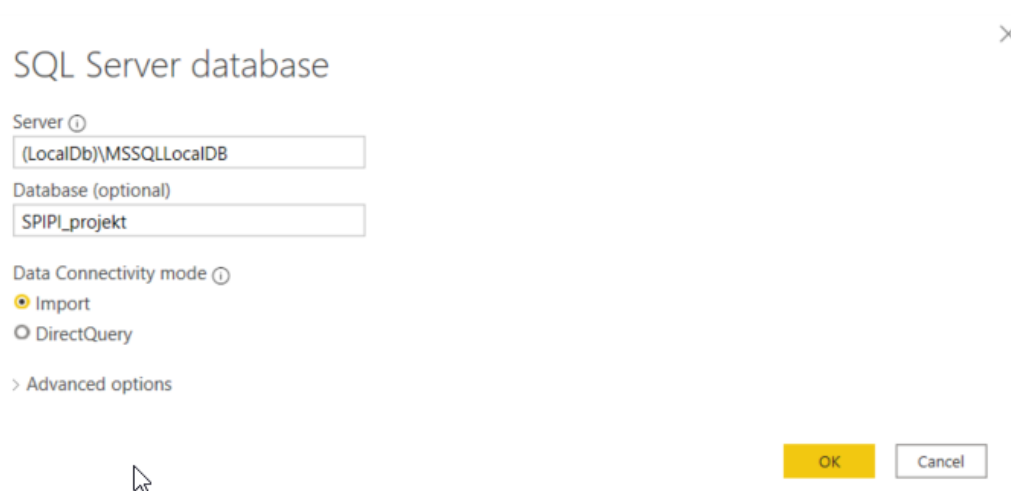


## 7. Prikaz izvještaja napravljenih u Microsoft Power BI alatu

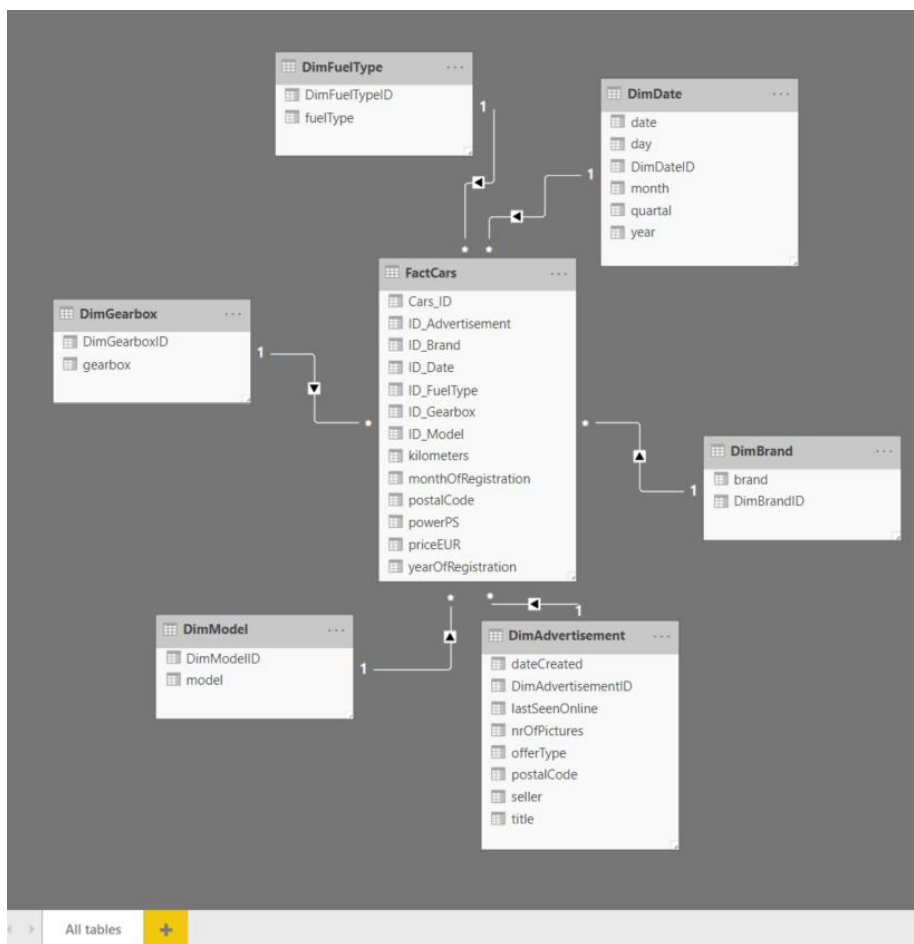
U ovom poglavlju prikazat ću izvještaje i postupak izrade istih. Power BI omogućuje nekolicinu načina dohvaćanja podataka kao što možemo vidjeti na slici 31. Odabiremo *SQL Server database* i otvara nam se prozor kao na slici 32. Unesemo podatke o serveru i samoj bazi podataka na koju se spajamo.



Slika 31. Opcije unosa podataka - Power BI



Slika 32. Spajanje na SQL Server bazu podataka



Slika 33. Model zvijezde unutar Power BI alata

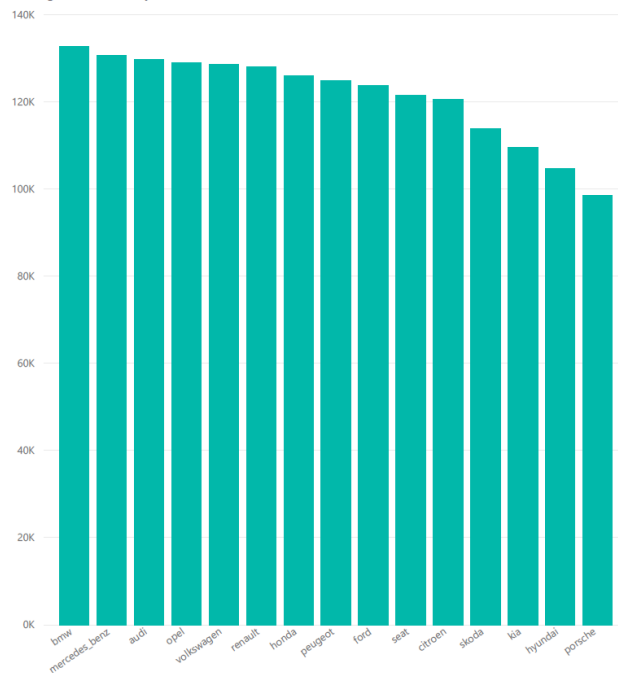
Na slici 33. vidljive su učitane tablice iz kreirane baze podataka koje su zapravo naš *Star model*. Na slikama ispod vidimo tri izvještaja koje sam kreirao, a svaki od njih ima nekoliko grafičkih vizualizacija.

Na prvome izvještaju vidimo model i cijenu najskupljeg automobila koji je naveden u oglasniku, vidimo prosječnu kilometražu automobila prema proizvođaču i vidimo prosječnu jačinu izraženu u konjskim snagama, također prema proizvođaču.

Na drugome izvještaju imamo tri grafa, u gornjem lijevom kutu vidljiva je prosječna automobila prema tipu goriva i godini, zatim prikaz najpopularnijih modela prema odabranim proizvođačima iz legende grafa, te graf s prosječnim vrijednostima cijene i kilometraže.

Na trećem izvještaju najstariju godinu prve registracije, prosječnu kilometražu svih automobila s oglasnika, odnos automatskih i ručnih mjenjača, te broj dostupnih automobila prema tipu goriva.

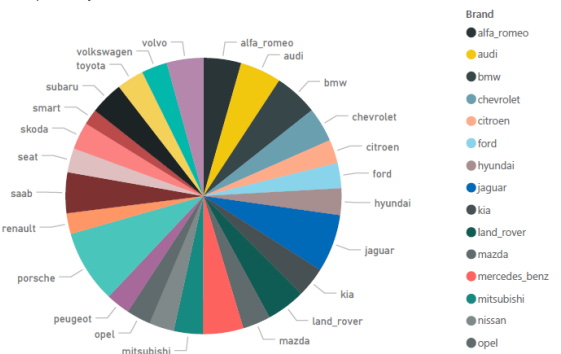
Average kilometers by brand



Most expensive used car available

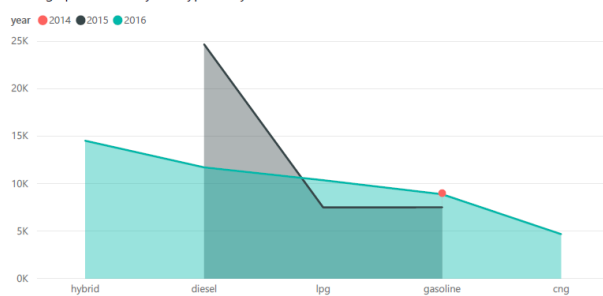


Average horse power by brand

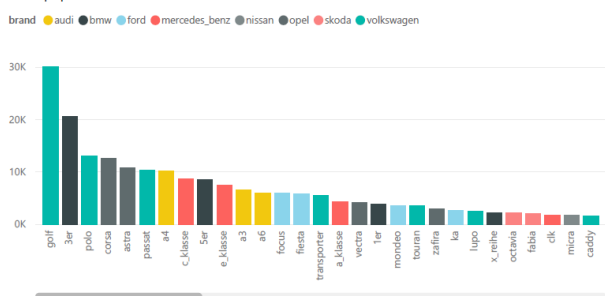


Slika 34. Prvi izvještaj

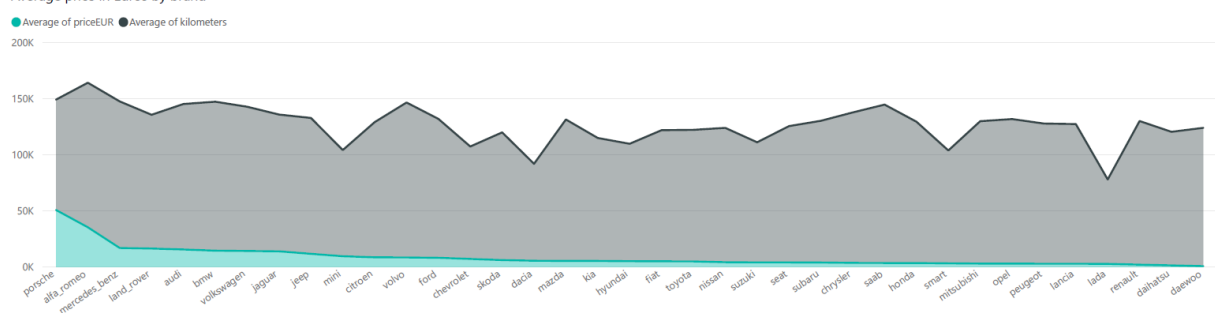
Average price of car by fuel type and year



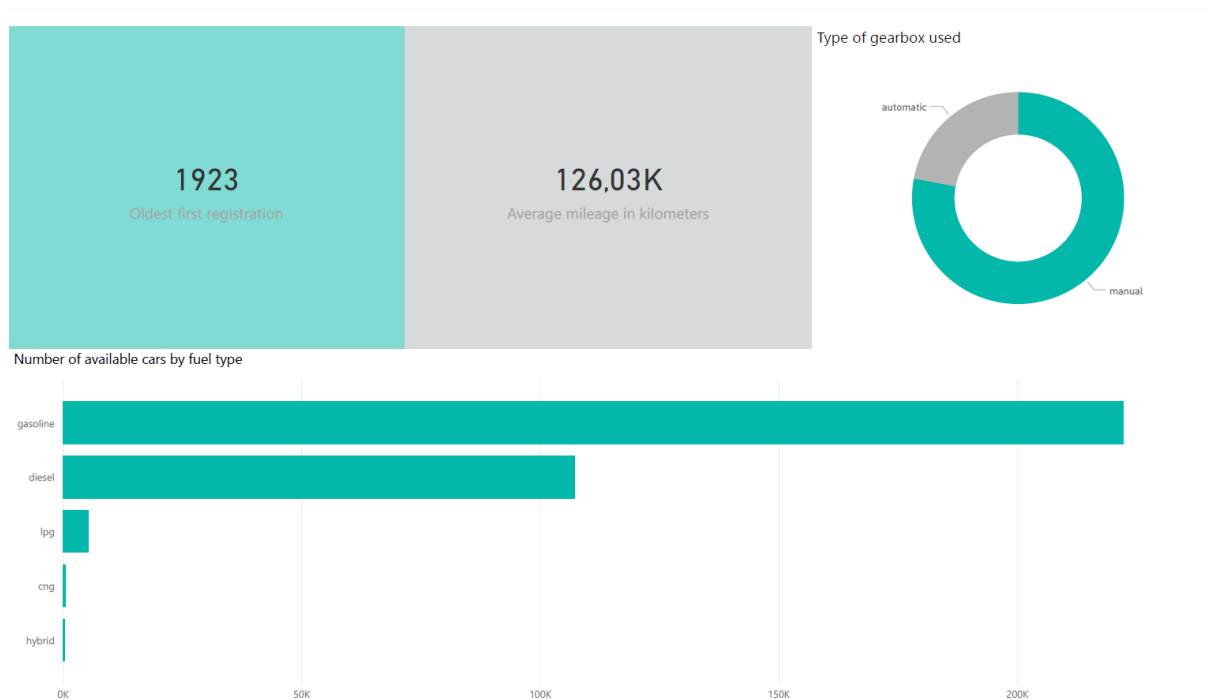
Most popular models



Average price in Euros by brand



Slika 35. Drugi izvještaj



Slika 36. Treći izvještaj

## 8. Zaključak

Kroz seminarski rad opisana je izrada i implementacija skladišta podataka vezanog za analiziranje rabljenih automobila, kao i izvještaja kreiranih na temelju dostupnih podataka koji su korisni za analizu istih. Rad u Microsoft SQL Server Management Studio-u i Power BI-u je bio vrlo zanimljiv, te koristan za znati. Radom na ovom projektu susreo sam se s alatima koje sam prije koristio, no isto tako susreo sam se s alatom koji nisam do sada koristio, Power BI, a svakako ću ga početi implementirati sve više u daljnjem radu i preporučam ga svakome tko je željan vizualno imerzivnih i interaktivnih uvida u podatke.

## Popis literature

- [1] „Data Visualization | Microsoft Power BI“. [Na internetu]. Dostupno na: <https://powerbi.microsoft.com/en-us/>. [Pristupljeno: 04-lip-2020].
- [2] „What is Power BI? - Power BI | Microsoft Docs“. [Na internetu]. Dostupno na: <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>. [Pristupljeno: 04-lip-2020].
- [3] O. Leka, „orgesleka/used-cars-scrapy-script: <https://www.kaggle.com/orgesleka/used-cars-database>“. [Na internetu]. Dostupno na: <https://github.com/orgesleka/used-cars-scrapy-script>. [Pristupljeno: 05-lip-2020].
- [4] „Understand star schema and the importance for Power BI - Power BI | Microsoft Docs“. [Na internetu]. Dostupno na: <https://docs.microsoft.com/en-us/power-bi/guidance/star-schema>. [Pristupljeno: 05-lip-2020].

## Popis slika

Slika 1. Model skladišta podataka.....	2
Slika 2. Prikaz originalne csv datoteke .....	3
Slika 3. Prikaz modificirane csv datoteke.....	3
Slika 4. Postupak odabira opcije učitavanja datoteke .....	5
Slika 5. Odabir putanje do datoteke.....	6
Slika 6. 'Preview' učitanih podataka.....	6
Slika 7. Dizajn tablice datasetCars .....	7
Slika 8. Kreirana tablica <i>datasetCars</i> .....	7
Slika 9. Upiti za popunjavanje datasetCars tablice .....	8
Slika 10. Struktura tablice DimGearbox .....	9
Slika 11. Upit za unos podataka u tablicu DimGearbox .....	9
Slika 12. Tablica DimGearbox .....	10
Slika 13. Struktura tablice DimFuelType .....	10
Slika 14. Upit za unos podataka u tablicu DimFuelType .....	10
Slika 15. Tablica DimFuelType .....	11
Slika 16. Struktura tablice DimBrand .....	11
Slika 17. Upit za unos podataka u tablicu DimBrand.....	11
Slika 18. Tablica DimBrand .....	12
Slika 19. Struktura tablice DimModel .....	13
Slika 20. Upit za unos podataka u tablicu DimModel .....	13
Slika 21. Tablica DimModel .....	13
Slika 22. Struktura tablice DimDate .....	14
Slika 23. Upit za unos podataka u tablicu DimDate .....	14
Slika 24. Tablica DimDate .....	15
Slika 25. Struktura tablice DimAdvertisement .....	16
Slika 26. Upit za unos podataka u tablicu DimAdvertisement .....	16
Slika 27. Tablica DimAdvertisement .....	16

Slika 28. Struktura tablice FactCars.....	17
Slika 29. Tablica FactCars.....	18
Slika 30. Model zvijezde .....	19
Slika 31. Opcije unosa podataka - Power BI.....	20
Slika 32. Spajanje na SQL Server bazu podataka .....	20
Slika 33. Model zvijezde unutar Power BI alata .....	21
Slika 34. Prvi izvještaj.....	22
Slika 35. Drugi izvještaj .....	22
Slika 36. Treći izvještaj.....	23



## Popis tablica

Tablica 1. Atributi korištenog skupa .....	4
--	---