

Fakultet tehničkih nauka

Inženjerstvo informacionih sistema

Projektovanje skladišta podataka

- Projekat ATP World Tour tennis -

Sadržaj:

1. Zadatak i ciljevi projekta.....	3
2. Opis postupaka projektovanja DW sistema	3
3. Specifikacija zahteva korisnika.....	4
4. Specifikacija modela	4
a. Specifikacija izvora podataka, OLTP šema	4
b. Specifikacija ciljanog DataWarehous sistema	5
5. Opis ETL procesa.....	7
6. Prikaz izveštaja kao odgovora na pitanja.....	48
7. Zaključak.....	58

1. Zadatak i ciljevi projekta

Projekat ima za zadatak da kreira skladište podataka i time omogući olakšano pohranjivanje podataka.

Set podataka koji sam odabrala za temu obezbeđuje jako veliki broj podataka iz više različitih izvora stoga je pravljenje odgovarajućeg skladišta podataka ključno za poslovanje i dalji razvoj ove oblasti. Rezultati mogu daleko doprineti jednostavnijem kreiranju potrebnih izveštaja i vršenju analiza.

Jedan od vodećih razloga za opredeljenje da navedena tema bude iz sportske oblasti jeste to što u slobodno vreme pratim sport i analiziram rezultate, kao i velika mogućnosti postavljanja raznih upita.

2. Opis postupaka projektovanja DW sistema

Originalni podaci su dobijeni putem interneta i deo su javnog seta podataka kome svi imaju pristup. Jedan od prvih zadataka bio je razumevanje podataka i prečišćavanje istih, zatim kreiranje DDL naredbi kojima se kreiraju tabele dimenzija za svaki tip entiteta iz realnog sistema. Jedan od težih zadataka jeste definisanje OLTP šeme na osnovu koje sam dalje kreirala sve tabele i njihove međusobne zavisnosti. Potrebno je bilo odlučiti se koji su to atributi od veće važnosti i na smisleni način ih ubacivati u tabele.

Aktivnosti koje trebaju biti sprovedene za kompletno projektovanje DW sistema:

1. Analiza i specifikacija korisničkih zahteva
2. Projektovanje šeme DW BP
3. Projektovanje arhitekture DW sistema
4. Projektovanje ETL softverske podrške
5. Projektovanje softverske podrške za izveštavanje i analizu podataka

3. Specifikacija zahteva korisnika

Sledi spisak pitanja koji predstavlja listu korisničkih zahteva na koje je, nakon detaljne analize ulaznih podataka, odgovoreno kroz izveštaje:

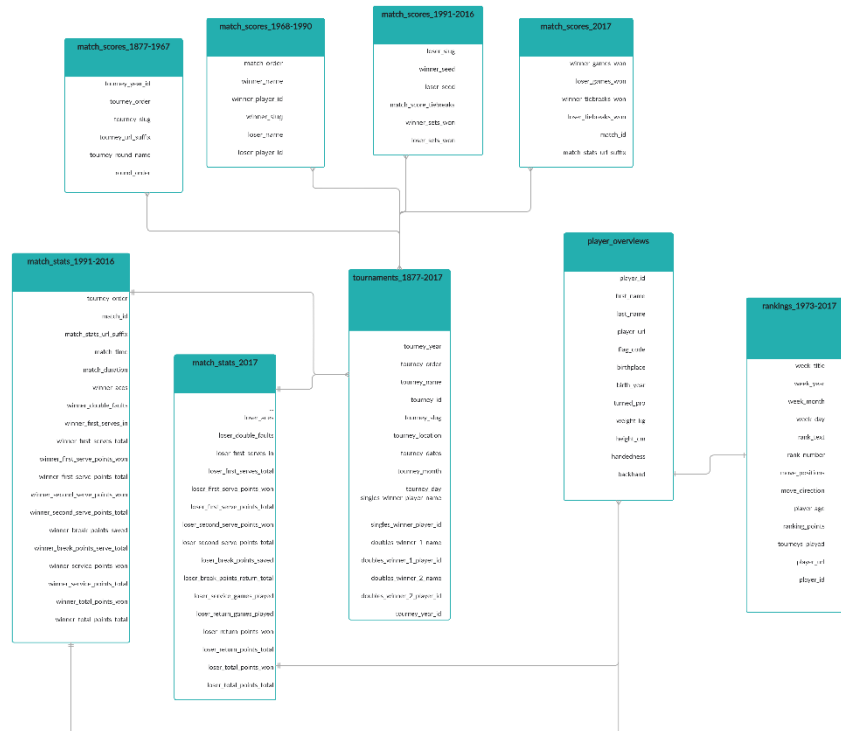
- Procenat mečeva koji su trajali duže od 3,5H
- Pregled statistike tri najbolje odigrana meča novaka djokovica
- Procenat od ukupnog broja igrača onih koji su backhand: Two-Handed i handedness: Left
- Svi igrači sa imenom Ricardo kako su rangirani na ATP listi
- Grupisanje na nivou lokacije turnira (Pariz, Barselona, Miami) i ispisivanje u okviru grupe uslova turnira (prebrojavanje posebno uslova?) i pobednika
- Poređenje broja igrača koji igraju za svaku od država i broj igrača koji se od njih smatraju profesionalcima

4. Specifikacija modela

a. Specifikacija izvora podataka, OLTP šema

Izvor podataka je relacionalna baza podataka koju čini 9 tabela, pritom se neke od tabela ponavljaju samo sa razlikom o kom vremenskom intervalu se radi. Neki podaci vremenski obuhvaju period od čak 140 godina.

Prikaz sistema kroz dijagram klasa:



Opis poslovnih procesa

Osvajanje seta – igrač osvaja poene koji će predstavljati jedan osvojen set za igrači koji ih prvi skupi

Prolazak u naredni krug – proces u kojem igrač stiže pobedom pravo odlaska u sledeću rundu turnira

Igranje petog seta – ukoliko je predviđeno da igrač mora osvojiti tri seta za pobedu, može doći pri rezultatu 2:2 do igranja petog seta

Skok na ATP listi – podrazumeva da se igrač plasirao do finala na turniru i time stekao bodove

Kraj meča - nakon što je odigran i poslednji poen, znamo pobednika i meč se završava

b. Specifikacija ciljanog DataWarehous sistema

- Specifikacija zahtevanih dimenzija

Uočavamo da je za projektovanje jednog DW sistema neophodno napraviti izmene izvorne šeme kako bi se odgovori na piTanja lakše i brže dobijali. Stoga je neophodno napraviti jednu glavnu tabelu, tabelu činjenica, za ovaj sistem to je tabela MATCH. Ova tabela kao svoje dimenzije ima sledeće TOURNAMENT, ROUND, PLAYER, TIME I DATE.

Dimenzija PLAYER takodje uzima podatke iz BIRTHPLACE I RANK dimenzija, pored toga ima podatke poput imena, prezimena, datuma rođenja i ostalih osnovnih ličnih informacija.

ROUND dimenzija ima samo naziv runde i numerički prikaz iste.

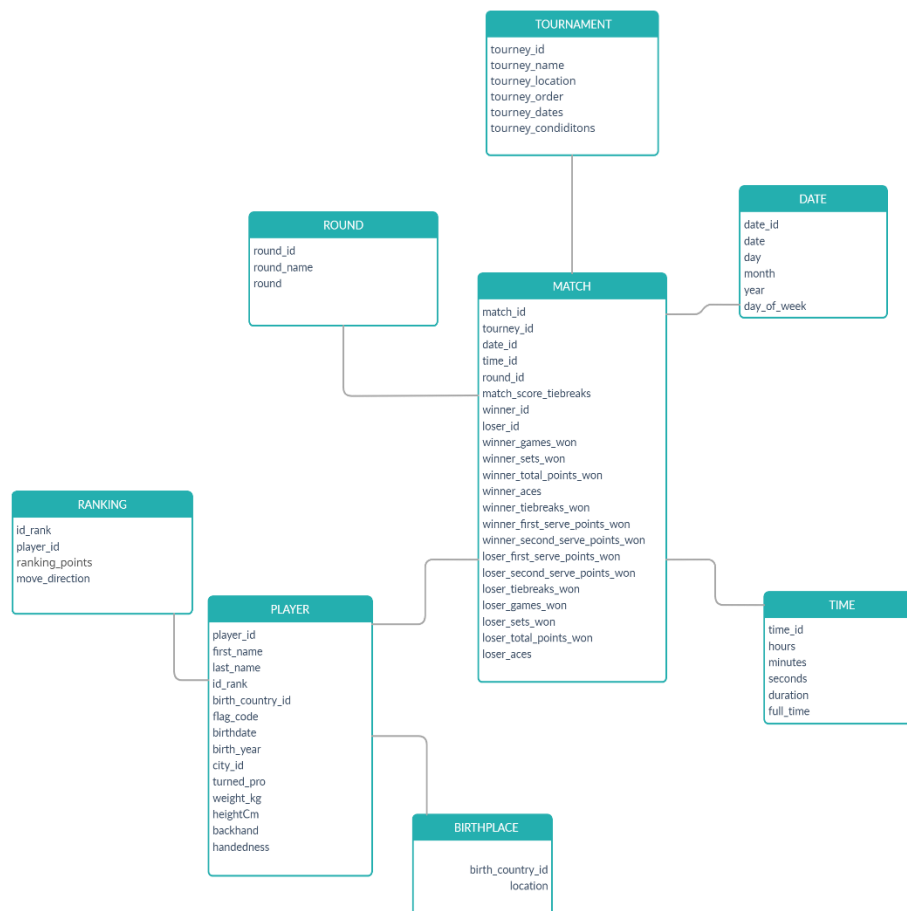
TIME dimenzija ima podatak o trajanju meča, može se posmatrati na različitim nivoima granulacije, sati, minuti...

DATE sadži podatke o datumu meča i takođe može da se posmatra na različitim nivoima – kvartal, mesec, dan u nedelji, godina...

- Specifikacija zahtevanih mera

Nisu postojali nikakvi zahtevi za posebnim merama, stoga nisu ni implementirane. Svakako usled budućih zahteva se struktura skladišta može izmeniti i dodati se neke potrebne mere.

Prikaz šeme skladišta:



5. Opis ETL procesa

U ovom poglavlju biće opisan ETL proces kao i alati korišćeni kako bi se kreirao. Jedan od osnovnih alata je dodatak za Microsoft Office Visual Studio pod nazivom SSDT (SQL Server Data Tools), kao i sam SQL Server i SSMS (SQL Server Management Studio).

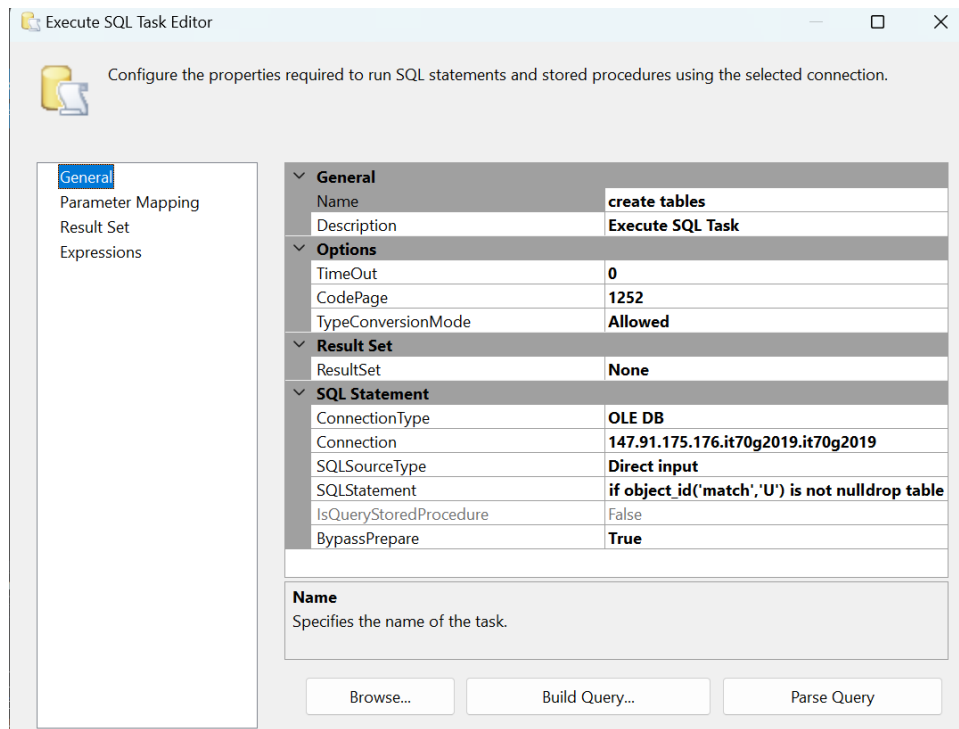
ETL proces je započet korišćenjem ručno pravljenog DDL-a za novu bazu (upit za kreiranje nekih od tabela):

```
if object_id('match','U') is not null
drop table match
if object_id('player','U') is not null
drop table player
if object_id('birthPlace','U') is not null
drop table birthPlace
if object_id('time','U') is not null
drop table time
if object_id('tournament','U') is not null
drop table tournament
if object_id('round','U') is not null
drop table round
if object_id('ranking','U') is not null
drop table ranking

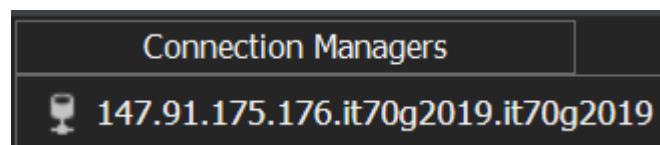
create table round(
    round_id int not null primary key identity,
    roundName varchar(100) not null,
    round int not null
)

create table tournament(
    id int not null primary key identity,
    tourney_id nvarchar(20),
    tourneyName varchar(100) not null,
    tourneyLocation varchar(100) not null,
    tourneyOrder numeric(3) not null,
    tourneyDates varchar(100) not null,
    tourneyConditions varchar(50) not null
```

Zatim, koristeći opciju Execute SQL Task, dodat je konekcion string kao i već pripremljene DDL naredbe. Na narednim slikama je prikazan proces kreiranja konekcije na već postojeću studentsku bazu.



Konekcija je kreirana.



Nakon uspešnog izvršavanja DDL naredbi, ETL proces se nastavlja korišćenjem Data Flow Task-a sa ciljem da se popune novokreirane tabele podacima.

Kreiranje i unos podataka u dateDimension

Za kreiranje date dimenzije korišćen je script sa interneta i unutar njega je napisan početni i krajnji datum koji se pojavljuju u setu podataka.

Slede slike koje prikazuju deo navedenog scripta:

```
IF object_id ('dateDimension','U') is not null
DROP TABLE dateDimension;
GO
CREATE TABLE dateDimension
(
    [DateKey] INT primary key,
    [Date] DATETIME,
    [FullDate] CHAR(10), -- Date in dd-MM-yyyy format
    [DayOfMonth] VARCHAR(2), -- Field will hold day number of Month
    [DayName] VARCHAR(9), -- Contains name of the day, Sunday, Monday
    [DayOfWeek] CHAR(1), -- First Day Sunday=1 and Saturday=7
    [DayOfWeekInMonth] VARCHAR(2), --1st Monday or 2nd Monday in Month
    [DayOfWeekInYear] VARCHAR(2),
    [DayOfQuarter] VARCHAR(3),
    [DayOfYear] VARCHAR(3),
    [WeekOfMonth] VARCHAR(1), -- Week Number of Month
    [WeekOfQuarter] VARCHAR(2), --Week Number of the Quarter
    [WeekOfYear] VARCHAR(2), --Week Number of the Year
    [Month] VARCHAR(2), --Number of the Month 1 to 12
    [MonthName] VARCHAR(9), --January, February etc
    [MonthOfQuarter] VARCHAR(2), -- Month Number belongs to Quarter
    [Quarter] CHAR(1),
    [QuarterName] VARCHAR(9), --First,Second..
    [Year] CHAR(4), -- Year value of Date stored in Row
    [YearName] CHAR(7), --CY 2012,CY 2013
    [MonthYear] CHAR(10), --Jan-2013, Feb-2013
    [MMYYYY] CHAR(6),
    [FirstDayOfMonth] DATE,
```

U ovom delu su navedeni početni i krajnji datum:

```
}/*****
--Specify Start Date and End date here
--Value of Start Date Must be Less than Your End Date
DECLARE @StartDate DATETIME = '01/01/1877' --Starting value of Date Range
DECLARE @EndDate DATETIME = '12/31/2017' --End Value of Date Range
```

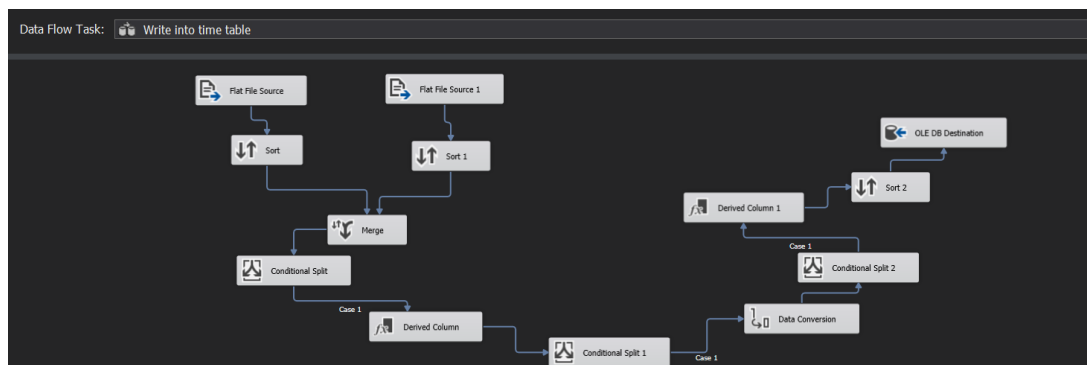
Rezultat upisa u tabelu izgleda ovako:

	DateKey	Date	FullDate	DayOfMonth	DayName	DayOfWeek	DayOfWeekInMonth	DayOfWeekInYear	DayOfQuarter	DayOfYear	WeekOfMonth	Week
1	18770101	1877-01-01 00:00:00.000	01/01/1877	1	Monday	1	1	1	1	1	1	1
2	18770102	1877-01-02 00:00:00.000	01/02/1877	2	Tuesday	2	1	1	1	2	1	1
3	18770103	1877-01-03 00:00:00.000	01/03/1877	3	Wednesday	3	1	1	1	3	1	1
4	18770104	1877-01-04 00:00:00.000	01/04/1877	4	Thursday	4	1	1	1	4	1	1
5	18770105	1877-01-05 00:00:00.000	01/05/1877	5	Friday	5	1	1	1	5	1	1
6	18770106	1877-01-06 00:00:00.000	01/06/1877	6	Saturday	6	1	1	1	6	1	1
7	18770107	1877-01-07 00:00:00.000	01/07/1877	7	Sunday	7	1	1	1	7	2	1
8	18770108	1877-01-08 00:00:00.000	01/08/1877	8	Monday	1	2	2	2	8	2	2

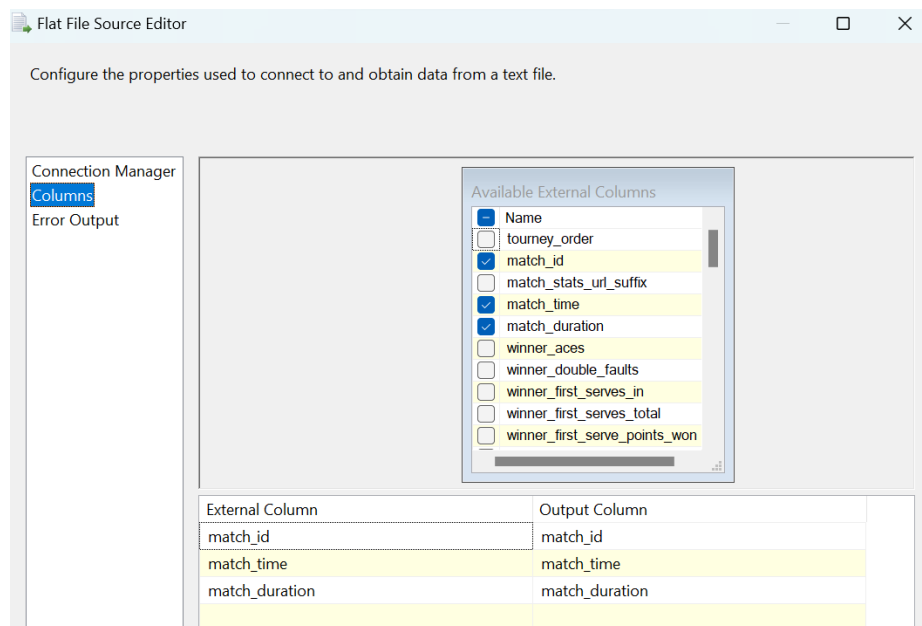
Query executed successfully. 147.91.175.176 (12.0 SP2) it70g2019 (71) it70g2019 00:00:01 51,498 rows

Unos podataka u tabelu time

Prikaz Data Flow Task-a za time dimenziju:



Prvo su izabrana dva izvora podataka koja su u stvari tabele sa istim kolonama ali su u pitanju razlicite godine. Iz oba izvora su izabrane iste kolone koje su prikazane na narednoj slici.



Zatim oba izvora sortiram po istoj koloni kako bih bez problema mogla da uradim spajanje odnosno merge ove dve tabele.

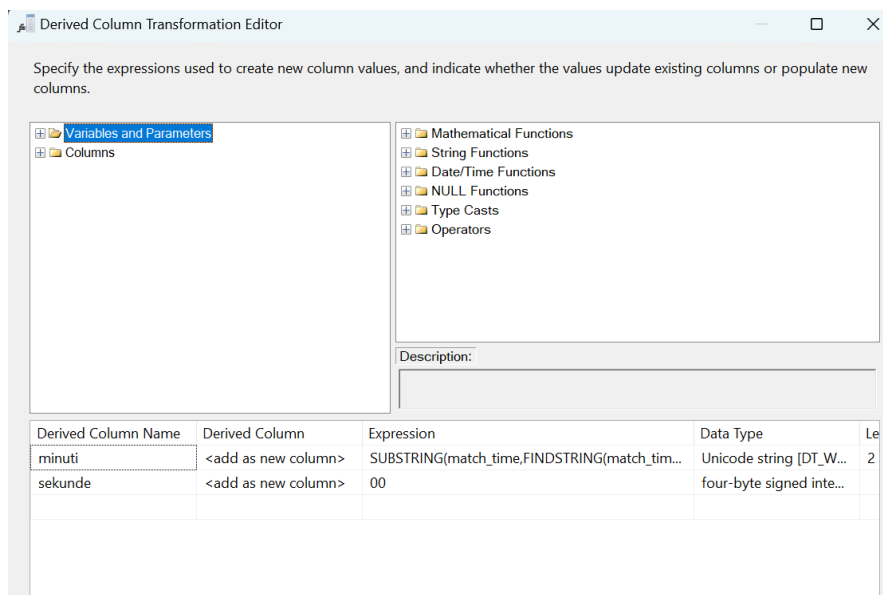
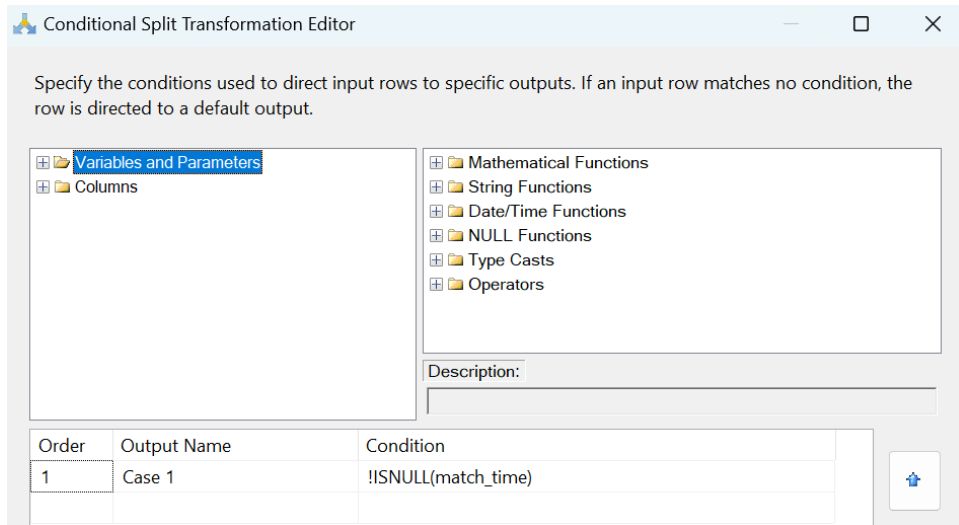
The first screenshot shows the 'Sort Transformation Editor' window. It contains an 'Available Input Columns' dialog box with a list of columns: 'Name', 'match_id', 'match_time', and 'match_duration'. The 'match_id' column is selected with a checkmark. Below the dialog is a table with the following data:

Input Column	Output Alias	Sort Type	Sort Order	Condition
match_id	match_id	ascending	1	

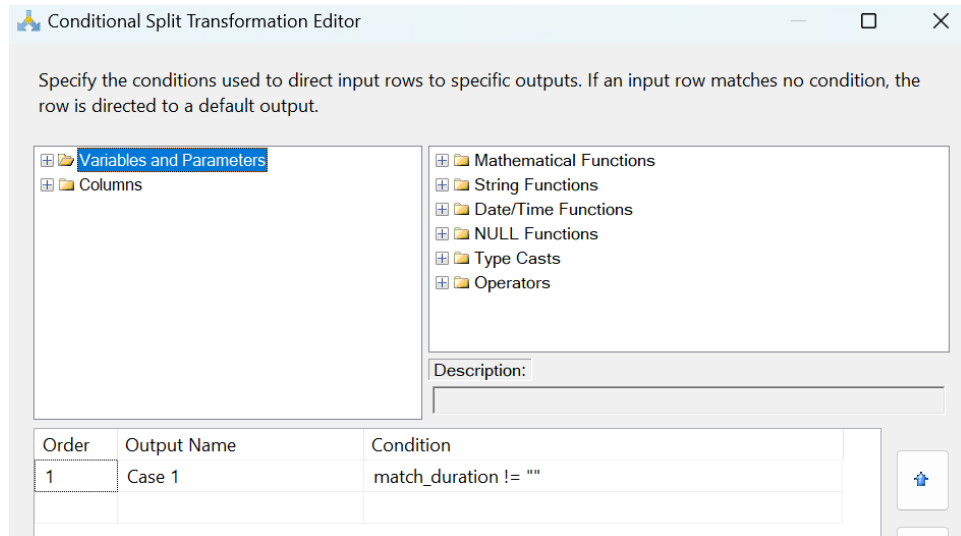
The second screenshot shows the 'Merge Transformation Editor' window. It contains a table with the following data:

Output Column Name	Merge Input 1	Merge Input 2
match_id (Sort key: 1)	match_id (Sort key: 1)	match_id (Sort key: 1)
match_time	match_time	match_time
match_duration	match_duration	match_duration

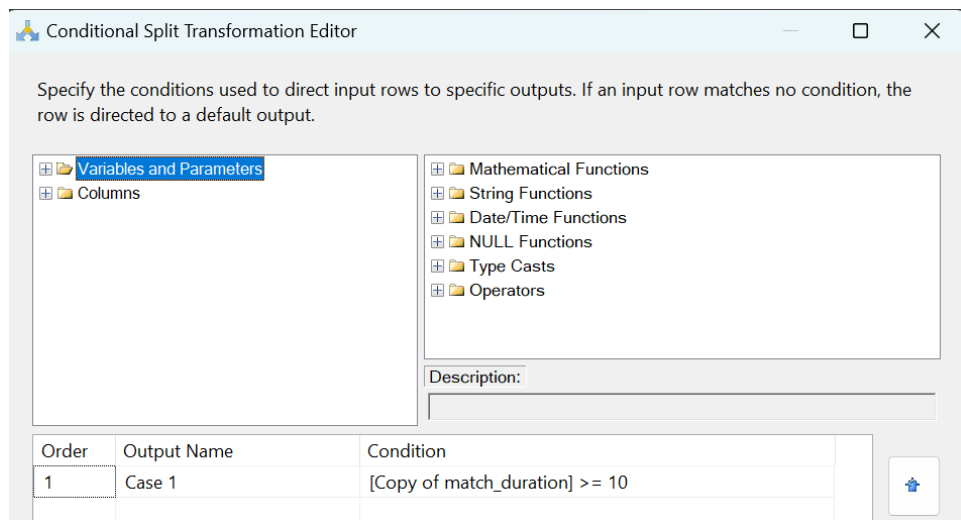
Nakon što su podaci spojeni, odrađen je conditional split kako bismo izuzeli null vrednosti. Potom je preko derived column-a izvučena vrednost za minute i sekunde.



Zatim je obrađan ponovo conditional split kako bih iz druge kolone odbacila null vrednosti:



Odлучila sam se da izbacim i vrednosti da je meč trajao kraće od 10 minuta:



Potom se vrši konverzija podataka u željeni tip:

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

☒ Name
☐ match_id
☒ match_time
☒ match_duration
☒ minuti
☒ sekunde

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
minuti	Copy of minuti	string [DT_STR]	5			1252 (ANSI - Latin I)
sekunde	Copy of sekunde	string [DT_STR]	10			1252 (ANSI - Latin I)
match_duration	Copy of match_durat...	four-byte signed intege...				
match_time	Copy of match_time	string [DT_STR]	100			1252 (ANSI - Latin I)

Podaci se sortiraju i izbacuju se duplikati vrednosti:

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

Name	Pass T...
<input type="checkbox"/> match_id	<input checked="" type="checkbox"/>
<input type="checkbox"/> match_time	<input checked="" type="checkbox"/>
<input type="checkbox"/> match_duration	<input type="checkbox"/>
<input type="checkbox"/> minuti	<input type="checkbox"/>
<input type="checkbox"/> sekunde	<input type="checkbox"/>
<input type="checkbox"/> Copy of minuti	<input type="checkbox"/>
<input type="checkbox"/> Copy of seku...	<input type="checkbox"/>

Column name: match_time
Data type: string [DT_STR]
Length: 50
Scale: 0
Precision: 0
Code Page: 1252 (ANSI - Latin I)
Source Component: Merge

Input Column	Output Alias	Sort Type	Sort Order	Com
Copy of match_time	Copy of match_time	ascending	1	

☒ Remove rows with duplicate sort values

OK Cancel Help

Odabrana je željena destinacija odnosno kreirana tabela u bazi i potom su namapirane sve kolone.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager

Mappings

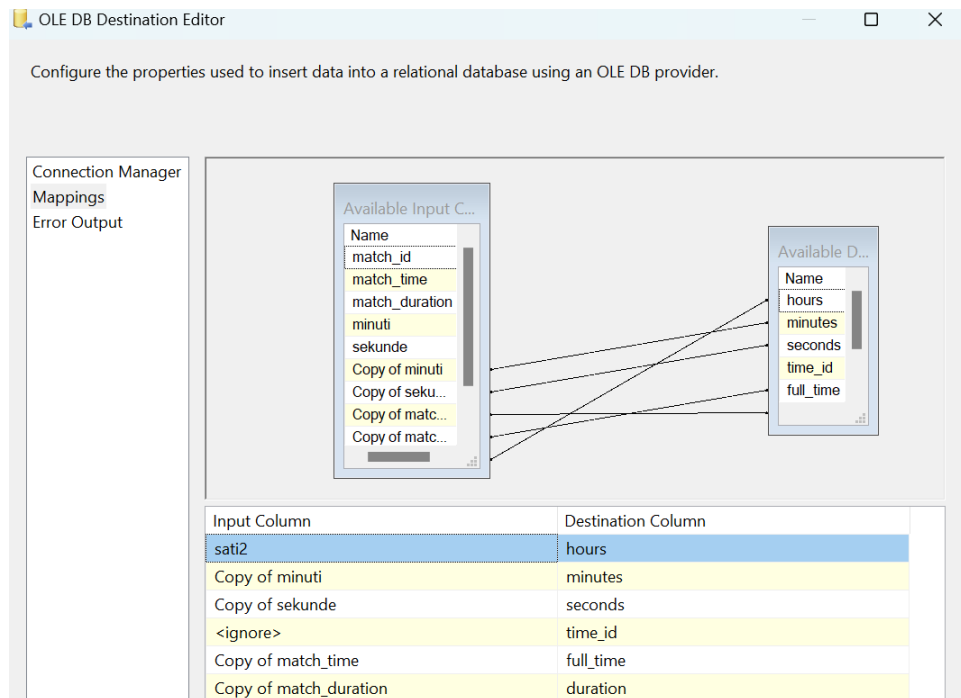
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:
147.91.175.176.it70g2019.it70g2019 New...

Data access mode:
Table or view - fast load

Name of the table or the view:
[dbo].[time] New...



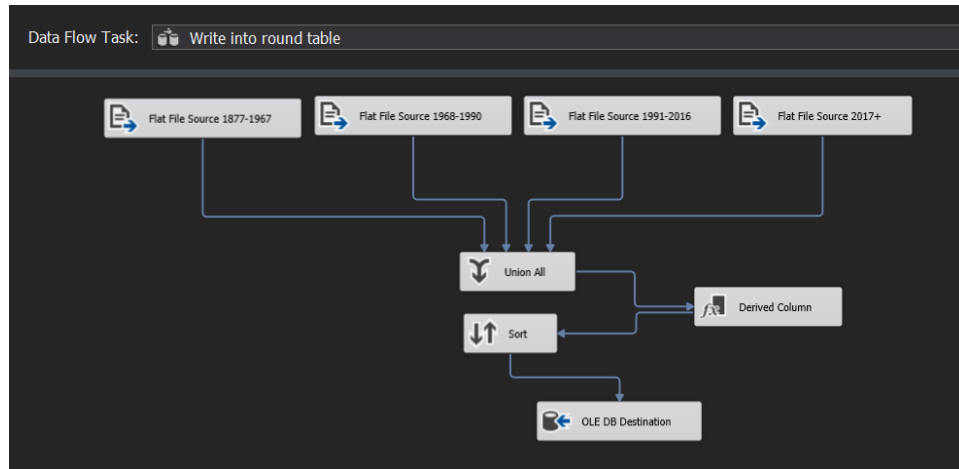
Prikaz tabele time nakon pokretanja Data Flow Task-a:

Results		Messages				
	time_id	hours	minutes	seconds	duration	full_time
107	107	1	56	0	116	01:56:00
108	108	1	57	0	117	01:57:00
109	109	1	58	0	118	01:58:00
110	110	1	59	0	119	01:59:00
111	111	2	00	0	120	02:00:00
112	112	2	01	0	121	02:01:00
113	113	2	02	0	122	02:02:00
114	114	2	03	0	123	02:03:00
115	115	2	04	0	124	02:04:00

✓ Query executed successfully.

Unos podataka u tabelu round

Naredna tabela je round. Kao i u prethodnom primeru koristi se Flat File Source kako bi se izvukli podaci iz CSV fajla. U ovom slučaju imamo 4 izvorne tabele sa istim podacima ali za različite vremenske intervale, te ih spajam pomoću Union All funkcije.



Potom se preko Derived Column spajaju vrednosti iz dve kolone – naziv i redni broj runde, kako bi se potom na osnovu novodobijene kolone sortirala tabela i izbacile se duple vrednosti.

Derived Column Transformation Editor

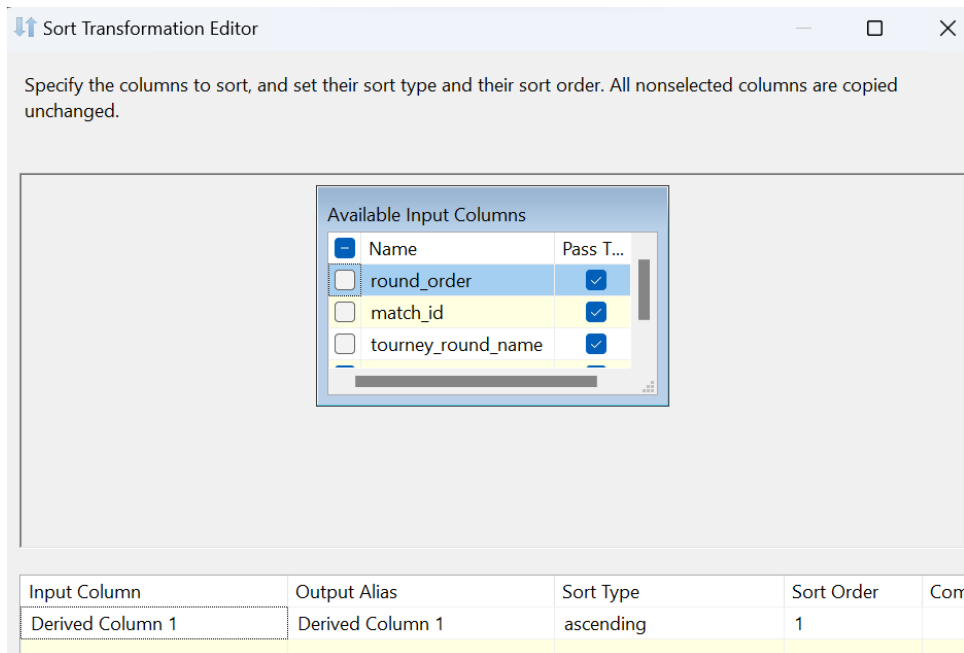
Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters
Columns

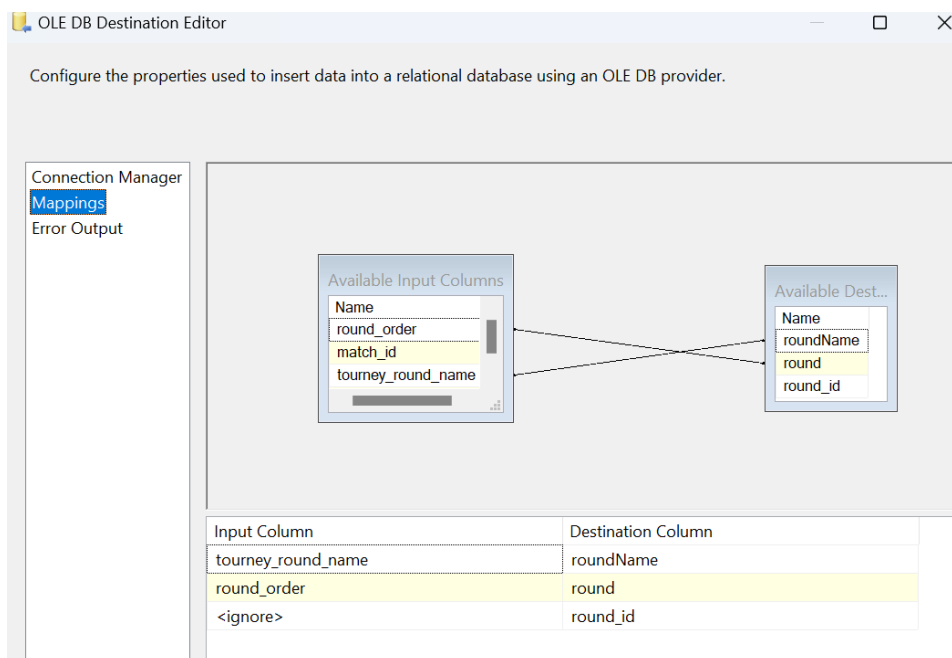
Mathematical Functions
String Functions
Date/Time Functions
NULL Functions
Type Casts
Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length
Derived Column 1	<add as new column>	tourney_round_name + "-" + round_order	Unicode string [DT_W...	10

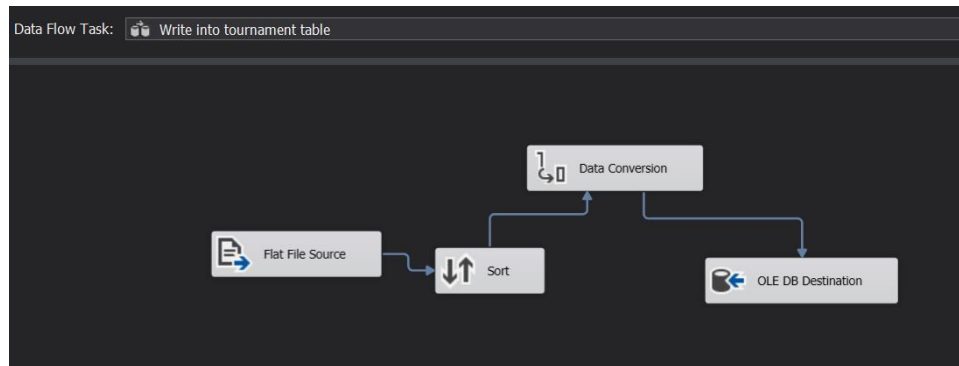


Na samom kraju, nakon što je izabrana tabela u bazi za destinaciju, vrši se mapiranje kolona.



Unos podataka u tabelu tournament

Kao i u prethodnim tabelama, prvo se preko Flat File Source odredio izvor podataka.



Prikupljeni podaci su sortirani kako bi se izgubile duple vrednosti I potom je u Data Conversion dodeljen trazeni tip podatka kako bi sve bilo spremno za unos podataka u OLE DB Destination.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

- ☒ Name
- ☐ tourney_name
- ☐ tourney_location
- ☐ tourney_order
- ☐ tourney_dates
- ☐ tourney_conditions
- ☒ tourney_year_id

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
tourney_year_id	Copy of tourney_year...	Unicode string [DT_WST...	50			

Izabrana je odgovarajuća tabela za destinaciju nakon čega su kolone mapirane i unos je izvršen.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:
147.91.175.176.it70g2019.it70g2019 New...

Data access mode:
Table or view - fast load

Name of the table or the view:
[dbo].[tournament] New...

☐ Keep identity ☒ Table lock
☐ Keep nulls ☒ Check constraints

Rows per batch:

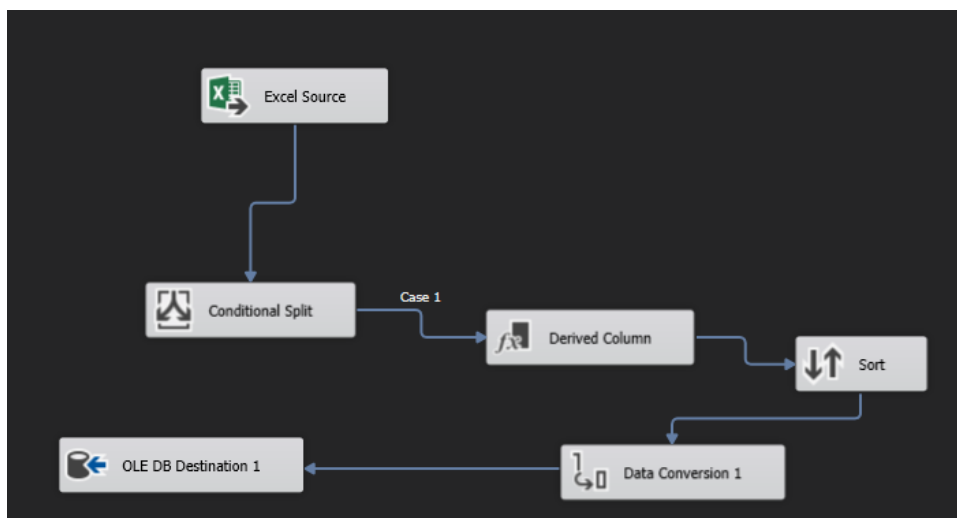
Maximum insert commit size: 2147483647

Prikaz tabele tournament nakon pokretanja Data Flow Task-a:

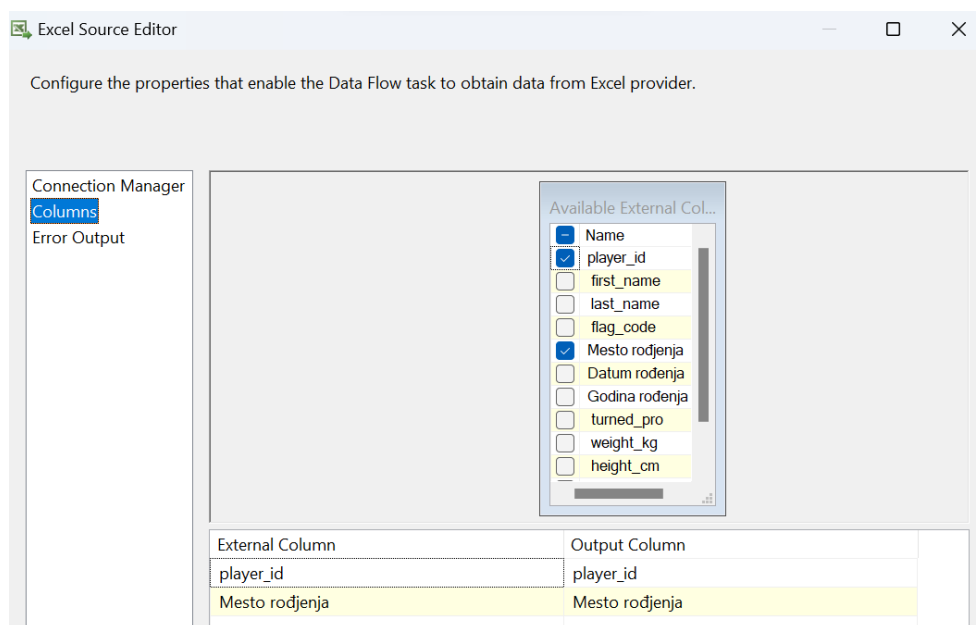
Results		Messages					
	id	tourney_id	tourneyName	tourneyLocation	tourneyOrder	tourneyDates	tourneyConditions
1	1	s938,2017-807	Abierto Mexicano Telcel	Acapulco Mexico	16	2017.02.27	Hard
2	2	marcel-granollers,g7	ABN Amro World Tennis Tournament	Rotterdam Netherlands	12	2017.02.13	Indoor
3	3	v258,2007-807	Acapulco	Acapulco	16	2007.02.26	Outdoor
4	4	u024,2002-7308	Adelaide	Adelaide	1	2001.12.31	Outdoor
5	5	1982-393	Adelaide-2	Adelaide	1	1982.01.04	Outdoor
6	6	bruno-soares,s938,20	Aegon Championships	London Great Britain	37	2017.06.19	Outdoor
7	7	b589,2017-741	Aegon International	Eastbourne Great Britain	39	2017.06.26	Outdoor
8	8	n480,2017-891	Aircel Chennai Open	Chennai India	2	2017.01.02	Outdoor

Unos podataka u tabelu birthPlace

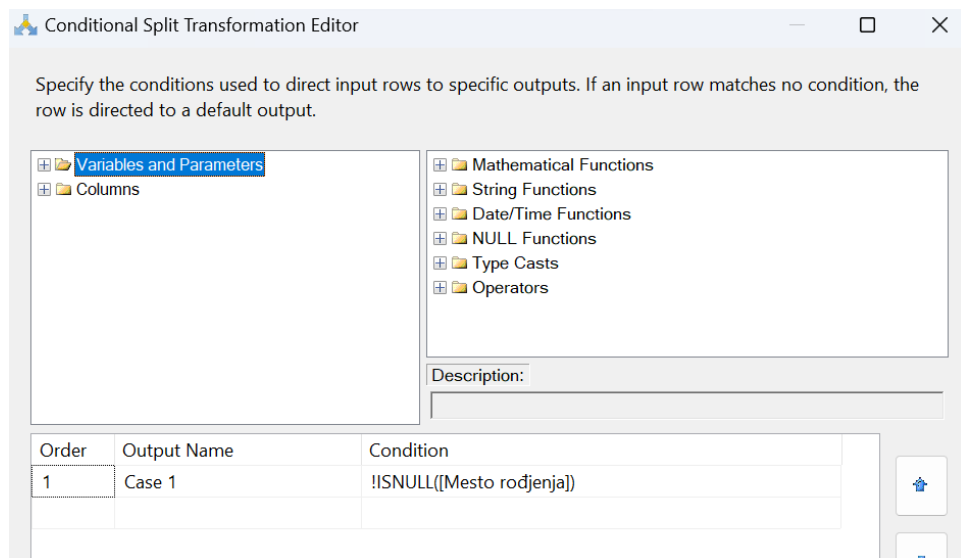
Za tabelu birthPlace korišćen je Excel Source.



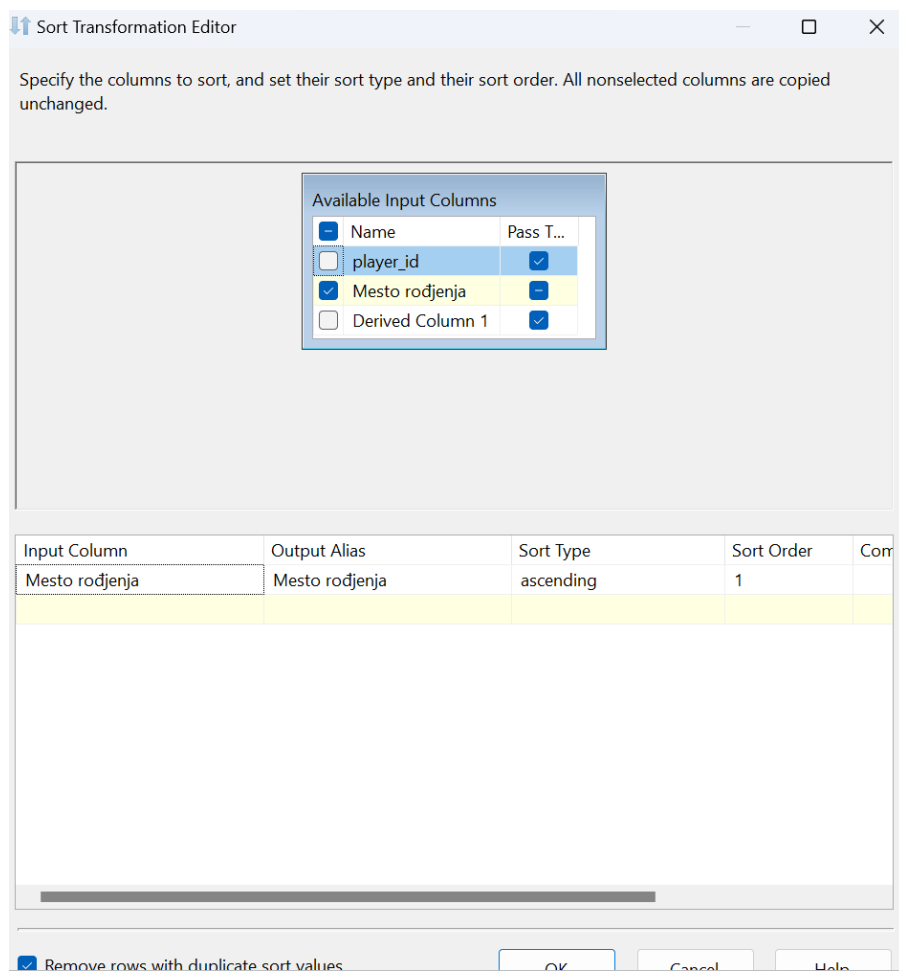
Kao i u prethodnim tabelama, na samom početku odabrane su kolone iz izvorne table koje su nam potrebne.



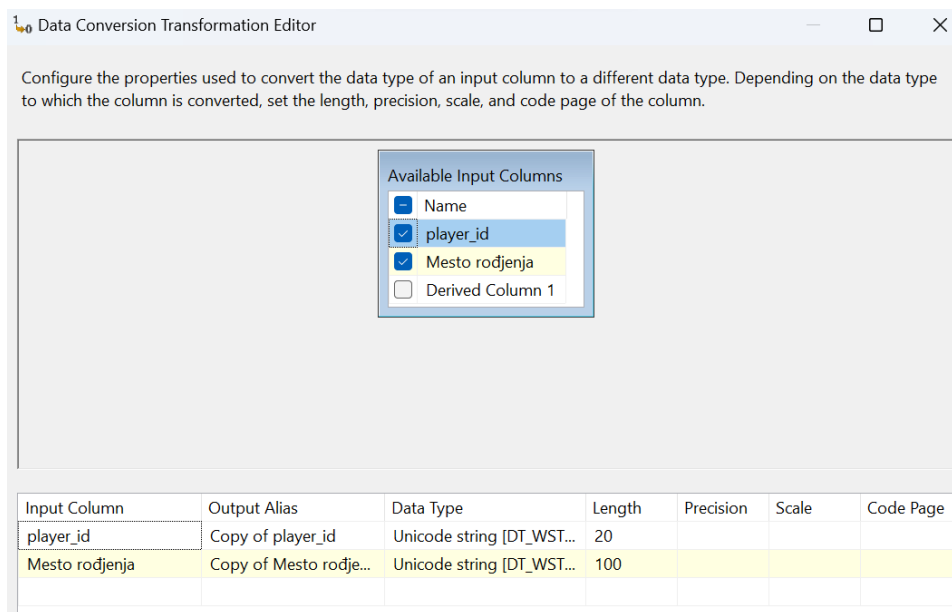
Potom, pomoću Conditional Split-a odbacujemo vrednosti koje su null.



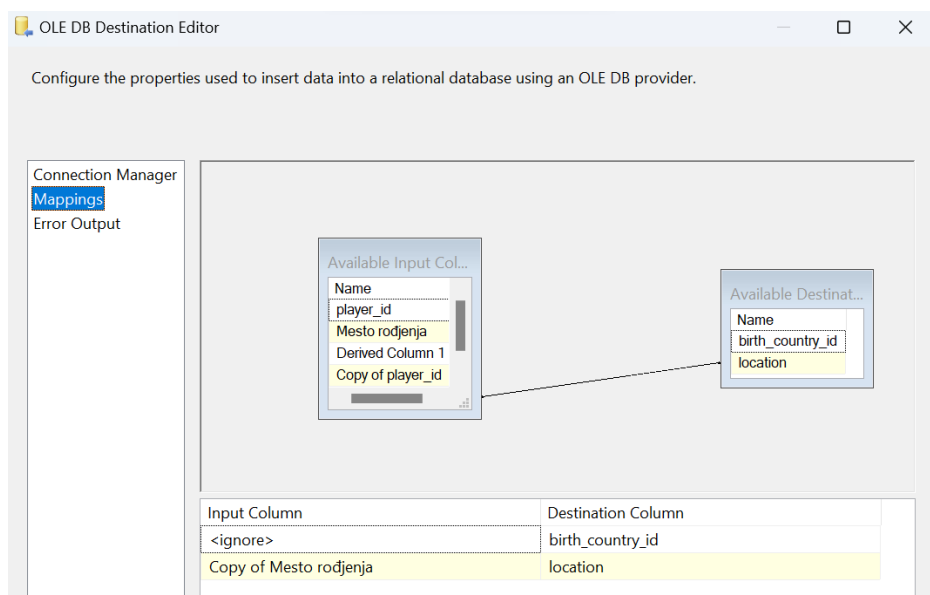
Podaci su sortirani i odbačene su duple vrednosti.



Pomoću Data Conversion su promenjeni tipovi podataka na odgovarajuće poput navedenih u tabeli destinacije.



Sledi prikaz mapiranja kolona pred upis u destinaciju. ID polje je ignorisano jer se samo generiše kroz tabelu.

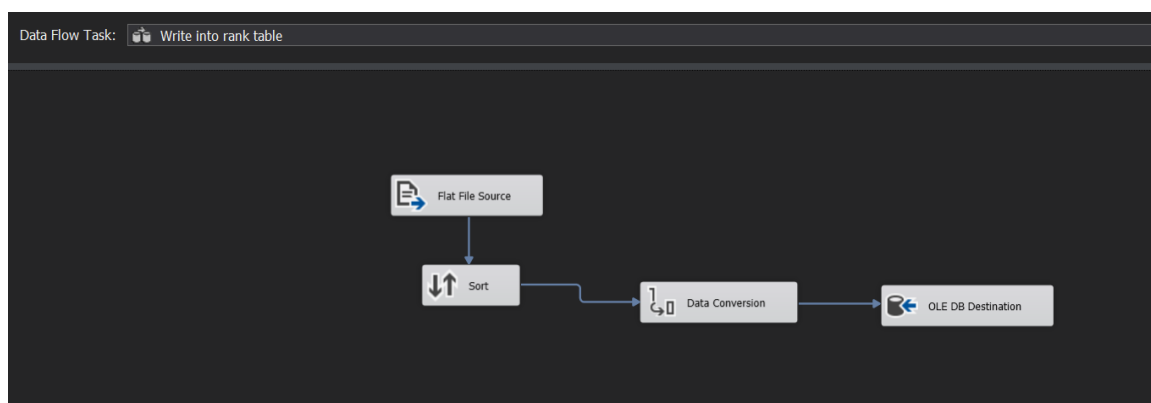


Prikaz tabele birthPlace nakon pokretanja Data Flow Task-a:

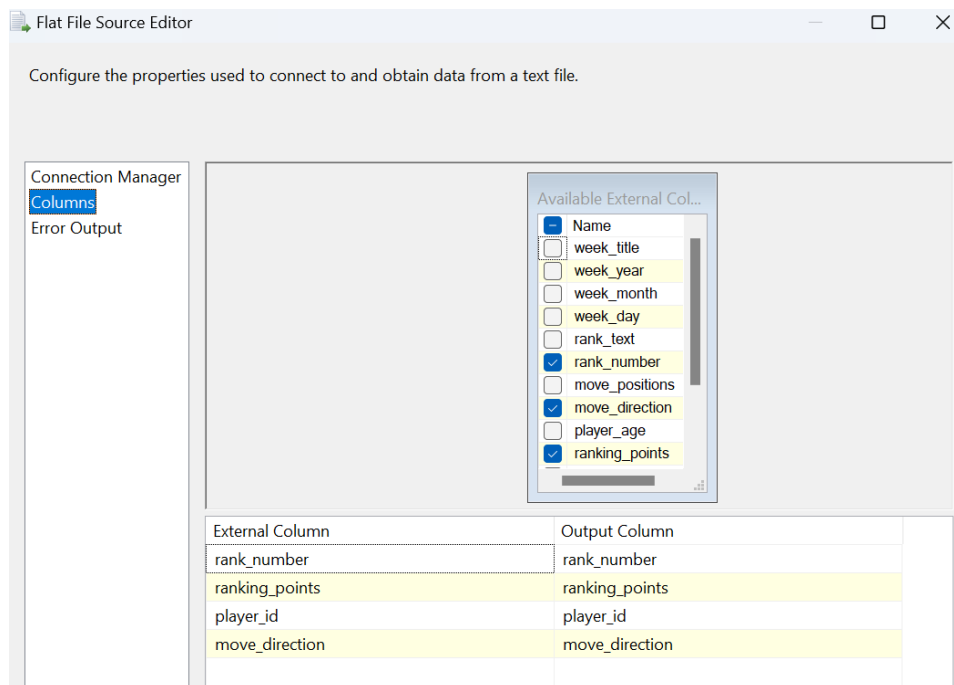
Results			Messages		
	birth_country_id	location			
59	59	Lancashire			
60	60	Larnaca			
61	61	Levin			
62	62	Linz			
63	63	London			
64	64	Los Angeles			
65	65	Lyon			
66	66	Madrid			

Unos podataka u tabelu ranking

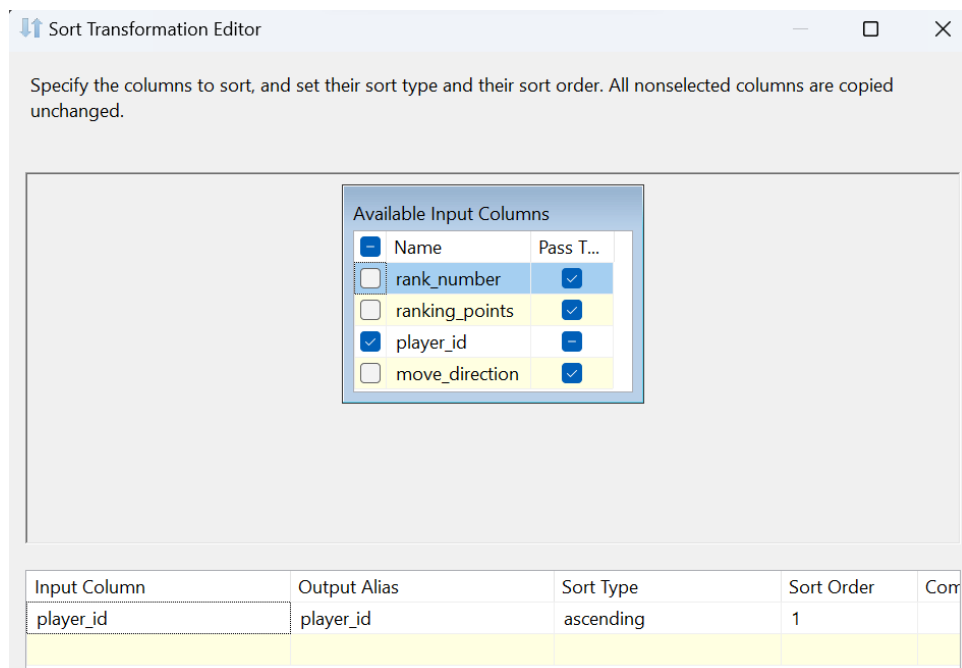
Prikaz Data Flow Task-a za ranking tabelu:



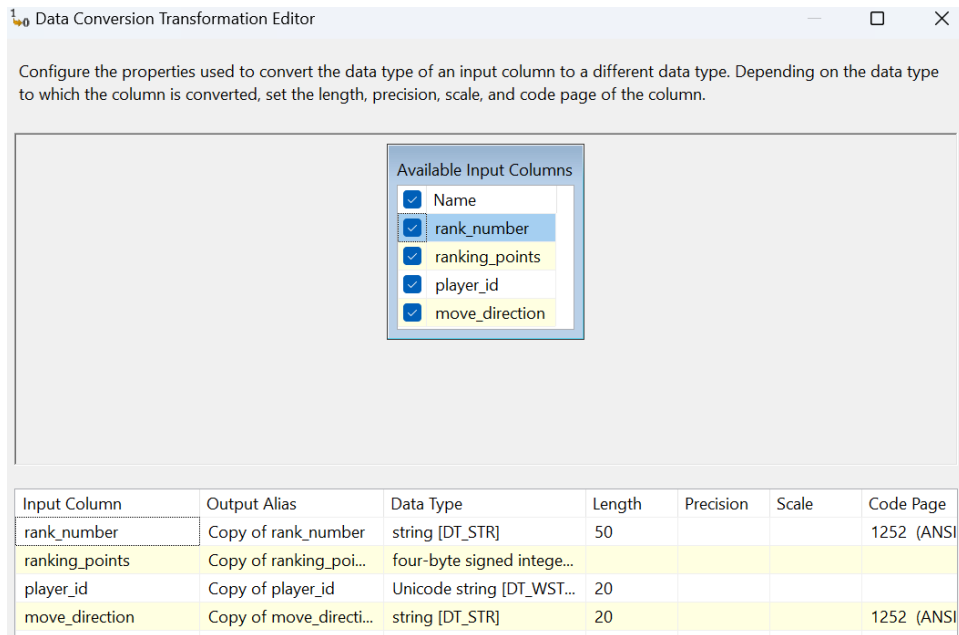
Iz odabranog Flat File izvora biram sledeće kolone:



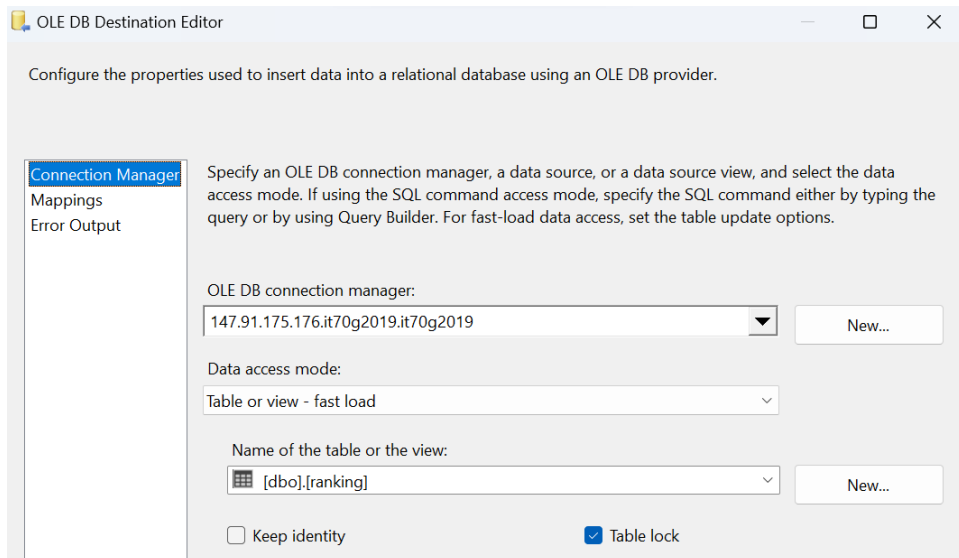
Odabrane kolone sortiram kako bih izbrisala duplirane podatke.



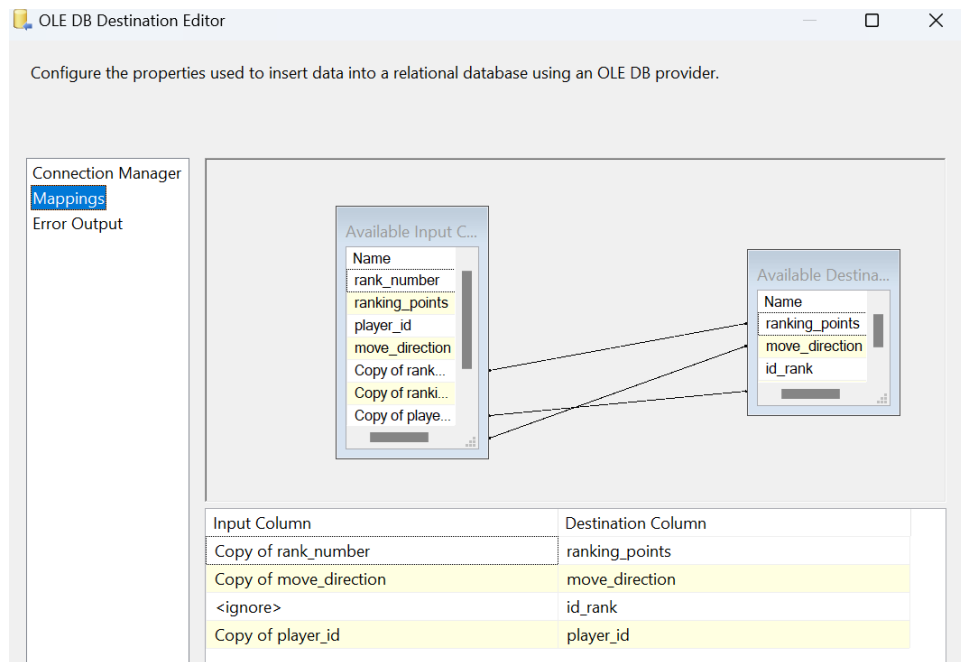
Zatim, preko Data Conversion-a sve izabrane kolone transformišem u željene tipove podataka.



Nakon toga preko OLE DB Destination-a biram tabelu iz baze koja će biti destinacija.



Na samom kraju odrađeno je mapiranje kolona tako da svaki podatak ode na ispravno mesto u destinacionoj tabeli.

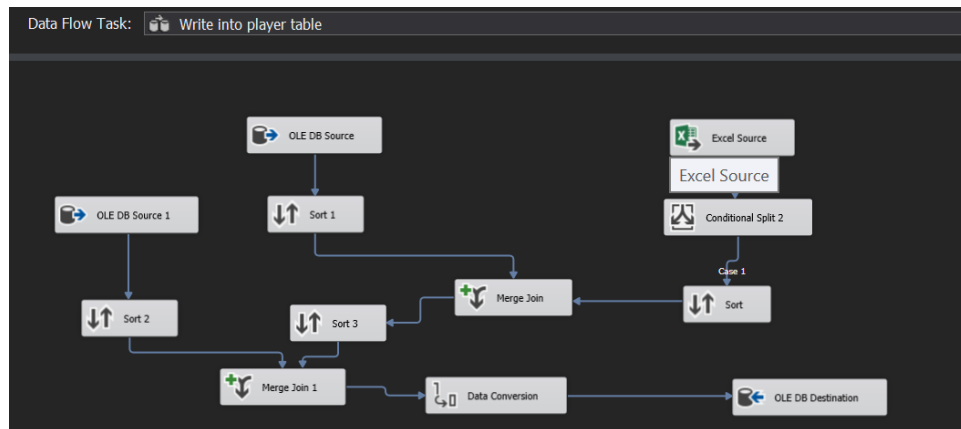


Prikaz tabele ranking nakon pokretanja Data Flow Task-a:

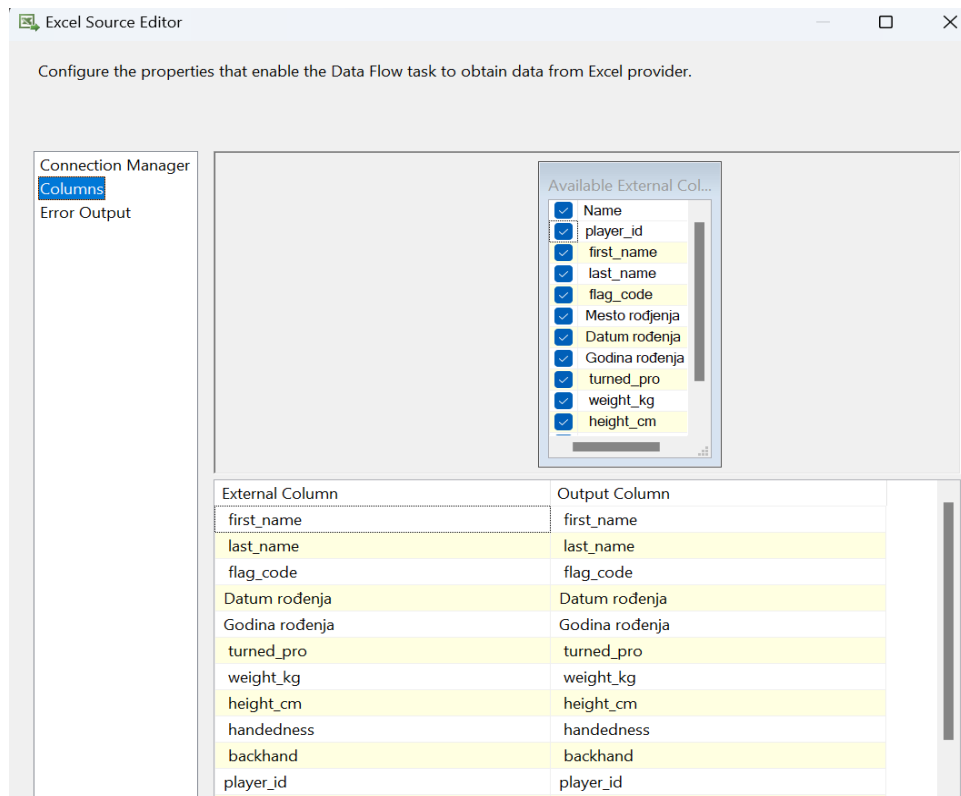
Results		Messages		
	id_rank	player_id	ranking_points	move_direction
25	25	a025	838	up
26	26	a026	745	down
27	27	a027	1209	up
28	28	a028	1240	down
29	29	a029	444	down
30	30	a030	774	down
31	31	a031	714	down
32	32	a032	994	down

Unos podataka u tabelu player

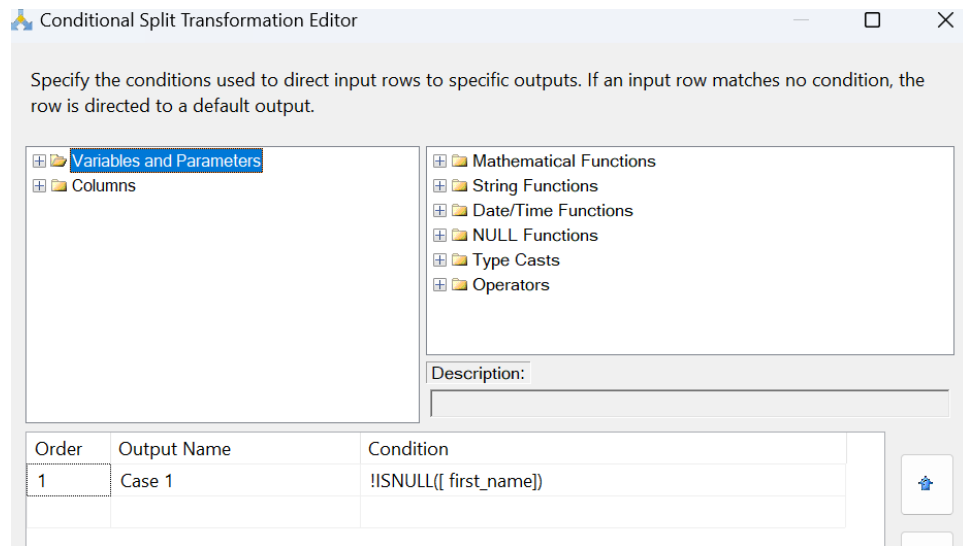
Prikazan je Data Flow Task za player tabelu:



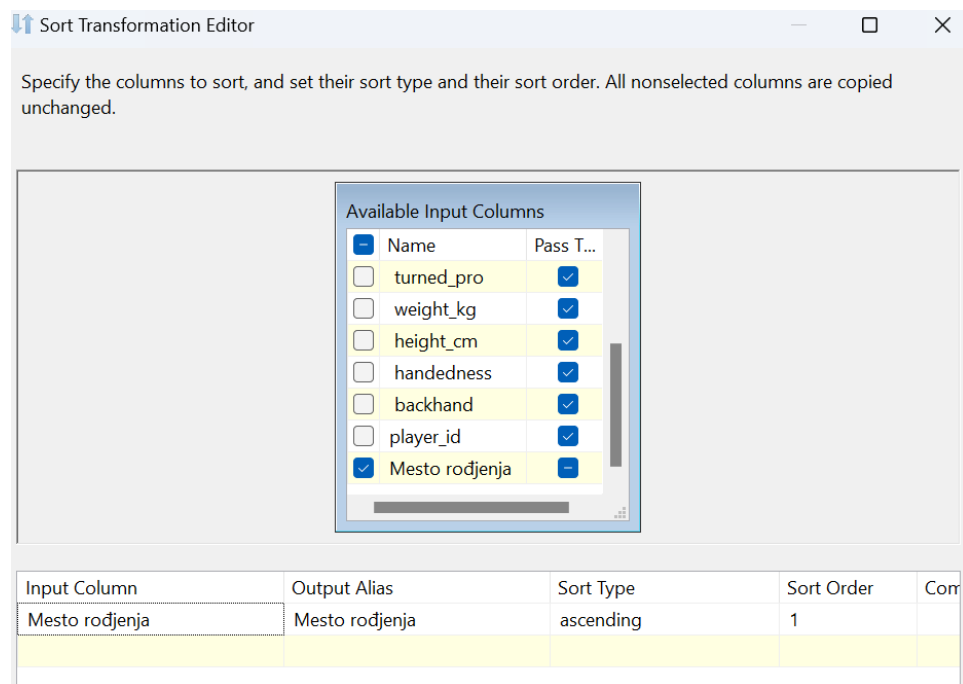
Jedan od izvora je Excel Source. Nakon što sam izabrala izvor, biram koje kolone su mi potrebne kao što je prikazano na slici.



Potom odbacujem one redove koji u koloni first_name nemaju unešenu vrednost.

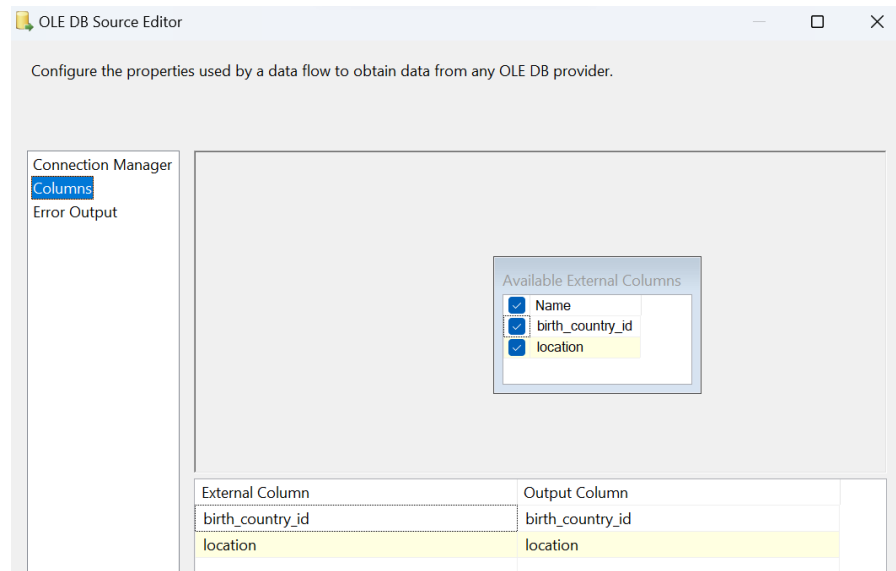


Zatim sortiram unešene podatke po koloni Mesto rođenja kako bih mogla da je spojim sa drugom tabelom koja sadrži te podatke.

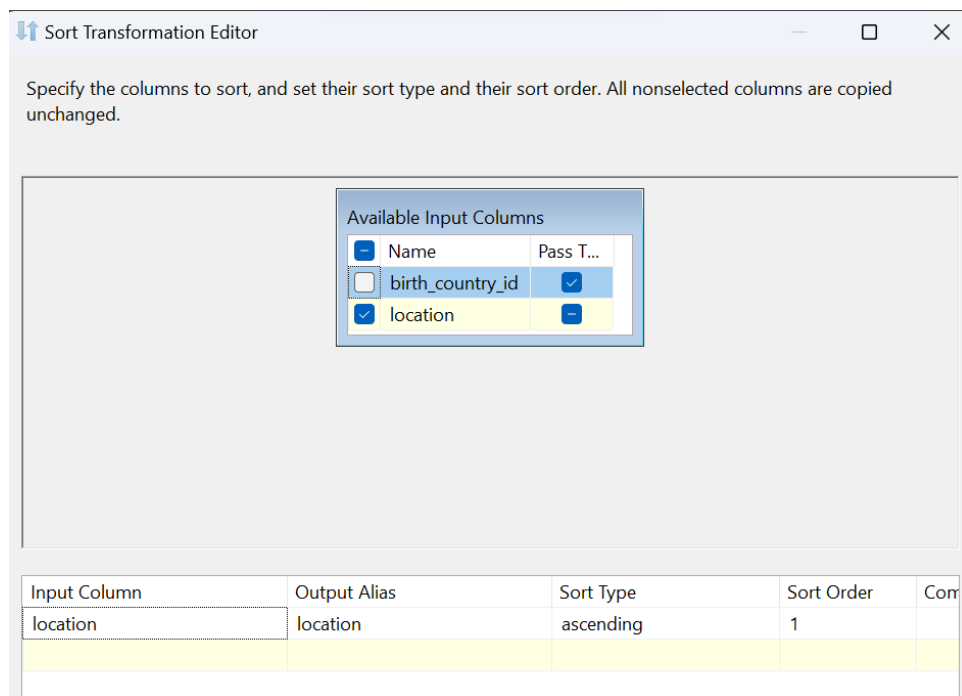


Sledeći izvor koji koristim za ovu tabelu je OLE DB Source, odnosno ranije kreirana birthPlace tabela.

Iz nje uzimam sve dostupne kolone.



Sortiram podatke i iz ove tabele na osnovu location kolone, kako bih mogla da spojim sa prethodno sortiranom tabelom.



Koristim Merge Join transformaciju da bih na osnovu location i mesto rođenja kolona spojila dve source tabele. Taj podatak odnosno kolonu location ne uzimam već samo ključ te tabele – birth_country_id kako bih mogla da joj pristupam i iz nje čitam podatke o lokaciji.

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type: Inner join Swap Inputs

Sort 1

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	birth_country_id	0	<input type="checkbox"/>
<input type="checkbox"/>	location	1	<input checked="" type="checkbox"/>

Sort

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	turned_pro	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	weight_kg	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	height_cm	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	handedness	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	backhand	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	player_id	0	<input type="checkbox"/>
<input type="checkbox"/>	Mesto rođenja	1	<input checked="" type="checkbox"/>

Input	Input Column	Output Alias
Sort	first_name	first_name
Sort	last_name	last_name
Sort	flag_code	flag_code
Sort	Datum rođenja	Datum rođenja
Sort	Godina rođenja	Godina rođenja
Sort	turned_pro	turned_pro
Sort	weight_kg	weight_kg
Sort	height_cm	height_cm
Sort	handedness	handedness
Sort	backhand	backhand

Nakon toga, novonastale podatke sortiram preko player_id kako bih mogla da ih spojim sa sledecom source tabelom.

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass T...
<input type="checkbox"/>	turned_pro	<input checked="" type="checkbox"/>
<input type="checkbox"/>	weight_kg	<input checked="" type="checkbox"/>
<input type="checkbox"/>	height_cm	<input checked="" type="checkbox"/>
<input type="checkbox"/>	handedness	<input checked="" type="checkbox"/>
<input type="checkbox"/>	backhand	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	player_id	<input type="checkbox"/>
<input type="checkbox"/>	birth_country_...	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	Com
player_id	player_id	ascending	1	

Treća source tabela će biti OLE DB kreirana ranking tabela. Biram kolone koje su mi potrebne.

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager

Columns

Error Output

Available External Co...

<input checked="" type="checkbox"/>	Name
<input checked="" type="checkbox"/>	id_rank
<input checked="" type="checkbox"/>	ranking_points
<input checked="" type="checkbox"/>	move_direction

External Column	Output Column
id_rank	id_rank
ranking_points	ranking_points
move_direction	move_direction
player_id	player_id

Zatim i ovu tabelu sortiram preko player_id kako bih mogla da omogućim spajanje tabela.

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass T...
<input type="checkbox"/>	id_rank	<input checked="" type="checkbox"/>
<input type="checkbox"/>	ranking_points	<input checked="" type="checkbox"/>
<input type="checkbox"/>	move_direction	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	player_id	<input type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	Com
player_id	player_id	ascending	1	

Merge join-ijem kao i prethodni put samo na osnovu drugih kolona. Iz ranking tabele kupim id_rank kako bih mogla preko player tabele da joj pristupam.

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type:
Inner join
Swap Inputs

Sort 2

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	id_rank	0	<input type="checkbox"/>
<input type="checkbox"/>	player_id	1	<input checked="" type="checkbox"/>
<input type="checkbox"/>	ranking_points	0	<input type="checkbox"/>
<input type="checkbox"/>	move_direction	0	<input type="checkbox"/>

Sort 3

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	turned_pro	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	weight_kg	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	height_cm	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	handedness	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	backhand	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	player_id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	birth_country_...	0	<input type="checkbox"/>

Input	Input Column	Output Alias
Sort 2	id_rank	id_rank
Sort 3	first_name	first_name
Sort 3	last_name	last_name
Sort 3	flag_code	flag_code
Sort 3	Datum rođenja	Datum rođenja
Sort 3	Godina rođenja	Godina rođenja
Sort 3	turned_pro	turned_pro
Sort 3	weight_kg	weight_kg
Sort 3	height_cm	height_cm
Sort 3	handedness	handedness

Pred sam kraj, izvršila sam Data Conversion transformacije svih kolona koje upisujem u destinacionu tabelu tako da se poklapaju sa navedenim formatima u ddl skriptu za kreiranje player tabele.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

- ☒ Name
- ☒ id_rank
- ☒ first_name
- ☒ last_name
- ☒ flag_code
- ☒ Datum rođenja
- ☒ Godina rođenja
- ☒ turned_pro

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
first_name	Copy of first_name	string [DT_STR]	30			1252 (A)
last_name	Copy of last_name	string [DT_STR]	30			1252 (A)
flag_code	Copy of flag_code	string [DT_STR]	10			1252 (A)
Datum rođenja	Copy of Datum rođe...	string [DT_STR]	20			1252 (A)
Godina rođenja	Copy of Godina rođe...	string [DT_STR]	20			1252 (A)
turned_pro	Copy of turned_pro	string [DT_STR]	10			1252 (A)
weight_kg	Copy of weight_kg	numeric [DT_NUMERIC]		5	0	
height_cm	Copy of height_cm	numeric [DT_NUMERIC]		5	0	
handedness	Copy of handedness	string [DT_STR]	30			1252 (A)
backhand	Copy of backhand	string [DT_STR]	30			1252 (A)
player_id	Copy of player_id	Unicode string [DT_WST...]	20			

Potom sam u OLE DB Destination editoru izabrala kreiranu player tabelu.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager

Mappings

Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:

147.91.175.176.it70g2019.it70g2019 New...

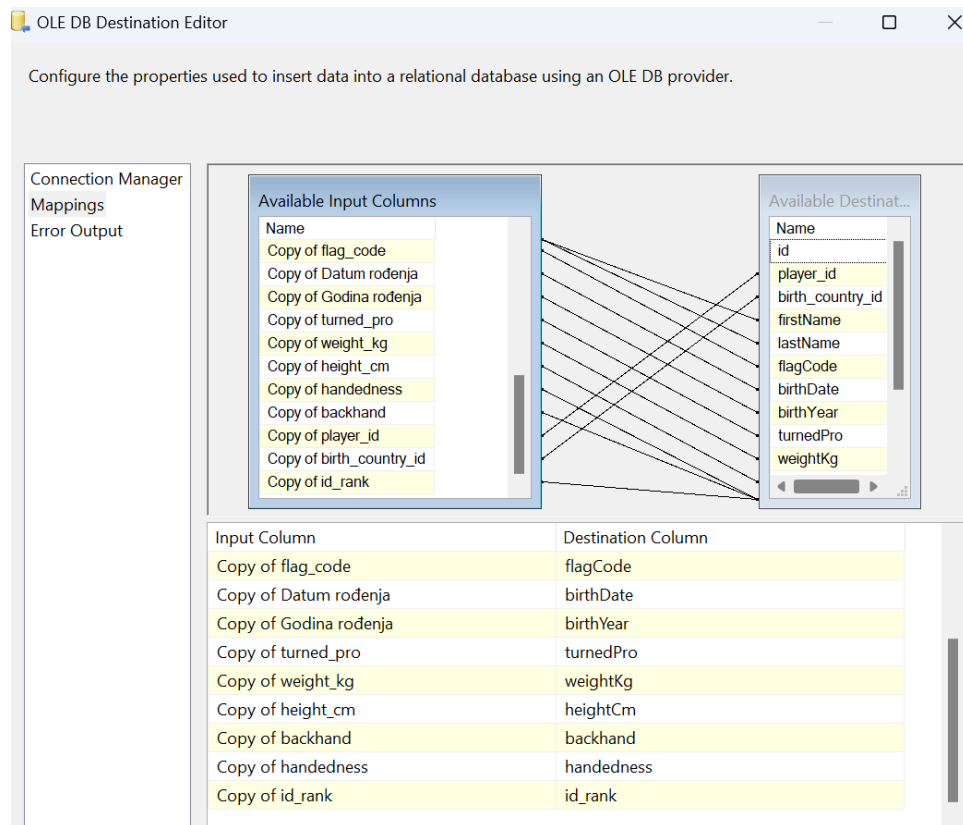
Data access mode:

Table or view - fast load

Name of the table or the view:

[dbo].[player] New...

I na kraju sam izvršila mapiranje kolona, kao što je prikazano na slici:

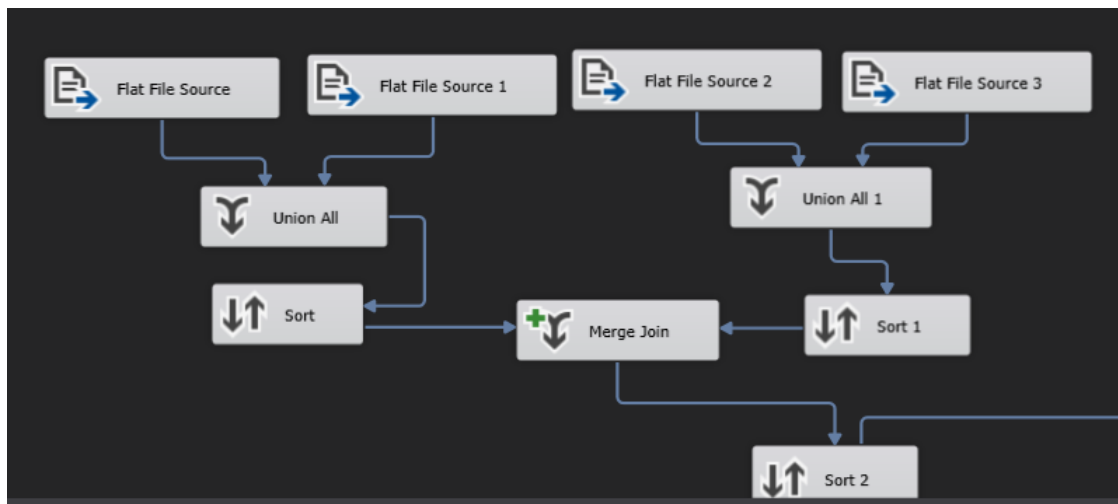


Prikaz tabele player nakon pokretanja Data Flow Task-a:

Results Messages													
	player_id	birth_country_id	id_rank	firstName	lastName	flagCode	birthDate	birthYear	turnedPro	weightKg	heightCm	backhand	handedness
8	a020	1564	20	Mehl	Ampon	USA	11/30/1956	1956		79	181	Two-Handed Backhand	Right-Handed
9	a022	1042	22	Vijay	Amritraj	IND	12/14/1953	1953		67	195	Two-Handed Backhand	Right-Handed
10	a025	2284	25	Matt	Anger	USA	6/20/1963	1963		95	194	Two-Handed Backhand	Right-Handed
11	a027	2028	27	Paul	Annac...	USA	3/20/1963	1963	3/20/1963	76	192	Two-Handed Backhand	Left-Handed
12	a028	2260	28	Alex	Antonit...	AUT	2/8/1966	1966		87	202	One-Handed Backhand	Left-Handed
13	a029	1078	29	Corrado	Aprili	ITA	11/13/1964	1964		67	197	Two-Handed Backhand	Right-Handed
14	a031	470	31	Jimmy	Arias	USA	8/16/1964	1964		72	182	One-Handed Backhand	Left-Handed
15	a035	636	35	Pablo	Arraya	PER	10/21/1961	1961		77	184	One-Handed Backhand	Left-Handed

Unos podataka u tabelu match

Prikazan je Data Flow Task za player tabelu:



Na početku uzimam po dva Flat File Sourca koja imaju iste podatke samo iz drugih vremenskih perioda. Spajam ih preko Union All funkcije:

Union All Transformation Editor

Configure the properties used to merge multiple inputs into one output by creating mappings between columns.

Output Column Name	Union All Input 1	Union All Input 2
match_id	match_id	match_id
match_time	match_time	match_time
winner_aces	winner_aces	winner_aces
winner_double_faults	winner_double_faults	winner_double_faults
winner_first_serves_in	winner_first_serves_in	winner_first_serves_in
winner_break_points_saved	winner_break_points_saved	winner_break_points_saved
winner_service_points_total	winner_service_points_total	winner_service_points_total
winner_break_points_return_total	winner_break_points_return_total	winner_break_points_return_total
winner_return_points_won	winner_return_points_won	winner_return_points_won
winner_total_points_total	winner_total_points_total	winner_total_points_total
loser_aces	loser_aces	loser_aces
loser_double_faults	loser_double_faults	loser_double_faults
loser_first_serves_in	loser_first_serves_in	loser_first_serves_in
loser_break_points_saved	loser_break_points_saved	loser_break_points_saved
loser_service_points_total	loser_service_points_total	loser_service_points_total
loser_return_points_won	loser_return_points_won	loser_return_points_won
loser_break_points_return_total	loser_break_points_return_total	loser_break_points_return_total
loser_total_points_total	loser_total_points_total	loser_total_points_total

Union All Transformation Editor

Configure the properties used to merge multiple inputs into one output by creating mappings between columns.

Output Column Name	Union All Input 1	Union All Input 2
loser_name	loser_name	loser_name
winner_player_id	winner_player_id	winner_player_id
winner_name	winner_name	winner_name
loser_player_id	loser_player_id	loser_player_id
match_id	match_id	match_id
match_score_tiebreaks	match_score_tiebreaks	match_score_tiebreaks
tourney_year_id	tourney_year_id	tourney_year_id
round_order	round_order	round_order
tourney_round_name	tourney_round_name	tourney_round_name

Zatim sortiram po istoj koloni kako bih mogla da spojim tabele.

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type: Inner join Swap Inputs

Sort

	Or...	Joi...
<input checked="" type="checkbox"/> Name		<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> match_id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> match_time	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_aces	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_double_faults	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_first_serves_in	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_break_point...	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_service_poin...	0	<input type="checkbox"/>

Sort 1

	Or...	Joi...
<input checked="" type="checkbox"/> Name		<input type="checkbox"/>
<input checked="" type="checkbox"/> loser_name	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_player_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> winner_name	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> loser_player_id	0	<input type="checkbox"/>
<input type="checkbox"/> match_id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> match_score_ti...	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> tourney_year_id	0	<input type="checkbox"/>

Input	Input Column	Output Alias
Sort	match_id	match_id
Sort	match_time	match_time
Sort	winner_aces	winner_aces
Sort	winner_double_faults	winner_double_faults
Sort	winner_first_serves_in	winner_first_serves_in
Sort	winner_break_points_saved	winner_break_points_saved
Sort	winner_service_points_total	winner_service_points_total
Sort	winner_break_points_return_total	winner_break_points_return_total

Potom sortiram po sledecoj vrednosti kako bih mogla da spojim sa sledecim izvorom podataka.

Sort Transformation Editor

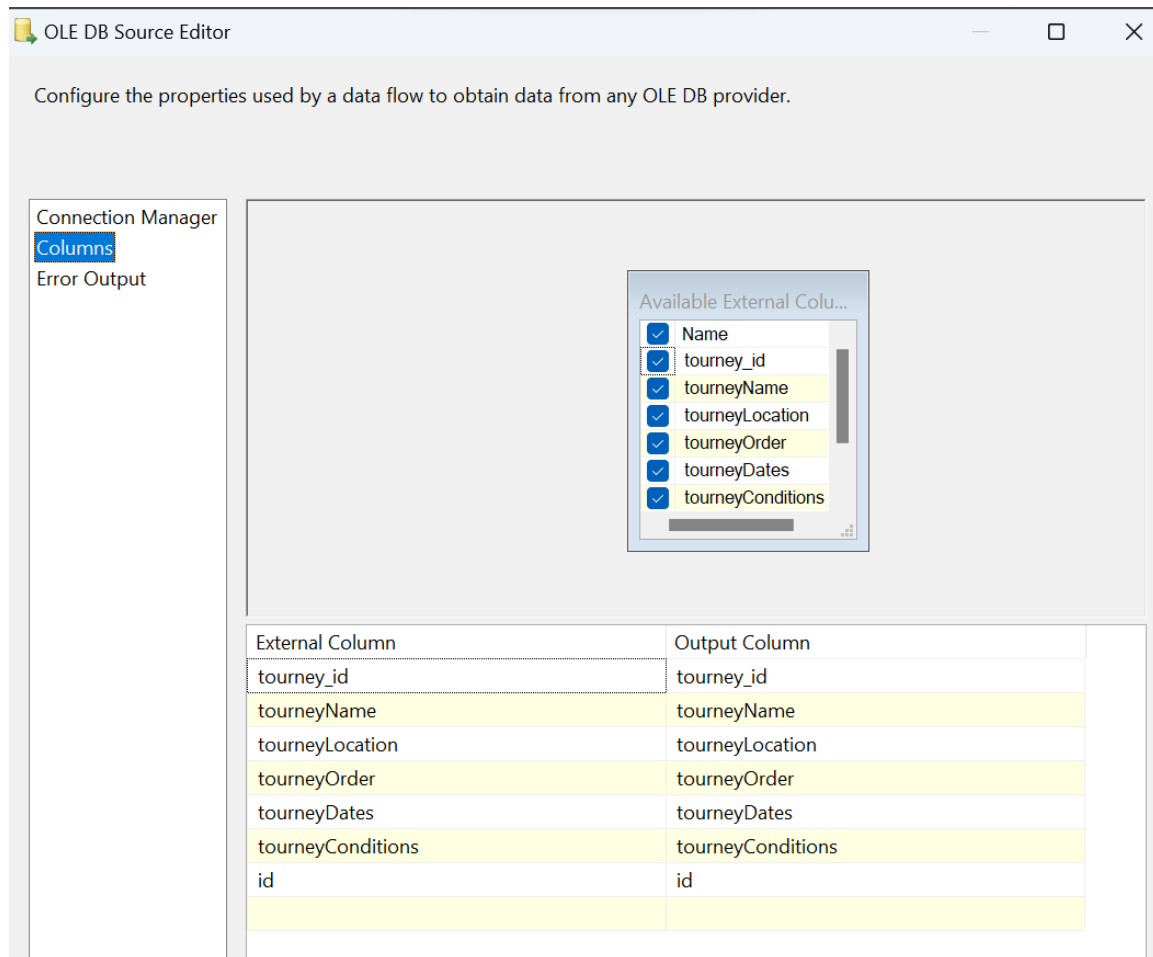
Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

	Pass T...
<input checked="" type="checkbox"/> Name	<input checked="" type="checkbox"/>
<input type="checkbox"/> match_id	<input checked="" type="checkbox"/>
<input type="checkbox"/> match_time	<input checked="" type="checkbox"/>
<input type="checkbox"/> winner_aces	<input checked="" type="checkbox"/>
<input type="checkbox"/> winner_double_faults	<input checked="" type="checkbox"/>
<input type="checkbox"/> winner_first_serves_in	<input checked="" type="checkbox"/>
<input type="checkbox"/> winner_break_points_saved	<input checked="" type="checkbox"/>
<input type="checkbox"/> winner_service_points_total	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	Com
tourney_year_id	tourney_year_id	ascending	1	

Sledeći izvor podataka je OLE DB Source kreirane tabele tournament. Sortiram i tu tabelu po tourney_id kako bih uspela da korisim merge join.



Prilikom spajanja tabela biram da se samo ključ tabele prenese, bez ostalih obeležja.

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type: Inner join Swap Inputs

Sort 2

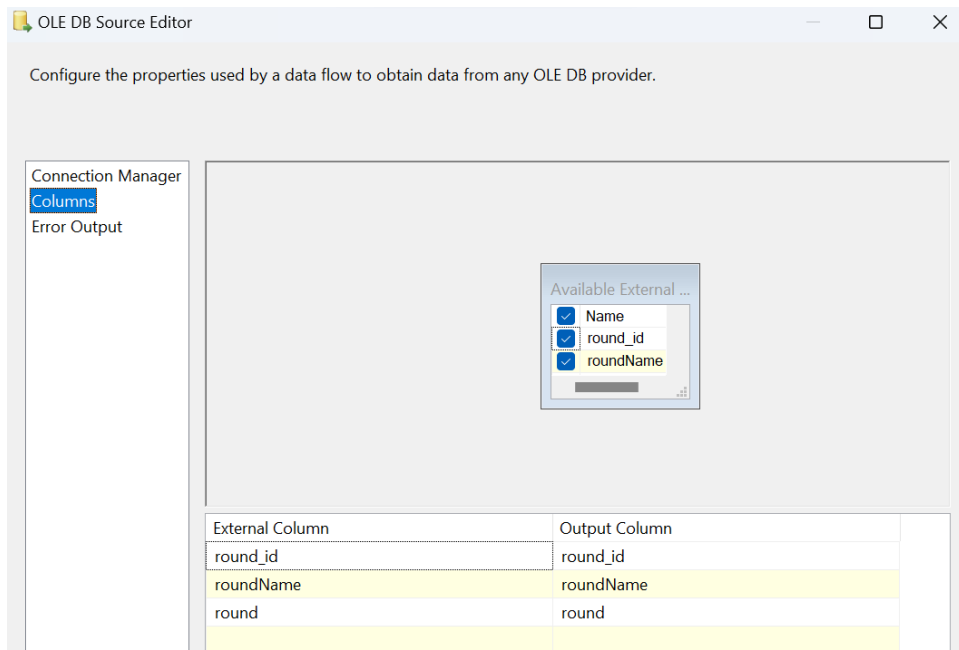
<input checked="" type="checkbox"/>	Name	Or...	Joi...
<input checked="" type="checkbox"/>	winner_player_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	winner_name	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	loser_player_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	match_score_tiebre...	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	tourney_year_id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	round_order	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	tourney_round_name	0	<input type="checkbox"/>

Sort 3

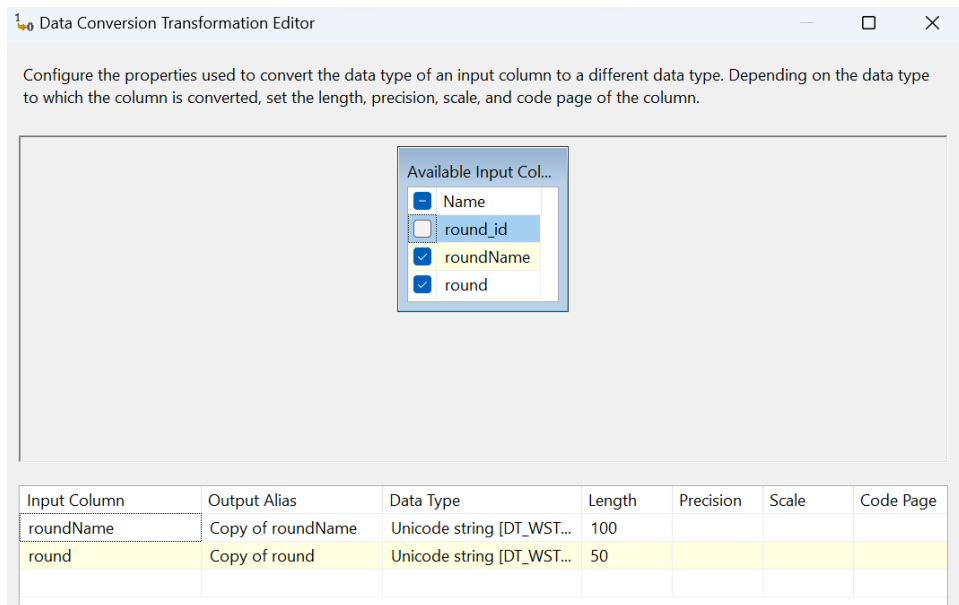
<input checked="" type="checkbox"/>	Name	Or...	Joi...
<input type="checkbox"/>	tourney_id	1	<input checked="" type="checkbox"/>
<input type="checkbox"/>	tourneyName	0	<input type="checkbox"/>
<input type="checkbox"/>	tourneyLocation	0	<input type="checkbox"/>
<input type="checkbox"/>	tourneyOrder	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	tourneyDates	0	<input type="checkbox"/>
<input type="checkbox"/>	tourneyCondit...	0	<input type="checkbox"/>

Input	Input Column	Output Alias
Sort 2	match_id	match_id
Sort 2	match_time	match_time
Sort 2	winner_aces	winner_aces
Sort 2	winner_double_faults	winner_double_faults
Sort 2	winner_first_serves_in	winner_first_serves_in
Sort 2	winner_break_points_saved	winner_break_points_saved
Sort 2	winner_service_points_total	winner_service_points_total
Sort 2	winner_break_points_return_total	winner_break_points_return_total
Sort 2	winner_return_points_won	winner_return_points_won
Sort 2	winner_total_points_total	winner_total_points_total

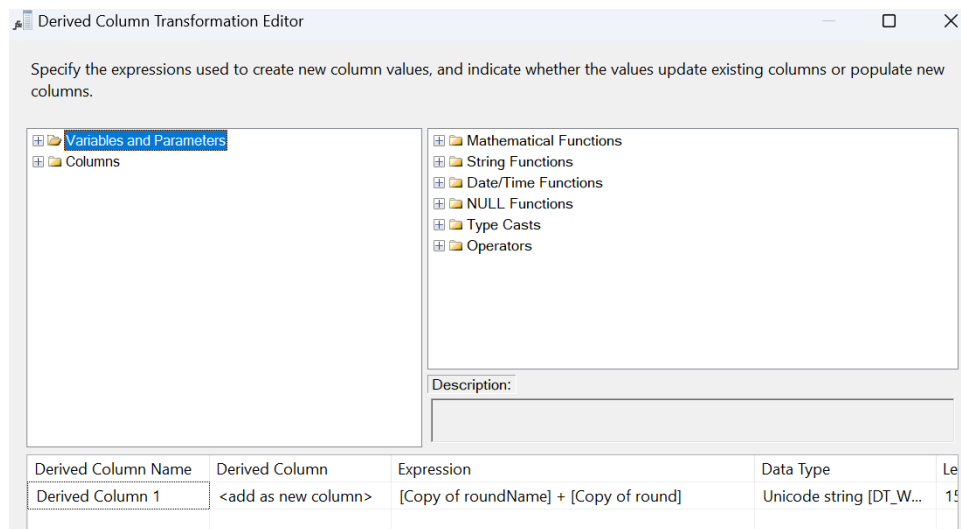
Sledeći izvor je takođe OLE DB Source, kreirane tabele round. Biram potrebne kolone.



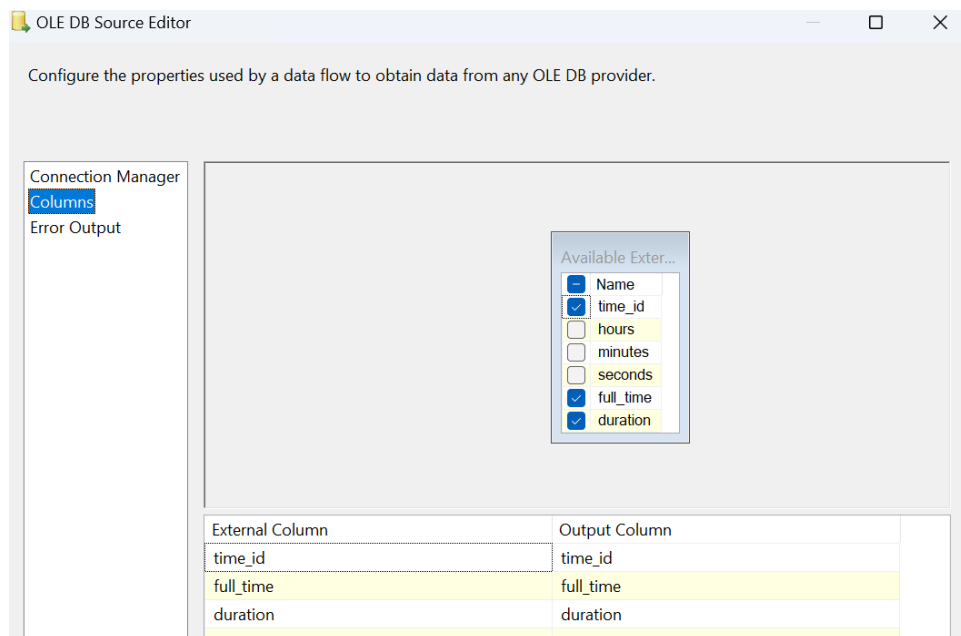
Konvertujem obeležene kolone kako bih mogla da ih spajam.



Pomoću Derived Column sam spojila naziv runde i numeričku oznaku iste. Potom sam sortirala tabelu preko dobijene kolone i na osnovu nje ponovo odradila merge join.



Uzimam OLE DB Source kako bih pristupila kreiranoj tabeli time. Biram potrebne kolone.



Vršim konverziju izabrane kolone i zatim sortiram po njenoj vrednosti.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input ...
☒ Name
☐ time_id
☒ full_time
☐ duration

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
full_time	Copy of full_time	string [DT_STR]	50			1252 (ANSI)

Spajam tabele i uzimam samo ključ time tabele kako bih kasnije mogla dajoj pristupam.

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type:

Sort 6

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	time_id	0	<input type="checkbox"/>
<input type="checkbox"/>	full_time	0	<input type="checkbox"/>
<input type="checkbox"/>	Copy of full_time	1	<input checked="" type="checkbox"/>

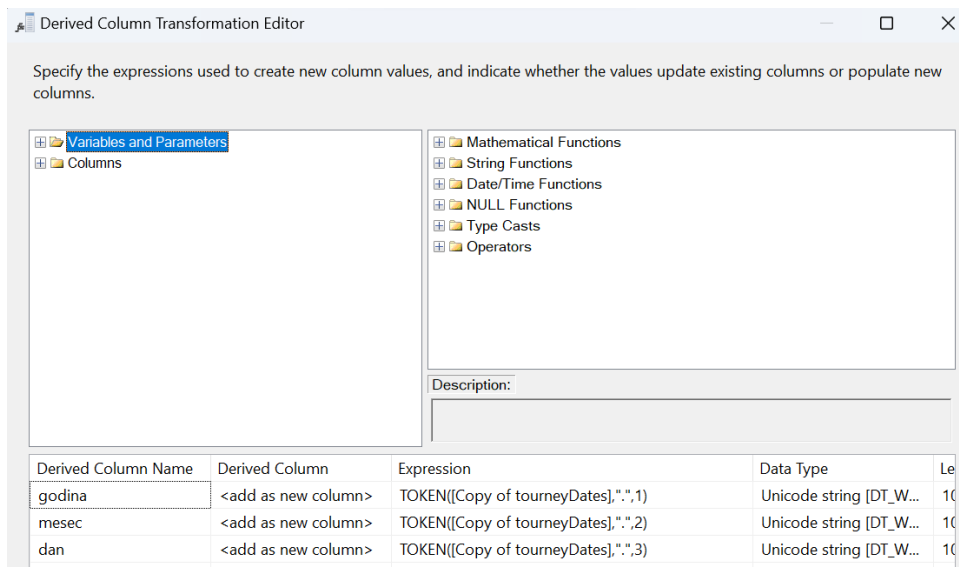
Sort 7

	Name	Or...	Joi...
<input checked="" type="checkbox"/>	match_id	0	<input type="checkbox"/>
<input type="checkbox"/>	match_time	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	winner_aces	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	winner_double_faults	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	winner_first_serves_in	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	winner_break_point...	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	winner_service_poin...	0	<input type="checkbox"/>

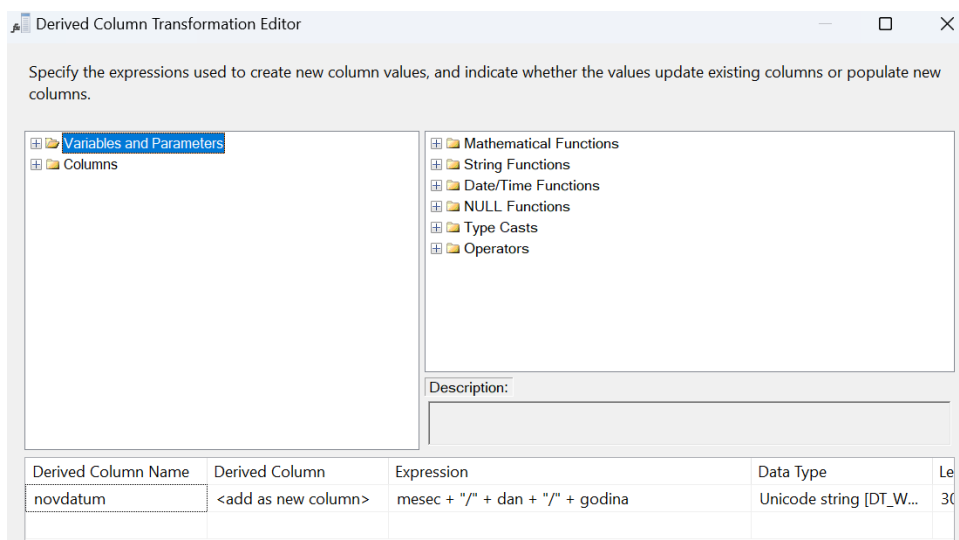
Input	Input Column	Output Alias
Sort 6	time_id	time_id
Sort 7	match_id	match_id
Sort 7	winner_aces	winner_aces
Sort 7	winner_double_faults	winner_double_faults
Sort 7	winner_first_serves_in	winner_first_serves_in
Sort 7	winner_break_points_saved	winner_break_points_saved
Sort 7	winner_service_points_total	winner_service_points_total
Sort 7	winner_break_points_return_total	winner_break_points_return_total
Sort 7	winner_return_points_won	winner_return_points_won
Sort 7	winner total points total	winner total points total

Isti postupak se ponavlja i za date dimenziju. Podaci se konvertuju u isti tip, zatim sortiraju kako bi bilo moguće njihovo spajanje. Potom iz date tabele kupim ključ tabele. Jedina razlika je što datumi u dve tabele nisu bili u istim formatima. To je rešeno na sledeći način:

Korićenjem Derived Column-a sam prvo korićenjem TOKEN-a razdvojila datum na kolone koje predstavljaju odvojene delove – godinu, mesec i dan.



Zatim sam dodala kosu crtu i odredila kojim redosledom se ubacuju prethodno kreirane kolone da bih dobila format datuma koji je korišćen i u date dimenziji.



Na samom kraju ostalo je samo da se konvertuje u tražene tipove podataka.

1 Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

☐ Name

☒ time_id

☒ match_id

☒ winner_aces

☒ winner_double_faults

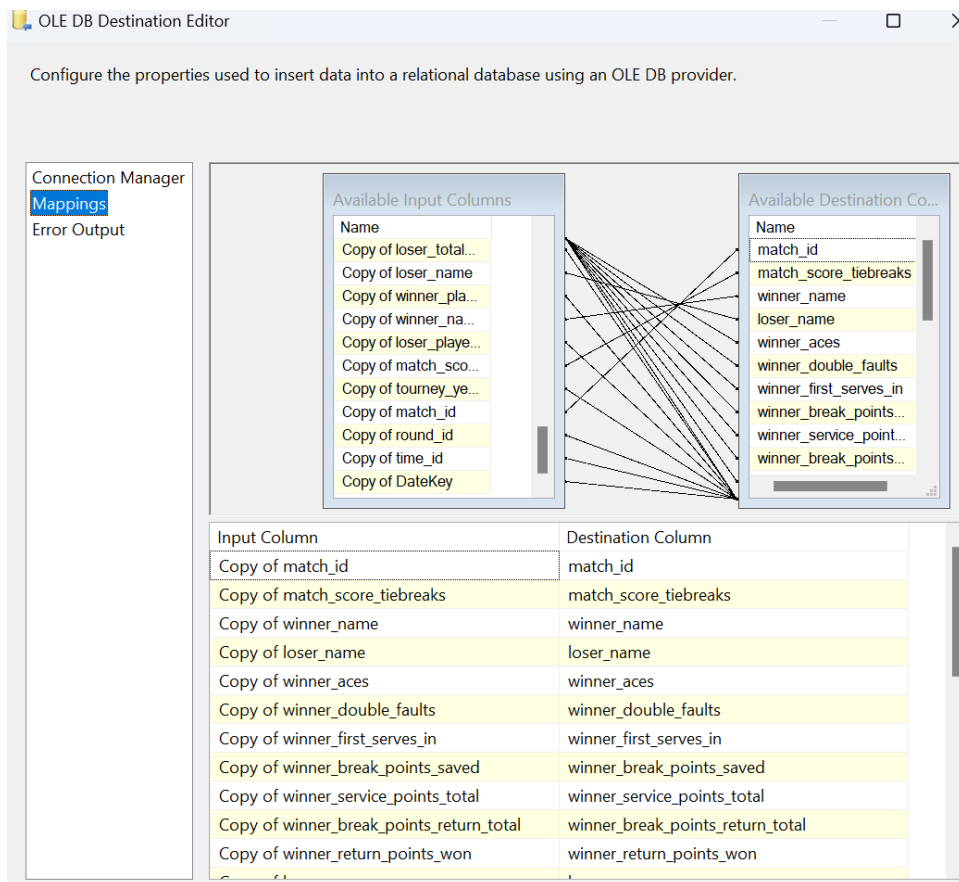
☒ winner_first_serves_in

☒ winner_break_points_saved

☒ winner_service_points_total

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
winner_aces	Copy of winner_aces	four-byte signed intege...				
winner_double_faults	Copy of winner_dou...	four-byte signed intege...				
winner_first_serves_in	Copy of winner_first_...	four-byte signed intege...				
winner_break_points_...	Copy of winner_brea...	four-byte signed intege...				
winner_service_points...	Copy of winner_servi...	four-byte signed intege...				
winner_break_points_...	Copy of winner_brea...	four-byte signed intege...				
winner_return_points_...	Copy of winner_retur...	four-byte signed intege...				
winner_total_points_t...	Copy of winner_total...	four-byte signed intege...				
loser_aces	Copy of loser_aces	four-byte signed intege...				
loser_double_faults	Copy of loser_double...	four-byte signed intege...				
loser_first_serves_in	Copy of loser_first_se...	four-byte signed intege...				

Odabrala sam koristeći OLE DB Destination tabelu koju sam prethodno kreirala i odradila mapiranje koje je prikazano na slici.



Prikaz tabele match nakon pokretanja Data Flow Task-a:

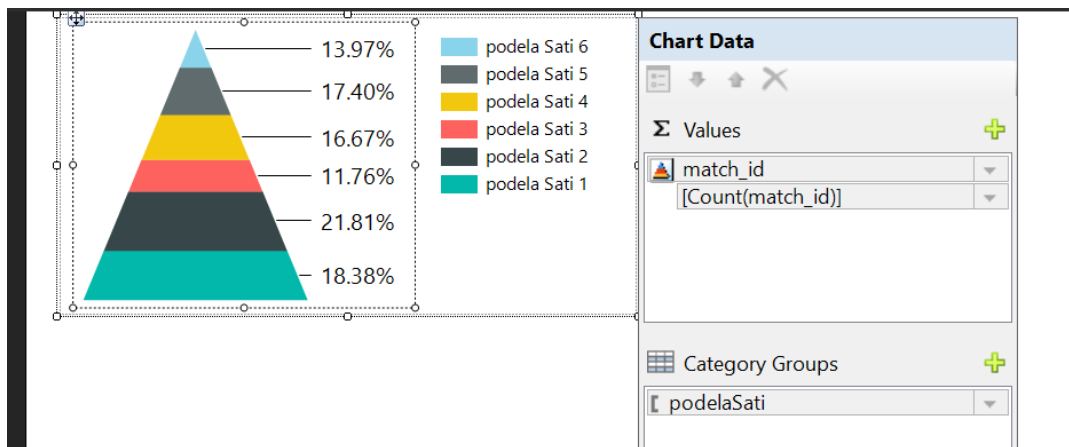
	match_id	date_id	tourney_id	round_id	time_id	winner_id	loser_id	match_score_tiebreaks
1	2012-96-f324-d683	20120725	2012-96	40	257	f324	d683	36 76(5) 1917
2	2012-96-b837-g628	20120725	2012-96	29	65	b837	g628	64 64
3	2012-96-d402-s694	20120725	2012-96	33	89	d402	s694	64 63
4	2012-96-r485-k966	20120725	2012-96	33	90	r485	k966	75 64
5	2012-96-n289-d616	20120725	2012-96	33	64	n289	d616	63 61
3	2012-96-p624-l797	20120725	2012-96	33	64	p624	l797	76(5) 61
7	2012-96-d683-sa93	20120725	2012-96	29	93	d683	sa93	63 76(2)

winner_break_points_saved	winner_service_points_total	winner_break_points_return_total	winner_return_points_won	winner_total_points_won	loser_aces	loser_dou
5	174	13	63	366	11	5
0	57	7	20	117	5	1
5	63	8	29	133	6	1
0	53	13	36	143	2	2
0	41	12	38	106	3	3
5	74	2	20	129	4	2
10	76	5	30	152	6	2

6. Prikaz izveštaja kao odgovora na pitanja

Procenat mečeva koji su trajali duže od 3,5H

Prvi izveštaj koristi podatke iz time dimenzije. Odabrala sam kategorije i vrednosti koje želim da se prikažu kroz izabrani chart.



Napravljeno je posebno polje korišćenjem Fields u DataSet properties-u.

Dataset Properties

Query
Fields
Options
Filters
Parameters

Change query and calculated fields.

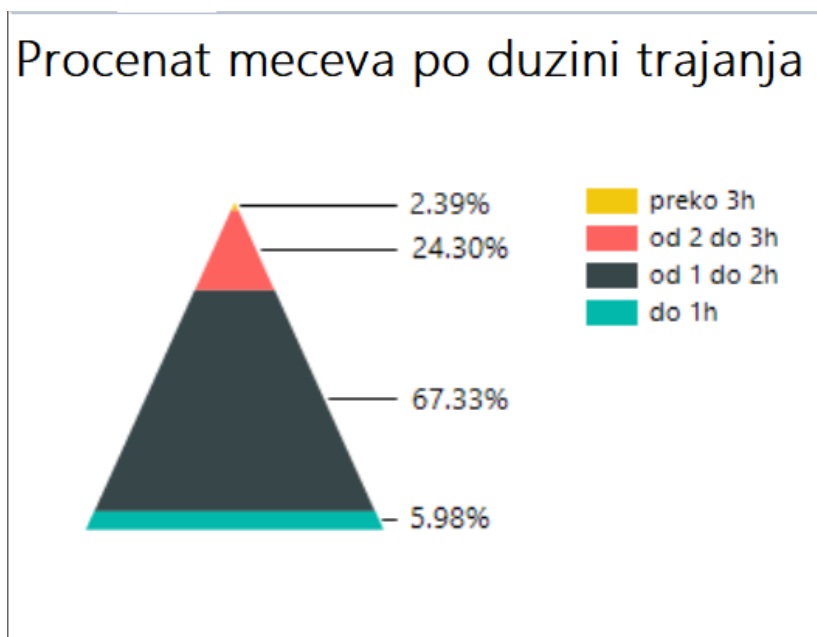
Add Delete Up Down

Field Name	Field Source
full_time	full_time
duration	duration
match_id	match_id
podelaSati	«Expr»

Kako bih izdvojila vreme trajanja meča koristila sam funkciju switch i podelila trajanje filma na sledeće kategorije:

```
Expression
Set expression for: Value
=Switch(
Fields!duration.Value<60,"do 1h",
Fields!duration.Value>=60 and Fields!duration.Value <120 ,"od 1 do 2h",
Fields!duration.Value>=120 and Fields!duration.Value <180 ,"od 2 do 3h",
Fields!duration.Value>=180, "preko 3h",
True, "Terminated")
```



Nakon pokretanja izveštaj izgleda ovako. Možemo zaključiti da najveći broj mečeva traje između 1-2 časa.



Pregled statistike tri najbolje odigrana meča novaka djokovica

Kreirala sam zabelu sa željenim kolonama i formatirala njen izgled kao na slici. U ovom primeru korišćena su parametarska polja.

Player's first name:	<input type="text"/>		
Player's last name:	<input type="text"/>		

Statistika 3 najbolje odigrana meča za izabranog igrača:						
[firstName][lastName]						
	aces	total points won	double faults	first serves in	break points saved	return points won 
winner aces	er total points	inner double faults	ner first serves	er break points saved	er return points	

Prilikom kreiranja parametra biram data type, dajem naziv parametru i upisujem tekst koji želim da se prikazuje korisniku nakon pokretanja izveštaja.

Report Parameter Properties

General

Available Values

Default Values

Advanced

Change name, data type, and other options.

Name:

ReportParameter2

Prompt:

Player's last name:

Data type:

Text

☐ Allow blank value ("")

☐ Allow null value

☐ Allow multiple values

Select parameter visibility:

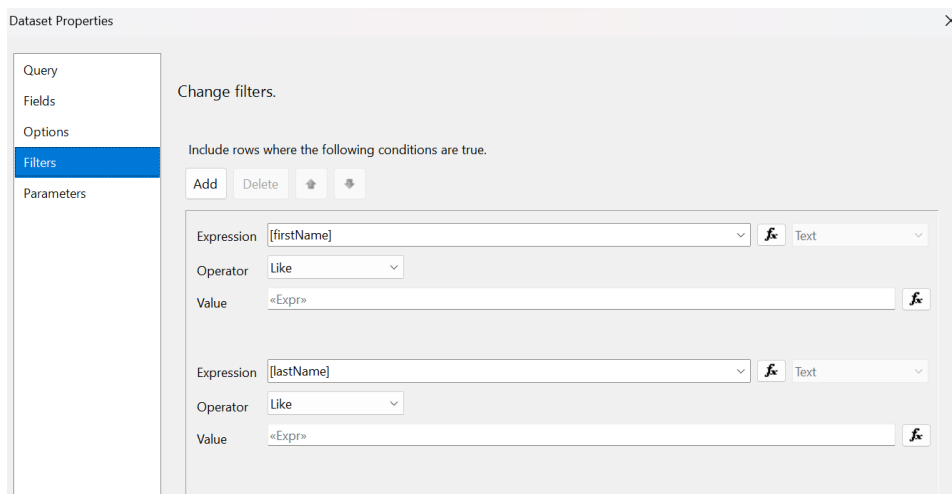
☒ Visible

☐ Hidden

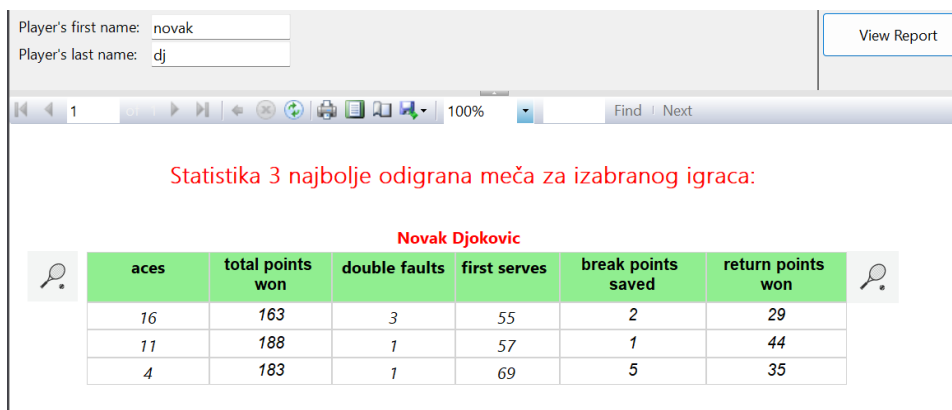
☐ Internal

Zatim u DataSet Properties-u dodajem dva filtera za dva napravljena parametra.

Operacija za koju se odlučujem je like.



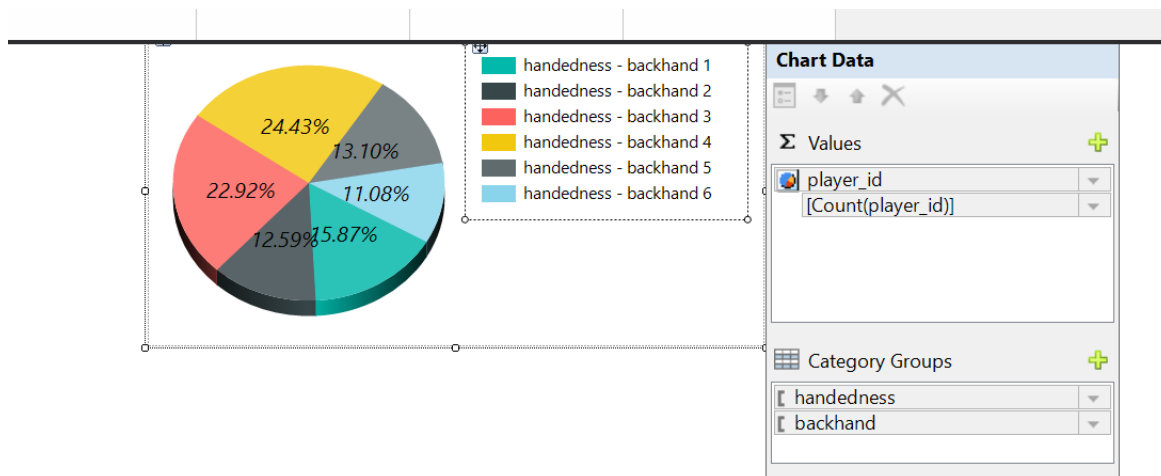
Prikaz kako izveštaj izgleda nakon pokretanja i unosa potrebnih parametara.



Na kraju su dodate slike i textBox koji prikazuje puno ime i prezime igrača čiju statistiku posmatramo.

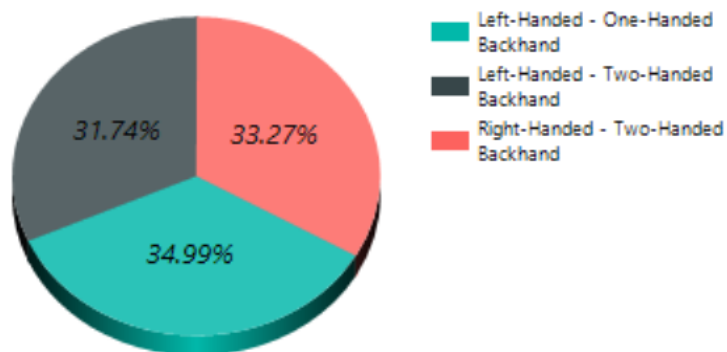
Procenat od ukupnog broja igrača onih koji su backhand: Two-Handed i handedness: Left

U narednom izveštaju korišćen je pie chart. Odabrani su podaci za prikaz kao na slici:



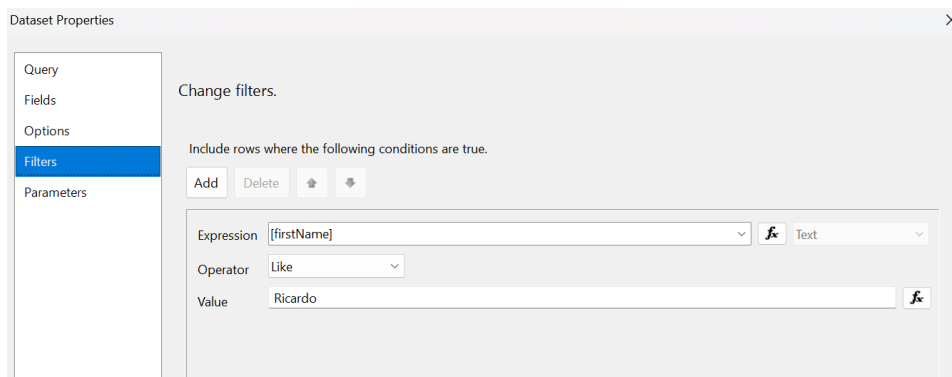
Nakon pokretanja izveštaja prikaz je ovakav:

Procenat od ukupnog broja igrača:



Svi igrači sa imenom Ricardo kako su rangirani na ATP listi

S obzirom da sam primetila da ima igrača koji se isto zovu, u ovom slučaju čak 6 različitih ljudi, odlučila sam se za ovaj izveštaj. Odradila sam filter na osnovu imena kako bih izlistala samo željenje igrače.



Dataset Properties

Query
Fields
Options
Filters
Parameters

Change filters.

Include rows where the following conditions are true.

Add Delete

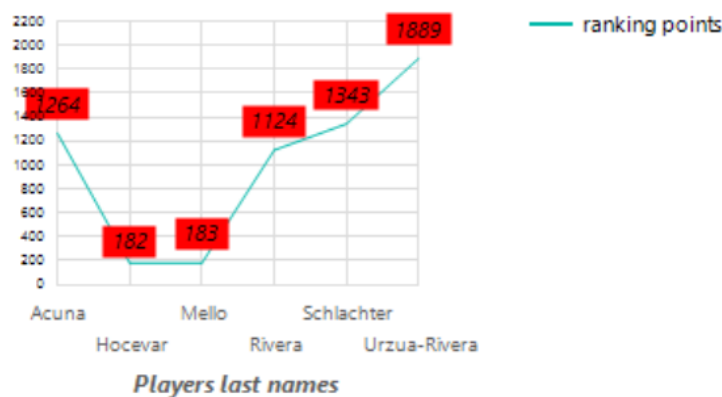
Expression: [firstName] Text

Operator: Like

Value: Ricardo

Izgled izveštaja nakon sređivanja dataLabel properties-a:

Rangiranje na ATP listi svih igrača pod imenom Ricardo



Poređenje broja igrača koji igraju za svaku od država i broj igrača koji se od njih smatraju profesionalcima

Za kreiranu tabelu odradila sam grupisanje na nivou flagCode-a.

Izvestaj [ExecutionTime]

	[flagCode]	
	first Name	last Name
	[firstName]	[lastName]
	turned Pro	[turnedPro]
	Od ukupnog broja igrača za ovu	Profesionalnih

Row Groups

(table1_flagCode)

(table1_Details_Group)

Column Groups

Ubacila sam Header koji prilikom pokretanja ispisuje datum pokretanja izveštaja.

Placeholder Properties

General

Number

Alignment

Font

Action

Change label, value, and markup options.

Label:

Value:

[ExecutionTime]

Tooltip:

Markup type

Izvestaj [<i>&ExecutionTime</i>]			
	[flagCode]		
	first Name	last Name	turned Pro
	[firstName]	[lastName]	[turnedPro]
	Od ukupnog broja igraca za ovu		Profesionalnih

Row Groups




(table1_flagCode)	
(table1_Details_Group)	


Column Groups

Ubacila sam dodatno polje koje će prebrojavati različite ID- igrača I stavljati vrednost u posebno polje. Isto je odrađeno i za prebrojavanje onih koji imaju vrednost datuma u Pro koloni.

Od ukupnog broja igraca za ovu drzavu: [CountDistinct(player_id)]	Profesionalnih
---	----------------

Za redove unutar grupe je podešen row visibility tako da je konačan izgled izveštaja sledeći:

	ARG	
	ARM	
	first Name	last Name turned Pro
	Sargis	Sargsian
Od ukupnog broja igrača za ovu državu: 1		Profesionalnih igrača ima: 0
	AUS	
	first Name	last Name turned Pro
	Trevor	Allan
	Terry	Addison
	Mal	Anderson
	Johan	Anderson

	Luke	Saville
	Akira	Santillan
	Desmond	Tyson
	Peter	Thrupp
	Peter	Tramacchi 11/8/1970
	Michael	Tebbutt 12/22/1970
	Bernard	Tomic
	Mark	Verryth
	Gavin	Van Peperzeel
	Aleksandar	Vukic
	Laurie	Warder
	Greg	Whitecross
	Mark	Woodforde
	Todd	Woodbridge 4/2/1971
	Andrew	Whittington
	Simon	Youl
Od ukupnog broja igrača za ovu državu: 123		Profesionalnih igrača ima: 19
	AUT	

Prikaz turnira po lokacijama sa grupisanjem na osnovu uslova turnira

U poslednjem izveštaju odrađeno je grupisanje unutar grupisanja. Odabrala sam kolone i formatirala tabelu. Zatim sam dodala Header sa vrednostima ukupnog broja strana i trenutne strane.

Ukupan broj strana izvestaja: [&TotalPages] Trenutna strana: [&PageNumber]	
[tourneyLocation]	
Name of tourney	Order
[tourneyConditions]	
[tourneyName]	[tourneyOrder]

Row Groups

(table1_tourneyLocation)

(table1_tourneyConditions)

(table1_Details_Group)

Column Groups

Dodala sam mogućnost interaktivnog sortiranja za sledeću kolonu:

Text Box Properties

General

Number

Alignment

Font

Border

Fill

Visibility

Interactive Sorting

Action

Change interactive sort options for the text box.

☒ Enable interactive sorting on this text box

Choose what to sort:

☒ Detail rows

☐ Groups



Sort by:

[tourneyOrder]

☒ Apply this sorting to all groups and data regions in:

DataSet1

Izveštaj nakon pokretanja izgleda ovako:

Wellington		
Name of tourney	Order	
Outdoor		
Wellington	2	
Washington		
Name of tourney	Order	
Indoor		
Washington-2	18	
Washington WCT	25	
Washington Indoor WCT	29	
Outdoor		
Washington WTC	48	

7. Zaključak

Ovaj dokument sadrži detaljno opisan način izrade projekta iz Projektovanja Skladišta Podataka. Opisani su motivacija i cilj izrade projekta, korisnički zahtevi, njihova ispunjenost, detaljan ETL proces kao i proces kreiranja izveštaja za korisničke zahteve.

Podaci koji se čuvaju u skladištu uvek će biti dostupni za poredjenje sa onim koji nadolaze u budućnosti, zarad praćenja poboljšanja, kao i za dalji razvoj ovog skladišta podataka.