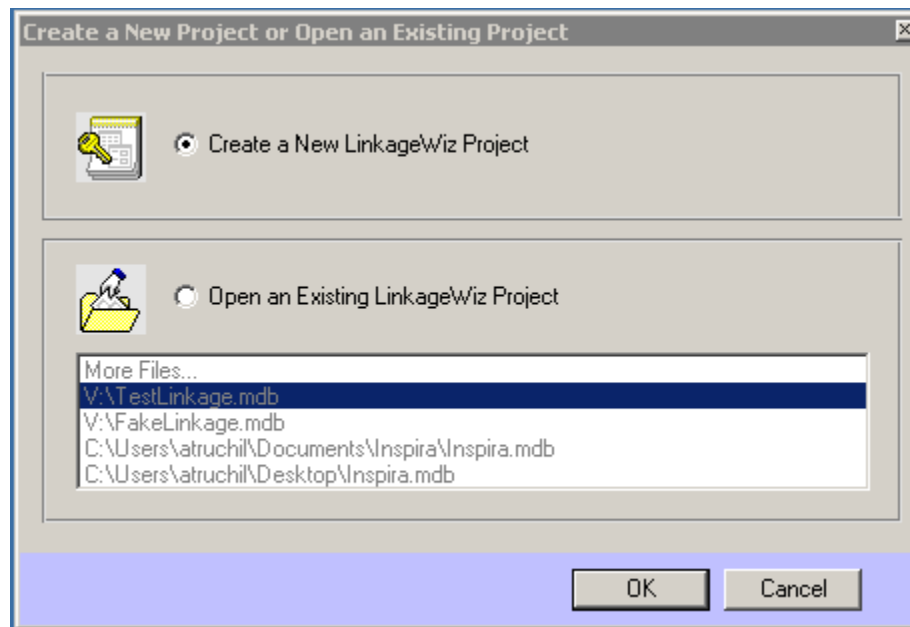
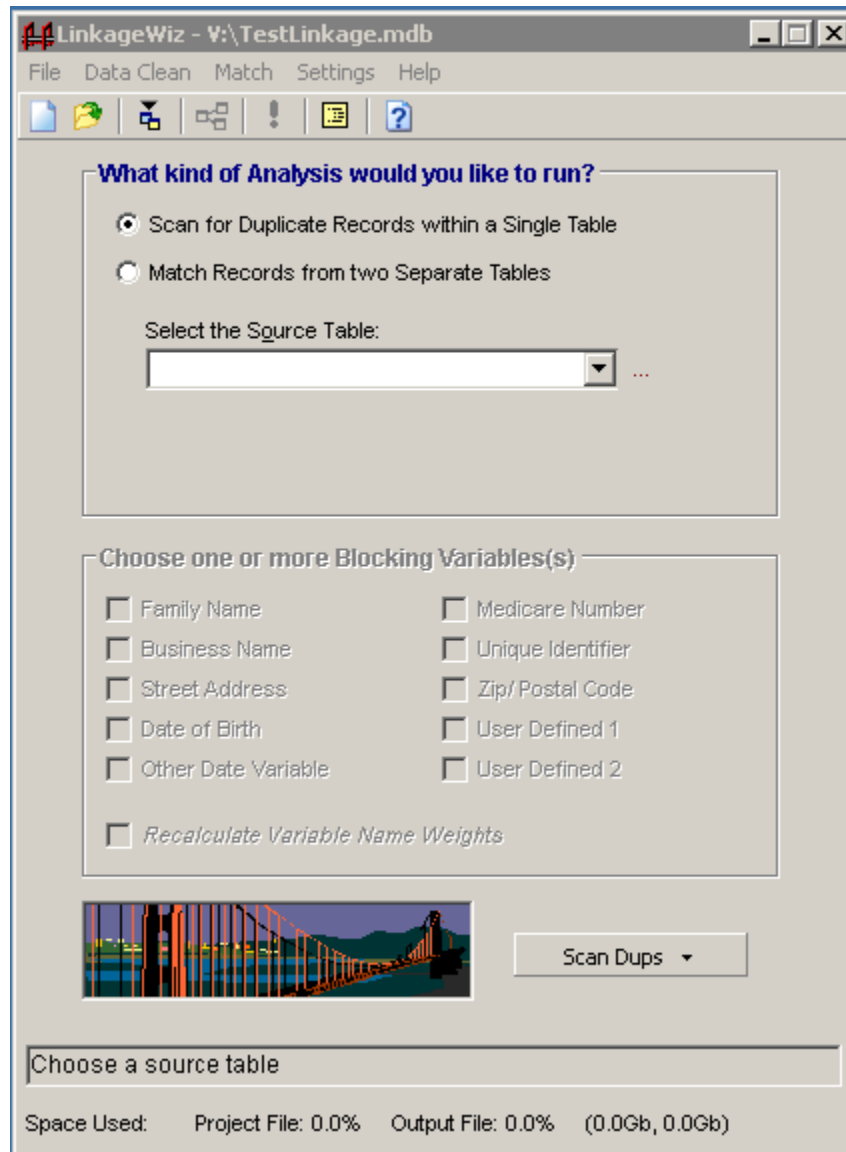


1) Create a new project



## 2) Import Data

You have the option to “Scan for duplicates in a single table” or “Match records from two separate tables”. For this example, we will choose “Scan for duplicates in a single table”.



The screenshot shows the LinkageWiz application window titled "LinkageWiz - V:\TestLinkage.mdb". The menu bar includes "File", "Data Clean", "Match", "Settings", and "Help". The toolbar contains icons for file operations, a funnel, a match icon, a warning icon, a list icon, and a help icon.

The main dialog box is titled "What kind of Analysis would you like to run?". It contains two radio button options:

- Scan for Duplicate Records within a Single Table
- Match Records from two Separate Tables

Below these options is a label "Select the Source Table:" followed by a dropdown menu and an ellipsis button "...".

The next section is titled "Choose one or more Blocking Variables(s)". It contains a grid of checkboxes for various variables:

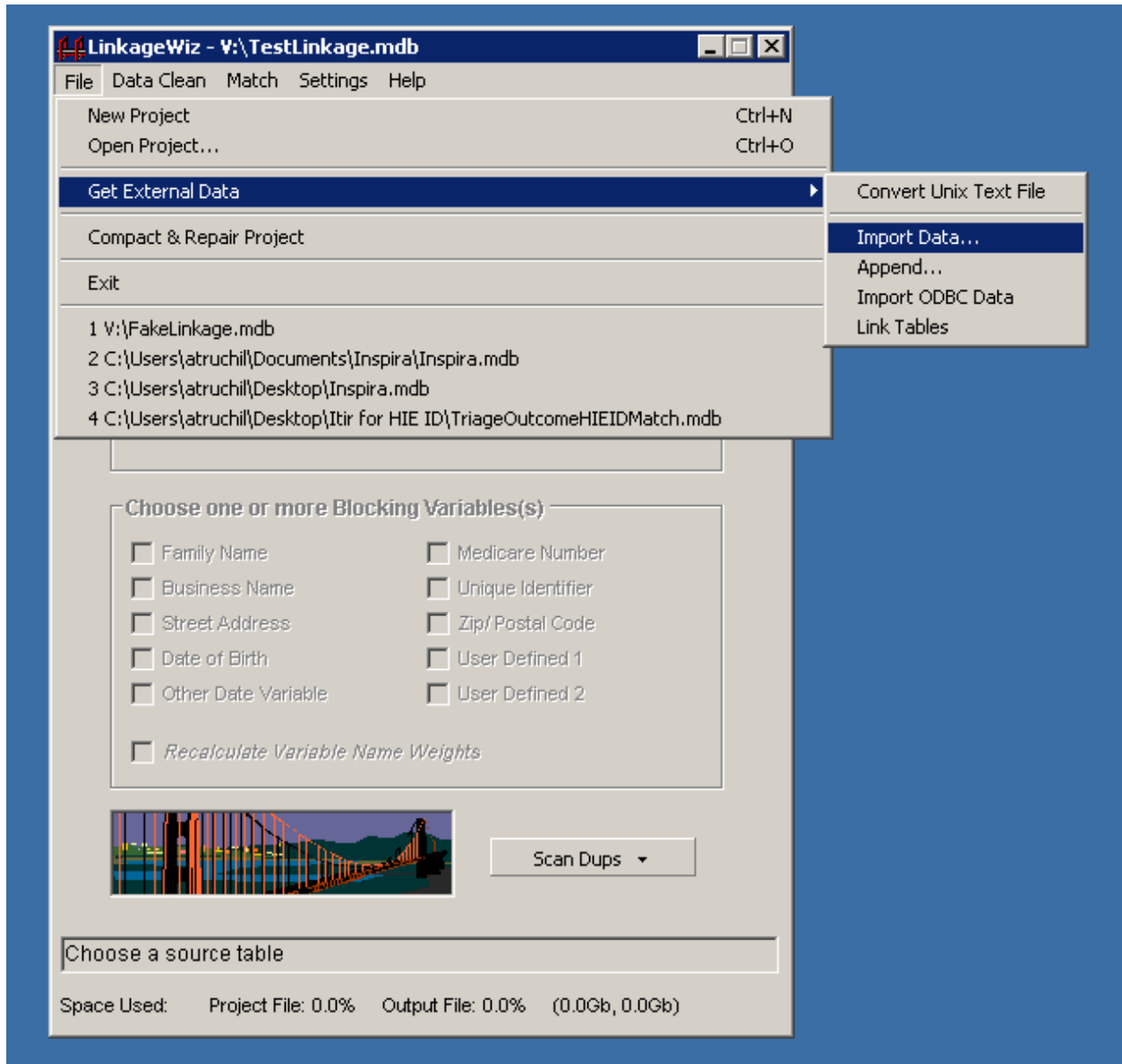
<input type="checkbox"/> Family Name	<input type="checkbox"/> Medicare Number
<input type="checkbox"/> Business Name	<input type="checkbox"/> Unique Identifier
<input type="checkbox"/> Street Address	<input type="checkbox"/> Zip/Postal Code
<input type="checkbox"/> Date of Birth	<input type="checkbox"/> User Defined 1
<input type="checkbox"/> Other Date Variable	<input type="checkbox"/> User Defined 2
<input type="checkbox"/> Recalculate Variable Name Weights	

At the bottom of the dialog, there is a small image of a suspension bridge and a button labeled "Scan Dups" with a dropdown arrow.

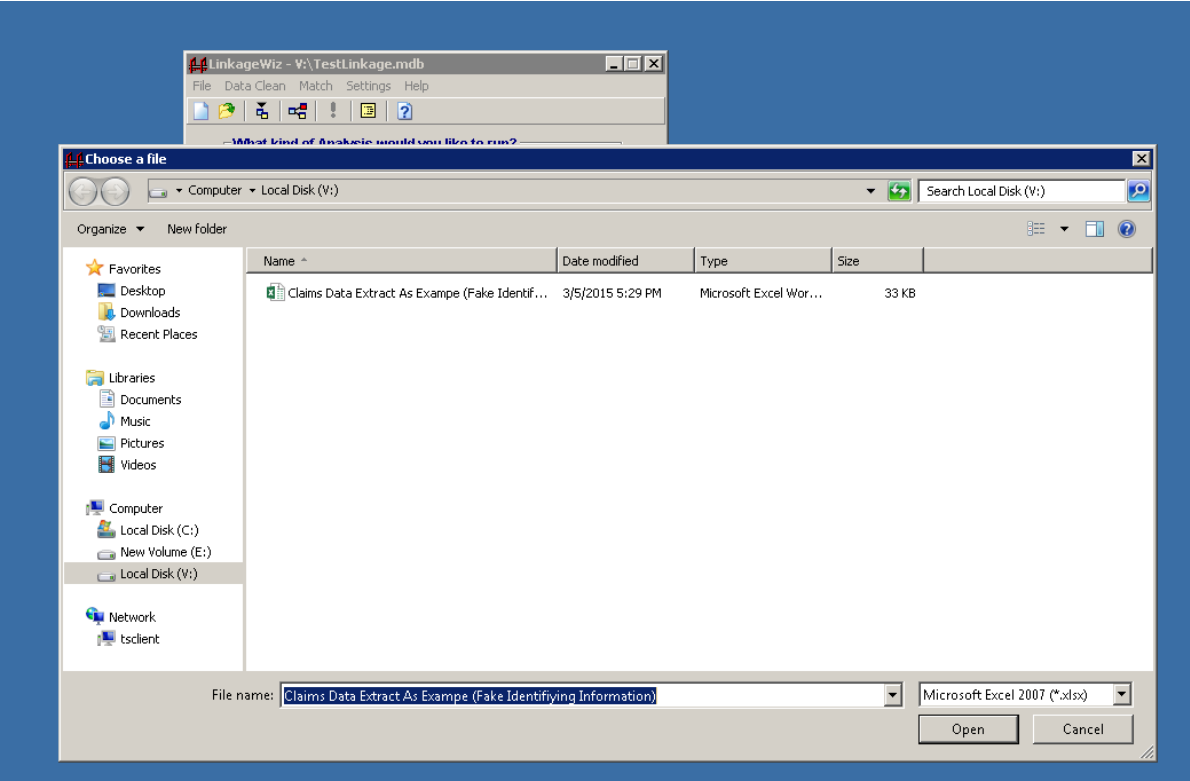
Below the dialog is a text box labeled "Choose a source table".

At the very bottom, there is a status bar showing "Space Used: Project File: 0.0% Output File: 0.0% (0.0Gb, 0.0Gb)".

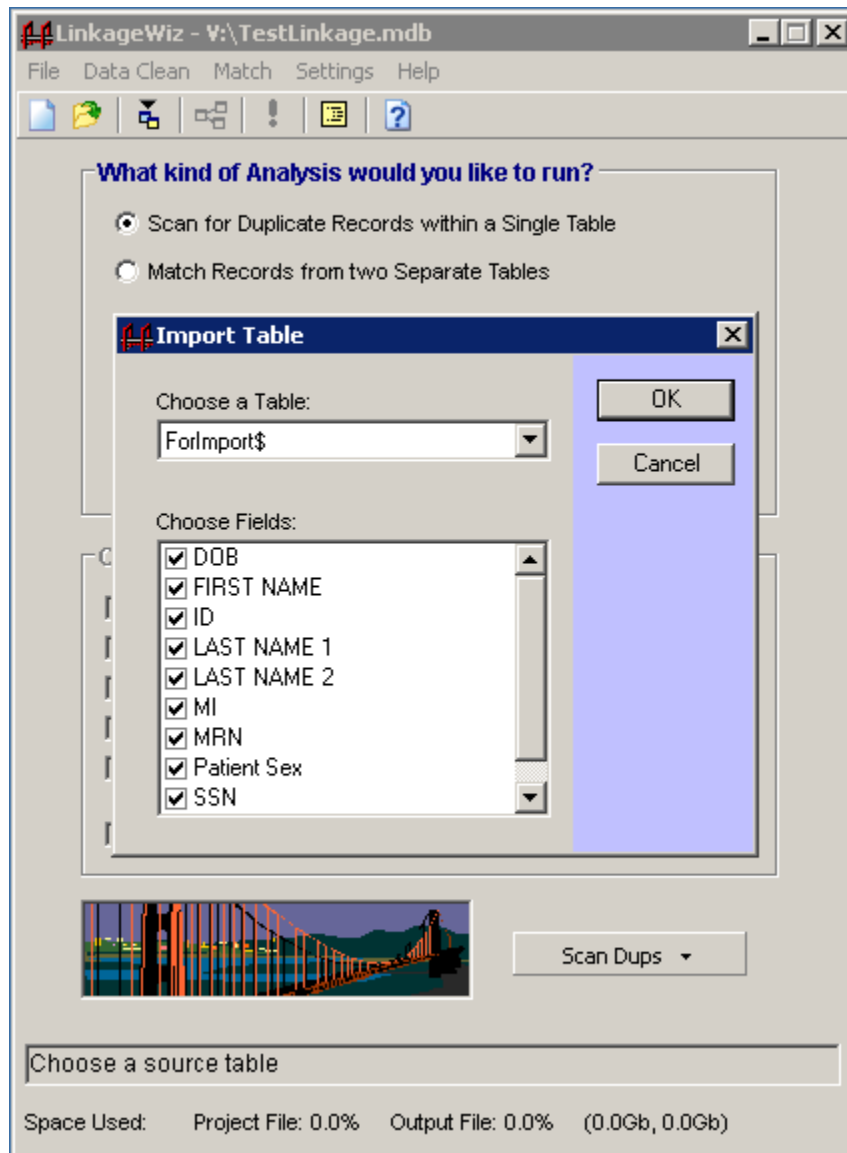
To import your table, navigate to File>Get External Data> Import Data...



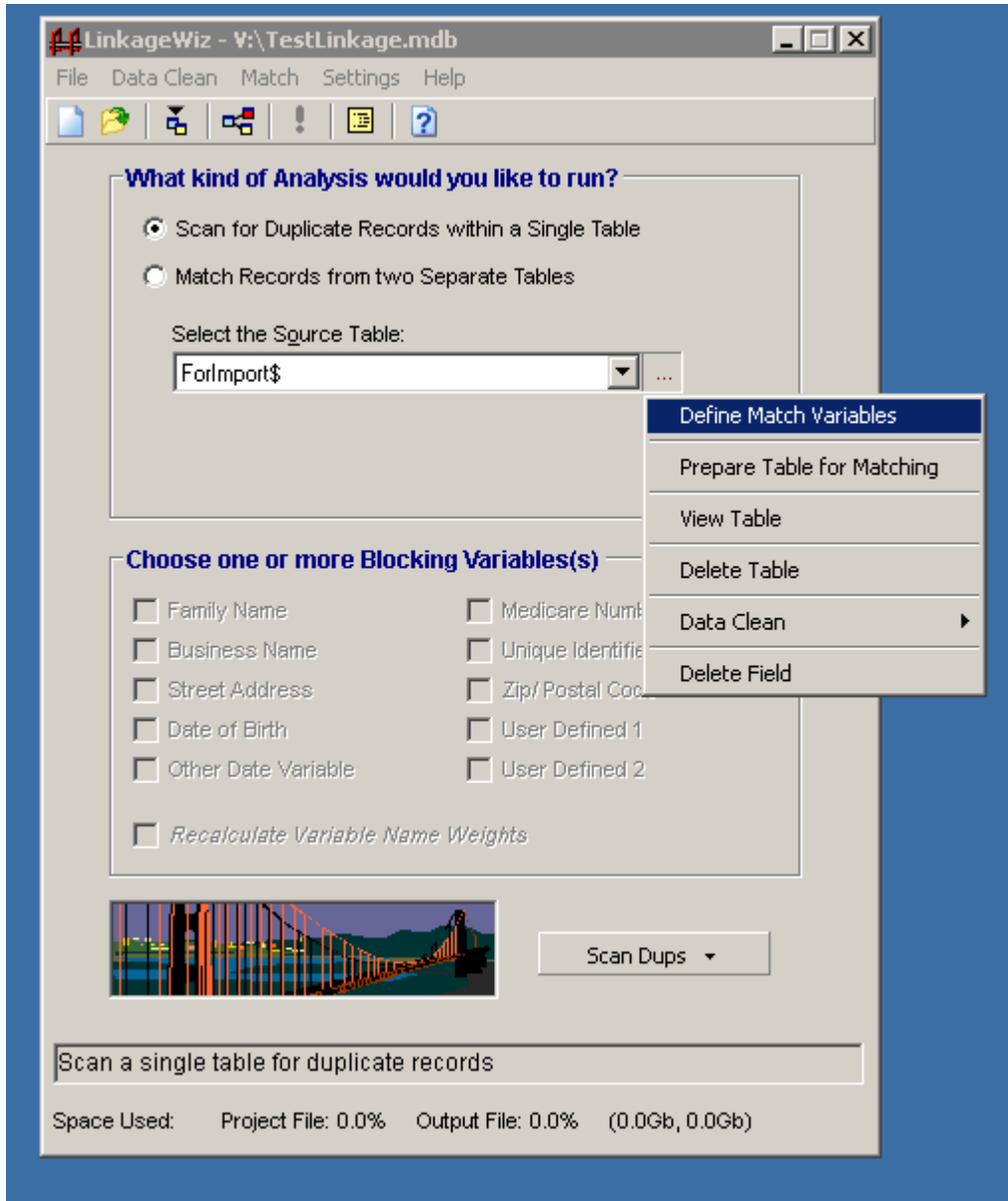
Navigate to your data:



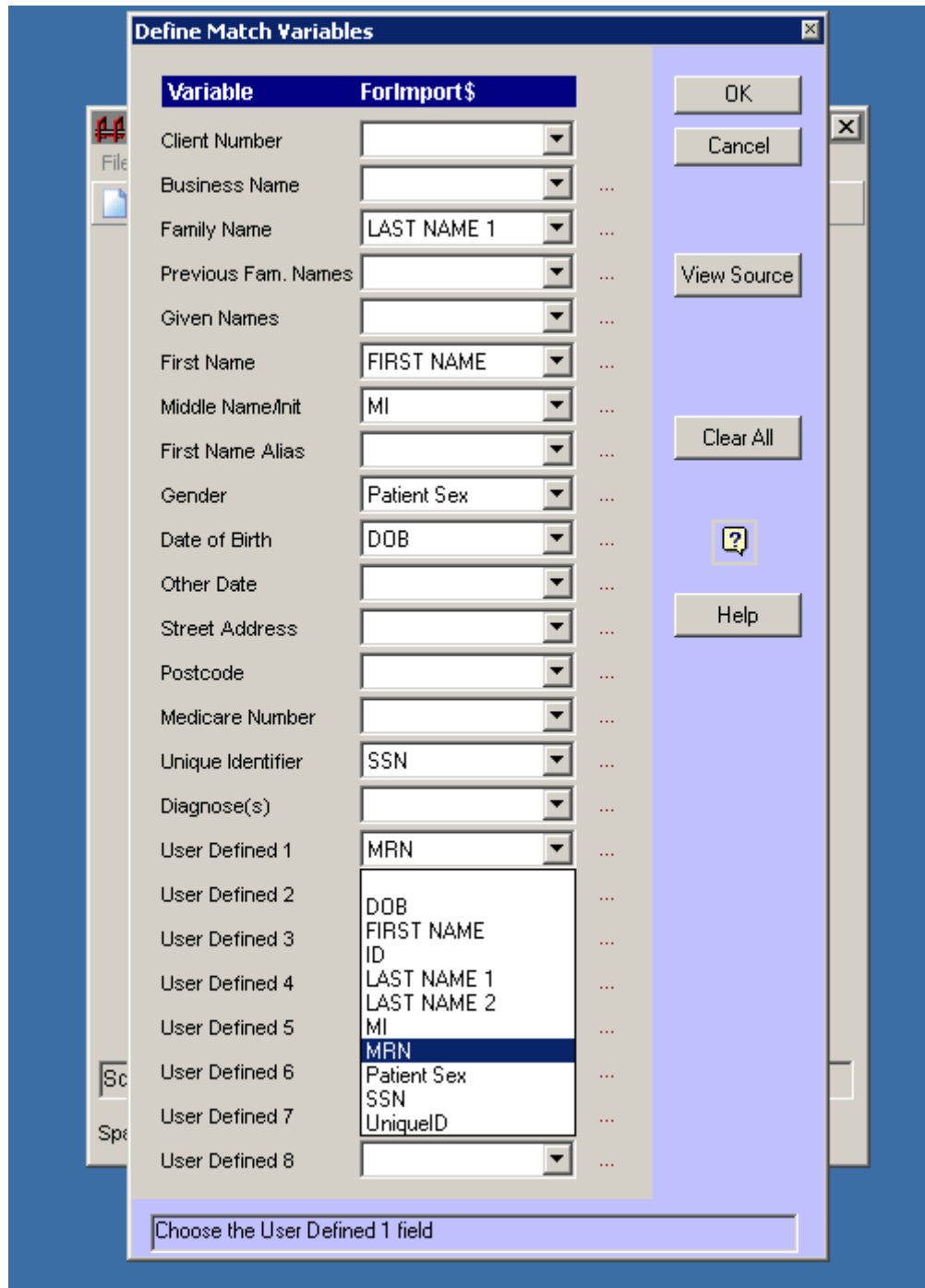
Choose the variables you will use for linking. Import all fields if you prepared your table as specified in (see previous section on preparing data).



3) Navigate to Define Match Variables



Assign variables that correspond to preset variables in LinkageWiz. You may need to add categories for unique identifiers used by your hospital system. Considering your population's demographics, decide which part of the patient's last name you will put in the primary Family Name category and which will be user defined.



To define weights for match variables click on the ellipses next to each variable. You may leave the default weights for those variables that fall into standard categories, but you must choose weights for user defined variables. For your first run, the default weights are generally adequate. You will most likely perform the entire linking process multiple times, tweaking the weights to increase match accuracy.

The image shows a software interface for defining match variables. It consists of a list of variables on the left, each with a dropdown menu and an ellipsis button to its right. The variables and their current values are:

Variable	Value	Action
Business Name	[Dropdown]	...
Family Name	LAST NAME 1	...
Previous Fam. Names	[Dropdown]	...
Given Names	[Dropdown]	...
First Name	FIRST NAME	...
Middle Name/Init	MI	...
First Name Alias	[Dropdown]	...

On the right side of the interface, there are two buttons: "View Source" and "Clear All".



When defining weights for each variable, assess the risks and cases for false positives and false negatives. If certain variables disagree between records, should the match score be penalized or unaffected?

The screenshot shows a software interface for defining match weights. The main window is titled "Define Match Variables" and contains a sub-dialog titled "Set Match Weight for Date of Birth".

**Choose a variable:**

- Family Name
- Business Name
- Given Names
- Gender
- Date of Birth**
- Other Date
- Address
- Identifiers
- User Defined
- Other

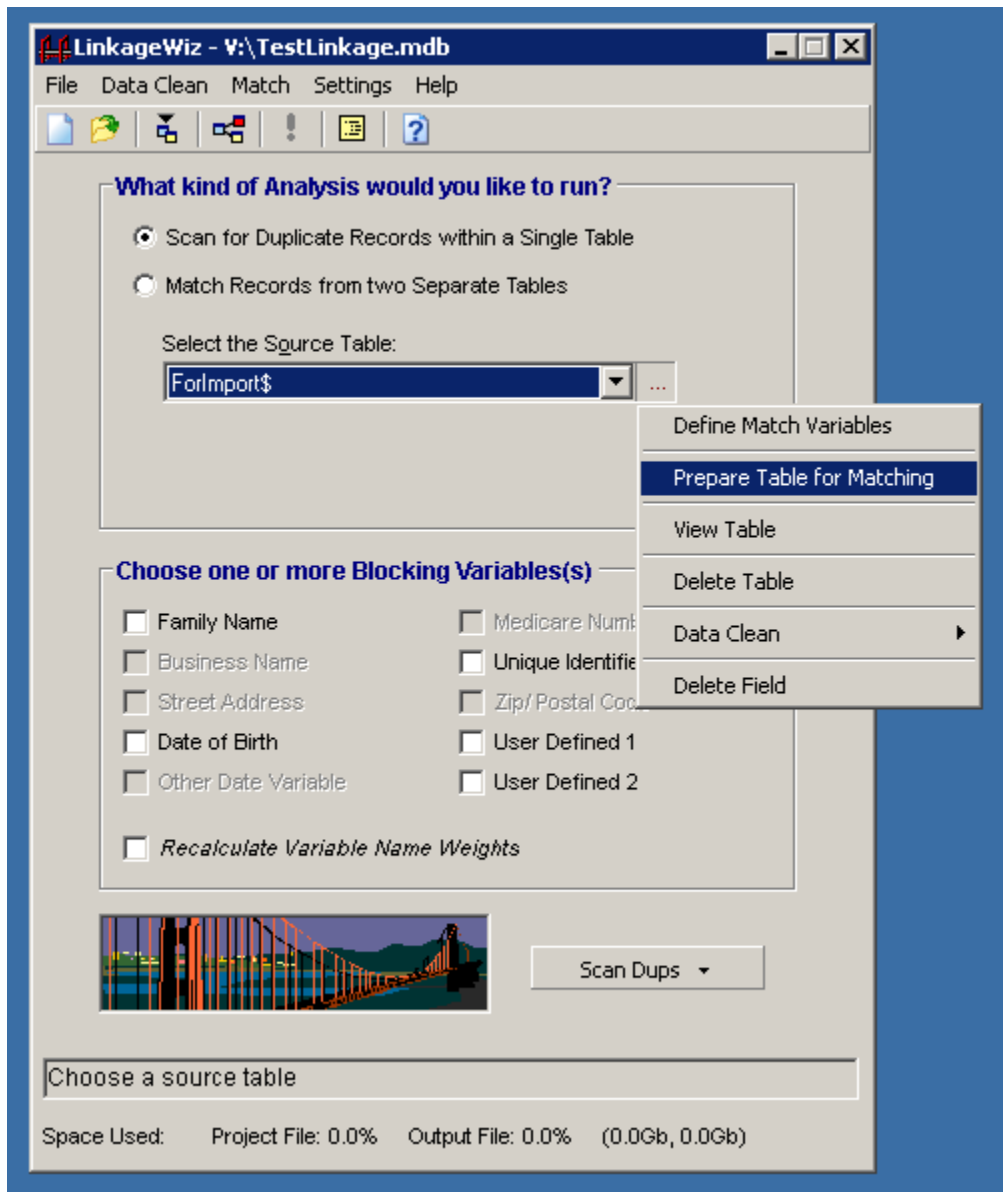
**Linkage Weight:**

Dates of Birth agree (exact):	10.0
Months of Birth agree:	2.5
Months of Birth disagree:	-2.0
Days of Birth agree:	3.0
Days of Birth disagree:	-2.0
Day and month of birth are transposed:	2.0
Years of Birth agree:	4.0
Years of Birth differ by 1 yr :	-0.5
Years of Birth differ by 2-3 yrs:	-2.5
Years of Birth differ by 4-6 yrs:	-3.0
Years of Birth differ by 7-10 yrs:	-4.0
Years of Birth differ by more than 10 yrs:	-9.0
Day of Birth = 01 and Month of Birth = 01	0.0
Date of Birth Missing	0.0

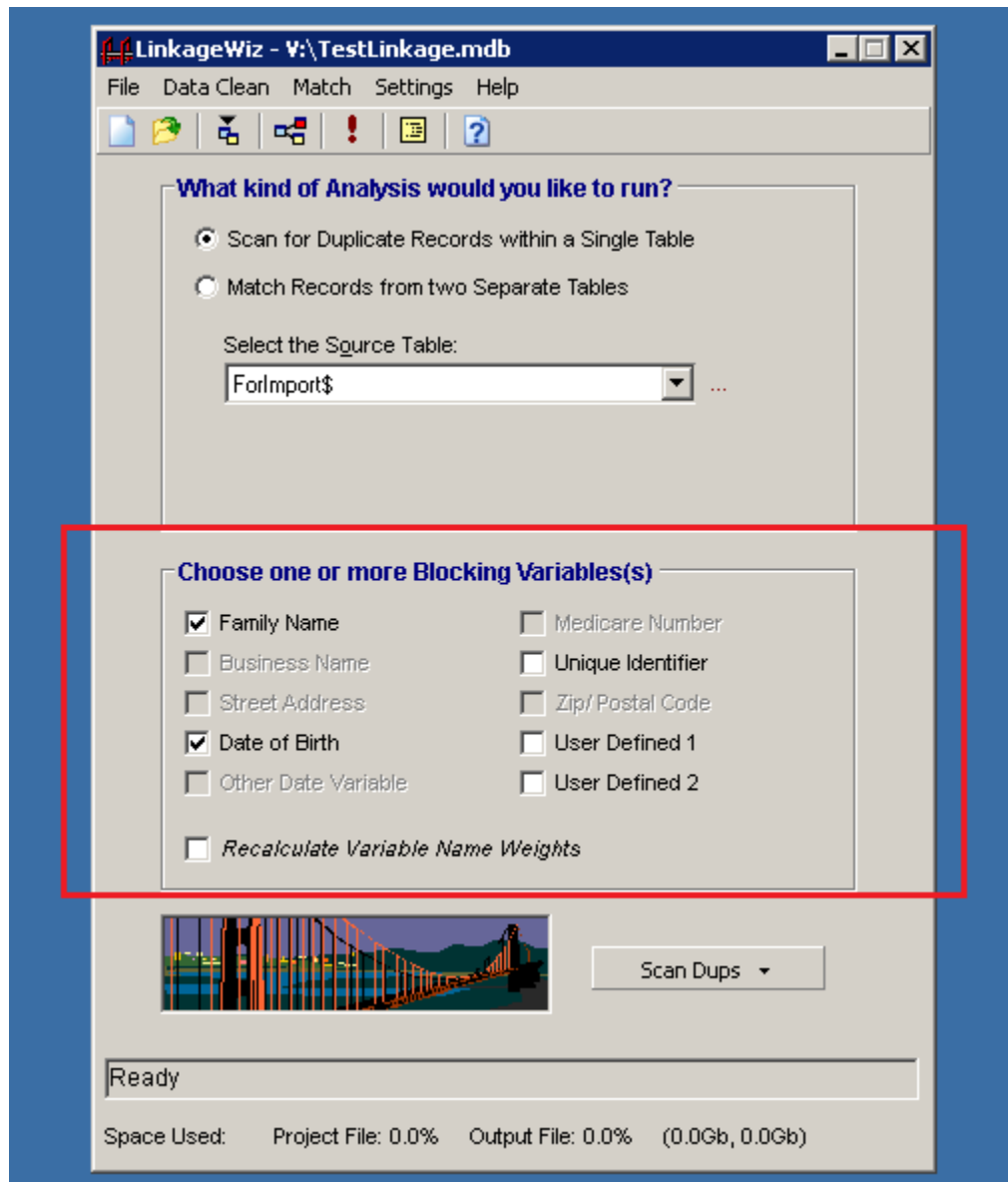
Buttons: OK, Cancel, Set Default

Footer: Choose the User Defined 2 field

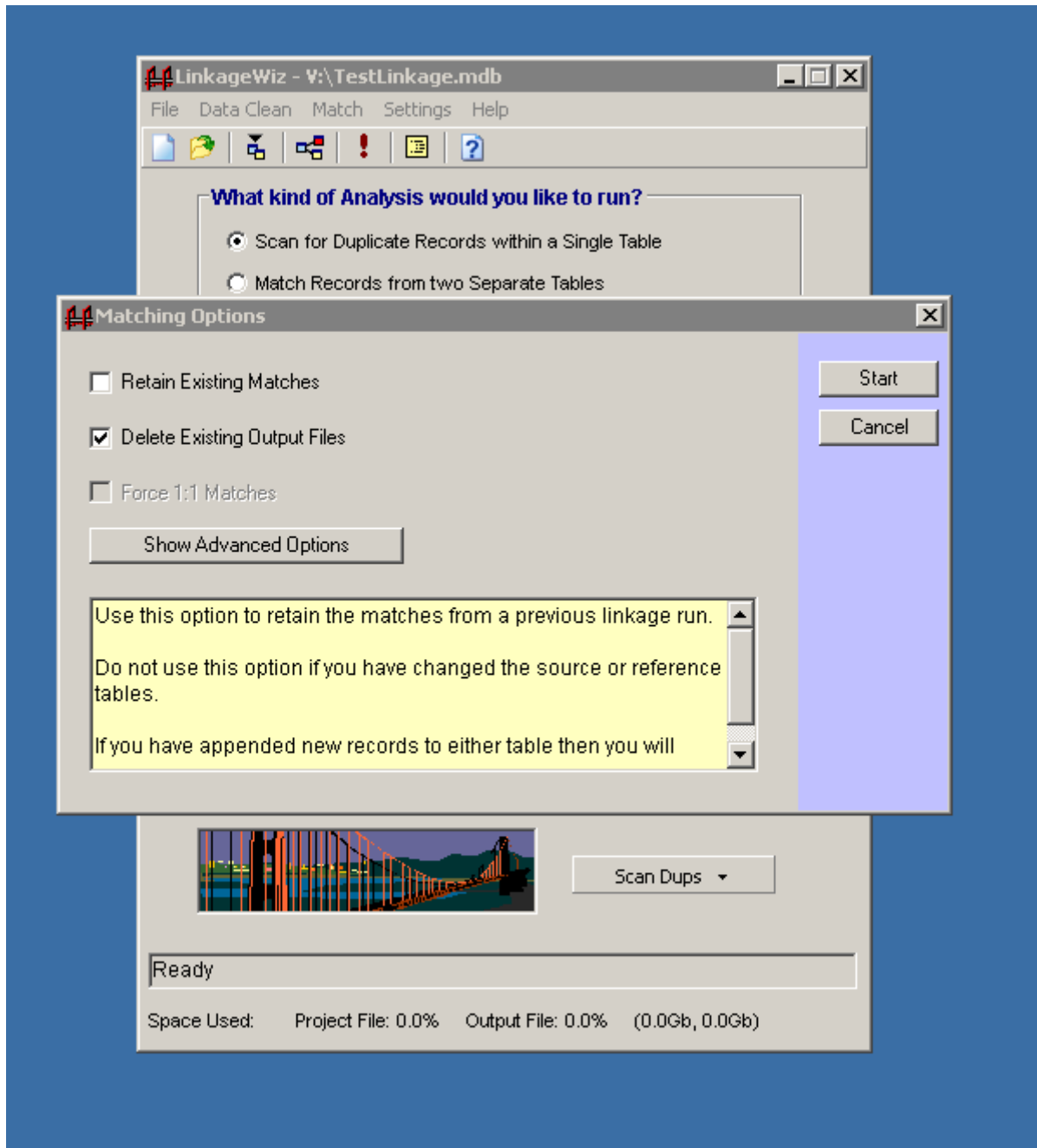
4) Prepare your table for matching:



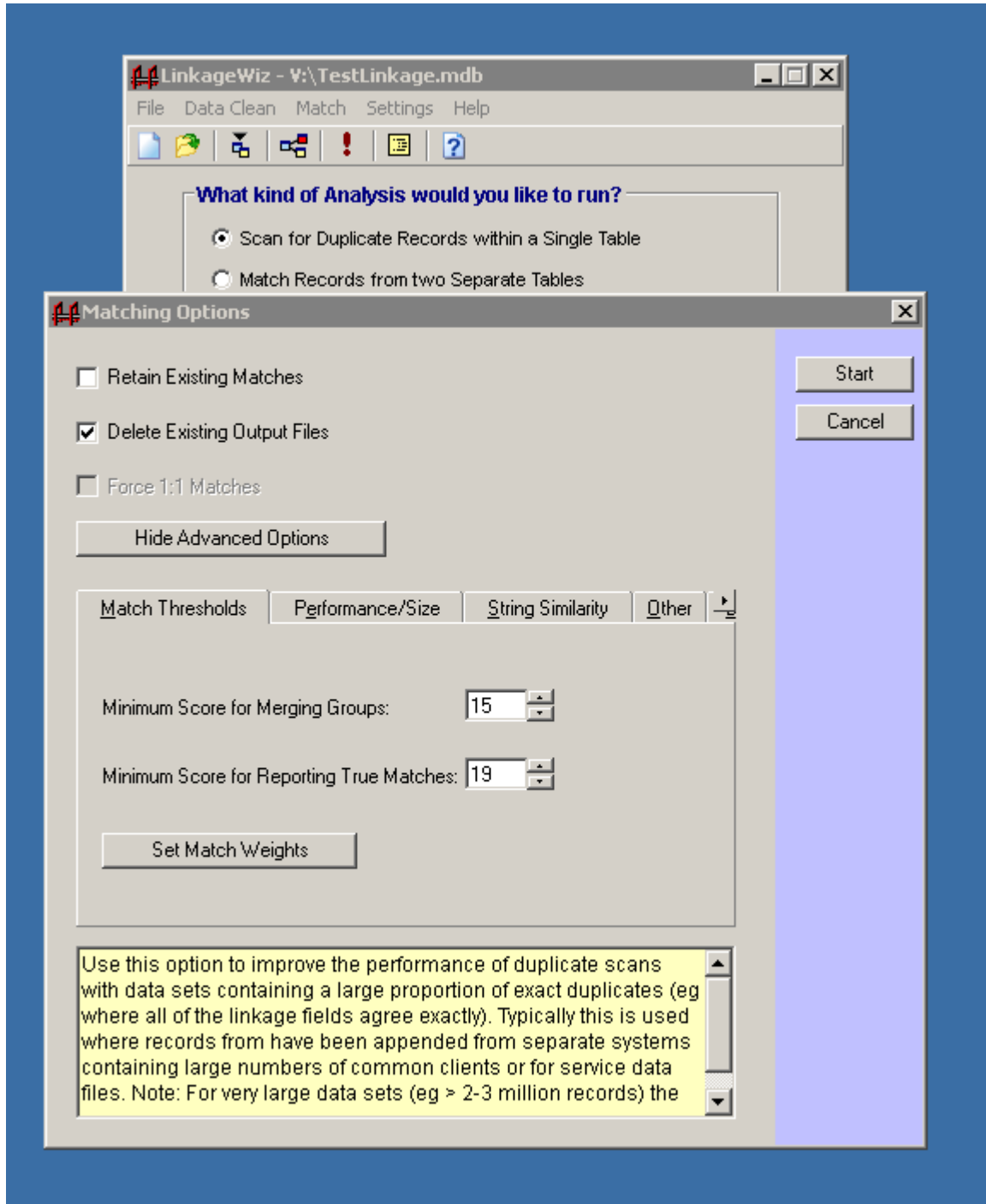
5) Choose blocking variables:



6) Press “Scan Dups” to run the linkage process. Choose whether you would like your matches from the previous linkage run saved. If this is your first run, either choice is fine.

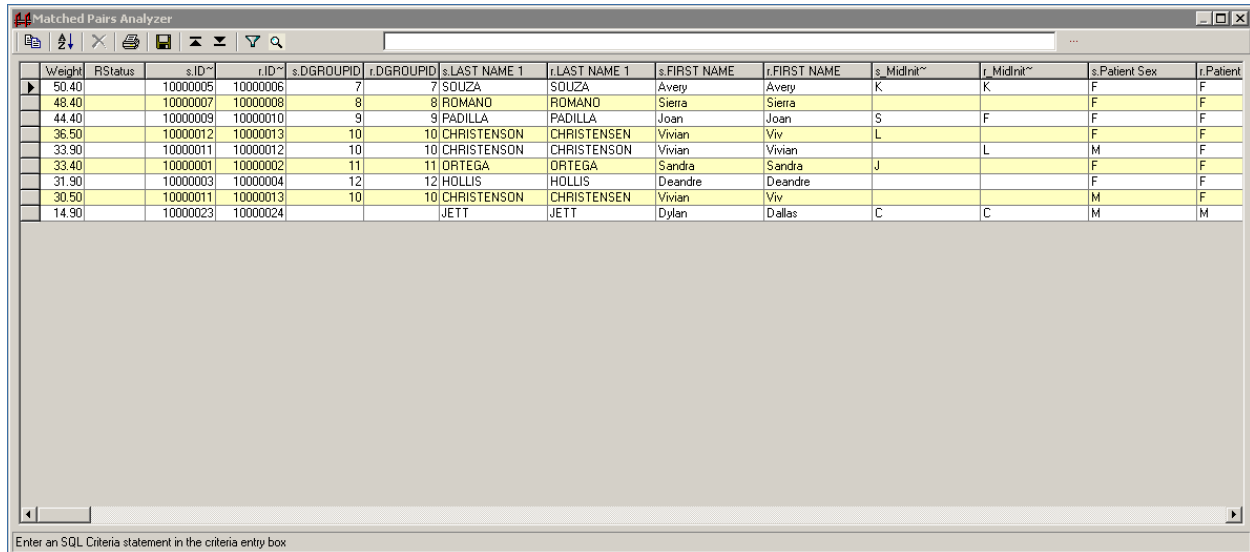


If you press “Show Advanced Options”, a menu opens that provides added control over how matches are determined. This includes setting “Match Thresholds”. The default settings will be used in this example, but these options can be particularly helpful in future iterations.



## 7) Review output:

LinkageWiz provides the Matched Pairs Analyzer as the default interface to review your matches.



The screenshot shows the Matched Pairs Analyzer window. The table displays the following data:

Weight	RStatus	s.ID~	r.ID~	s.DGROUPID	r.DGROUPID	s.LAST NAME 1	r.LAST NAME 1	s.FIRST NAME	r.FIRST NAME	s.Midlnit~	r.Midlnit~	s.Patient Sex	r.Patient
50.40		10000005	10000006	7	7	SOUZA	SOUZA	Avery	Avery	K	K	F	F
48.40		10000007	10000008	8	8	ROMANO	ROMANO	Sierra	Sierra			F	F
44.40		10000009	10000010	9	9	PADILLA	PADILLA	Joan	Joan	S	F	F	F
36.50		10000012	10000013	10	10	CHRISTENSON	CHRISTENSEN	Vivian	Vivian	L		F	F
33.90		10000011	10000012	10	10	CHRISTENSON	CHRISTENSEN	Vivian	Vivian		L	M	F
33.40		10000001	10000002	11	11	ORTEGA	ORTEGA	Sandra	Sandra	J		F	F
31.90		10000003	10000004	12	12	HOLLIS	HOLLIS	Deandre	Deandre			F	F
30.50		10000011	10000013	10	10	CHRISTENSON	CHRISTENSEN	Vivian	Vivian			M	F
14.90		10000023	10000024			JETT	JETT	Dylan	Dallas	C	C	M	M

Enter an SQL Criteria statement in the criteria entry box

If you prefer, you can analyze the results in Microsoft Access.

On your computer, navigate to the location you chose when you started this project and you will find the output saved as a Microsoft Access Database file. You can see the output in the table with the same name as the table you imported into LinkageWiz. All records with a match are given DGroupID that is the same for all of its matches. The value for DConfidence represents the *Match Score* at which the record was brought into the group; a higher value indicates a greater likelihood that this record belongs in this group.

DGROUPID	DCONFIDEN	FIRSTEXACT	SND_SURNA	STD_SURNAI	STD_SURNAI	STD_SURNAI	STD_SURNAI	STD_SURNAI	STD_SURNAI	STD_SURNAI	STD_MIDDLE	STD_FIRS
11	33.4	0	ORTAG				ORTEGA	SANDRA	S536		J	S
11	33.4	0	ORTAG				ORTEGA	SANDRA	S536			S
12	31.9	0	HAL				HOLLIS	DEANDRE	D536			D
12	31.9	0	HAL				HOLLIS	DEANDRE	D536			D
7	50.4	0	SAS				SOUZA	AVERY	A16	K		A
7	50.4	0	SAS				SOUZA	AVERY	A16	K		A
8	48.4	0	RANAN				ROMANO	SIERRA	S6			S
8	48.4	0	RANAN				ROMANO	SIERRA	S6			S
9	44.4	0	PADAL				PADILLA	JOAN	J5	S		J
9	44.4	0	PADAL				PADILLA	JOAN	J5	F		J
10	33.9	0	CRASTA				CHRISTENSON	VIVIAN	V15			V
10	36.5	0	CRASTA				CHRISTENSON	VIVIAN	V15	L		V
10	36.5	0	CRASTA				CHRISTENSEN	VIV	V1			V
		0	CALDAR				CALDERON	KYLA	K4			K
		0	FARTAN				FORTUNE	JAIDEN	J35			J
		0	RADRAG				RODRIGUEZ	CESAR	C26			C
		0	TAT				TOTH	EDEN	E35			E
		0	CRAS				KRUSE	IMAN	I5			I
		0	LACAT				LOCKETT	ALEX	A42			A
		0	STARLA				STERLING	MAXWELL	M24	M		M
		0	TRATAR				TROTTER	JOHANNA	J5			J
		0	WAD				WARD	KEITH	K3			K
		0	JAT				JETT	DYLAN	D45	C		D
		0	JAT				JETT	DALLAS	D42	C		D
		0	BLAC				BLOCK	JULIANA	J45			J
		0	WALAD				WILLARD	DORIAN	D65			D
		0	BRADLY				BRADLEY	ADAM	A35			A
		0	CANDLA				CHANDLER	TALON	T45	F		T
		0	FALAR				FULLER	CHRISTINA	C623			C
		0	CTAD				CTURC	ANDREW	A536			A