

VERSION 1.1
MAY 14, 2019



DUPLICATE DETECTOR

USER GUIDE

PRESENTED BY: **JEFF RICE**
CHURCH CHAMPIONS
WWW.CHURCHCHAMPIONS.COM
JRICE@CHURCHCHAMPIONS.COM
817-917-5215

OVERVIEW

The Duplicate Detector was designed to fill a large gap in the FellowshipOne (F1) feature set: the ability to **find** duplicate people records. These are two independent people records that refer to the same person in real life. The current reports available for this purpose, such as **A1004 v1.1 - Duplicate Individuals with Optional Filters**, are too limited to identify all the likely duplicate records.

Using the knowledge gained after years of resolving these duplicate records for many churches, Church Champions was able to create a new, highly sophisticated program to detect all of them that exist in a database. The Excel output contains simple record pairings along with a score that indicates the likelihood of the pair being the same person.

Armed with this comprehensive output, it's now possible to confidently know exactly how many duplicates exist in your data and then eliminate them all.

CRITERIA

Generally, two records cannot be considered a match unless the First and Last Names match along with some other piece of information such as an email address, phone number or physical address. However, there are other data elements that either support a match when they are the same or detract from a being a match when they are different. The Duplicate Detector considers all of these to produce its score:

- First Name - (Full Match, 3 Character Match, First Initial Match, Mismatch)
- Last Name - (Full Match, 3 Character Match, First Initial Match, Mismatch)
- Middle Name – (Full Match, First Initial Match, Mismatch)
- Suffix Name – (Full Match, Mismatch)
- Date of Birth – (Full Match, Partial Match: Year/Month, Month/Day, or Year/Day, Mismatch)
- Household Position – (Match, Mismatch)
- First Record Date – (Match, Mismatch)
- Phone Number – (Personal Match, Match to someone else in same Household, Mismatch)
- Email Address - (Personal Match, Match to someone else in same Household, Mismatch)
- Physical Address – (Primary Address Match, Previous/Statement/Incorrect Address Match)
- Relationships – (Looks for matching names of other people in the households being compared. If both households have a Jim, Susan and Tom, the likelihood of the Jims being a match increases. When none of the names are the same between the households, the likelihood decreases.)

OUTPUT

	A	B	C	D	E	F	G	H	I	J	K			
1	Score	PersonID	1st Record	Name	Status	SubStatus	DOB	Age	Address	Phone	EMail			
2	95	62368515	4/20/2014	Adams, Daisy	Attendee	OLW	8/	2005	13	719 E Missouri St, Evansville	812-	068	slam	mail.com
3		60266022	1/19/2014	Maxheimer, Daisy	Attendee	OLW	8/	2005	13	218 N Barker Ave, Evansville	812-	274	zachi	r@gmail.com
3	310	92222882	11/13/2018	Ball, Jacob	Visitor	OLE	12	/1995	22	200 N Park Dr, Evansville	812-	131	jacob	gmail.com
4		89474594	3/4/2018	Bell, Jacob	UnConnected	OLE	12	/1995	22		812-	131	jacob	gmail.com
4	110	77240506	1/4/2016	Berry, Brody	Dropped	OLE				1726 Greenfield Rd, Evansville				
5		42614309	2/26/2012	Berry, Brody	UnConnected	OLE	5/	2003	15	1726 Greenfield Rd, Evansville	812-	556		
5	105	90572557	5/23/2018	Curlin, Cadence	Attendee	OLH	8/	2005	13	343 Tree Line Rd, Henderson	270-	019		
6		50094564	1/13/2013	Woolfork, Cadence	UnConnected	OLH	8/	2005	13	799 Lakeview Dr, Henderson	270-	019		
6	160	90926761	7/8/2018	De La Luz, Miguel	Attendee	OLH	8/	2009	9	2106 N Elm St, Henderson	270-	130	amil	gmail.com
7		39356474	9/4/2011	De La Luz, Miguel	UnConnected	OLH	8/	2010	8	1127 Clay St, Henderson	270-	419		
7	205	62086818	4/7/2014	Duncan, Sheena	Dropped	OLH				270-	248	sdun	@hotmail.com	
8		57592331	10/6/2013	Duncan, Shana	Attendee	OLH	12	1975	42	947 Julianne Pl, Henderson	270-	248	sdun	@hotmail.com
8	250	90727137	6/12/2018	Espenlaub, Kameryn	Connected	OLE	4/	1994	24	5316 Gravenstein Ct, Evansville	270-	319	kami	r@gmail.com
		35163052	3/3/2010	Schmelzer, Kami	Connected	OLE	4/	1994	24		270-	319	Schr	i@yahoo.com

An Excel file consisting of one row per duplicate candidate pairing is produced. The number of rows is the number of likely duplicates. It is sorted alphabetically by last name of the first person listed for each pair, which is the newest record of the two. The columns include:

1. Score – A higher score means it’s more likely the two records are the same person.
 - a. 95 – The two are **Possibly** the same person
 - b. 100 to 120 – The two are **Probably** the same person
 - c. 125 & higher - The two are **Positively** the same person
2. PersonID – The unique F1 identifier assigned to that person.



This can be seen in the URL when viewing the person’s profile.

3. First Record Date – The date the record was created in F1.
4. Person’s Last & First Name
5. Status
6. SubStatus (if available)
7. Date of Birth (if available)
8. Age (if available)
9. Street & City Address
10. Phone Number
11. Email Address

The last 5 values are the same as is shown in the F1 search results for easy comparison to identify the correct person.

IGNORED VALUES

The Duplicate Detector can ignore duplicate candidate pairings in these specific cases:

SKIP STATUSES

If both persons have an identified Status to be “skipped”, the pairing will be ignored. Each church can specify which statuses to be included in this skipped list. Typically, this includes **Dropped** and **Inactive Member**. The idea behind this concept is to not spend time on eliminating duplicates when both persons have left and are unlikely to re-engage the church.

KNOWN DUPLICATE STATUS

Many churches have records in their database that they know are a duplicate and wish to retain them anyway. Usually these records are marked with a specific status such as “Known Duplicate” or “Intentional Duplicate”. Each church can specify the status they are using for this purpose. The Duplicate Detector will ignore the candidate pairing if either person has this status.

DON’T MERGE PEOPLE LIST

If a candidate pairing is examined and determined not to be a duplicate, they can be placed in a People List of the church’s choosing. The Duplicate Detector will ignore that pairing in the future when both persons are in that people list.