# Record Linkage Match Logic v1.1

6/29/2023

Updates:

* Include simple geography criteria

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pass** | **Blocking** | **Strict** | **Moderate** | **Relaxed** |
| 0 | Fname, lname, DOB, common\_id | Any | Any | Any |
| 1 | Common\_id | High-similarity fname & high-similarity lname & DOB partial match\*  OR High-similarity fnamelname & high-similarity lnamefname & DOB partial match\* | Any Strict matches  OR High-similarity fname & high-similarity lname  OR High-similarity fnamelname & high-similarity lnamefname  OR High-similarity fname & byear within 1  OR (High-similarity fname | high-similarity lname) & DOB is partial match\* | Any Strict or Moderate matches  OR any high-similarity fname  OR any DOB partial match\* |
| 2 | Xf, xl | **If common\_id is close:**  Fname, lname, DOB: 2/3 exact, 1/3 high-similarity  **If common\_id is null:**  Fname, lname, DOB: 2/3 exact, 1/3 high-similarity  & (minitial good match\* | high-similarity altlname)  OR fname, lname, DOB all exact  & mname is null  & same zip/county  **If common\_id is not close:**  Fname, lname, DOB are exact  & (mname initial is a good match\* | atlname is exact) | Any Strict matches  OR High-similarity fname  & high-similarity lname  & high-similarity DOB  & one of the matching identifiers:  - similar common ID edit distance  - similar mname/altlname  - same zip/county | Any Strict or Moderate matches  OR High-similarity fname  & high-similarity lname  & (high-similarity DOB | similar common ID edit distance) |
| 3 | Xf/xl, xf/xl  [Inverted names, logic is mostly same as p2 with exceptions in strict and moderate when common\_id is null] | **If common\_id is close:**  Fname/lname, lname/fname, DOB: 2/3 exact, 1/3 high-similarity  **If common\_id is null:**  Fname/lname, lname/fname, DOB: 2/3 exact, 1/3 high-similarity  & (minitial good match\* | high-similarity altlname)  **If common\_id is not close:**  Fname/lname, lname/fname, DOB are exact  & (minitial is a good match\* | atlname is exact) | Any Strict matches  OR High-similarity fname/lname  & high-similarity lname/fname  & high-similarity DOB  & one of the additional matching identifiers:  - similar common ID edit distance  - similar mname/altlname  - same zip/county  OR **if common\_id is null:**  fname/lname, lname/fname, DOB all exact  & mname is null  & same zip/county | Any Strict or Moderate matches  OR High-similarity fname/lname  & high-similarity lname/fname  & (high-similarity DOB | similar common ID edit distance) |
| 4 | DOB  [Pairs that don’t agree on ID, xf xl (incl. Inverted)] | **If common\_id is close:**  high similarity fname & high similarity lname  **If common\_id is null:**  exact fname & high-similarity lname  & (minitial good match\* | altlname is highly similar)  **If common\_id is not close:**  exact fname & high-similarity lname  &(minitial good match\* | altlname is exact) | Any Strict matches  OR High-similarity fname  & high-similarity lname  & one of the additional matching identifiers:  - similar mname  - similar altlname  - same zip/county [Note: close common\_ids are already accepted in Strict] | Any Strict or Moderate matches  OR High-similarity in 2/3: fname, lname, common\_id |

Similarity definitions:

* High Similarity name components: jarowinkler score >= 0.88
* Close common ID: levenshtein score >= 0.75
* \***Similar/partial DOB match**:
  + bmonth matches & bday matches & byear w/in 1, OR
  + byear matches & bmonth matches, OR
  + byear matches & bday matches, OR
  + byear matches & bmonth/bday are inverted
* \***Minitial good match:**
  + minitial matches & one of mnames is a single letter, OR
  + minitial matches & neither of mnames is a single letter & near exact mname (i.e. jarowinkler >= 0.90)