Vendor address histories are first analyzed using the 01\_MatchAddresses1.sas code. As part of this step, potential duplicates for an individual in a given timeframe are output for review. These comparison guidelines are designed to facilitate an accurate result from that manual review.

The purpose of the review is to identify pairs of addresses that represent the same residence for a given individual as opposed to two different residences. The goal is to identify addresses where differences are likely due to typos, transpositions, or an omitted or extra character. The question to answer is: “are these two addresses likely to be the same residence?”

Terminology: A street address consists of a address number and a street name. The street name is often followed by street type (e.g., Street, Avenue, Place, Court). The street name can also have a directional prefix or suffix (e.g., N. Main Street, Virginia Avenue NW). Many cities have street names that are numeric (e.g. 48th Street, 5th Avenue) and are referred to as “numeric street names”.

1. If the street name yields a match, but the street type differs (e.g. Road versus Place), classify it as a match if the address number also matches along with city and ZIP code. If the address number does not match, but is off by one digit that is a likely typographical mistake or contains a likely transposition of two digits or is of differing length caused by presence/absence of a final digit, classify it as a match.
2. If the street name and street type match, but the directional prefix or suffix (e.g. North versus having no directional prefix or suffix), classify it as a match if the address number also matches along with city and ZIP code. If the address number does not match, but is off by one digit that is a likely typographical mistake or contains a likely transposition of two digits or is of differing length caused by presence/absence of a final digit, classify it as a match.
3. If the address number and street type yields a match, examine the street names assuming a match of city and ZIP code. If much of the spelling and context of the street names is the same, classify it as a match (e.g. Hampton versus Homtan). If a street name is composed of multiple parts where one part matches and the other part is missing (e.g. Bird Hill Lane versus Bird Lane or 201 Bird Lane versus Bird Lane where the 201 is part of the street name and not the address number), classify it as a match.
4. Similar numeric street names (e.g. 44th Street versus 48th Street) often appear as false positive matches since they differ by only one character. Similar numeric street names should only be classified as a match if the address number and directional prefix or suffix are likely matches along with matches being found for city and ZIP code.
5. If a street name is composed of a PO Box and the number of the box matches, classify it as a match even if there are variations in the designation of PO (e.g. P.O. or P O or po) assuming a match with city and ZIP code.
6. If a street number and street name match, but the city does not match, examine the ZIP code or ZIP+4 if available. If the ZIP code or ZIP+4 matches, classify it as a match.
7. Addresses numbers that differ by two digits or more in what is not a transposition paired with street names that share some similar element(s) but are not a contextual match (e.g. Sycamore has no contextual match to Sundown) should not be classified as a match.
8. Addresses numbers that differ by two digits or more in what is not a transposition paired with street names that lack matching elements should not be classified as a match even in the context of matching on city and ZIP code.
9. Street names composed of a PO Box where the numbers of the box do not match should not be classified as a match even in the context of matching on city and ZIP code.