**Wikipedia Connected People Analysis**

**Problem Statement**

There are more than 6M articles on Wikipedia (English). This data can be analyzed to get lot of interesting insights. For this use-case we want to analyze Wikipedia “people” pages and create a connected people graph (reference of person x on person y’s page) with the help of graph database (e.g. Neo4J). We should be able to query this data in different ways – e.g. to find key people (based on number of incoming connections), etc.

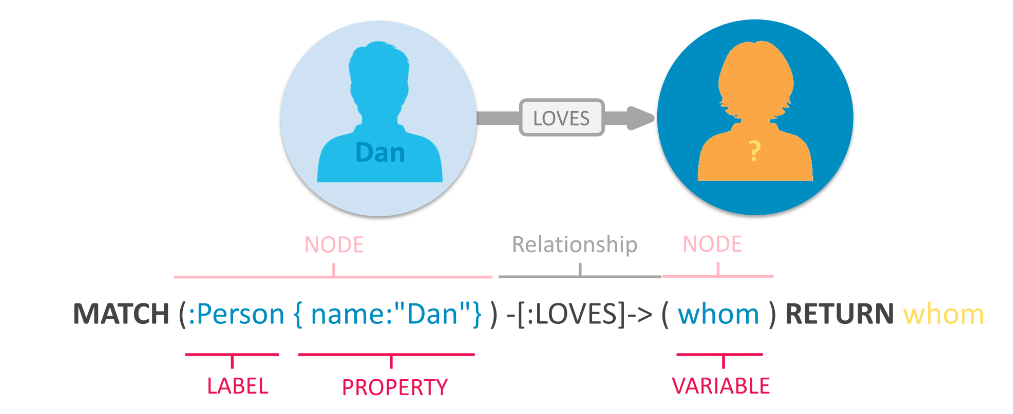
**Introduction**

Wikipedia Connected People Analysis is a process of analysing the relationships between people who are connected through Wikipedia articles. Wikipedia is a vast online encyclopedia that contains articles about a wide range of subjects, including notable individuals. These individuals may be connected to each other through personal or professional relationships, such as family, colleagues, friends, or mentors.

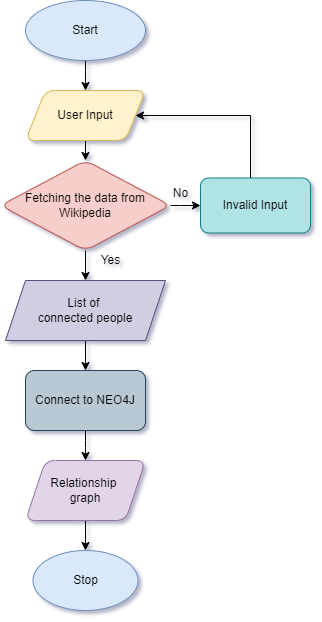
**Pre-requisites: -**

1. Python
2. Neo4J

**Neo4j :**

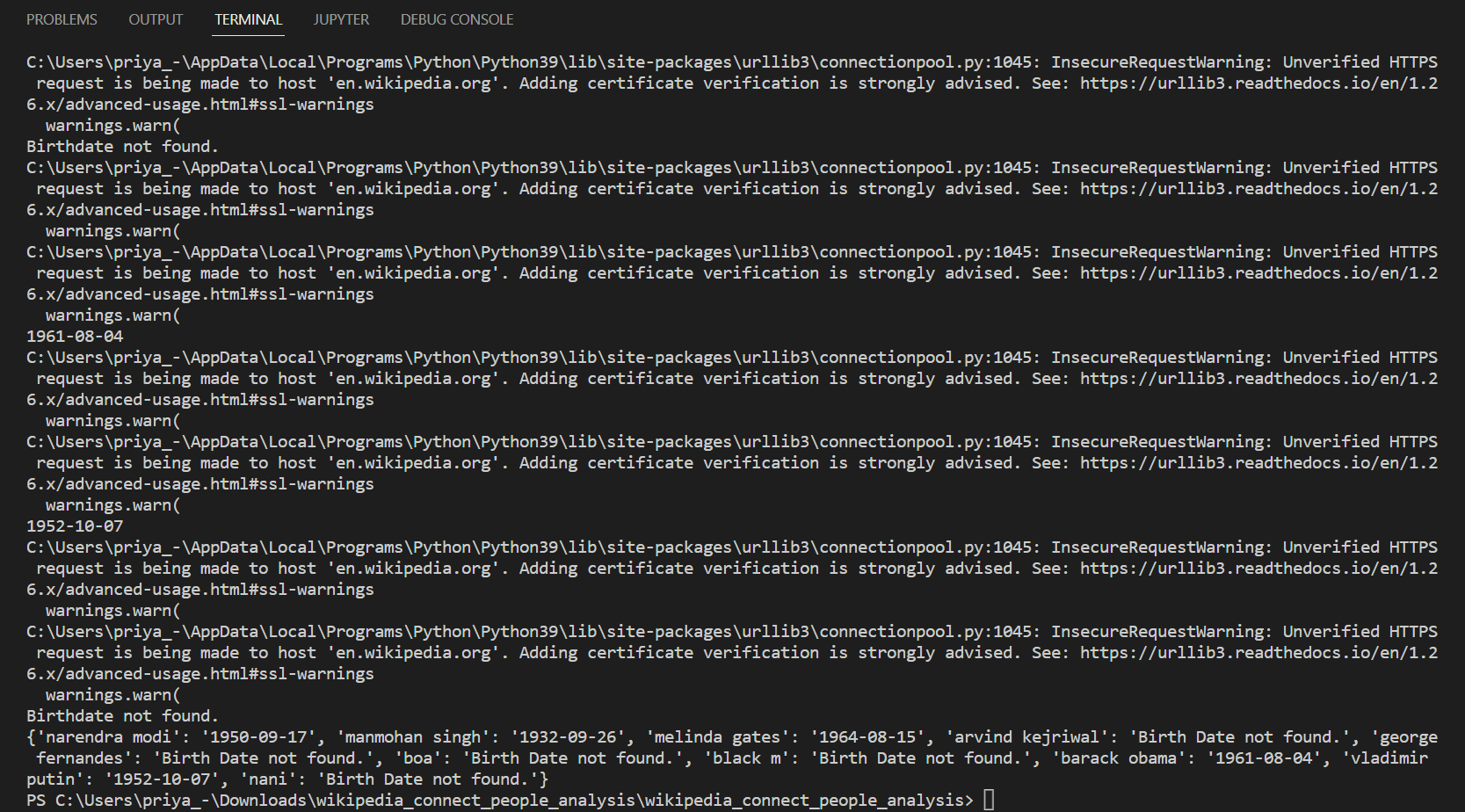


**Workflow**

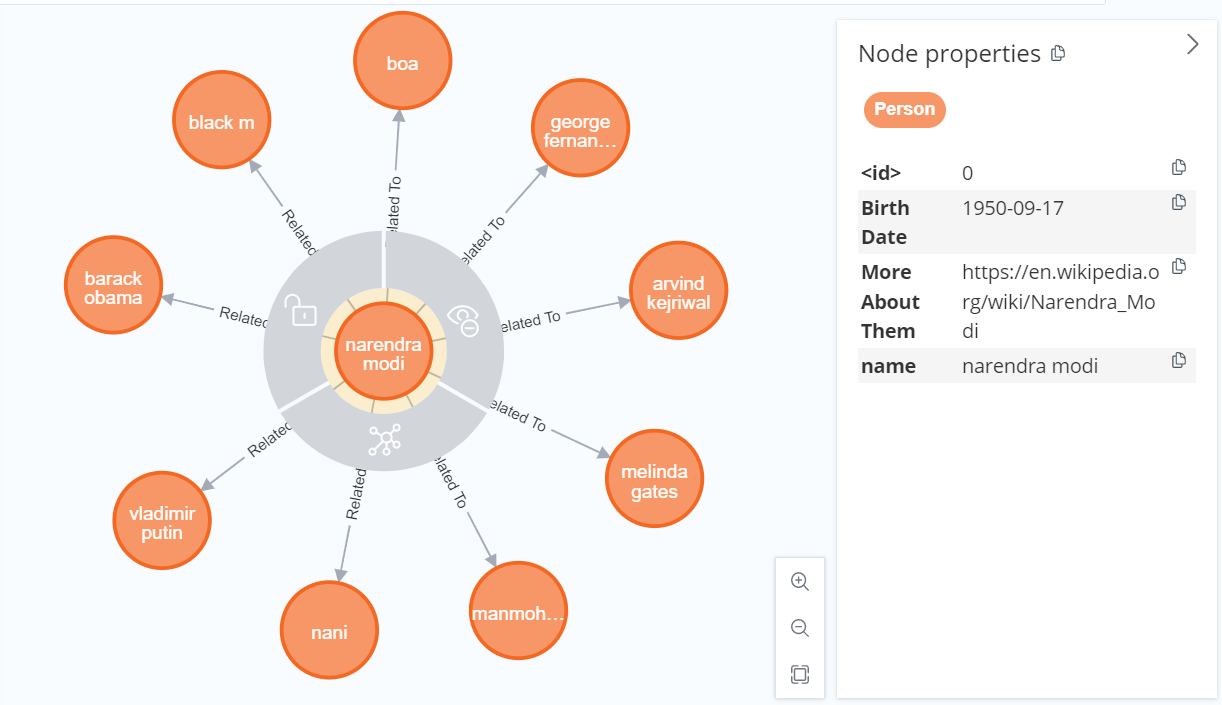
****

**Screenshots: -**

1. When we enter the name Narendra Modi (For sample) It’ll fetch the data from Wikipedia and list the name in the terminal who connected to Narendra Modi and after this result, it’ll connect to Neo4j to create a graph.



1. Created a Relationship graph. Here we can check a particular person’s birth date and if you require you can click on the link for more information.



1. When we enter a name whose name is not there in Wikipedia then we’ll not the get result again it’ll ask for the name.

