Assignment 6 - GraphX

1. Launch the Spark shell.
2. Import the GraphX libs, which are `org.apache.spark.graphx._` and `org.apache.spark.rdd.RDD`.
3. Build the property graph shown in Figure 1.

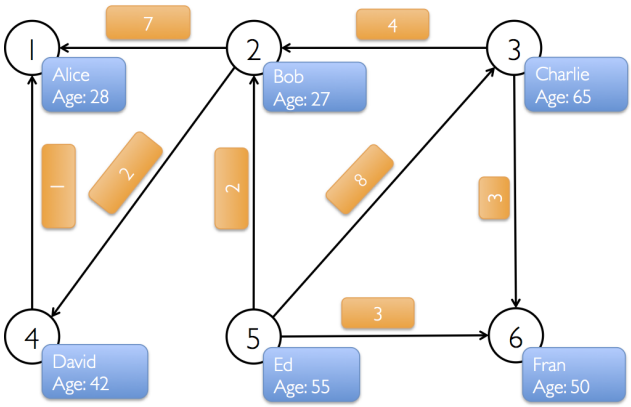

Figure 1: The property graph.

4. Display the name of the users older than 30 years old.
5. Display the in-degree of each vertex.
6. Display who follows who (through the edges direction).
7. Display who likes who (if the edge value is greater than 5).
8. Make a user graph such that each vertex stores the number of its incoming and outgoing links.
9. Display the name of the users who are followed by the same number of people they follow. For example Bob follows two persons, and two persons follow Bob.

10. Display the oldest follower for each user.

11. Find the average age of the followers of each user.

12. Make a subgraph of the users that are 30 or older.

13. Compute the connected components and display the component id of each user.

14. Write a standalone application to make a graph from the *followers file* (shown below), measure the page rank of the graph, and display the page rank of each vertex along with its user name. Each row of the *users* file is assigning a name to a vertex.

    followers file:
    2 1
    4 1
    1 2
    6 3
    7 3
    7 6
    6 7
    3 7

    users:
    1,Seif,SICS
    2,Amir,SICS
    3,Jim,KTH
    4,Ahmad,SICS
    6,Vlad,KTH
    7,Fatemeh,SICS
    8,Anonsys