Users analysis:

It has been widely believed that users data is highly important for the community analysis. Apparently, both the users set of stackoverflow and ru.stackoverflow provide a number of information.

Part 1: Location and AboutMe

First of all, the intersection of these two sets is the foundation of this part of research. A point worth mentioning is that not only people from Russia use Russian for introduction and answers, but also some other countries like UKraine and Belarus. Stackoverflow owns 6,438,658 users, while Ru.stackoverflow has 58,095 users. The number of intersection is 31,440 (nearly 54.118%), so the basic point is these 31,440 users owns accounts of both two communities in the meantime. Figure 1 shows the Russian Alphabet can be useful for the location query and nationality judgement. There must be an assumption that the user who has a Russian location is a Russian, and the result is 18,472 stackoverflow users have Russian locations and 4,727 of them owns accounts of ru.stackoverflow.

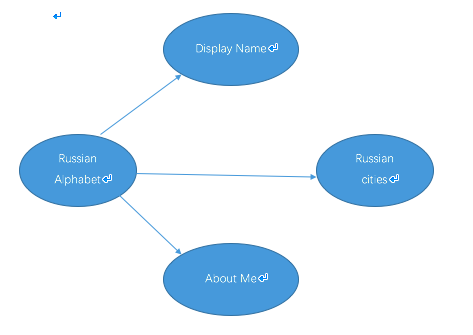


Figure 1

Union all the distinct information of users who have a Russian location , who use a Russian name and who use Russian for their self-introduction. 38,873 users can be found in the stackoverflow set, and 8,333 of them owns a ru.stackoverflow account at the same time.

Part 2: Creation Date& Reputation

Data used for this part of research is the set of intersection of stackoverflow users and ru.stackoverflow users. Figure 2 present the data structure of the intersection. Apparently, figure 3 shows which community is the first choice of these 31,440 people.

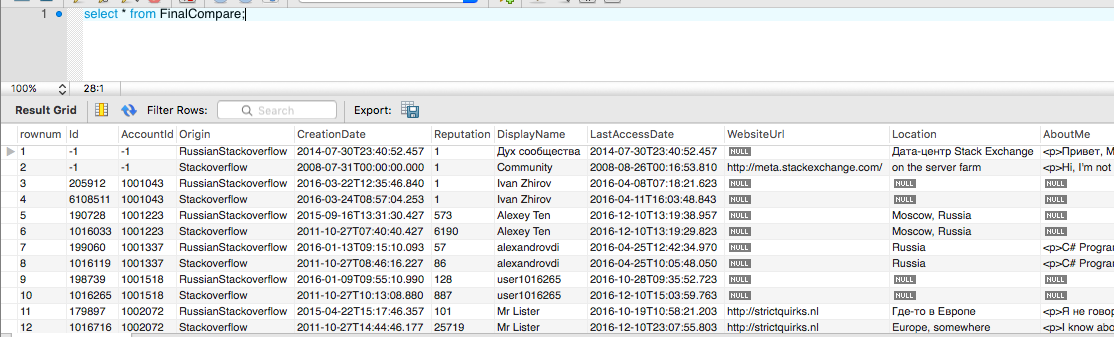


Figure 2

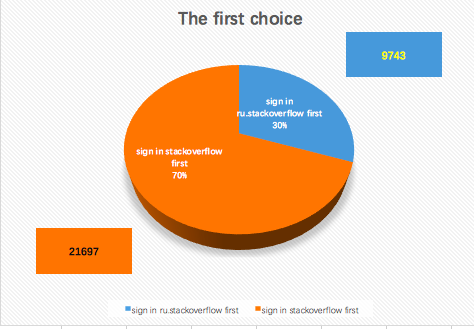
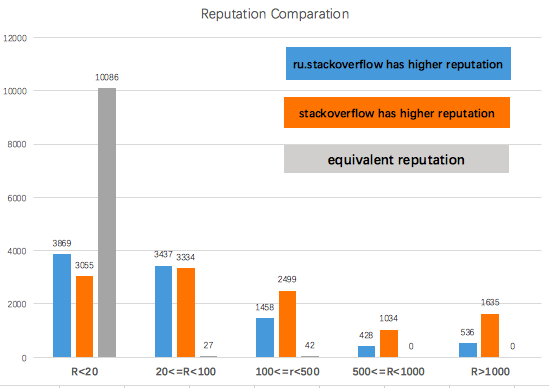


Figure 3

Similarly, Reputation can be figured out in this way. And the result is in Figure 4. A huge number of inactive users, which includes a number of barely used accounts that the account reputation is 1. And the valid data shows there are more users’stackoverflow account owns a higher reputation than their Russian ones.



Figure