

Report on gender structure on Serbian Wikipedia

Rationale

The reason behind this little “research”¹ stems from one of the Board meetings of Wikimedia Serbia. Namely, in an effort to produce a new Strategy of the organization, the Board has faced a problem of not knowing the basic statistic information which would serve as baseline values for tracking changes.

Lately, the most sought after and appreciated activity of various entities from the Wikimedia world is “the impact”, i.e. measurable effect that’s produced by a project or program. That measurable impact mostly represents statistical data regarding relevant parameters of the project (for instance, number of created/improved articles, number of participants, number of uploaded photographs, etc). In accordance with the trend to quantify as much data as possible and present it to the public, in a purpose of a better understanding of the projects, we also submit Wikimedia Serbia projects to metrics and measuring in order to have an idea of what works, and what doesn’t.

Seeing as the gender gap is a large focus of the entire Wikimedia community, especially Wikimedia Foundation, one of our strategic goals will certainly be working on that problem. However, the problem we have faced is that we had virtually no data regarding gender structure of the editors of Serbian Wikipedia. With no baseline values to start with, there is no point in setting strategic goals, because if we don’t know where we started from, we can’t project goal values, and we won’t know the amount of progress (or regress) we have had. In order to determine the initial referential values, the Board has set a goal to acquire data that would serve as referential values, and hence this document.

In the following text, information collected from three different sources will be considered.

Note: *Collecting and analyzing the data was not executed by the entire Board, but only Filip Maljković. The poll was conducted by Sanja Pavlović and Filip Maljković.*

User settings

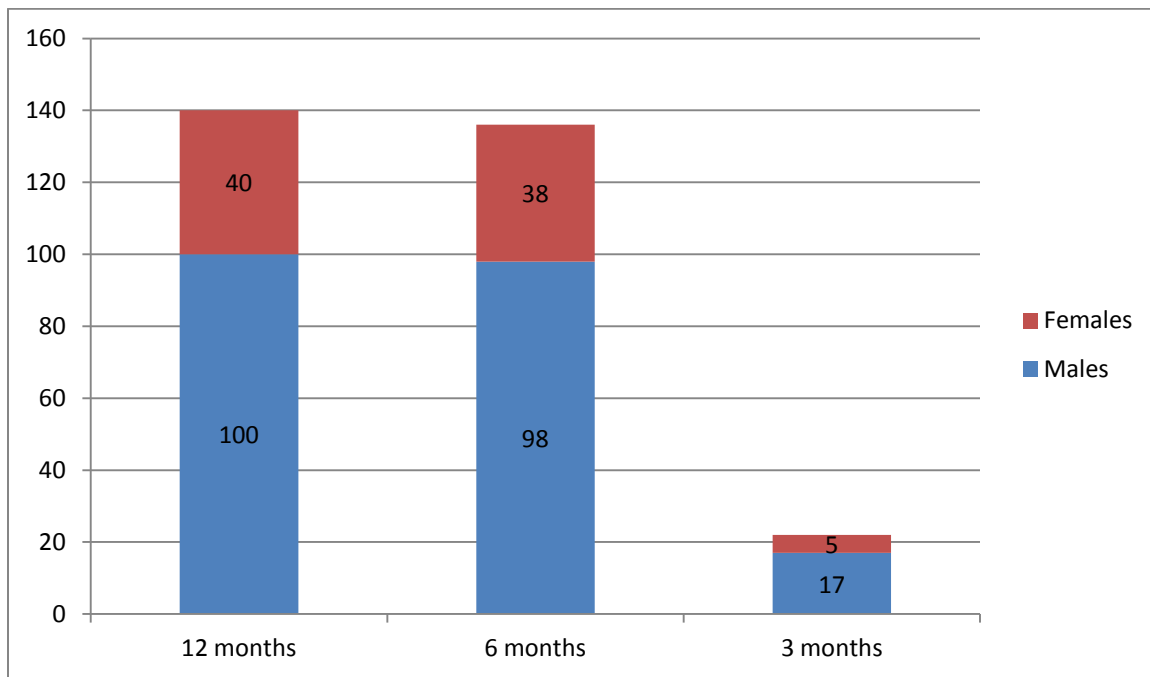
The first thing that came to our minds is the user settings on Wikipedia, where editors can select their gender. In practice, that setting only affects the writing of certain gender-sensitive system messages, but apart from that, it doesn’t have a particular significance in the MediaWiki software. We have consulted certain employees at the Foundation and they have instructed us in obtaining that data, seeing as it is publicly available. However, they mentioned that those numbers should be taken with caution, because they revolve around user settings that don’t have great significance, and that it is entirely possible that

¹ This is certainly not a real scientific research because we didn’t use standard methods, but a layman’s gathering of information through various channels, in order to determine the basic numbers about gender structure.

the information drawn from this doesn't necessarily represent the true state of things. We drew the information that is interesting to us, and it's up to the reader to interpret it.

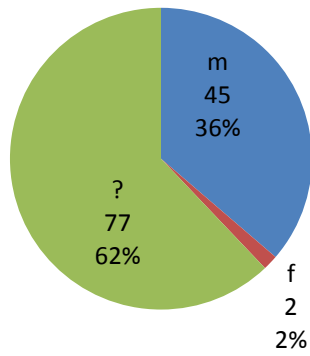
Out of a total 157,353 registered users (at the moment of making this document), 4,202 users have the gender setting on, out of which 3,409 with the selected option "male", and 793 "female". Therefore, according to that setting, the percentage of women is about 18.9%.

In order to see the trend, we had a look at the settings for recently registered users. Below is a chart that displays the number of males and females, for users that have joined Wikipedia 12, 6, and 3 months ago. To make things simpler, in the following charts that take time into consideration, the date set for present time is April 1st, 2015, which is the date from which "3 months ago" or "12 months ago" have been calculated.

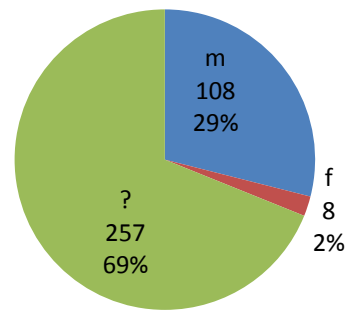


Below are the charts that display the gender setting selection for most active users by number of edits, in different time periods (from the beginning, in the last 5 years, 12, 6, 3 months ago).

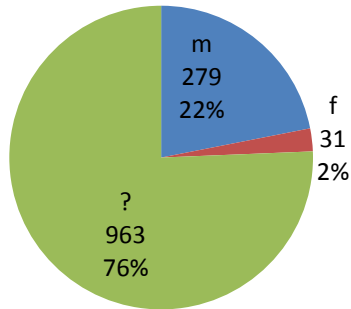
10,000+ edits, from start



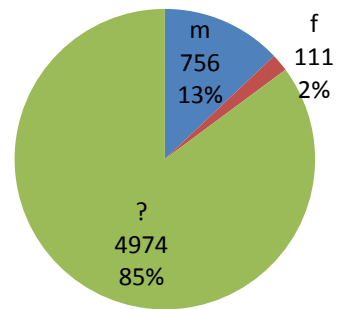
1,000+ edits, from start



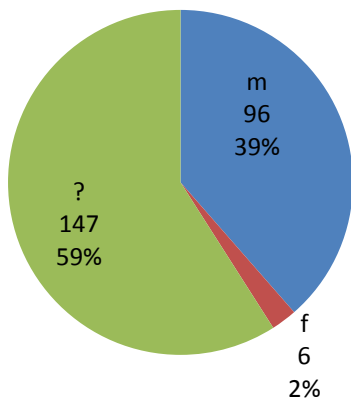
100+ edits, from start



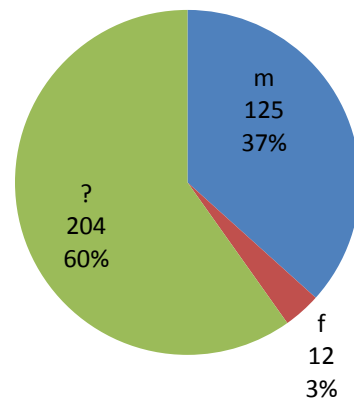
10+ edits, from start



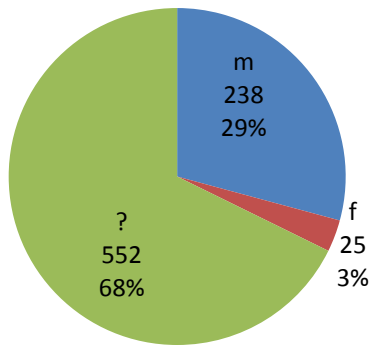
1,000+ edits, last 5 years



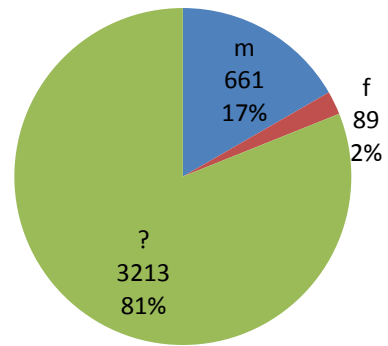
500+ edits, last 5 years



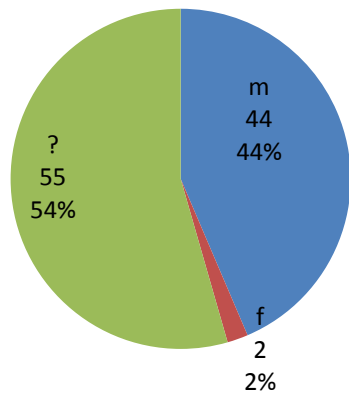
100+ edits, last 5 years



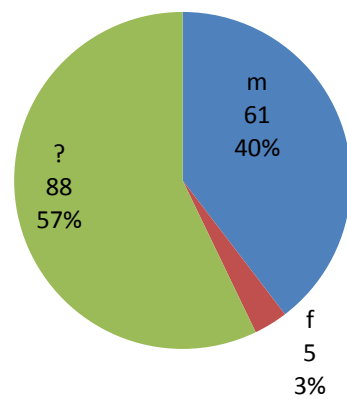
10+ edits, last 5 years



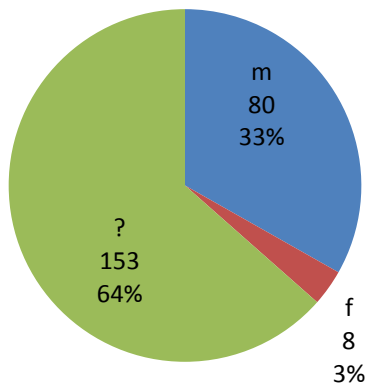
500+ edits, last 12 months



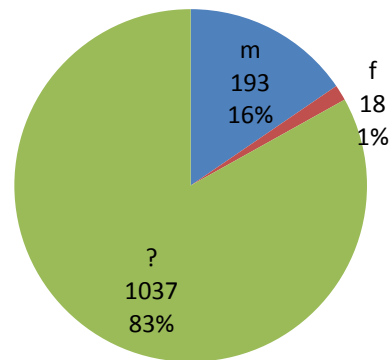
200+ edits, last 12 months



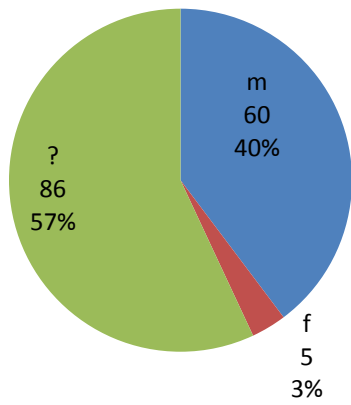
100+ edits, last 12 months



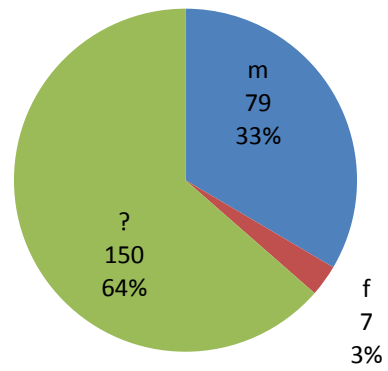
10+ edits, last 12 months



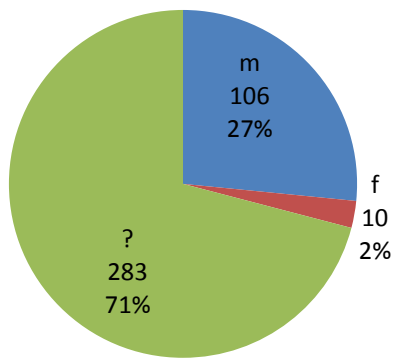
200+ edits, last 6 months



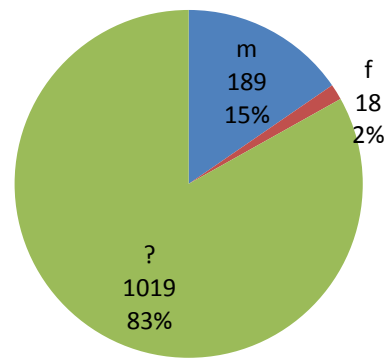
100+ edits, last 6 months



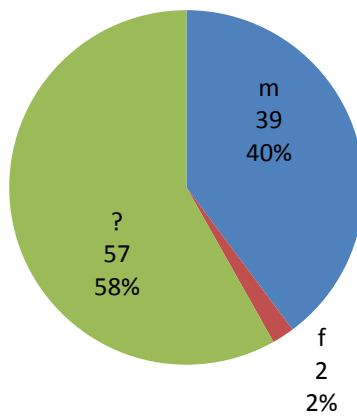
50+ edits, last 6 months



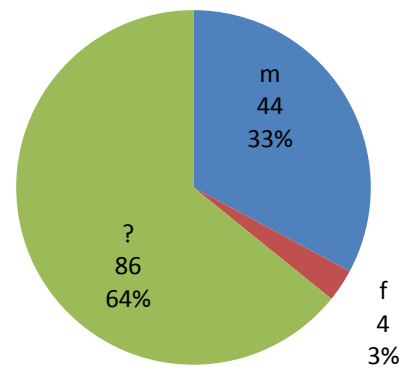
10+ edits, last 6 months



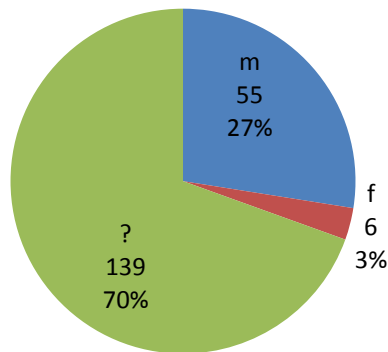
100+ edits, last 3 months



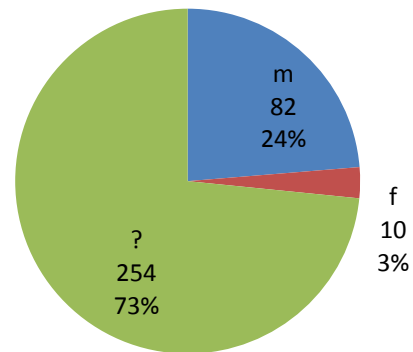
50+ edits, last 3 months



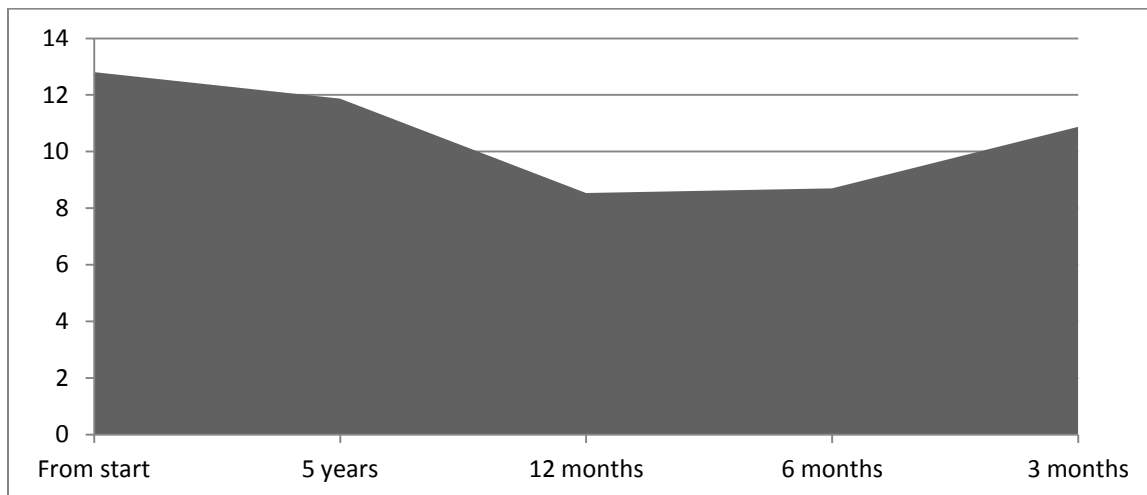
25+ edits, last 3 months



10+ edits, last 3 months



If the undecided users are exempt from the analysis, the ratio of males and females is somewhat different then. The chart below displays the percentage of women in relation to decided users with at least 10 edits, for the same time periods as in the previous charts.



Wikimedia Foundation research from 2011

Запослени у Задужбини су нам указали и на [истраживање](#) које је Задужбина спровела 2010. и 2011. године, где су такође прикупљани подаци о роду уредника. Међутим, како је веома мали узорак (ред величине 10) учесника са Википедије на српском језику, такви подаци нису статистички значајни.

The employees at the Wikimedia Foundation have pointed us toward a [research](#) that the Foundation conducted in 2011, which also included data about the gender of editors. However, seeing as the sample of participants from Serbian Wikipedia is very small (order of magnitude 10), that type of data is statistically insignificant.

Survey

We have also come up with the idea to create a simple survey which would ask about the information on gender and age² of the editors, and which would take less than a minute to fill in. The survey was created and set in the Sitenotice³ of Serbian Wikipedia on March 11, 2015 at around 9:15 p.m. The banner was removed, and the poll closed on March 21st at around 10:20 a.m. The initial intention was for the survey to be available to registered users only, but it was visible to all Wikipedia readers by mistake. However, once we realized this, we didn't want to change the strategy so we wouldn't have to eliminate some of the previously gathered information, but also because it is useful to know the gender and age structure of visitors.

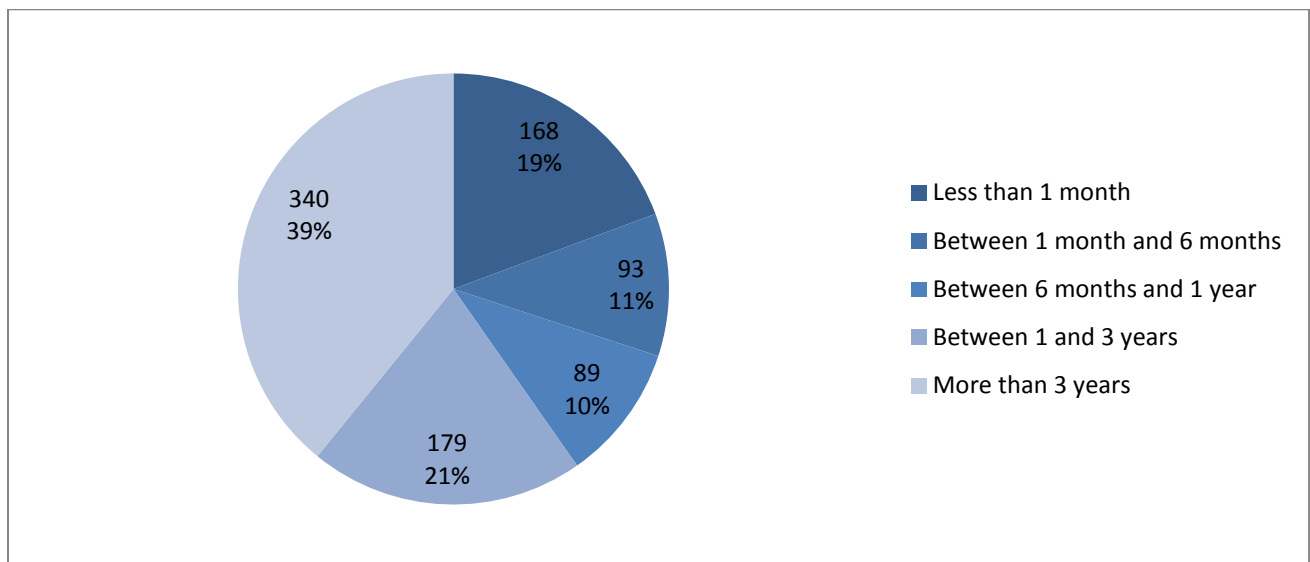
The poll consisted of three questions:

1. How long have you been editing Wikipedia?
2. How do you define yourself by gender?
3. What is your age?

All three questions were mandatory, but the second and third question also had an option "I would rather not say" (they are marked "?" in the charts below). The first and third question had timeframes as answer options.

Out of 4,402 people that took part in the survey, 3,533 claim they have never edited Wikipedia, and 869 claim they have. In other words, only about one fifth of the participants are editors.

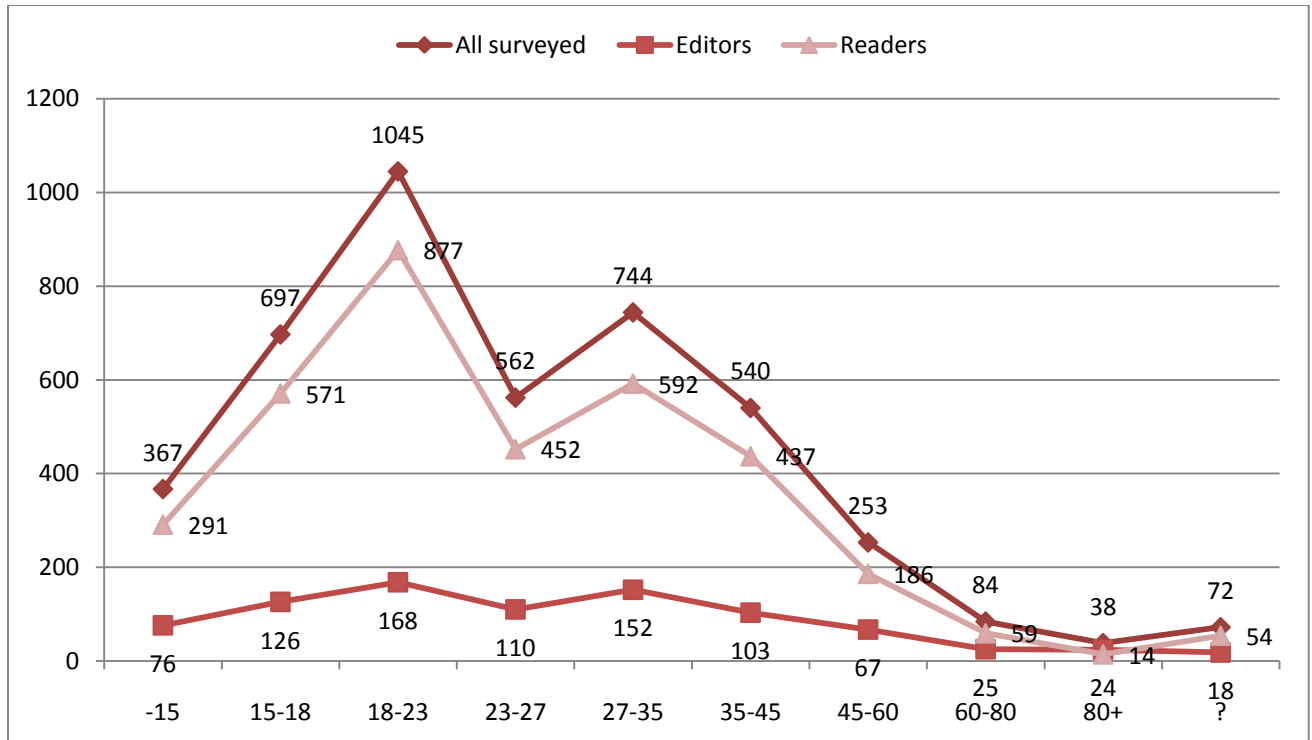
The distribution of editors based on the length of editing is displayed in the chart below.



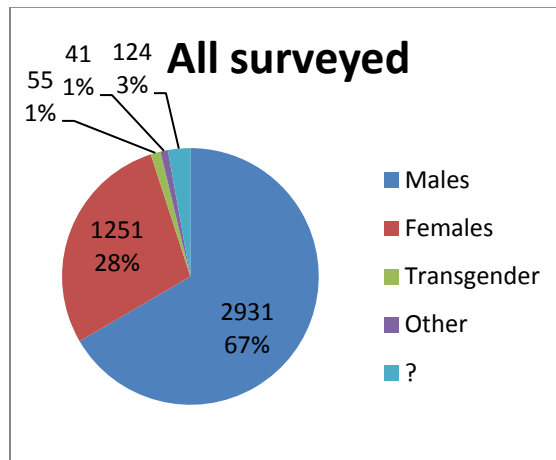
The age structure of the surveyed participants is displayed in the chart below.

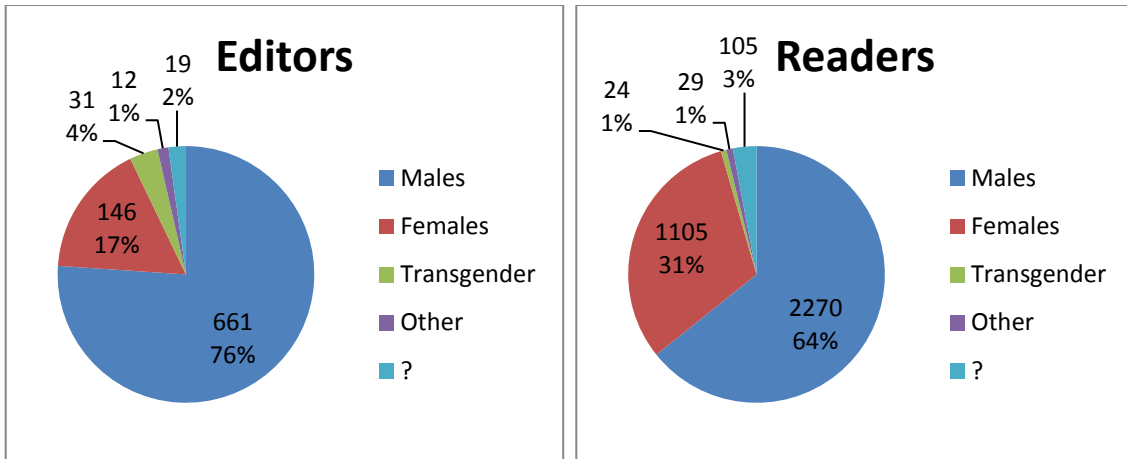
² Initially, the poll was devised to ask only about the gender information, but later we realized that it would make sense to ask for another "simple" piece of information – what age group they belong to.

³ It is located on top of every Wikipedia page; editors can hide the banner at will.

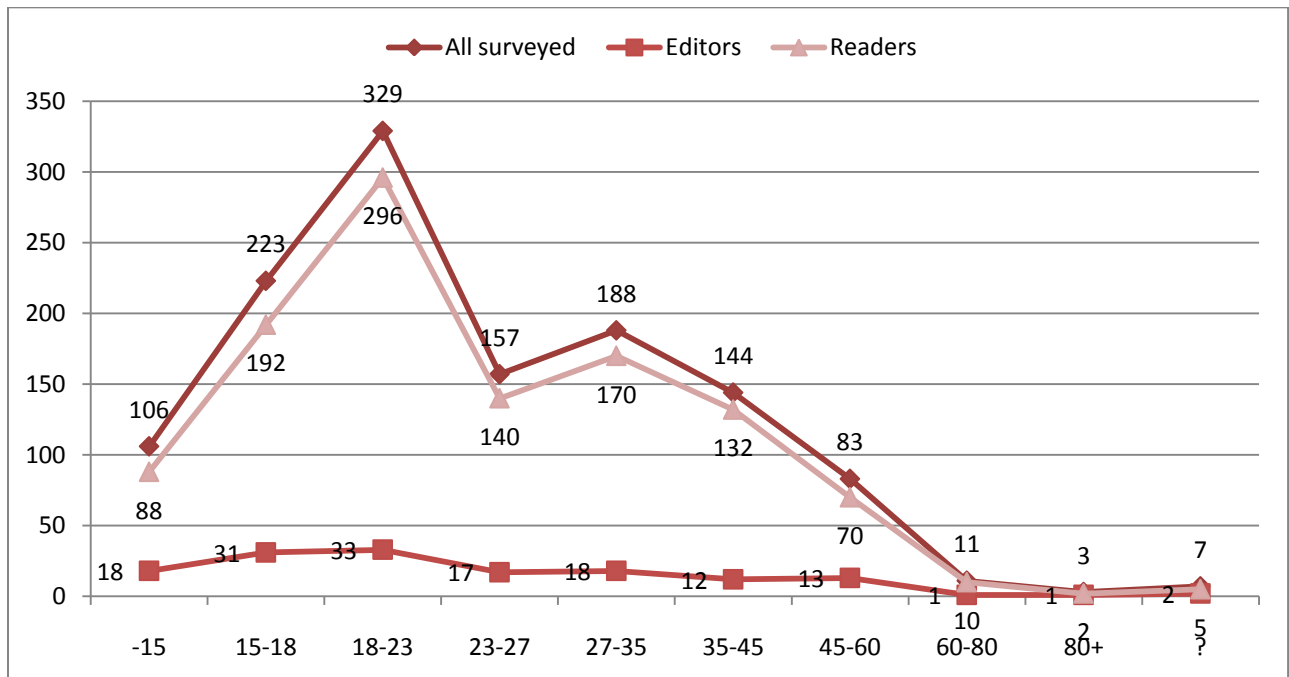


The gender structure of the surveyed participants is displayed in the charts below.





The age structure of women is displayed in the chart below.



In the following three charts, the gender-age structures are displayed for 1) all surveyed participants, 2) editors, and 3) readers, respectively.

