White, Man, and Highly Followed: 
Gender and Race Inequalities in Twitter

Johnnatan Messias  
Universidade Federal de Minas Gerais  
Belo Horizonte, Brazil  
johnnatan@dcc.ufmg.br

Pantelis Vikatos  
University of Patras  
Rio, Greece  
vikatos@ceid.upatras.gr

Fabrício Benevenuto  
Universidade Federal de Minas Gerais  
Belo Horizonte, Brazil  
fabricio@dcc.ufmg.br

ABSTRACT
Social media is considered a democratic space in which people connect and interact with each other regardless of their gender, race, or any other demographic factor. Despite numerous efforts that explore demographic factors in social media, it is still unclear whether social media perpetuates old inequalities from the offline world. In this paper, we attempt to identify gender and race of Twitter users located in U.S. using advanced image processing algorithms from Face++. Then, we investigate how different demographic groups (i.e. male/female, Asian/Black/White) connect with each other. We quantify to what extent one group follow and interact with each other and the extent to which these connections and interactions reflect in inequalities in Twitter. Our analysis shows that users identified as White and male tend to attain higher positions in Twitter, in terms of the number of followers and number of times in user’s lists. We hope our effort can stimulate the development of new theories of demographic information in the online space.

CCS CONCEPTS
• Human-centered computing → Social media;  • Applied computing → Sociology;

1 INTRODUCTION
Nowadays, millions of people constantly use online social networking sites, such as Facebook and Twitter, sharing content about their daily lives and things that happen around them. These systems have revolutionized the way we communicate, by organizing our offline social relationships in a digital form.

Increasingly, our society has been treating actions in the online and offline space in an indistinguishable way. It is now common to see cases in which content posted on Facebook or Twitter is used to vet job candidates, support divorce litigation, or terminate employees. Although these platforms provide a democrat space for conversations and debates, they also open space for unsolved issues of our society, such as gender and race inequalities.

Indeed, inequalities and the main barriers that keep a given demographic from rising beyond a certain level in a hierarchy have been perceived and quantified in many online systems. For example, recent efforts show gender disparities on content production in Wikipedia [17, 42]. A recent study shows that Black tenants have much fewer chances of renting a place on Airbnb1 [12]. Also, some drivers for both UberX2 and Lyft3 discriminate on the basis of the perceived race of the traveler [15]. There is also evidence that users in Twitter perceived as female experience a ‘glass ceiling’, similar to the barrier females face in attaining higher positions in companies [30].

Overall, identifying inequalities and asymmetries in demographics in the online world is crucial for the development of features that can promote equality and diversity in these systems. However, despite the recent surge of interest in demographic aspects of users from social networks, inferring such demographic attributes have been usually challenging and limited due to the difficult to identify them from the users in these systems.

Particularly, in Twitter, a number of efforts have explored gender differences by identifying gender based on the user’s name [2, 18, 30]. However, studies of race inequality in Twitter are still limited. In this paper, we focus on filling this gap by answering the following research question: Do Gender and Race affect online interactions and impact people’s connections in Twitter? To do that, we crawled a large scale sample of active Twitters and then we identify the gender and race of about 1.6 million users located in U.S by using Face++4, a face recognition software able to recognize gender and race of identifiable faces in the user’s profile pictures.

Our findings reinforce previous observations about disadvantages against female users in Twitter and also identify advantages for White users, in comparison with those identified as Black and Asian. Our analysis also characterizes how these different demographic groups interact among themselves, providing several insights about the perceived inequality causes. We believe our findings not only contribute to the social science literature, but we also hope that they can contribute to the design of link recommendation systems that can act to diminish such inequalities. As a final contribution, we plan to make our dataset available to the research community by the time of publication of this article.

The rest of the paper is organized as follow. The Section 2 briefly reviews related efforts in the literature. Then, in Section 3, we present the Twitter and demographic dataset. Next, Section 4 characterizes gender and race inequalities in visibility and in Section 5 in terms of connections and interactions in Twitter. Finally, Section 6 concludes the paper and offers directions of future work.

1http://www.airbnb.com
2http://www.uber.com
3http://www.lyft.com
4http://www.faceplusplus.com
2 RELATED WORK

In this section, we review the related literature along two axes. First, we discuss the methodology used by efforts that measure demographic factors in Twitter. Then, we discuss the main efforts that have explored gender and race inequalities in social networks.

2.1 Demographics in Social Media

There are several studies that retrieve information from social media to analyze and predict real-world phenomena such as stock market [3], migration [24], election [31, 40], and also political leaning of Twitter users [19, 33].

In the field of demographics, one of the first efforts in demographics compares geographic and demographic distribution of users of U.S. population [26]. After that, there have been many studies that extract demographic information, using different strategies and different purposes [2, 18, 21, 30, 34, 41]. Particularly, Chen et al. [7] focus on demographic inference using a dataset of Twitter profiles and profile images. They compare the five signals in order to categorize users to their demographic status. They use users’ names, self-descriptions, tweets, social networks, and profile images to infer demographic attributes (gender, age). [11] describe an alternative methodology assuming that the demographic profiles of visitors to a website are correlated with the demographic profiles of followers of that website on social network, e.g. Twitter; and propose a regression model to declare demographic attributes (gender, ethnicity, education, income, and child status). Meanwhile, [29] show that female streamers receive significantly more objectifying comments while male streamers receive more game-related comments.

Another study, An et al. [1] examine the correlation of hashtags that are used in different demographic groups, using Face++ and the user’s profile pictures. They show that this strategy is reliable and provides accurate demographic information. More recently, Chakraborty et al. [6] analyze the demographic distribution of Twitter Trending Topics, using Face++ as well. They use an extensive data collected from Twitter in order to make the first attempt to quantify and explore the demographic biases in the crowdsourced recommendation system. They identify that a large fraction of trends are promoted by groups whose demographics are significantly different from the overall Twitter population. More worryingly, certain demographic groups are systematically under-represented among the promoters of the trending topics.

Our effort uses the similar strategy to gather demographic information, but we investigate very different research questions as we focus on inequalities and demographic aspects of users connected. We note that this strategy to gather demographic aspects, applied to active users within a country, allows us to take a step further in many existing studies on the field as it allows us to study race inequalities and also allows us to investigate how connections are established among demographic groups (i.e. male/female, Asian/Black/White). We believe our new and large-scale dataset may largely contribute to this research field.

2.2 Inequalities in Twitter Visibility

Homophily states that similarity breeds social connections [22]. In our society, people tend to characterize and categorize others in terms of demographic aspects, such as gender and race [43], making them intuitively natural concepts through which people might account similarities. The existing literature is still limited in explaining the inequalities that arise from the establishments of these social connections in Twitter.

Most of the existing efforts in this space attempt to quantify and understand other factors that are highly correlated with visibility (i.e. number of followers). Particularly, Morales et al. [27] show that the reputation and the impact on the society of a user constitute a significant factor of the online visibility growth. Other efforts show that the attention that users gain is correlated with his/her online behavior in terms of the frequency of their interactions [9, 35, 44]. Freitas et al. [14] created 120 socialbots in Twitter, with different behaviors and demographics features, to show that users who post more, post about specific topics, reply, favorite or mention other frequently, may acquire more followers, and consequently, more visibility. Despite the extreme importance of these efforts, they do not investigate the correlation of demographic issues with visibility. Thus, our effort is complementary to theirs.

There have been efforts quantifying differences and inequalities in many other social media, including Wikipedia [17, 42] and Pinterest [16]. Also, inequalities in between conversation and objectification in different genders have been analyzed in online gaming such as Twitch [29].

However, the study that is closest related to ours explores gender inequalities in Twitter [30]. Authors show that gender may allow inequality to persist in terms of online visibility. Our work further explores race as a new demographic dimension. More importantly, unlike previous studies, our work focuses on investigating how different demographic groups (i.e. male/female, Asian/Black/White) connect with each other. We quantify to what extent one group follows and interacts with each other and the extent to which these connections and interactions reflect in inequalities in Twitter.

3 DEMOGRAPHIC INFORMATION DATASET

This section describes the data collection and the method to infer demographic information of individual Twitter users. Our ultimate goal consists of gathering a dataset containing active U.S. Twitter users, their demographic information, a sample of their connection graph (friends), and their tweets. Next, we describe our steps to create this dataset and also discuss its main limitations.

3.1 Twitter Dataset Gathered

To identify active Twitter users we gathered data from the 1% random sample of all tweets, provided by the Twitter Stream API 3. Our dataset covers a period of three complete months, from July to September 2016. In total, we collected 341, 457, 982 tweets posted by around 50, 270, 310 users.

We restricted our demographic studies to Twitter users located in the United States. As geographic coordinates are available on Twitter only for a limited number of users (i.e. < 2%) [5], our strategy to identify U.S. Twitter users was based on part of the methodology used in previous efforts [6, 20]. We have used the timezone information as a first filter and then we attempted to remove from this filtered dataset those users that provided free text location

3https://dev.twitter.com/streaming/public
indicating they are not U.S. (i.e. Montreal, Vancouver, Canada). We end up with a dataset containing 6,286,477 users likely located in the United States.

3.2 Crawling Demographic Information
Most of the existing studies related to demographics of users in Twitter have looked into gender and age. Some efforts attempt to infer the user’s gender from the user name [2, 18, 21, 26]. However, some users may not use proper names, consequently their gender could not be inferred properly [21]. Others have attempted to identify patterns like ‘25 yr old’ or ‘born in 1990’ in Twitter profile description to identify the user age [37].

Here, we use a different strategy that allows us to study another demographic dimension: the user’s race. To do that, we crawl the profile picture Web link of all Twitter users identified as located within the United States. In December 2016, we crawled the profile picture’s URLs of about 6 million users, discarding 4,317,834 (68.68%) of them. We discarded users in two situations, first when the user does not have a profile picture and second when the user has changed her picture since our first crawl. When users change their picture, their profile picture URL changes as well, making it impossible for us to gather these users in a second crawl.

From the remaining 1,968,643 users, we submitted the profile picture Web links into the Face++ API. Face++ is a face recognition platform based on deep learning [13, 45] able to identify the gender (i.e. male and female) and race (limited to Asian, Black, and White) from recognized faces in images.

We have also discarded those users whose profile pictures do not have a recognizable face or have more than one recognizable face, according to Face++. Our baseline dataset, also used by Chakraborty et al. [6] work, contains 1,670,863 users located in U.S. with identified demographic information. Table 1 shows the demographic distribution of users in our dataset across the different demographic groups. We use this dataset to study inequalities in visibility that will be present in Section 4.

Table 1: Demographic distribution of 1.6 million users in baseline dataset.

<table>
<thead>
<tr>
<th>Race</th>
<th>Gender</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Asian</td>
<td>120,950 (7%)</td>
<td>177,205 (11%)</td>
</tr>
<tr>
<td>Black</td>
<td>130,954 (8%)</td>
<td>107,827 (6%)</td>
</tr>
<tr>
<td>White</td>
<td>538,625 (32%)</td>
<td>595,302 (36%)</td>
</tr>
<tr>
<td>Total</td>
<td>790,529 (47%)</td>
<td>880,334 (53%)</td>
</tr>
</tbody>
</table>

3.3 Gathering Social Connections and Interactions
Ideally, to study how different demographic groups are connected and interact with each other, we would like to have at our disposal the followers (others who are following the user) and friends (others who are followed by the user) of all users from our dataset. However, the number of followers and friends of the 1.6 million users in our dataset surpasses 6.4 billion users, which is unfeasible to be crawled through the Face++ API. Next, we discuss a sampling procedure we used to create a dataset of users, their connections and their interactions with demographic information.

For this specific analysis, we randomly selected a total of 6,000 users from our dataset, 1,000 users of each demographic group (i.e. White male, White female, Black male, Black female, Asian male, and Asian female). Then, we gathered their friends. As some of these users have a prohibitive large number friends, we limited our crawl to gather only the most recent 5,000 friends of a given user, as this is the maximum number of user IDs the Twitter API returns per request. However, this strategy gathered the entire friends’ list for 98.51% of the users.

Then, we follow the same methodology we discussed before to gather the demographic information of users. Firstly, we remove users that are not located in U.S., and then we attempt to identify the demographic aspects of each user using Face++ API. It is undeniable that the aggregated number of friends is extremely high and it is difficult to gather due to the limitation in term of requests that Face++ API allows us to make. Thus we were limited to gather demographic of at least 5% of the friend’s list.

Our social connections and interactions dataset contains 428,697 users with the proper demographic information identified. Table 2 presents the total number of the friends gathered for each demographic group that we examined. The average and median percentage of friends of the 6,000 users for which we were able to gather demographic information are 10.15% and 9.40%, respectively. We note that these fractions are usually higher than 5% as some extra users were previously gathered in our 1.6 million demographic dataset.

Also, we worked in a similar way for the analysis of interactions between the demographic groups. More specifically, we check the users that retweet and mention the tweets of the randomly selected users. To do that, we gathered all tweets (limited to a maximum of 3,200 tweets due to Twitter REST API limitation4) of our 6,000 users. We then identified the users who were mentioned or retweeted and we limited to gather the demographic information of only 5% of retweeters and mentioned users for our analysis. Table 3 summarizes the amount of crawled users for each demographic group.

Our study about the connections among demographic groups (Section 5) is based on this specific dataset.

Table 2: Number of Friends in each Group

<table>
<thead>
<tr>
<th>Race</th>
<th>White</th>
<th>Black</th>
<th>Asian</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>151,840</td>
<td>52,437</td>
<td>24,299</td>
<td>228,576</td>
</tr>
<tr>
<td>Black</td>
<td>137,010</td>
<td>31,011</td>
<td>32,100</td>
<td>200,121</td>
</tr>
<tr>
<td>Total</td>
<td>288,850</td>
<td>83,448</td>
<td>56,399</td>
<td>428,697</td>
</tr>
</tbody>
</table>

3.4 Potential Limitations
The gender and race inference are challenge tasks, and as other existing strategies have limitations and the accuracy of Face++ inferences is an obvious concern in our effort. The limitations of our data is discussed next.

4https://dev.twitter.com/rest/reference/get/statuses/user_timeline
We are interested to analyze the association of demographic status with visibility and discover possible inequalities. These asymmetries can be derived from the prejudices and stereotypes in the selection of which user will follow based on gender or race. Nilizadeh et al. [30] provide evidence of inequalities and asymmetries in terms of visibility between males and females.

We focus only in the visibility in the social network and not in user’s influence of the audience which his/her is exposed. For our purpose we use two different features to measure a user’s visibility: its follower count and its listed count. In other words, we test whether the gender and race affect users’ number of followers and how many lists they are added to.

The number of followers measures the real size of the audience that someone is exposed to. Finally, listed count represents the amount of times a user was added into a specific list by others in Twitter. Twitter Lists allow users to group and organize Twitter accounts that tweet on a topic that is of interest to her, and follow their collective tweets. Many users carefully create Lists to include other Twitter users who they consider as experts on a given topic. Previous research efforts [36] attempt to gather Lists in large scale to find experts in Twitter.

As a baseline for comparisons, Table 1 shows the demographic breakdown of users in our dataset across the different demographic groups. We can note a prevalence of females (52.69%) in comparison to males (47.31%) and a predominance of Whites (67.86%) in comparison to Blacks (14.29%) and Asians (17.85%). This means that if we pick users randomly in our dataset, we would expect demographic groups with these proportions. We used these proportions next as a baseline for characterizing inequalities in Twitter.

4.1 Gender Inequality

We begin our analysis by sorting the number of followers of all users in our dataset and we check the fraction of males and females in an incremental proportion of top followed users. Figure 1 provides some insight on this point. Although, the number of females is higher than the males as shown in Table 1, males tend to be among the top followed users. The most significant difference exists in top 1% with a fraction of 57.34% of males and 42.65% of females, which is almost 15% and this difference is decreasing until the top 14%, where the fraction of females become higher than the fraction of males.

The same kind of observation can be made for the rank of lists, as shown in Figure 2. However, the discrepancy perceived is smaller in the lists basically restricted to the top 1% most listed users with 51.23% of males and 48.64% of females. In top 6% the fraction of females become higher than males. The same occurs in top 8%. The results of this analysis reflect the general idea that higher positions are usually taken by males [38].

![Figure 1: Distribution of the fraction of males and females in the rank of users with most followers. The intersection point happens in the Top 14%, meaning the fraction of females became higher than males after this point.](image)

Our observations also reinforce previous findings related to gender in Twitter [30], suggesting the existing of the so called ‘glass ceiling effect’, which poses that females face an invisible barrier at the highest levels of an organization [10]. One concern raised by authors on that work was that disparity may be driven by a small number of ‘elite’ users. Our findings diminish such concerns as our dataset is at least two order of magnitudes larger and we noted such discrepancies even in the top 10% positions (i.e. top 167,086 users).
We now turn our attention to the analysis of race inequalities. Similarly to what we have done in the previous section, we examine here the presence of each user race within the top positions in a rank of top followed users and top listed users. Figure 3 and Figure 4 provide the necessary information about race and top followed users. We can note that the amount of White users tend to be higher in the both top positions. In the rank of users with more followers and in the rank of most listed users.

Our results suggest that the race disparity in Twitter visibility occurs, meaning that at the highest levels of visibility, users perceived to be White come out on top position. This observation reinforces many previous observations related to race inequality in United States [4, 32].

4.3 Taking Together Gender and Race Inequalities

Finally, we attempt to quantify the existing inequality when we look into gender and race aspects at the same time. We note from Figure 5 that the amount of White males among the top 1% users with more followers is 42.27%. This represents an increase of 10.04% in the fraction of White males in comparison to the expected proportion of White males in our dataset, which is 32.23%. We can note that populations of Black females and Asian females experience a reduction of 3.60% and 3.84% in the top 1% most followed users, respectively. We note similar trends in terms of lists as we can see Figure 6.

To better analyze to what extent demographic groups in the top 1% is substantially higher or lower than what one would expect based on our demographic information dataset, Table 4 shows the increase or decrease in the proportion of each demographic group in comparison with such baseline expected population. Overall, this results quantify the perceived inequalities that six demographic groups experience.
groups experience at the highest levels of visibility (i.e. top 1% most followed and listed). The most privileged demographic group in the rank of followers is the group of White males, with more than 20% of users than a baseline expected random sample. The amount of White females is also higher than the baseline, but much smaller amount (only 3%) in comparison to males. The most unprivileged groups are Asian females and Black females, underrepresented in the top 1% by more than 31%.

Table 4: Relative proportion of each demographic group in the top 1% rank of users with more followers and the most listed in comparison to the baseline expected population.

<table>
<thead>
<tr>
<th>Race</th>
<th>Followers (%)</th>
<th>Listed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Asian</td>
<td>-12.70</td>
<td>-36.25</td>
</tr>
<tr>
<td>White</td>
<td>+31.16</td>
<td>-7.27</td>
</tr>
</tbody>
</table>

5 DEMOGRAPHIC GROUP INTERCONNECTIONS

Next, we analyze how demographic groups are connected with each other. Our ultimate goal is to examine the proportion of connections and interactions among each demographic group. Thus, for this analysis we used the dataset related to friends, followers and interactions discussed previously.

In order to represent connection probabilities among demographic groups we built a probabilistic graph and we compute the probability of a connection/interaction for every pair of users. We declare as probability the empirical probability of the existence of the specific event in our dataset.

To gain a holistic view, we aggregate all pairs of users making the nodes to represent demographic groups and each direct edge to represent the probability of a relationship happened, like friendship or interaction. The sum of all outgoing probabilities for each demographic group is 1.0.

5.1 Gender and its Interconnections

We begin by analyzing interconnection among groups of users separated by Gender. We analyze the probability of males and females users to be friends with (i.e. to follow) other males and females users. In our analysis, we can note that females tend to equally follow males and females in 50%, while males tend to follow 56% of males against 44% of females. If we take the numbers in Table 1 as a proxy for the expected population of males and females in a randomly created list of friends, we note that even the equally distributed list of friends of females is actually disproportional as the expected percentage of males and females are 47.31% and 52.69%, respectively. This means that males tend to follow 18.37% more other males than expected, whereas females tend to follow 5.7% more males than expected. This suggests that both, males and females connections, take their part of responsibility on these gender inequalities, but it also shows that males tend to take the larger part of the responsibility.

In terms of interactions, similar observations can be made. However, the proportion of males and females retweet/mentioned by females is quite close to the expected population (53% of females against 47% for males). For the group of males, we noted a much higher disparity in terms of interactions in comparison to friendship. Males tend to interact with 64% of other males and only 36% of females.

This result quantifies the perceived inequality in Twitter. These observations are inline with perceived inequalities in the offline world, including in academic job hiring, earnings, funding and academic rating as well as income differences between males and females in the population of high earners.

5.2 Race and its Interconnections

Next, we turn our attention to the analysis of interconnection among groups of users separated by race. Figure 7 shows these interconnections in terms of friendship and interactions, respectively.

![Figure 7: Probability of friendship and interaction for demographic groups of users separated by race](image)

We start by analyzing the self loops on each node. We note that White users tend to follow about 79% of other White users, Black users tend to follow 32% of Black users and Asians tend to follow 16% of Asians. This means that the key to analyze these Figures is to recall the reference distribution of users in our dataset according to their race, in which Whites represent 67.86%, Black accounts to 14.29%, and Asian accounts to 17.85%. Thus, we can note that, in comparison to the expected distribution of users, Whites tend to follow 16.41% more Whites than expect, Blacks tend to follow impressive 2.24 times more Blacks than expect, and Asians end up following less (10.26%) Asians than expect. In other words, the expected homophily was not clear for the case of Asians.

Similar observations can be made for interactions, only exception for among Asians, where the fraction of interactions is higher than the reference distribution. Although Asians tend to follow more White users (even more than the reference distribution), the interactions with White users are below the reference fraction. Overall, Whites and Blacks showed quite high numbers of endogenous connections and interactions as both groups seem to avoid following a different race. The stronger mutual interconnection is between Black users.
5.3 Demography of Interconnections
Finally, we characterize the interconnection among groups of users both separated by race and gender. Figures 8 and 9 show these interconnections in terms of friendship and interactions, respectively. However, given the high number of edges in these figures, instead of providing the transition probabilities, we computed their relative increase or decrease in comparison to the expected demographic population.

Figure 8: Relative probability of friendship for groups of users separated by race and gender. The value declares the relative increase (blue color) or decrease (red color) in comparison to the expected demographic population.

In addition to reinforce previous findings from the analysis of gender and race separately presented before, these results show interesting trends. First, we note that White males tend to receive only positive edges from White and Asian groups, meaning that only Black users (both males and females) tend to have a fraction of White males users as friends that are relatively low than the expected demographic distribution in Twitter. This suggests that race plays a more important role for connections with White males. In terms of interactions, similar observations hold, except that Asian females tend to have a relatively smaller number of interactions with White males. Interestingly, White females appear over-represented only in the list of friends of Asia females. Another interesting pattern is that the highest values for self-loops are for Black males (130%) and Black females (86%), which is even higher for interactions, 232% and 164%, respectively. The interconnections between these two groups are also impressively large. On the other hand, the lower values for self-loops are for Asians, with negative values for friendship, but positive for interactions, as discussed before. Interestingly, all incoming links to the two Asian groups are negative, meaning that the other two demographic groups tend to have an under-represented proportion of Asians as friends and in terms of interactions.

Figure 9: Relative probability of interactions for groups of users separated by race and gender. The value declares the relative increase (blue color) or decrease (red color) in comparison to the expected demographic population.

6 CONCLUDING DISCUSSION
This paper characterizes demographic information of users in Twitter, identifying inequalities and asymmetries in gender and race. To do that we gather gender and race of Twitter users located in U.S. using advanced image processing algorithms from Face++. We collect a large sample of 1,670,863 users located in U.S. with identified demographic information. We have also crawled a large sample of tweets and friends of these to study connections and interactions among demographic groups of users.

Our analysis reinforces evidence about gender inequality in terms of visibility and introduces race as a significant demographic factor, which reveals hiding prejudices between groups of different demographic status. More important, our effort is the first of a kind to quantify to what extent one demographic group follows and interact with each other and the extent to which these connections and interactions reflect in inequalities in Twitter.

Among our findings, we show that the Twitter glass ceiling effect, typically applied to females, also occurs in Twitter for males, if they are Black or Asians. Our analysis reinforces evidence about gender inequality in terms of visibility and introduces race as a significant demographic factor, which reveals hiding prejudices between groups of different demographic status. More important, our effort is the first of a kind to quantify to what extent one demographic group follows and interact with each other and the extent to which these connections and interactions reflect in inequalities in Twitter.

Among our findings, we show that the Twitter glass ceiling effect, typically applied to females, also occurs in Twitter for males, if they are Black or Asians. Our analysis reinforces evidence about gender inequality in terms of visibility and introduces race as a significant demographic factor, which reveals hiding prejudices between groups of different demographic status. More important, our effort is the first of a kind to quantify to what extent one demographic group follows and interact with each other and the extent to which these connections and interactions reflect in inequalities in Twitter.

Among our findings, we show that the Twitter glass ceiling effect, typically applied to females, also occurs in Twitter for males, if they are Black or Asians. Our analysis reinforces evidence about gender inequality in terms of visibility and introduces race as a significant demographic factor, which reveals hiding prejudices between groups of different demographic status. More important, our effort is the first of a kind to quantify to what extent one demographic group follows and interact with each other and the extent to which these connections and interactions reflect in inequalities in Twitter.

ACKNOWLEDGMENTS
This research is funded by projects InWeb (grant MCT/CNPq 573871/2008-6) and MASWeb (grant FAPEMIG/PRONEX APQ-01400-14), by grants from CAPES, CNPq,