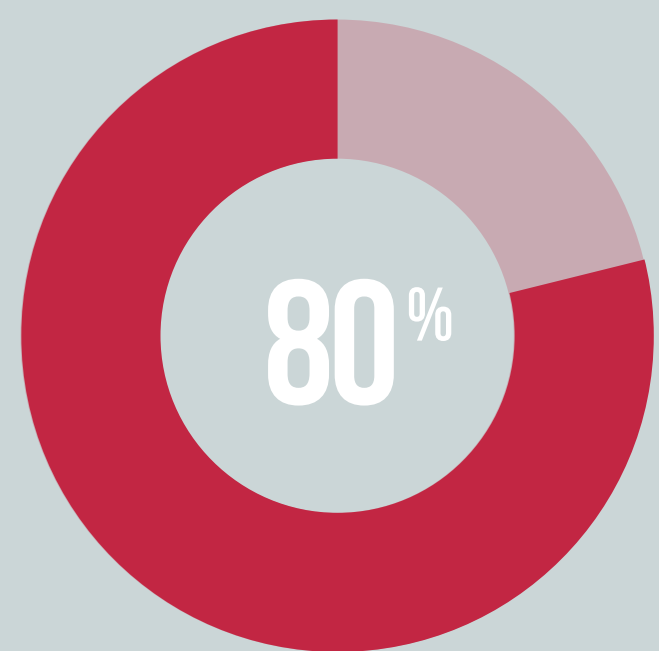


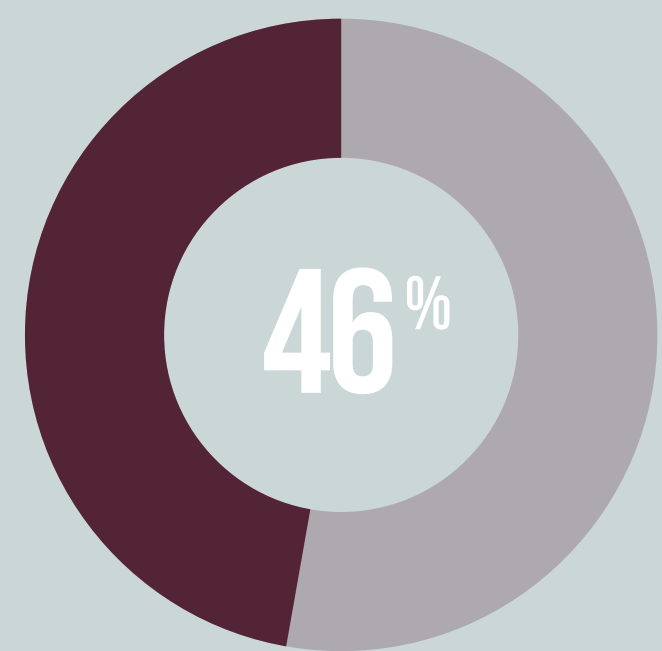
# WHAT IS YOUR OPINION WORTH?

## A STUDY OF HOW TO MEASURE THE INFLUENCE OF DIGITAL PERSONAS



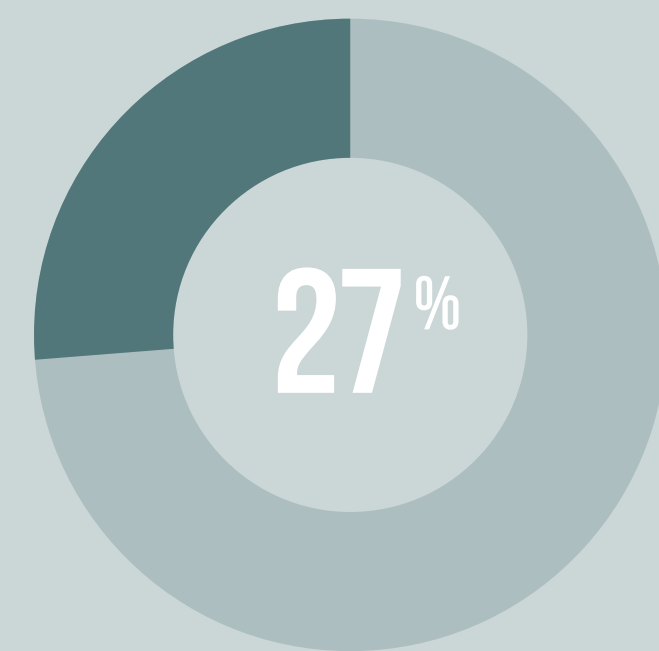
### PEOPLE LOOK TO SOCIAL MEDIA FOR BRAND PROMOTIONS

DIRECT MARKETING WITH COUPONS, PROMOTIONS, AND DISCOUNTS VIA SOCIAL MEDIA HAS OPENED A TWO-WAY CONVERSATION BETWEEN COMPANIES AND THEIR CUSTOMERS.<sup>1</sup>



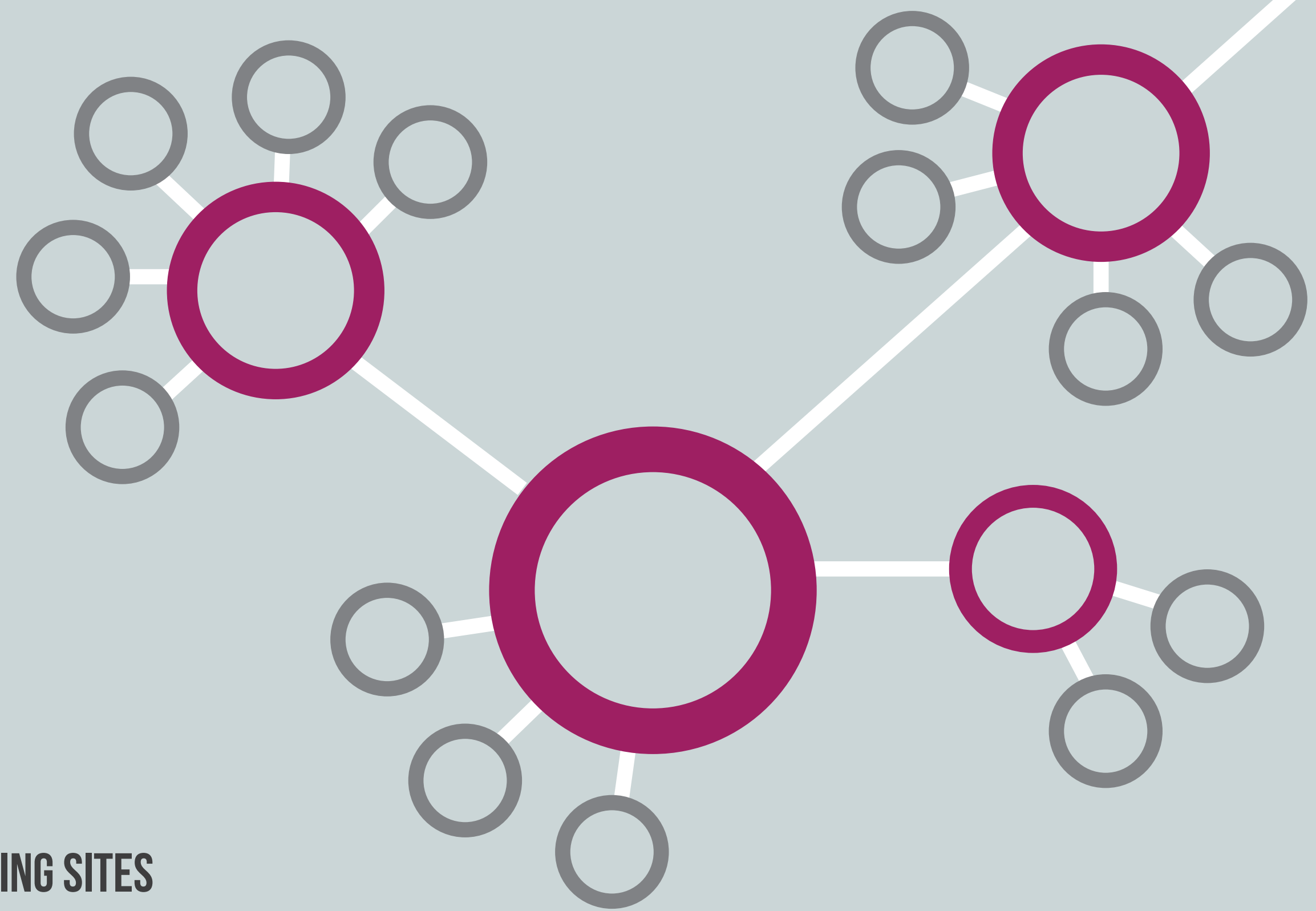
### USERS COUNT ON SOCIAL MEDIA WHEN MAKING A PURCHASE

AS CONNECTED DEVICES BECOME A WAY OF LIFE FOR MANY, DIGITAL COMMUNITIES ARE MORE RELIED UPON. PEOPLE LOOK TO SOCIAL MEDIA TO GATHER OPINIONS FROM TRUSTED SOURCES. UNDERSTANDING WHO HAS THE GREATEST PULL IN THESE COMMUNITIES BECOMES EXTREMELY VALUABLE.<sup>2</sup>



### TOTAL U.S. INTERNET TIME SPENT ON SOCIAL NETWORKING SITES

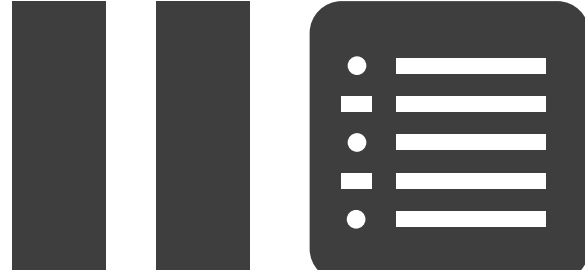
WHEN OVER A QUARTER OF THE TIME SPENT ON THE INTERNET IS ON SOCIAL NETWORKING SITES, ADVERTISING DOLLARS COMPETE FOR SPACE AT A PREMIUM. TARGETING USERS WITH GREATER INFLUENCE AMONG COMMUNITIES OF INTEREST ALLOWS FOR A MUCH GREATER RETURN ON INVESTMENT.<sup>3</sup>



## PROBLEM

OR PERHAPS BETTER PHRASED: OPPORTUNITY. FINDING A REPEATEDLY SUCCESSFUL MARKETING STRATEGY FOR SOCIAL MEDIA HAS BEEN ELUSIVE. DETERMINING A RELIABLE WAY TO GAUGE RETURN ON INVESTMENT HAS BEEN EVEN HARDER. THIS PROJECT PRESENTS FINDINGS TO UNDERSTAND SOCIAL INFLUENCE IN THE DIGITAL WORLD.

SOCIAL INFLUENCE IN THE PHYSICAL WORLD CAN BE CORRELATED WITH INTANGIBLE ATTRIBUTES LIKE CHARISMA, ATTRACTION, AND CHARM. IN TRANSLATION TO A DIGITAL CONTEXT, THE CONCEPT OF INFLUENCE IS STILL VERY MUCH CONNECTED TO THESE TYPES OF HUMAN ELEMENTS, BUT THEIR IMPACT BECOMES MEASURABLE. THE VALUE IN DETERMINING HOW MUCH IMPACT AN INDIVIDUAL HAS ON A SPECIFIC TOPIC OF INTEREST IS BECOMING CLEARER AS MORE COMPANIES DEPEND ON DIGITAL COMMUNITIES FOR PRODUCT AWARENESS.

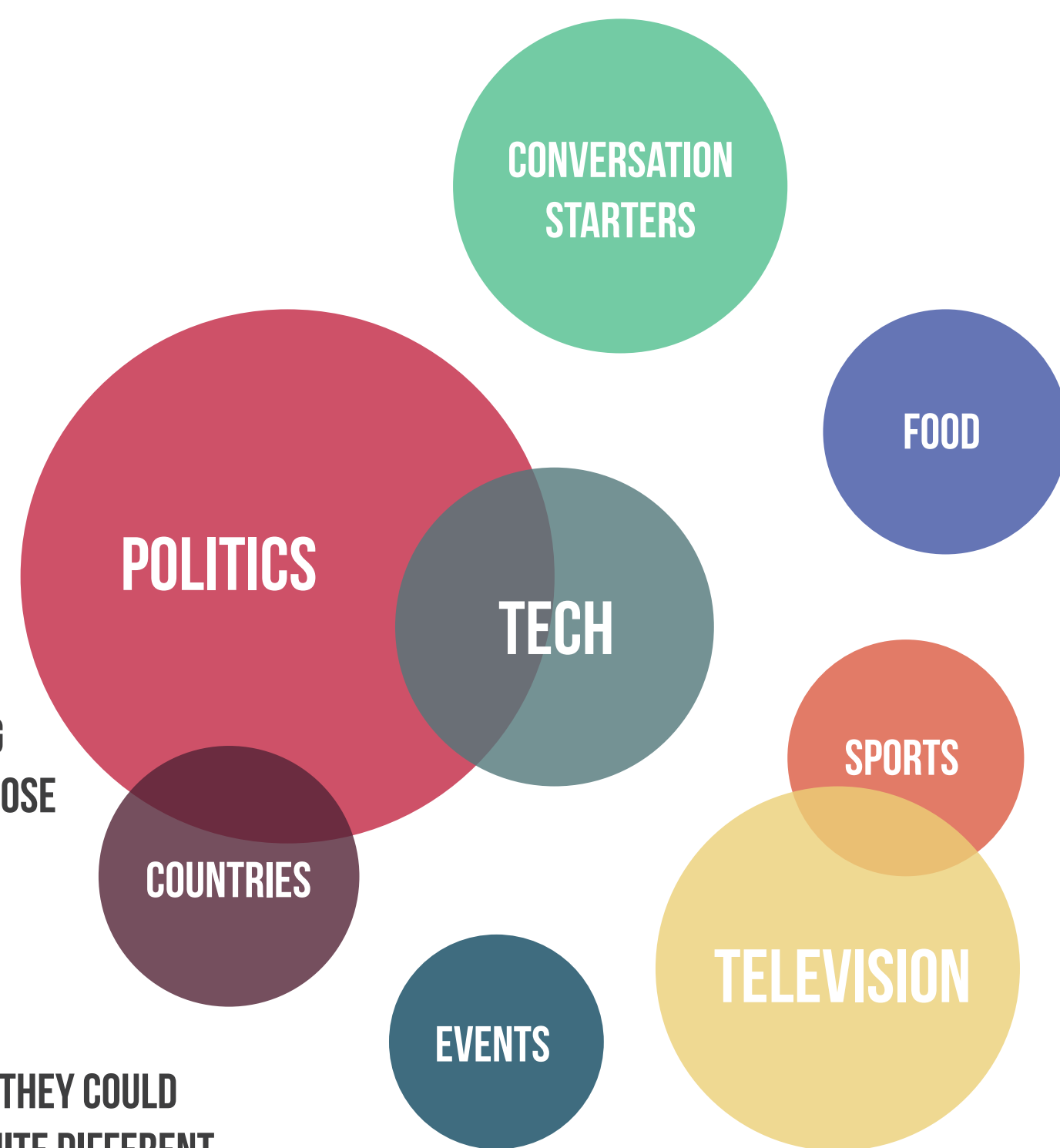


## TOPICS

THE FIRST STEP IN IDENTIFYING INFLUENCE IS TO FIND OUT WHAT AREA OF INTEREST AN INDIVIDUAL INTERACTS WITHIN. THIS TASK NARROWS THE FIELD TO A MORE MANAGEABLE SIZE FOR DETAILED ANALYSIS. GIVEN THAT 90% OF ALL DATA CREATED HAS BEEN DONE IN THE PAST TWO YEARS ALONE, THIS PROBLEM IS A COMPOUNDING ONE. CHUNKING DATA INTO INDIVIDUAL TOPICS DOESN'T MEAN THAT THERE AREN'T OVERLAPPING CORRELATIONS BUT IT PROVIDES THE FIRST FOCUS. IN THIS PROJECT WE CHOSE THE FOLLOWING THREE CATEGORIES FOR FURTHER ANALYSIS:

SPORTS, PHOTOGRAPHY, COOKING

IN THE CASE OF THESE THREE CATEGORIES, THERE ARE INSTANCES WHERE THEY COULD ALL OVERLAP WITH ONE ANOTHER. ON THEIR OWN, HOWEVER, THEY ARE QUITE DIFFERENT. THIS ALLOWS FOR AN INVESTIGATION TO UNCOVER THE MOST POPULAR SUBTOPICS AND WHETHER OR NOT THEIR OVERLAPS WILL SHOW UP IN A MEANINGFUL WAY.



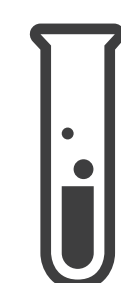
## ANALYSIS

AFTER ESTABLISHING A SET OF THREE SAMPLE TOPICS, THE NEXT STEPS IN THE PROCESS CAN BE OUTLINED AS FOLLOWS:



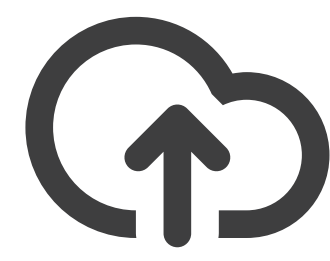
### KEYWORD DEFINITIONS

WITHIN EACH TOPIC, THERE ARE A HUGE NUMBER OF KEYWORDS THAT NEED TO BE CORRELATED WITH THE HIGHER LEVEL CATEGORY.



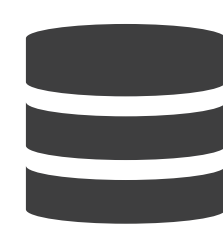
### ALGORITHM DEVELOPMENT

AN ALGORITHM WAS CREATED TO GATHER THE IMPACT RANKING OF INDIVIDUAL TWITTER ACCOUNTS. AMONG OTHERS, PRIMARY WEIGHT CONTRIBUTORS INCLUDE THE NUMBER OF FOLLOWERS, RETWEETS, AND OUTBOUND LINKS.



### INSTANCE IN THE CLOUD

AMAZON'S EC2 SERVICE WAS USED TO HOST A VIRTUAL MACHINE DEDICATED FOR RESEARCH USE.



### DATABASE CONFIGURATION

TWO RELATIONAL TABLES IN MYSQL WERE CREATED, ONE FOR TWEETS AND ANOTHER FOR FOLLOWER METRICS.



### DATA DOWNLOADING

A PYTHON SCRIPT WAS CONFIGURED TO DOWNLOAD DATA IN PARALLEL FROM TWITTER FOR 21 CONSECUTIVE DAYS.



### METRIC COMPARISON

RELATIONSHIP VISUALIZATIONS WERE GENERATED USING GEPHI. INITIAL WORD CLOUD EXPLORATIONS CREATED USING WORDLE.



## RESULTS

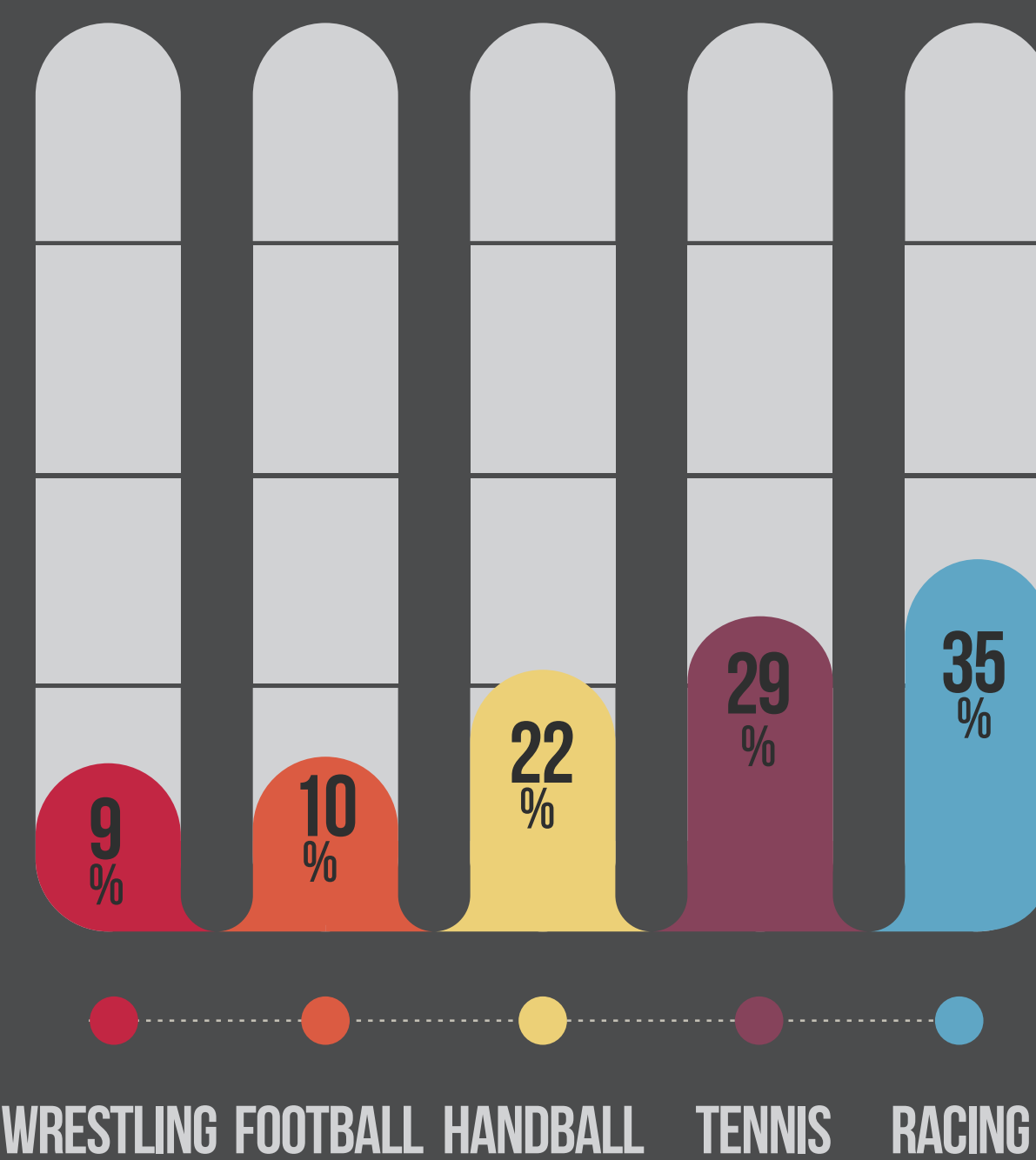
USING A DEFINED METRIC FOR DOMAIN-SPECIFIC IMPACT FACTOR IN CONJUNCTION WITH PREDICTIVE MODELLING, WE ARE ABLE TO VISUALIZE SOME INTERESTING TRENDS OF INFLUENCE AMONG TWITTER USERS. NOT ONLY CAN THESE FINDINGS ASSIST IN INTEGRATING SOCIAL MEDIA INTO AN EFFECTIVE MARKETING MODEL, THEY CAN POTENTIALLY OUTLINE A MORE DIRECT RETURN ON INVESTMENT.

THE FIRST STEPS ARE SHOWN TO THE RIGHT. FINDINGS OUTLINE SOME INTERESTING TRENDS AND TOPICS FOR FURTHER ANALYSIS. WE RAN A NUMBER OF STATISTICS ON THE GRAPH BASED ON THE DATA COLLECTED, INCLUDING AVERAGE DEGREE, AVERAGE WEIGHT DEGREE, MODULARITY AND PAGERANK. MODULARITY IS AN INTERESTING MEASURE WHICH HELPS US IDENTIFY COMMUNITIES IN OUR GRAPH, EVEN WITHIN THE DATA BEING COLLECTED USING THE SAME KEYWORD.

## TOPIC SNAPSHOT

FROM A SNAPSHOT IN TIME, THE MOST COMMON MENTIONS WITHIN THE TOPIC OF 'SPORTS' WAS OF RACING.

USING 'RACING' AS AN EXAMPLE, A KEYWORD WE ADOPTED TO FETCH DATA FOR TOPIC 'SPORTS' TO GENERATE THE RELATIONSHIP GRAPH. SEVERAL COMMUNITIES ARE FORMS. FOR EACH COMMUNITY, WE GOT THE TOP INFLUENCE BASED ON OUR IMPACT FACTOR RANKING ALGORITHM IN THE TIME PERIOD WE COLLECTED OUR DATA.



## TOP INFLUENCERS

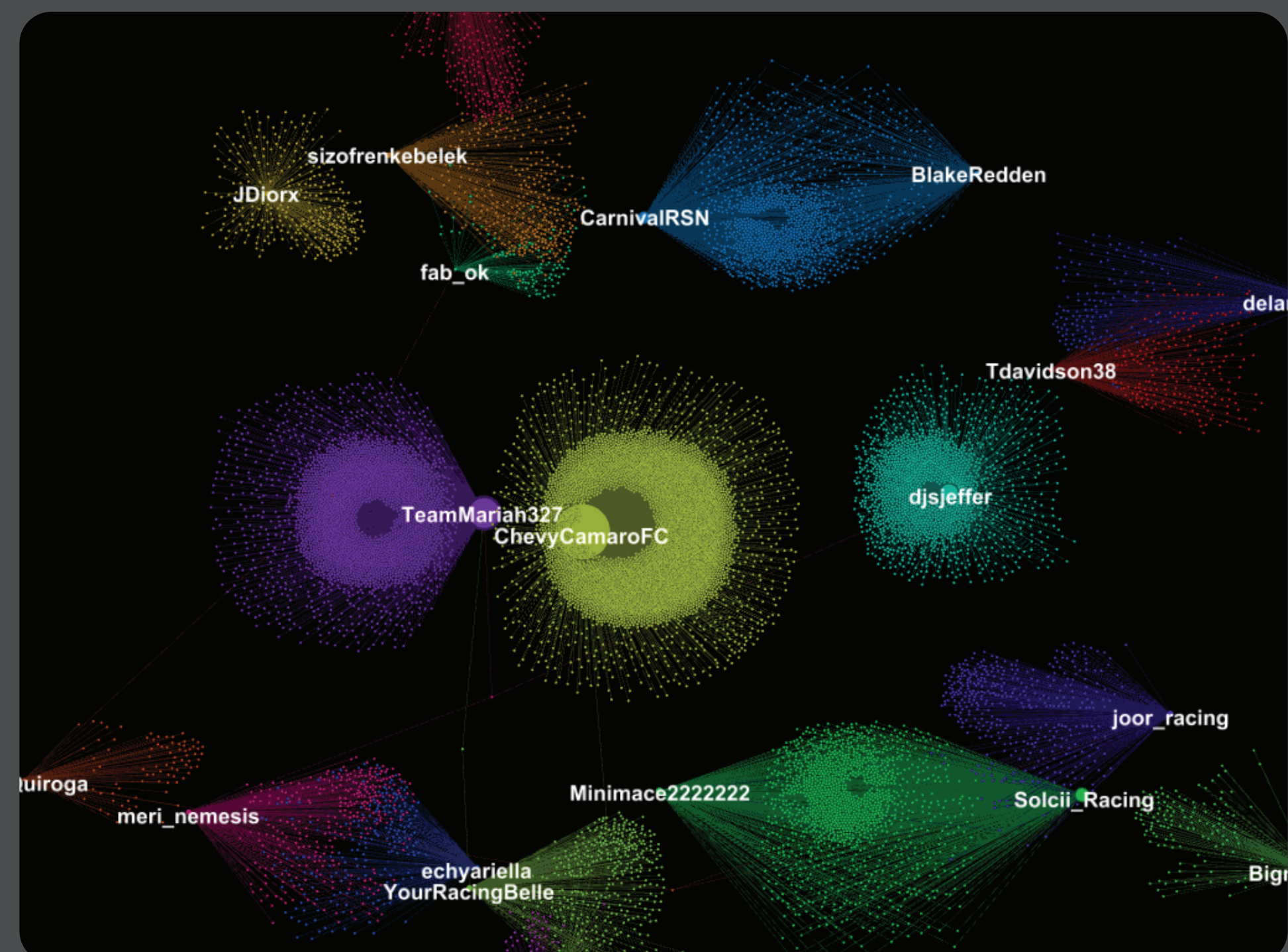
IN THE RELATIONSHIP GRAPH, EACH COLOR REPRESENTS A DISTINCT MODULARITY CLASS. THE TWITTER ACCOUNTS DISPLAYED AS LARGER NODES HAS A HIGHER IMPACT FACTOR THAN OTHERS WITHIN THE COMMUNITY THEY BELONG TO. WHEN WE DIVE INTO THE PROFILES WHICH REPRESENT EACH CLUSTER, WE FOUND TRAITS CAN DIFFERENTIATE THOSE COMMUNITIES, SUCH AS CONTEXT, REGION, LANGUAGE AND SO ON. THE TOP THREE INFLUENCERS ARE:

SCREEN NAME: CHEVY CAMARO (A COMMUNITY OF CAR RACING)  
HTTPS://TWITTER.COM/CHEVYCAMAROFC

SCREEN NAME: DJSJEFFER (A REAL RACING SPORT COMMUNITY)  
HTTPS://TWITTER.COM/DJSJEFFER

SCREEN NAME: SOLCIL\_RACING (A SPANISH SPOKEN COMMUNITY)  
HTTPS://TWITTER.COM/SOLCIL\_RACING

ACCORDING TO OUR RESULTS, TOP TOPIC-RELATED INFLUENCER ARE NOT NECESSARILY THOSE WHO ARE PRESUMED AS THE MOST WELL KNOWN CELEBRITIES ESPECIALLY WHEN NARROWED DOWN TO A SPECIFIC TOPIC.



<sup>1</sup> [HTTP://WWW.IMPACTBND.COM/15-REASONS-STATS-WHY-SOCIAL-MEDIA-MARKETING-IS-ESSENTIAL-IN-2015/](http://www.impactbnd.com/15-reasons-stats-why-social-media-marketing-is-essential-in-2015/)

<sup>2</sup> [HTTP://WWW.NIELSEN.COM/US/EN/NEWSROOM/2012/HOW-CONNECTIVITY-INFLUENCES-GLOBAL-SHOPPING.HTML](http://www.nielsen.com/us/en/newsroom/2012/how-connectivity-influences-global-shopping.html)

<sup>3</sup> [HTTP://PRESS.EXPERIAN.COM/UNITED-STATES/PRESS-RELEASE/EXPERIAN-MARKETING-SERVICES-REVEALS-27-PERCENT-OF-TIME-SPENT-ONLINE-IS-ON-SOCIAL-NETWORKING.ASPX](http://press.experian.com/united-states/press-release/experian-marketing-services-reveals-27-percent-of-time-spent-online-is-on-social-networking.aspx)

