Siva Devarakonda DS 501 May 9, 2015



# **Collecting Data from Twitter**

The topic of interest - Silk Accessories trends, data collected was hashtags related to fashion and silk to launch online marketplace for small business to support SILK ARTISANS:

tweets = searchTwitter('#silk+#fashion', n=1000, resultType='mixed')

The total number of tweets collected - 359

#### Why is the topic important to me

- Trying to help a silk artisan launch an online market place
- · Understand silk fashion trends to justify business proposition and investment
- Understand which avenues to target in the online marketplace

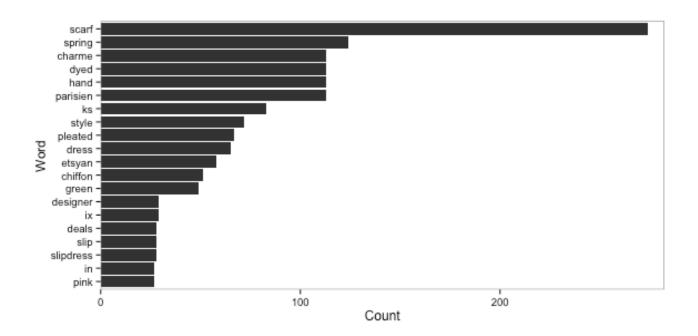
#### Approach to data analysis

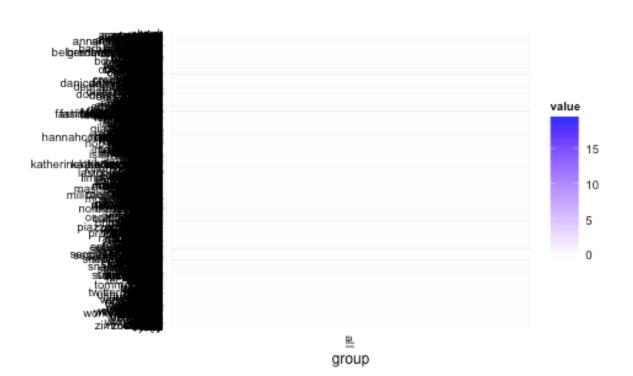
- 1. Preliminary search in twitter on the topic and identified the right #hashtags,
- 2. Utilized the twitter API to pull data, cleaning data took the largest amount of time(removing duplicates and special characters). Requested 1000 tweets but extracted 359

- 3. Used STRINGR/QDAP for data extraction and text mining
- Relied on simplistic visualization methods. Still learning to understand class types and how to visualize them with more advanced visualization packages.

#### **Early Wins and Actionable Insights**

Even simplistic data mining in Twitter indicates that SCARVES are a popular choice AND ETSY is one of the trending online marketplace





- Even though I used simplistic visualization(histograms and tables) with frequent terms, it yielded very useful insights
- · Data indicates clear high frequency patterns

#### **Top Ten Tweets**

value

- RT @GESRAY1: #ArtDeco Bobbie #Jerome Flower Patern #Silk #Clutch Crystal Clasp http://t.co/F4HOsnKYYF #eBay #dealydeal #gift #fashion #mo…
- #ArtDeco Bobbie #Jerome Flower Patern #Silk #Clutch Crystal Clasp http://t.co/ F4HOsnKYYF #eBay #dealydeal #gift #fashion #mothersday

RT @etsyan03: CHARME PARISIEN - #SILK #SCARF

http://t.co/gvlpCnFBsR

Emerald green pleated silk scarf - hand dyed

#spring #fashion #81ks #Tw…

CHARME PARISIEN - #SILK #SCARF

http://t.co/gvlpCnFBsR

Emerald green pleated silk scarf - hand dyed #spring #fashion #81ks #Twitterbrothers

RT @etsyan03: CHARME PARISIEN - #SILK #SCARF

http://t.co/gvlpCno0Bj

Emerald green pleated silk scarf - hand dyed

#spring #fashion #81ks #Twâ€!

CHARME PARISIEN - #SILK #SCARF

http://t.co/gvlpCno0Bj

Emerald green pleated silk scarf - hand dyed #spring #fashion #81ks #Twitterbrothers

RT @etsyan03: CHARME PARISIEN - #SILK #SCARF

http://t.co/U3vKWT5TYa

Gray pleated silk scarf - hand dyed on chiffon

#spring #fashion #sheffi…

CHARME PARISIEN - #SILK #SCARF

http://t.co/U3vKWT5TYa

- 8 Gray pleated silk scarf hand dyed on chiffon #spring #fashion #sheffield #bizitalk
- 9 RT @LeinaalaArt: #design for 'Skathi' #silk #print #clothing on #paom is #handpainted by moi! #fashion #style http://t.co/oaa6Ksw19F http:/…
- RT @muleties: Scarves are coming to town, here's our next line what do you think? #fashion #silk #scarve #style http://t.co/c7UklG1F2f
- Scarves are coming to town, here's our next line what do you think? #fashion #silk #scarve #style http://t.co/c7UklG1F2f

## **Top Ten User HashTags**

WORD	FREQ
1 scarf	141
2 spring	119
3 ks	81
4 style	71
5 dress	32
6 slipdress	28
7 deals	27
8 designer	26
9 design	23
10 etsy	19

## **Top Ten User Mentions**

	WORD	FREQ
14	etsyan	58
43	ukbusinessrt	8
42	ukbusi	7
32	muleties	6
11	deyanadeco	5
12	ebay	4
13	etsy	4
1	aimmea	3
6	beechwoodsc	3
18	gesray	3

# **Popular Personalities**

- Twitter API was super slow for Famous figures!
- I picked MC Mary Kom A huge inspiration for me and a role model for many women
- Friends 71, Followers = 7600, Mutual Friends and Followers = 17

# Table of Followers and Friends(evidently many indian celebrities are in the list)

V1	V2
1846626097	Riya Dey
3244825233	Pìñkú räj
360505512	Birjitkumar Khuman
1160507676	Prashant Kumar
2951005797	MonnA Mo
2397582806	Anjani Pandey
32520953	Kuldeep Kumar
482718719	Sujay Patil
3188979049	Pratik Bagadi
3243134446	Sankha Ghosh
3097711422	poornima rathinavelu
3189335485	Been
2938242128	Rona devi
3189859765	shreedevi s k
296682451	AKSHITA JAIN
3189677299	Jessi Walia
2782777135	Anurag Singh
3165378860	Oamik Eazam
22779056	Pam Sidhu
140057266	kamal medhi

	V1	V2
1	105710210	rishi kapoor
2	2838592766	Priya Shrivastava
3	2734909202	Onler Kom
4	334430754	Mangte C.Mary kom
5	57188977	Sonu Nigam
6	513067381	Satyamev Jayate
7	118024094	Arshad Warsi
8	109241992	Sunidhi Chauhan
9	57928790	Karan Johar
10	15093629	Shekhar Kapur
11	71201743	Virat Kohli
12	87174678	Sania Mirza
13	277434037	Ratan N. Tata
14	58173133	Mahesh Bhupathi
15	113352671	Dia Mirza
16	33801019	Parineeti Chopra
17	19895282	A.R.Rahman
18	230140702	Juhi Chawla
19	92724677	Virender Sehwag
20	471741741	PMO India

## **Mutual Followers and Friends**

	V1	V2
1	105710210	rishi kapoor
2	2838592766	Priya Shrivastava
3	2734909202	Onler Kom
4	334430754	Mangte C.Mary kom
5	57188977	Sonu Nigam
6	513067381	Satyamev Jayate
7	118024094	Arshad Warsi
8	109241992	Sunidhi Chauhan
9	57928790	Karan Johar
10	15093629	Shekhar Kapur

# Other Data Analysis

For the silk scarves business case, I tried to gather geo code information for the tweets but most of them were missing geocoded data. However text network analysis by geocode resulted in this which remains to be interpreted by overlaying with maps information:

