

# *The New York Social Graph*

## PROJECT OVERVIEW:

[New York Social Diary](#) provides a fascinating lens onto New York's socially well-to-do. The data forms a natural **social graph** for New York's social elite.

In this website, the photos have carefully annotated captions labeling those that appear in the photos. We can think of this as implicitly implying a social graph: there is a connection between two individuals if they appear in a picture together.

In this project, I assembled the social graph from photo captions for parties dated from 2007-2015. Using this graph, it is likely to guesses at the most popular socialites, the most influential people, and the most tightly coupled pairs.

## WEB SAMPLE



Tracey Pruzan, Elizabeth Sahlman, and Susan Scherr.



L. to r.: Amy Gold and Helaine Pruzan; Beth Rustin and Amy Rose Silverman.



Emily Sutton, Stephanie Spar, Deborah Raiman, Terri Goldberg, and Ruth Anne Kamins.



L. to r.: Jane Wilf and Rachel Wilf; Sara-Ann Wallerstein, Susan Scherr, and Marilyn Scherr.

## PROJECT OUTLINE:

This project is conducted in three phases:

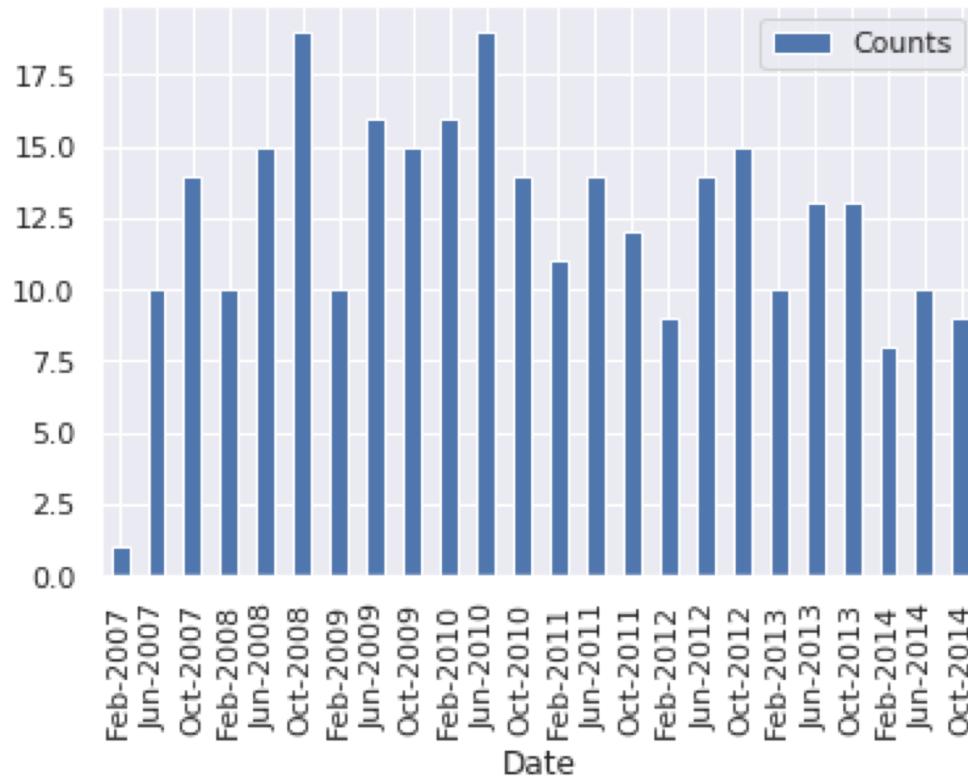
- Get a list of all the photo pages to be analyzed

```
https://web.archive.org/web/20150918040703/http://www.newyorksocialdiary.com/party-pictures/2014/the-fricks-annual-autumn-dinner-and-casita-marias-2014-fiesta  
https://web.archive.org/web/20150918040703/http://www.newyorksocialdiary.com/party-pictures/2014/new-york-and-beyond  
https://web.archive.org/web/20150918040703/http://www.newyorksocialdiary.com/party-pictures/2014/compassion-choices  
https://web.archive.org/web/20150918040703/http://www.newyorksocialdiary.com/party-pictures/2014/public-celebrations  
https://web.archive.org/web/20150918040703/http://www.newyorksocialdiary.com/party-pictures/2014/junior-councils-taking-the-lead
```

- Parse all of the captions on all page.
  - Via regular expression and spaCy
  - Title removal (i.e. "Dr./Mr./Mayor... Michael Bloomberg" to "Michael Bloomberg")
  - Couple recognition (i.e., "John and Mary Smith" to "John Smith" and "Mary Smith")
- Assemble the graph.
  - Via networkx

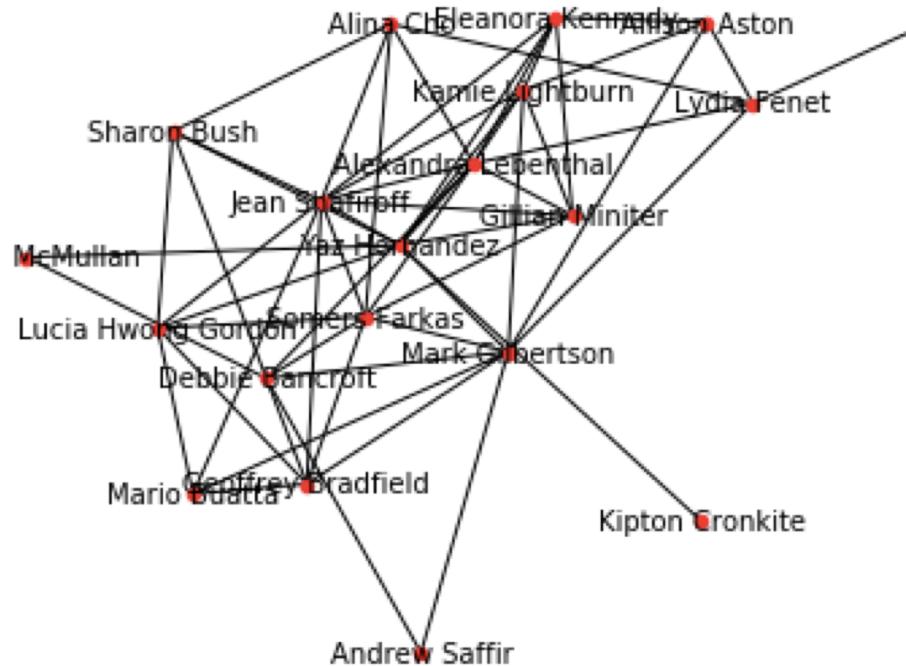
## RESULTS:

- Event frequency
  - More events happened during July ~ October .



## RESULTS:

- Build graph
  - Top 20 people with most connections.



## RESULTS:

- People with higher rank

```
[('Jean Shafiroff', 0.0005729404782509825),  
 ('Mark Gilbertson', 0.0005424701272511396),  
 ('Alexandra Lebenthal', 0.0003678972956520446),  
 ('Gillian Miniter', 0.00034085266753264),  
 ('Geoffrey Bradfield', 0.0003343690932123659),  
 ('Mario Buatta', 0.0003093290179459121),  
 ('Andrew Saffir', 0.00030919785864774897),  
 ('Kamie Lightburn', 0.00029474960937594914),  
 ('Patrick McMullan', 0.000283839960560991),  
 ('Alina Cho', 0.00027305428525433307),  
 ('Somers Farkas', 0.00026775089240886656),  
 ('Debbie Bancroft', 0.0002578625582786809),  
 ('Lydia Fenet', 0.0002473374371149404),  
 ('Sharon Bush', 0.0002471729705342972),  
 ('Yaz Hernandez', 0.00024625708721117155),  
 ('Lucia Hwong Gordon', 0.00024585525821305175),  
 ('Gala Co-', 0.00024159836495792116),  
 ('Allison Aston', 0.00023501704348384228),  
 ('Kipton Cronkite', 0.000235014754474167),  
 ('Eleanora Kennedy', 0.00023428228138834382),  
 ('friends', 0.00023202494329673467),  
 ('Martha Stewart', 0.0002288343722388108),  
 ('Barbara Tober', 0.00022046892131636104),  
 ('Leonard Lauder', 0.00021343140576846688),  
 ('Bettina Zilkha', 0.00021084647951754738),  
 ('Karen LeFrak', 0.00021068261639947501),  
 ('Lisa Anastos', 0.00020560860431604434),  
 ('Jamee Gregory', 0.00020334987836117304),  
 ('Amy Fine Collins', 0.00020086469976952806),  
 ('Michele Gerber Klein', 0.00019989880222665165),
```

- Best friends

```
[('Gillian Miniter', 'Sylvester Miniter'), 75),  
 ('Jamee Gregory', 'Peter Gregory'), 54),  
 ('Geoffrey Bradfield', 'Roric Tobin'), 50),  
 ('Bonnie Comley', 'Stewart Lane'), 49),  
 ('Daniel Benedict', 'Andrew Saffir'), 48),  
 ('Somers Farkas', 'Jonathan Farkas'), 40),  
 ('Alexandra Lebenthal', 'Jay Diamond'), 37),  
 ('Martin Shafiroff', 'Jean Shafiroff'), 36),  
 ('Barbara Tober', 'Donald Tober'), 33),  
 ('Chappy Morris', 'Melissa Morris'), 31),  
 ('Grace Meigher', 'Chris Meigher'), 31),  
 ('Margo Catsimatidis', 'John Catsimatidis'), 28),  
 ('Peter Regna', 'Barbara Regna'), 27),  
 ('Tatiana Platt', 'Campion Platt'), 27),  
 ('Deborah Norville', 'Karl Wellner'), 26),  
 ('Wilbur Ross', 'Hilary Geary Ross'), 26),  
 ('Arie Kopelman', 'Coco Kopelman'), 25),  
 ('Jonathan Tisch', 'Lizzie Tisch'), 25),  
 ('Julia Koch', 'David Koch'), 25),  
 ('Eleanora Kennedy', 'Michael Kennedy'), 24),  
 ('Guy Robinson', 'Elizabeth Stribling'), 24),  
 ('Richard Johnson', 'Sessa von Richthofen'), 24),  
 ('Valentin Hernandez', 'Yaz Hernandez'), 24),  
 ('Frederick Anderson', 'Douglas Hannant'), 21),  
 ('Janna Bullock', 'R. Couri Hay'), 21),  
 ('Mark Badgley', 'James Mischka'), 20),  
 ('Olivia Palermo', 'Johannes Huebl'), 20),  
 ('Jay McInerney', 'Anne Hearst McInerney'), 19),  
 ('Tommy Hilfiger', 'Dee Ocleppo'), 19),  
 ('Clo Cohen', 'Charles Cohen'), 18),  
 ('Marc Rosen', 'Arlene Dahl'), 18),
```