Social Networks

Lenddo, a Singaporean start-up, helps financial institutions collect users’ social network data. But why?

MARKETING SCIENCE
Articles in Advance, pp. 1–25
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0949

Credit Scoring with Social Network Data

Yanhao Wei
Department of Economics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, yanhao@sas.upenn.edu

Pinar Yildirim, Christophe Van den Bulte
Marketing Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104
{pyild@wharton.upenn.edu, vdbulte@wharton.upenn.edu}

Chrysanthos Dellarocas
Information Systems Department, Questrom School of Business, Boston University, Boston, Massachusetts 02215, dell@bu.edu

3

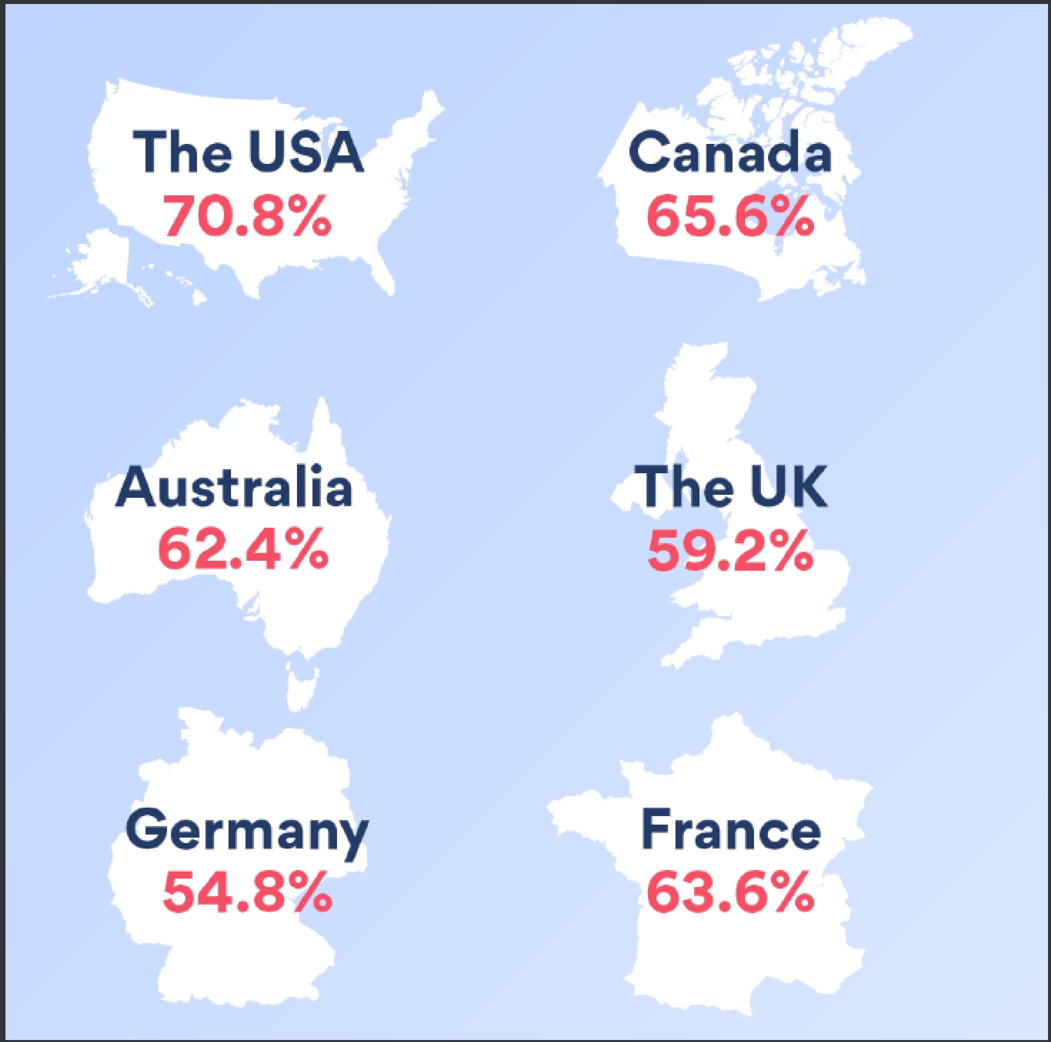
Obesity is an epidemic.

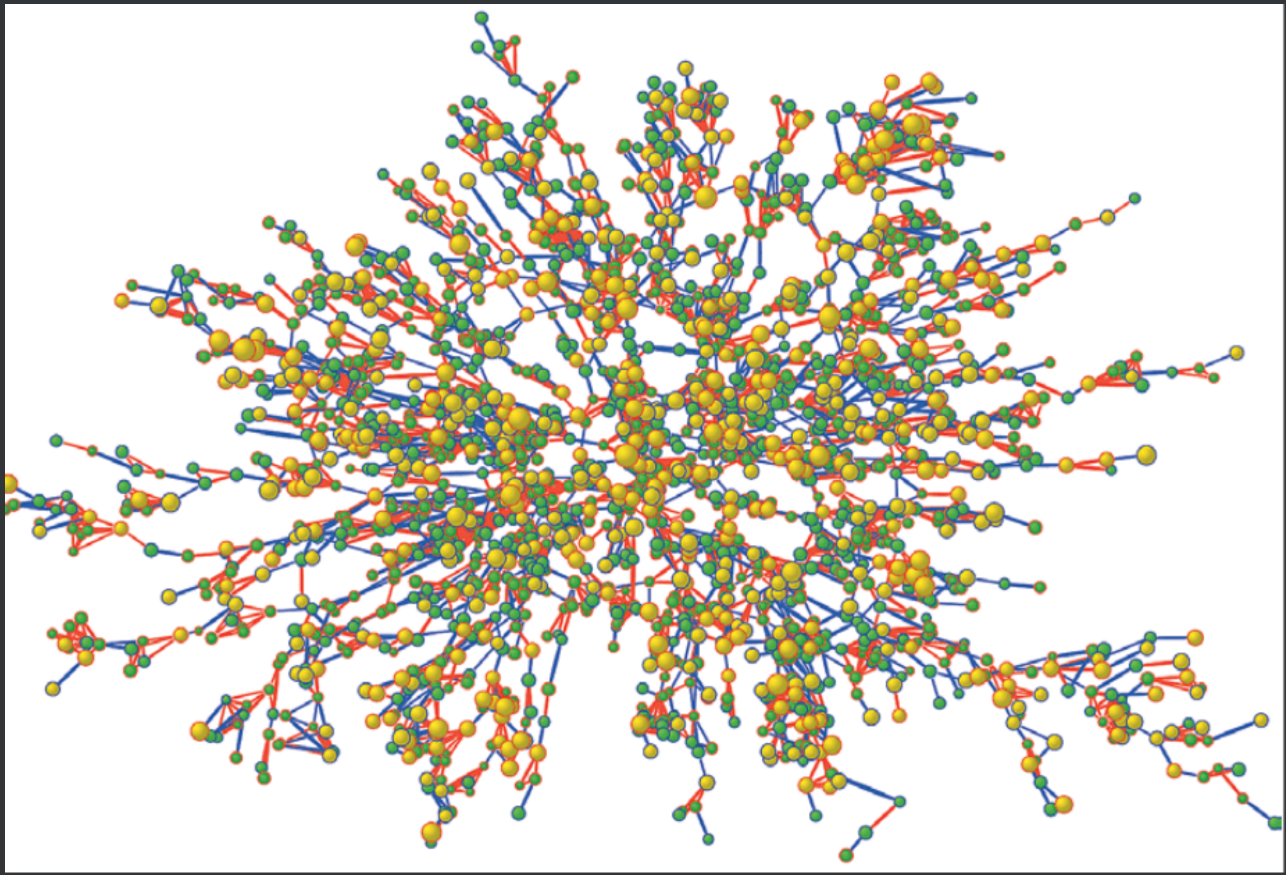
The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

The Spread of Obesity in a Large Social Network over 32 Years

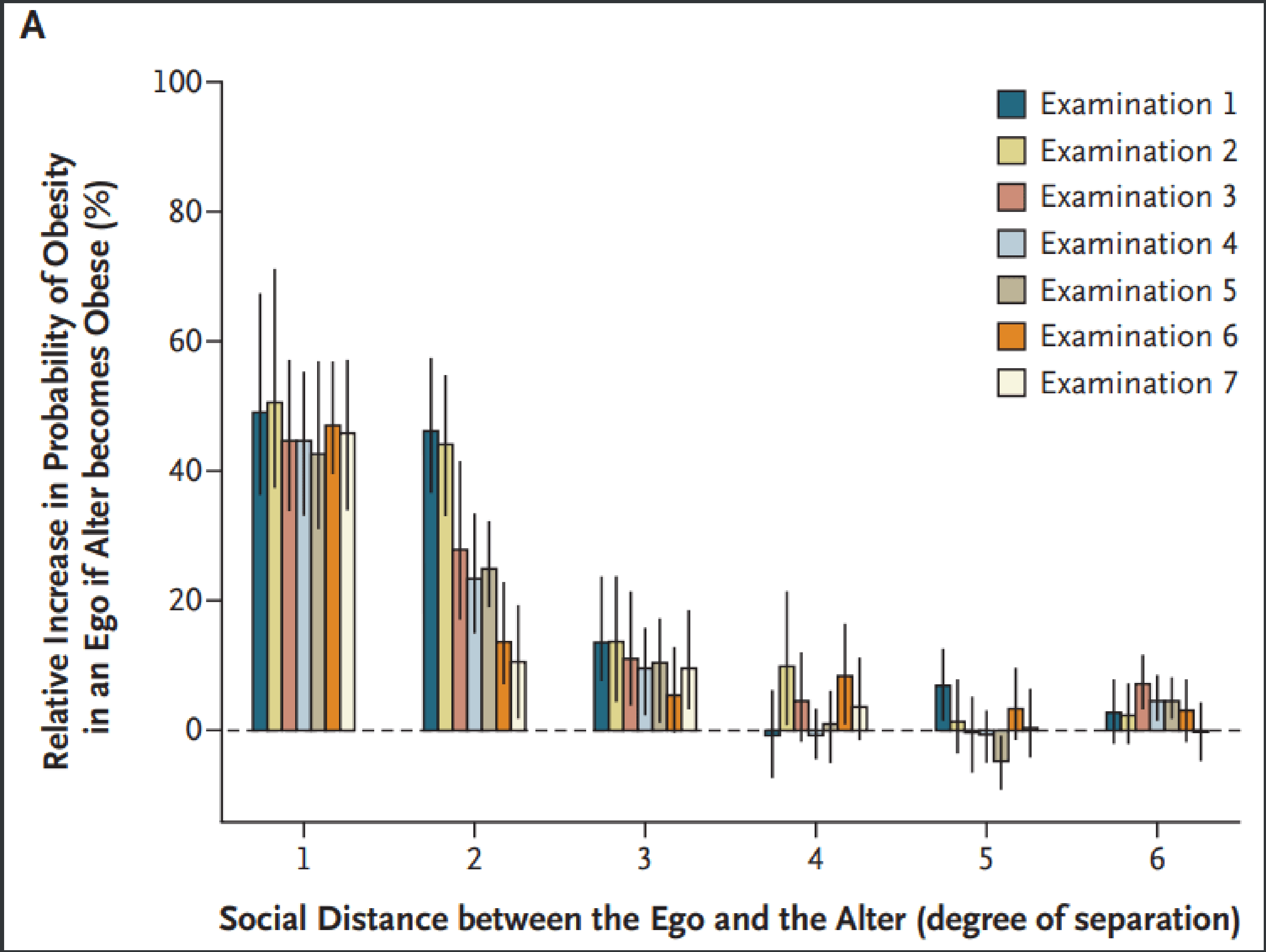
Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.


The USA
70.8%
Canada
65.6%
Australia
62.4%
The UK
59.2%
Germany
54.8%
France
63.6%



Node: individual; edge: connections; size of node: body mass index; yellow: obesity (i.e., $BMI > 30$.)

https://www.youtube.com/embed/pJfq-o5nZQ4?enablejsapi=1


A
Relative Increase in Probability of Obesity in an Ego if Alter becomes Obese (%)
100
80
60
40
20
0
Examination 1
Examination 2
Examination 3
Examination 4
Examination 5
Examination 6
Examination 7
1
2
3
4
5
6
Social Distance between the Ego and the Alter (degree of separation)

45%, 25%, and 10%

But why?

#1: Induction

"Hey, let's go and have muffins and beer!"

"Comparing with my friends, my weight sounds good."

CBS NEWS

OPINION

Gaining Weight? Blame Your Friends

The New York Times Magazine

Are Your Friends Making You Fat?

#2: Homophily

I make friends with you because we share the same body size.

#3: Confounding

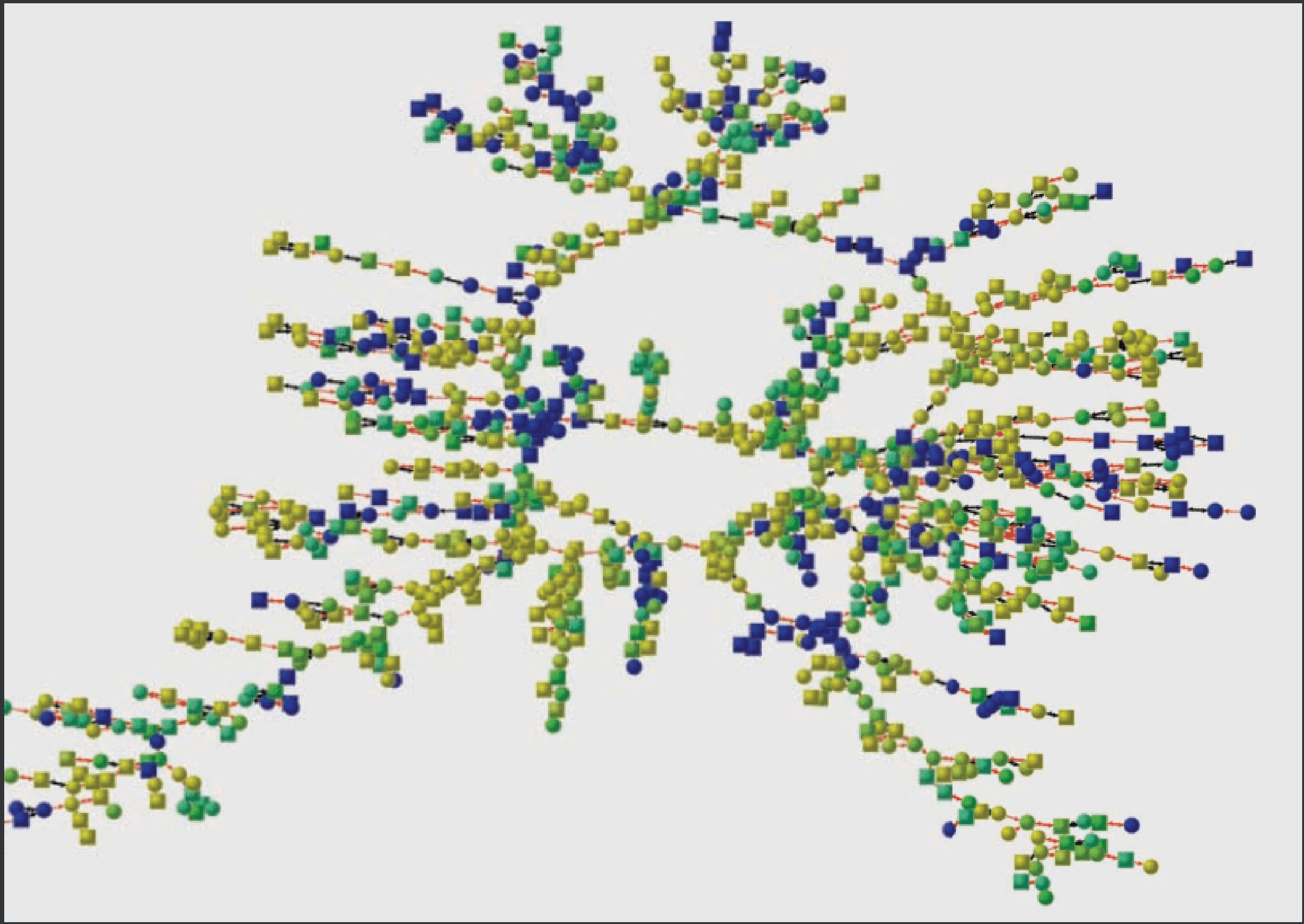
We share a common exposure to something, e.g., we are both visiting the same gym.

BMJ

RESEARCH

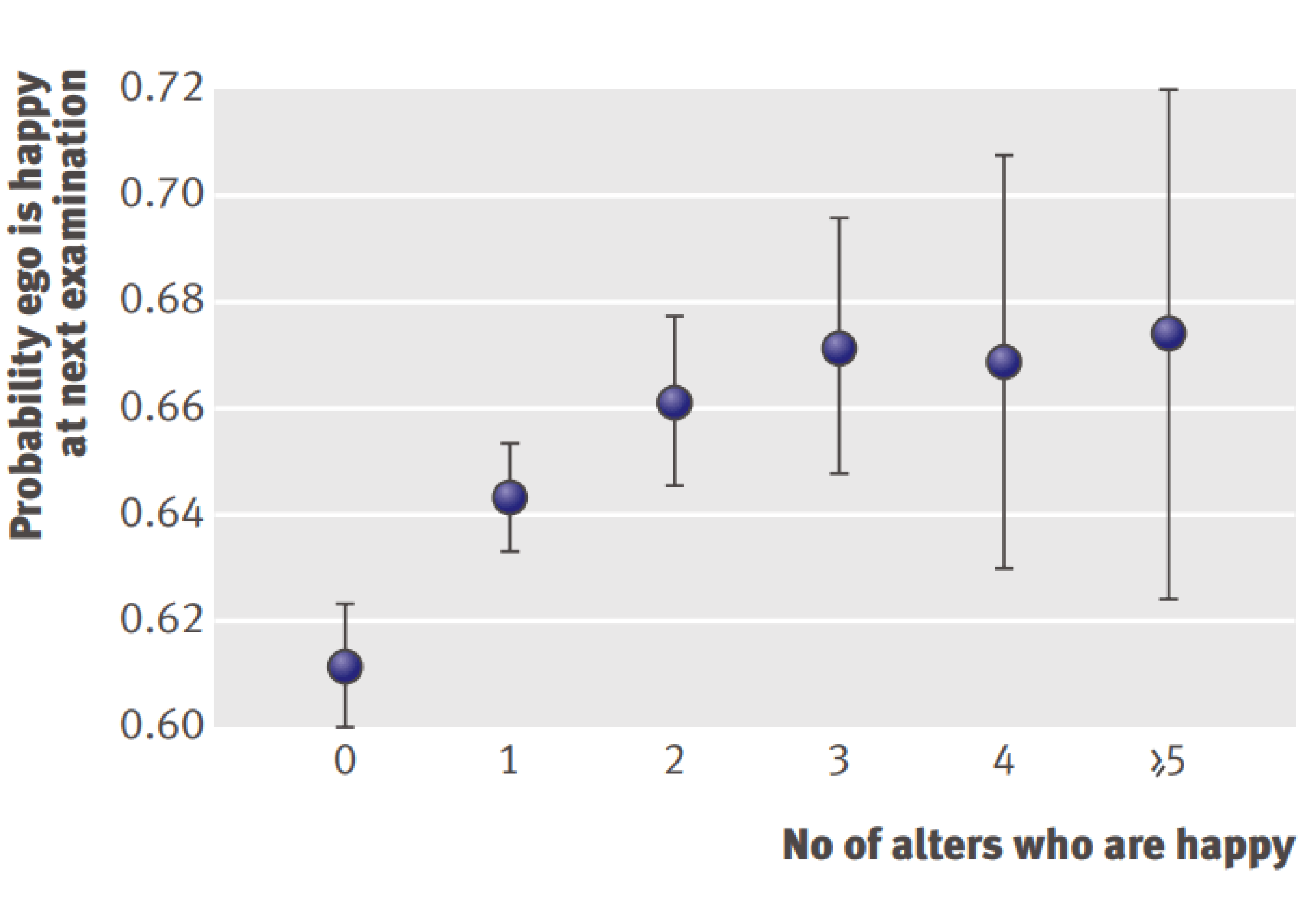
Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study

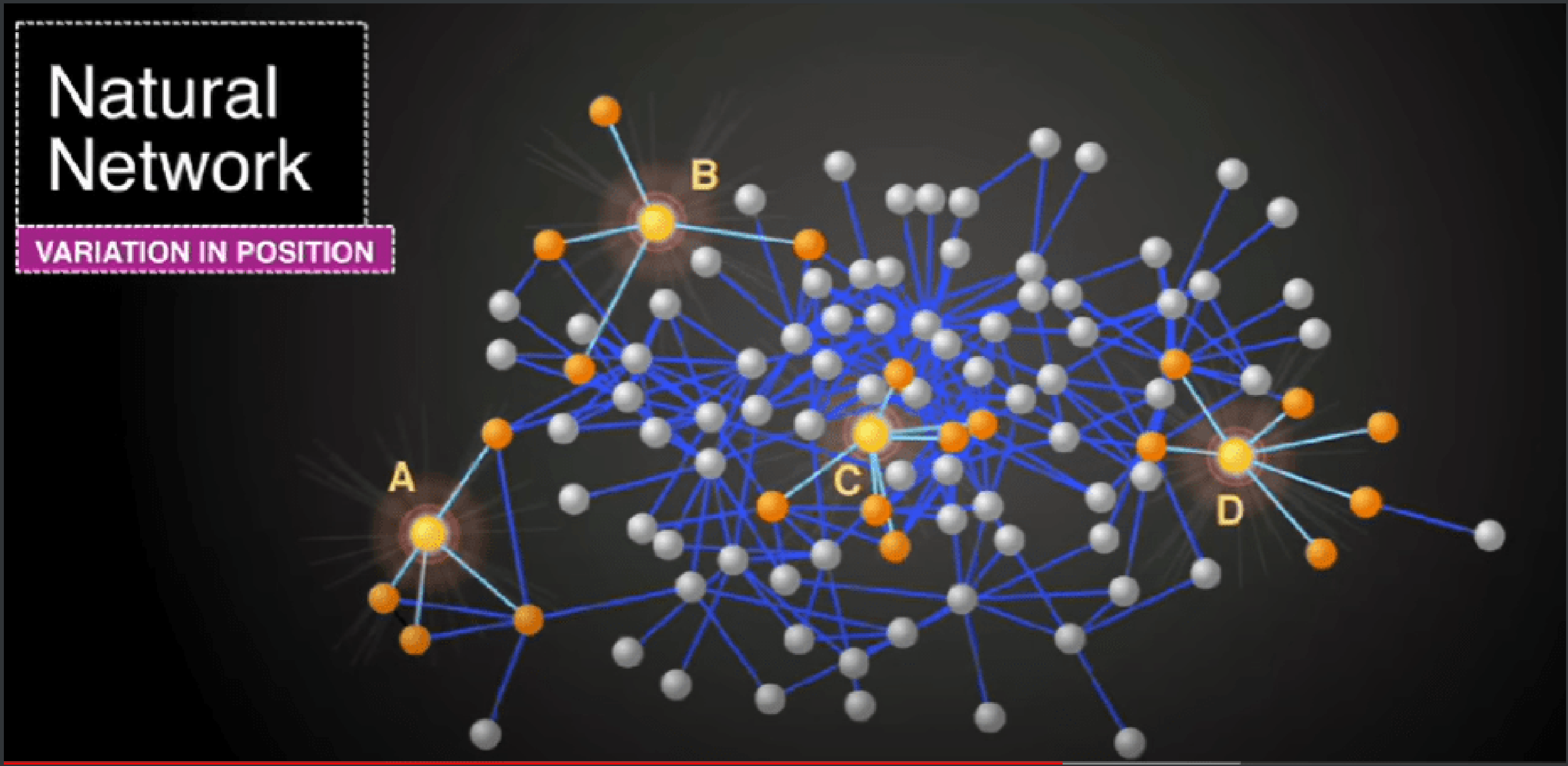
James H Fowler, associate professor,[1] Nicholas A Christakis, professor[2]

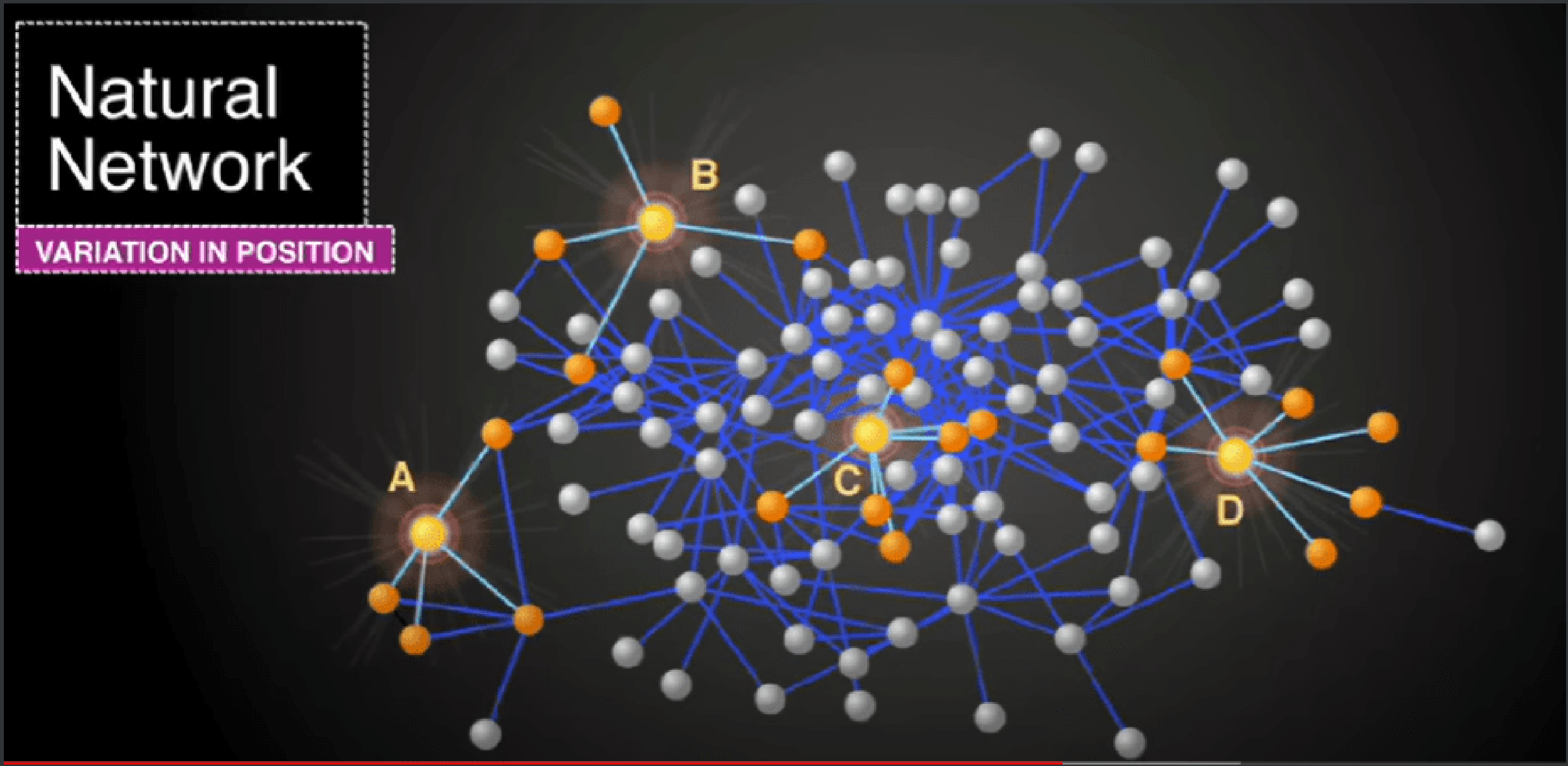


Happiness is contagious:

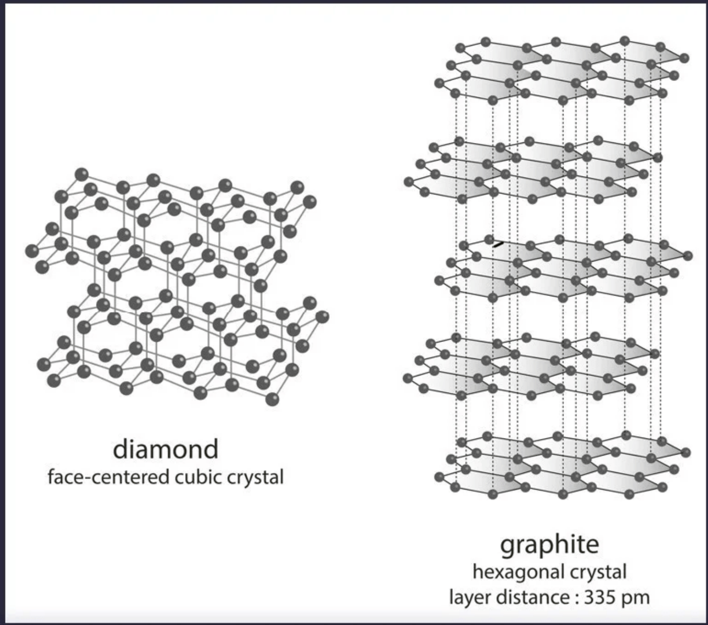
(square: male; circle: female; yellow: happy; blue: unhappy)


Probability ego is happy at next examination
0.72
0.70
0.68
0.66
0.64
0.62
0.60
0
1
2
3
4
≥5
No of alters who are happy


Natural
Network
VARIATION IN POSITION
A
B
C
D



If a deadly germ is going to spread in this social network, would you rather be person C or person D?



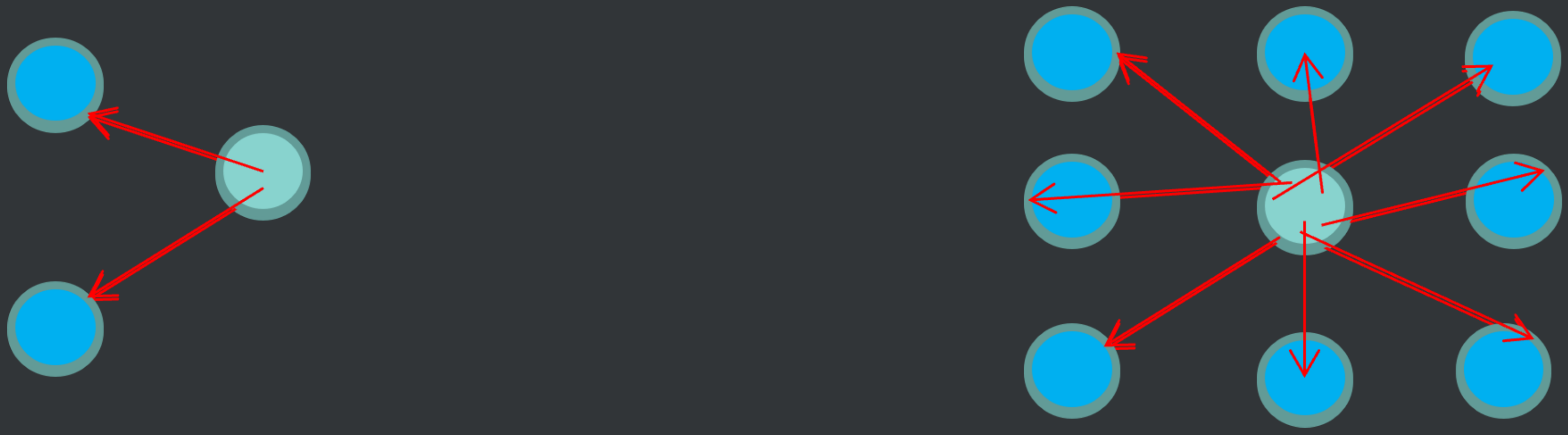
Network structure makes the difference.

What's the difference?

Web 1.0	Web 2.0 and beyond
Expedia	Twitter
Google	Snapchat
eBay	Instagram
Amazon.com	Pinterest
CNN.com	Reddit
WSJ.com	Wikipedia
	Facebook

Amplification Ratio

$$\text{amplification ratio} = \frac{\text{friends of fans exposed to}}{\text{fans exposed to}} = \frac{10}{2} = 5$$



Social Network Analysis: Theory

Key Metrics of a Social Network

Individual: Has meaning independently of social network
You live in Hong Kong island, HK

Connection: You are close friends with 10 people at HKU

Whole Network: On average, students know each other within 4 steps

Connection can be directed (e.g., follower and followee) or undirected (e.g., classmates)

Nodes and Edges

Vertex/Node: an end point, often a person

Edge/Link: What connects up the nodes, e.g., a relationship

Maximum number of edges in group of size $N(N-1)/2$.

- Where everyone connects to everyone else
- If undirected (my friends also have me as a friend)

Who is well-connected?

Degree (centrality): The number of linkages you have.

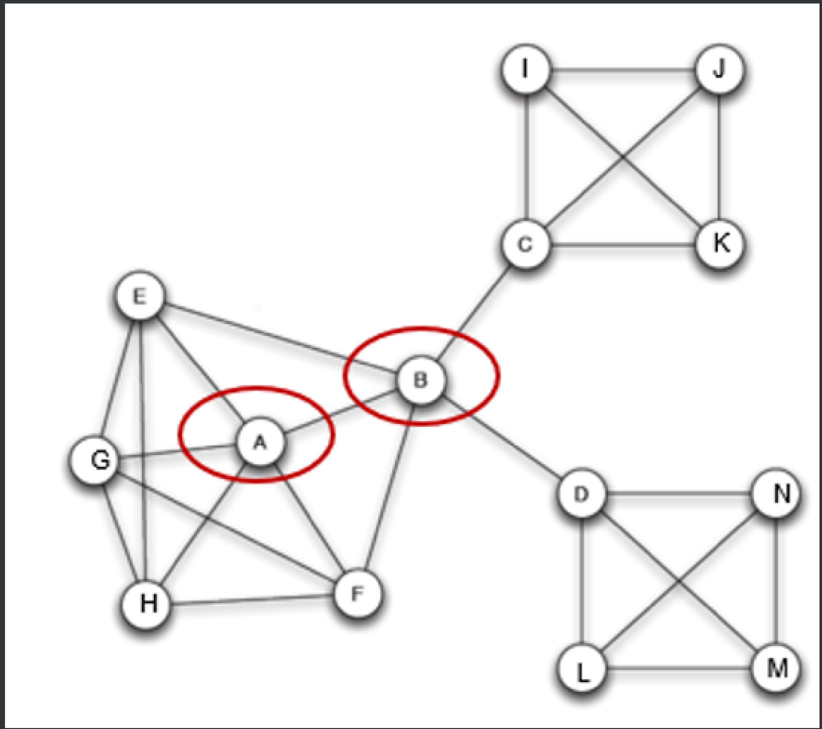
- "In-degree", e.g., someone that follows me.
- "Out-degree", e.g., I follow someone else.

Edge Weight

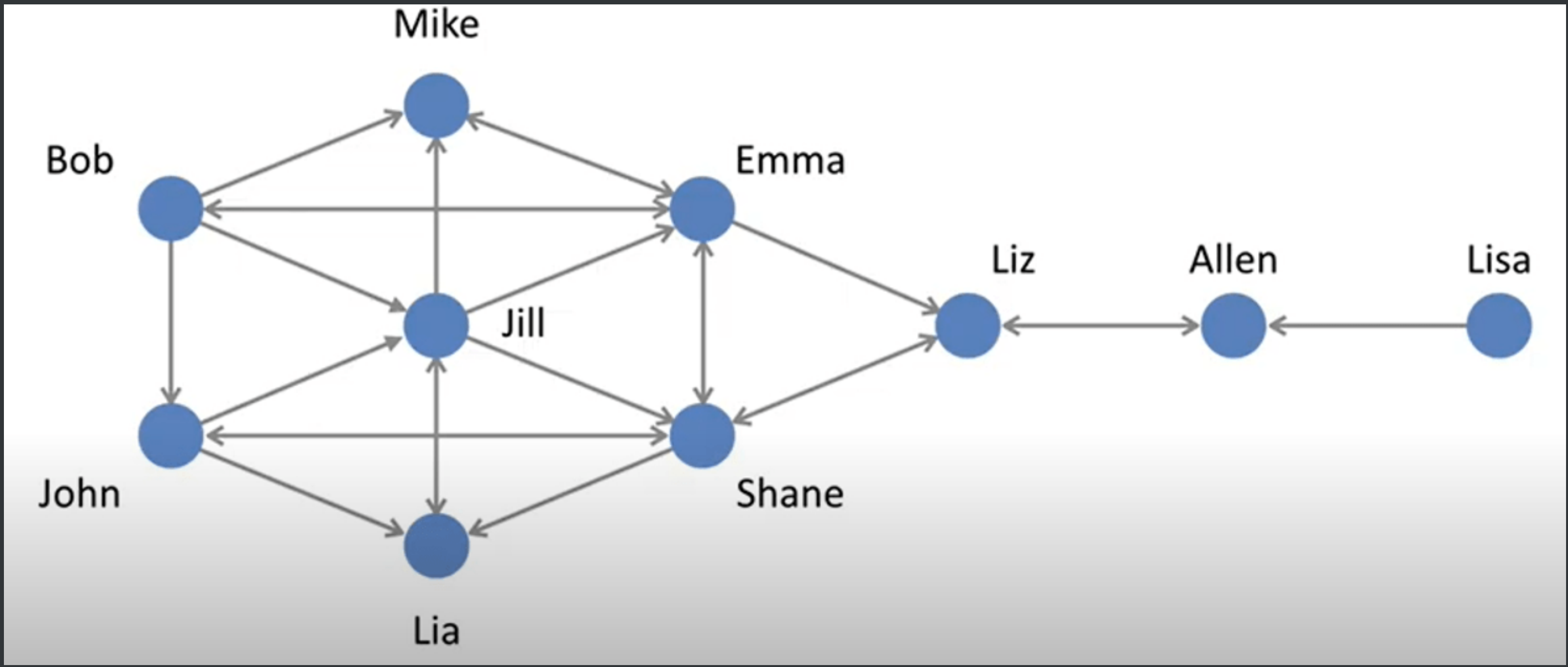
- Sometimes edge can also carry weight
- Can capture how deep the relationships are
- E.g., frequency of interactions between two nodes.

How to determine important persons in a social network?

Who is more important? Why?



Who is more important? Why?



https://www.youtube.com/embed/0aqvVbTyEmc?enablejsapi=1

Closeness Centrality

Only applies to a fully connected network (i.e., a path exists between any pair of nodes).

$$\text{closeness centrality}(x) = \frac{N-1}{\sum_y d(x,y)}$$

N: number of nodes in the network

$d(x,y)$: the shortest distance between nodes x and y.

Betweenness Centrality

Applies to disconnected networks as well.

$$\text{betweenness centrality}(x) = \sum_{y,z} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

σ_{yz} is the total number of shortest paths from y to z.
$\sigma_{yz}(x)$ is the number of shortest paths from y to z that go through x.

Strong ties vs. Weak Ties

Strong Ties vs. Weak Ties

A, B and C are currently iPhone users.

C has recently switched to Android system, and B still uses iPhone.

A is more likely to switch or stay, follow your friend or acquaintance?

Strength of strong ties.

Strong Ties vs. Weak Ties

A has recently changed job.

Is A more likely getting a lead from friend C or acquaintance B?

See a video here.

Strong Ties vs. Weak Ties

Although strong ties generally exert more normative influence, weak ties often have more informational influence.

Why?

Because different social circles have different info, i.e., you probably know what your good friends know. Most jobs are found through weak connections.

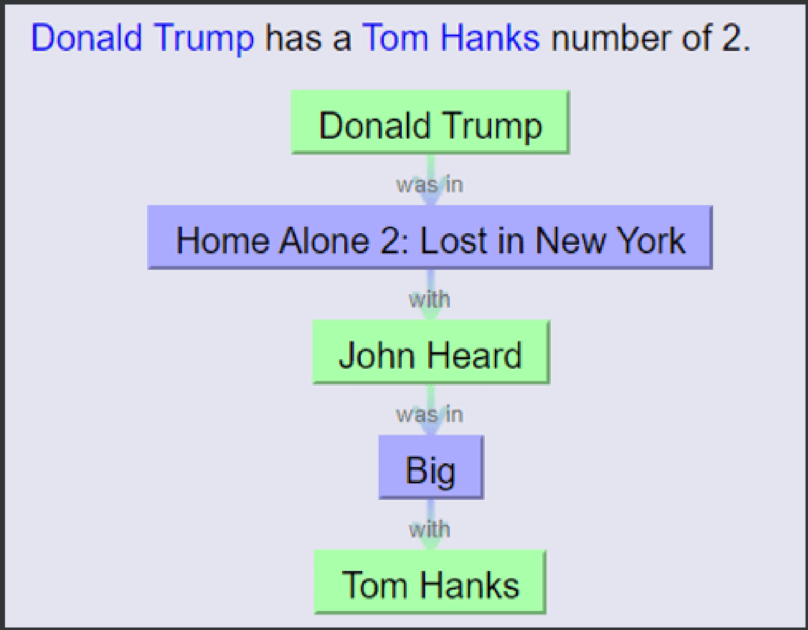
Degrees of Separation

Path of how many people are needed to connect people up

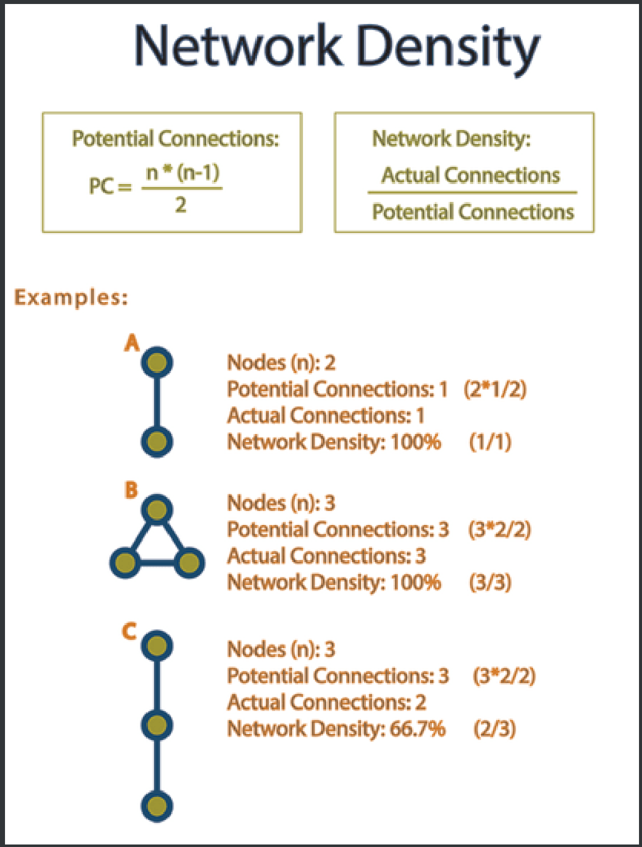
Technical name: Geodesic distance

6 is the magical number: Kevin Bacon game (Link)

Don't fixate on 6! It does not apply to all networks!



The Density of a Social Network



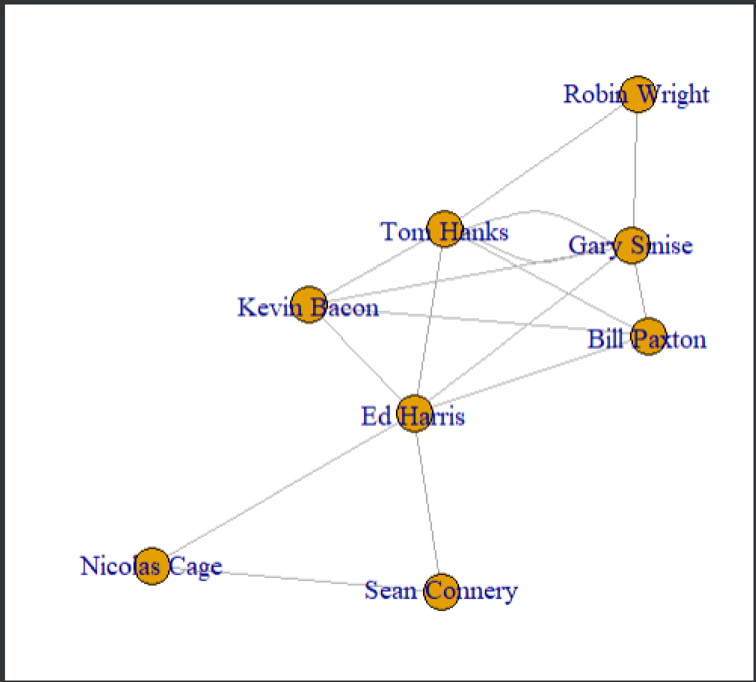
Network Analysis with R

Loading the Network Data

```r
library(igraph)
library(readr)
actors <- read_csv("https://ximarketing.github.io/class/DM//Actors.csv")
movies <- read_csv("https://ximarketing.github.io/class/DM/Movies.csv")
head(actors)
head(movies)
```

Constructing the Network

```r
actorNetwork <- graph_from_data_frame(d=movies, vertices=actors, directed=F)
plot(actorNetwork)
```



Coloring Your Network

```r
V(actorNetwork)$color <- ifelse(V(actorNetwork)$Gender == "Male", "lightblue", "pink")
plot(actorNetwork)
legend("topleft", c("Male","Female"), pch=21,
        col="#777777", pt.bg=c("lightblue","pink"), pt.cex=2, cex=.8)
```

Degree Centrality

```r
degree(actorNetwork, mode="all")
```

Closeness Centrality

```
closeness(actorNetwork, mode="all", weights=NA, normalized=T)
```

Betweenness Centrality

```
betweenness(actorNetwork, directed=F, weights=NA, normalized = T)
```

Density of Network

```
edge_density(actorNetwork)
```

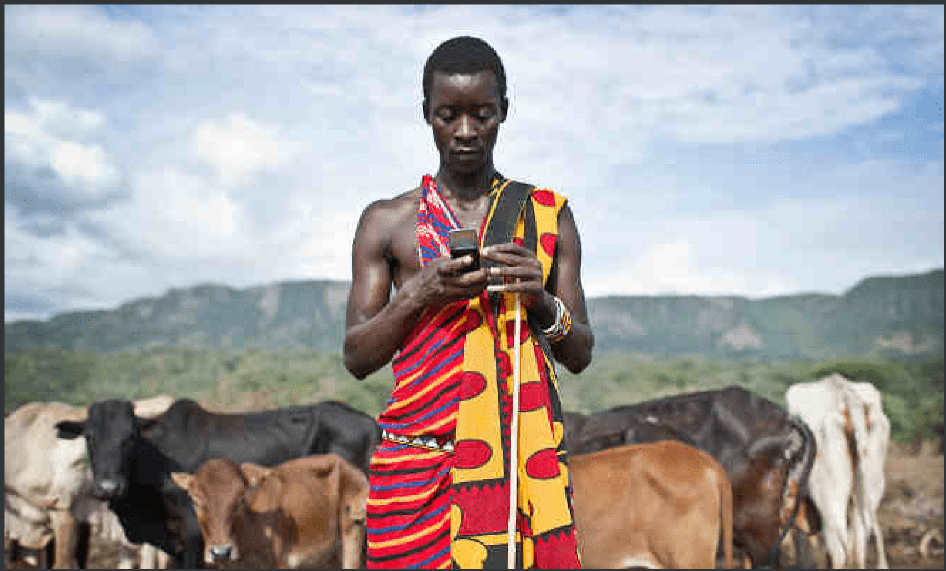
Exercise

```r
actors <- read_csv("https://ximarketing.github.io/class/DM//ActorsExercise.csv")
movies <- read_csv("https://ximarketing.github.io/class/DM/MoviesExercise.csv")
```

Exercise

```r
cities <- read_csv("https://ximarketing.github.io/class/DM/DirectedNodes.csv")
routes <- read_csv("https://ximarketing.github.io/class/DM/DirectedEdges.csv")
flightNetwork <- graph_from_data_frame(d=routes, vertices=cities, directed=T)
plot(flightNetwork)
degree(flightNetwork, mode="in")
degree(flightNetwork, mode="out")
```

Mobile

4.5 billion vs. 6.1 billion

Excluding your sleep, what is the percentage of time that you spend on screens?

The Average Screen Time

(% OF AWAKE HOURS) BY COUNTRY

Are you a slave to your screen? According to our research, people spend an average of 6 hours and 37 minutes in front of screens. The question is, how does each country differ?

The country with the **highest average** screen time is **South Africa**, spending more than half the day on screens (**58.21%**). Surprisingly, even as one of the first countries to harness the power of technology and explore its benefits, according to News On Japan, **Japan** reportedly spends the **least** time scrolling (**21.70%**).

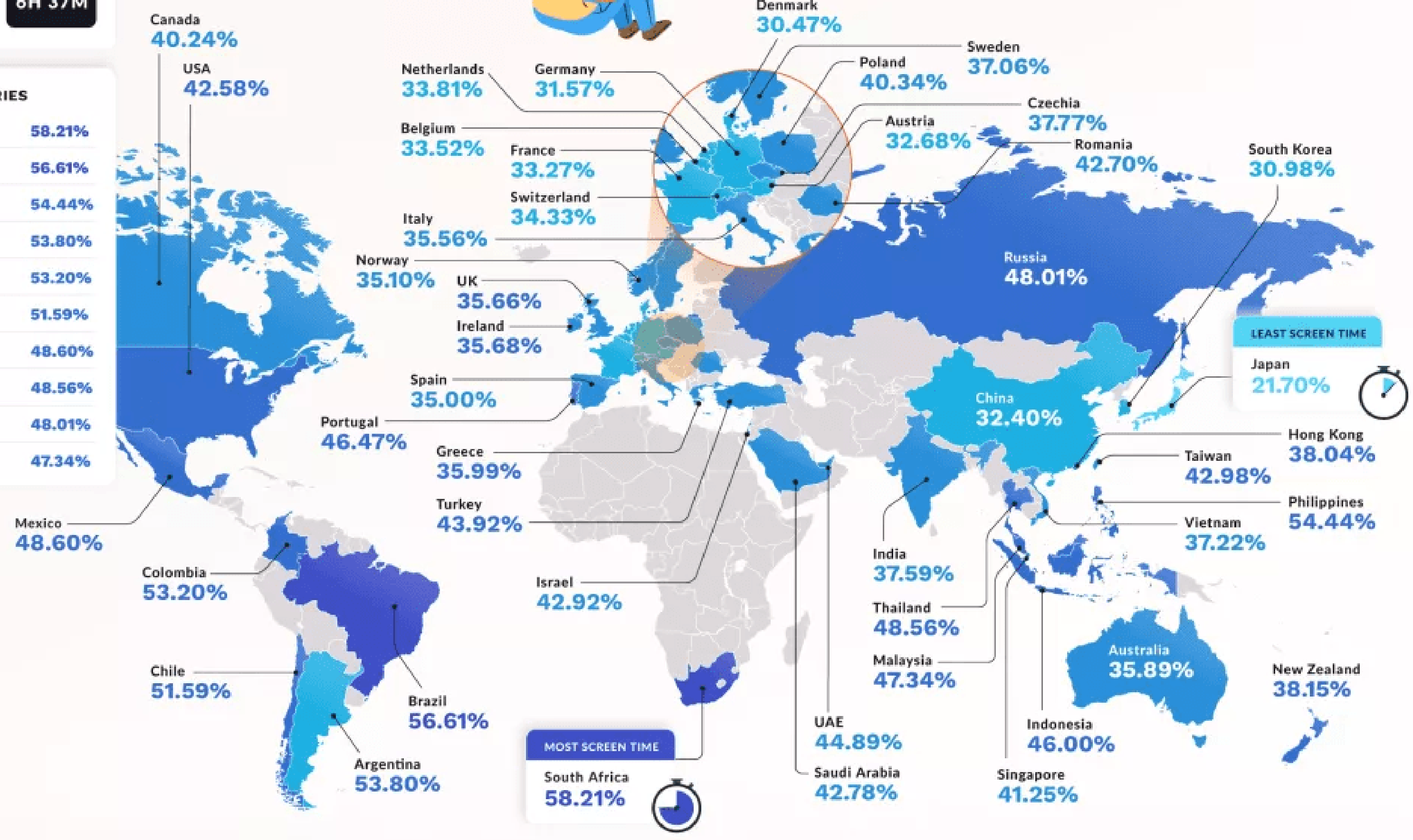
DAILY AVG SCREEN TIME (% OF AWAKE HOURS)

21.70% – 58.21% | No Data

WORLDWIDE AVG SCREEN TIME 6H 37M

TOP 10 COUNTRIES

1	South Africa	58.21%
2	Brazil	56.61%
3	Philippines	54.44%
4	Argentina	53.80%
5	Colombia	53.20%
6	Chile	51.59%
7	Mexico	48.60%
8	Thailand	48.56%
9	Russia	48.01%
10	Malaysia	47.34%



Methodology: We analyzed data from the Digital 2023: Global Overview Report by Datareportal and combined it with sleep patterns data from SleepCycle.com to calculate the % of awake hours each country (internet users aged 16-64) spends looking at screens for each category.

Electronics Hub

How is mobile different from PC? What new marketing opportunities are brought by mobile?

- Omnipresence: Always carried and always on.
- Reduced targeting errors: Unlike cookies, phone number and device ID cannot be deleted; mobile phones are usually not shared among households.
- The story of pies.

In United States, according to supermarket sales, among all 30-centimeter pies, apple pies are most popular.

However, among 11-centimeter pies, apple pies only rank the 5th. What makes the difference?

- Built-in payment system: Easily purchase at offline stores
- Location awareness: Location provides both proximity data and contextual information.

Location Based Targeting

Consumers search with their location and proximity in mind

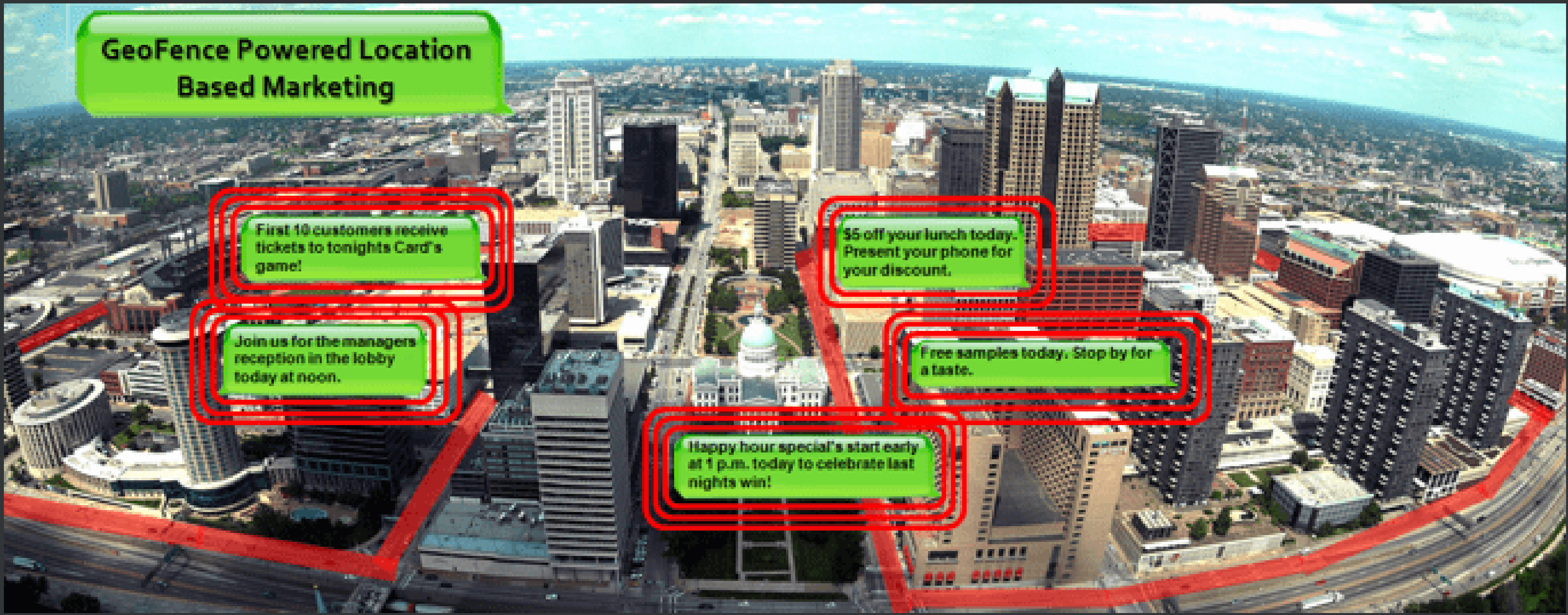
- 88% of consumers conduct local searches on smartphones.

Local searchers are more likely to take actions

- 50% of consumers who conducted a local search on their smartphone visited a store within a day.
- 18% of local searches on smartphone lead to a purchase within a day vs. 7% of non-local searches.

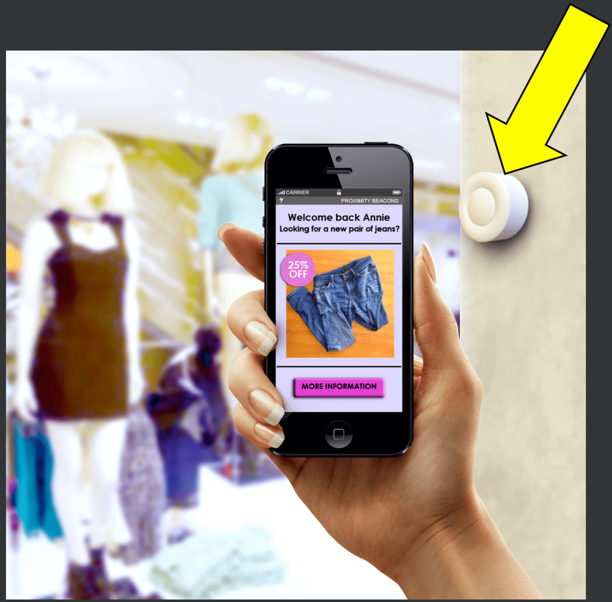
Geo-fencing

Geofencing is a location-based service that sends promotional messages to smartphone users who enter a defined geographic area such as a hotel, a mall, or a conference center.



Beacons

Beacons are small, often inexpensive devices that use Bluetooth to enable more accurate location within a narrow range than GPS, cell tower triangulation and Wi-Fi proximity.




Traditional Location Targeting
Works: Unknown
Lives: Unknown
Shops: Unknown
Age: Unknown
Income: Unknown
Interests: Unknown
Mobile Targeting
Works: Iowa City, IA
Lives: Iowa City, IA
Shops: McDonald's, Wal-Mart
Age: 25-29
Income: $50-75k+
Interests: Concerts
Mobile Targeting
Works: Midtown Manhattan
Lives: Garden City, NY
Shops at: Costco, Macy's
Age: 35-44
Income: $150k+
Travels for business

https://www.youtube.com/embed/nZ532wkhHYs?enablejsapi=1

- Personalize user experience
- Send mobile coupons
- Have high targetability such as demographics, timing, etc
- Be non-intrusive by giving users opt-out options
- Link with loyalty program