Introduction to the Workshop and Sharing Economy

James Pick and Avijit Sarkar

School of Business

University of Redlands

Redlands, California

SIGGIS Workshop on Locational Big Data and Analytics: Implications for the Sharing Economy

Americas Conference on Information Systems (AMCIS)

Boston MA

August 10, 2017

Spatial Analytics and Big Data

- *Spatial Analytics* refers to the contemporary concept of using specialized spatial analysis techniques to understand spatial arrangements, patterns, groupings and relationships in geographically referenced phenomena.
- These methods include overlays, buffers, hot spot analysis, spatial cluster analysis, spatial autocorrelation, proximity polygons, spatial econometrics and other techniques.
- *Spatial Big Data*. GIS data tend to be big, varied, and high-velocity, and that was true well before the buzzword "Analytics" became prominent.
  - Spatial data are large because mapping requires numerous points to be meaningful. A point needs to be identified with X,Y, and often Z coordinates and has additional attribute information associated with it. All this adds to the size of spatial data.

Definition of Spatial Big Data

- Big Data are "data sets that are so big they cannot be handled efficiently by common database management systems" (Dasgupta, 2013).
- Spatial Big Data represents Big Data in the form of spatial layers and attributes.
  - There is no standard threshold on minimum size of Big Data or Spatial Big Data, although big data in 2013 was considered one petabyte (1,000 terabytes) or larger (Dasgupta, 2013).

- Big Data are getting unbelievably large
  - More video is captured daily today than happened in the initial 50 years of television
  - Amount of data available today. More than 2.8 zettabytes (2.8 trillion gigabytes).

Example of a Spatial Big Data Problem

- Consider if you had the data on all the graduate students studying in the U.S.
- 1.7 million according to U.S. Dept. of Education.

- You are analyzing their recording with real-time updating, as the data change from day to day. You have 2 years of the data updated on a daily basis. For each graduate student you have location (lat.-long.), 25 attributes, photos, free-form audio recordings about the student's background and readiness, and a sample video in which the student discusses his/her graduate study goals.
- The analyst wishes to study trends in graduate student goals and interests could narrow the data down and do the necessary analytics to gain value. Keep in mind that the data are in varied formats (numbers, addresses (x-y), text, data-base, video, audio).

These types of problems are ones that today are able to be addressed with combinations of new and traditional analytics tools. Several example of contemporary tools in use for spatial analytics and big data include Tableau software, Esri's ArcGIS Insights and ArcPro, Stata's spmap, etc.

Sources of Spatial Big Data

- *Sources of Spatial Big Data include:*
  - GPS, including
  - GPS-enabled devices
  - Satellite remote sensing
  - Aerial surveying
  - Radar
  - Lidar
  - Sensor networks
  - Digital cameras
  - Location of readings of RFID

(Partially based on Dasgupta, 2013)

Five V's of Spatial Big Data

- Volume
  - Satellite imagery covers the globe so is vast.
  - Sensors are expanding worldwide at a rapid rate.
  - Digital cameras have reached several billion through spatially-reference cell phones.
  - One estimate indicates that 2.5 quintillion bytes are generated daily worldwide. (www.ibm.com). 2.5 with 18 zeros.

- Variety
  - The form of data is based on 2-D or 3-D points configured as vector or raster imagery. This is entirely different than conventional big data which is alphanumeric or pixel-based (similar to raster but not vector)
- Velocity
  - Velocity is very fast since imagery travels at speed of light.

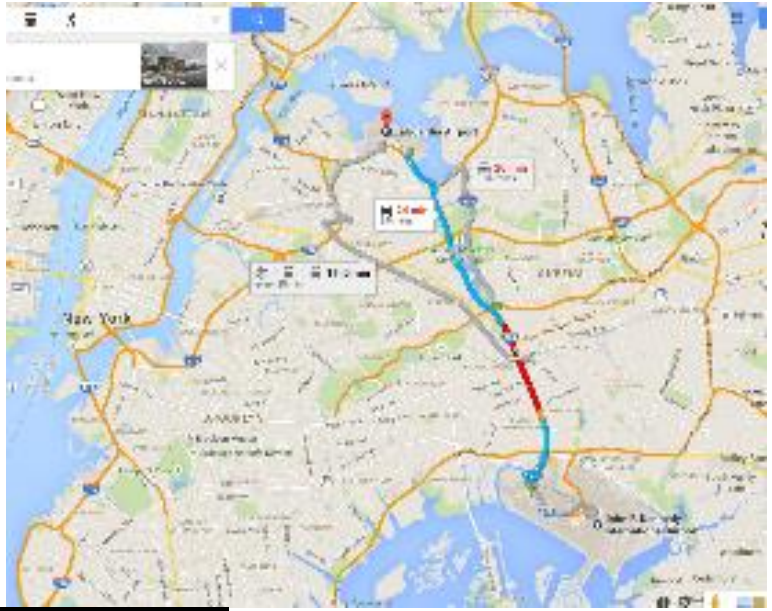
Five V’s of Spatial Big Data (cont.)

- Veracity
  - For vector data (points, lines, and polygons), the quality varies. It depends on whether the points have been GPS determined, or determined by unknown origins or manually. Also, resolution and projection issues can alter veracity.
  - For geocoded points, there may be errors in the address tables and in the point location algorithms associated with addresses
  - For raster data, veracity depends on accuracy of recording instruments in satellites or aerial devices, and on timeliness.

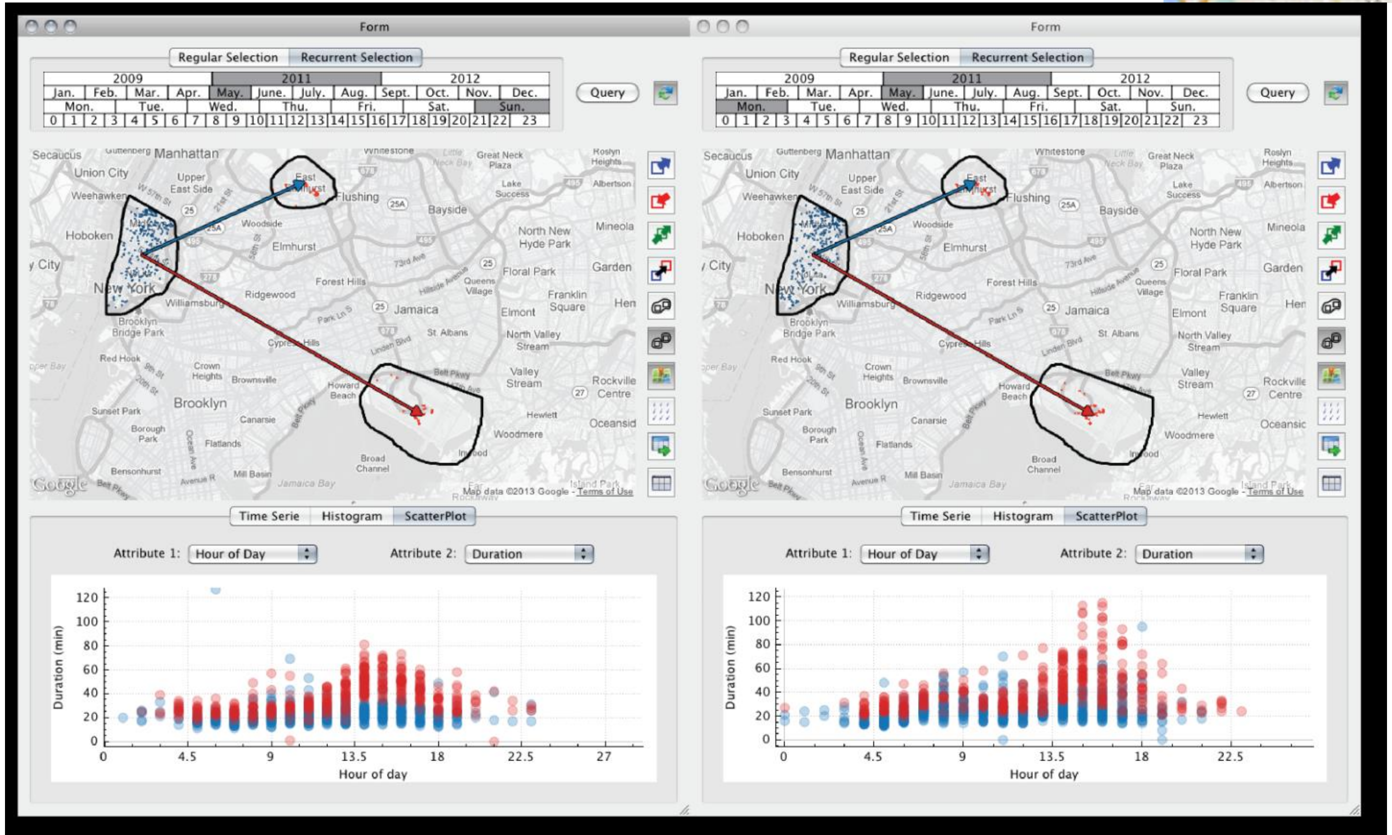
Five V’s of Spatial Big Data

- Value
  - For real-time spatial big data, decisions can be enhance through visualization of dynamic change in such spatial phenomena as climate, traffic, social-media-based attitudes, and massive inventory locations.
  - Exploration of data trends can include spatial proximities and relationships.
  - Once spatial big data are structured, formal spatial analytics can be applied, such as spatial autocorrelation, overlays, buffering, spatial cluster techniques, and location quotients.

Spatial Big Data – Example of Locations and Movement of Central New York City Taxicabs, based on space, time, and attributes



A user-friendly interface TaxiVis allows users to view and analyze the patterns and movements of 500,000 taxi trips daily in central NYC. The data from NY Taxi and Limousine Commission gives pickup and drop off locations, time, and attributes.
Commercial map rendering is done using Google Maps, Bing Maps and OpenStreet Map. Simple or complex queries can be done. Balance between simplicity and expressiveness.

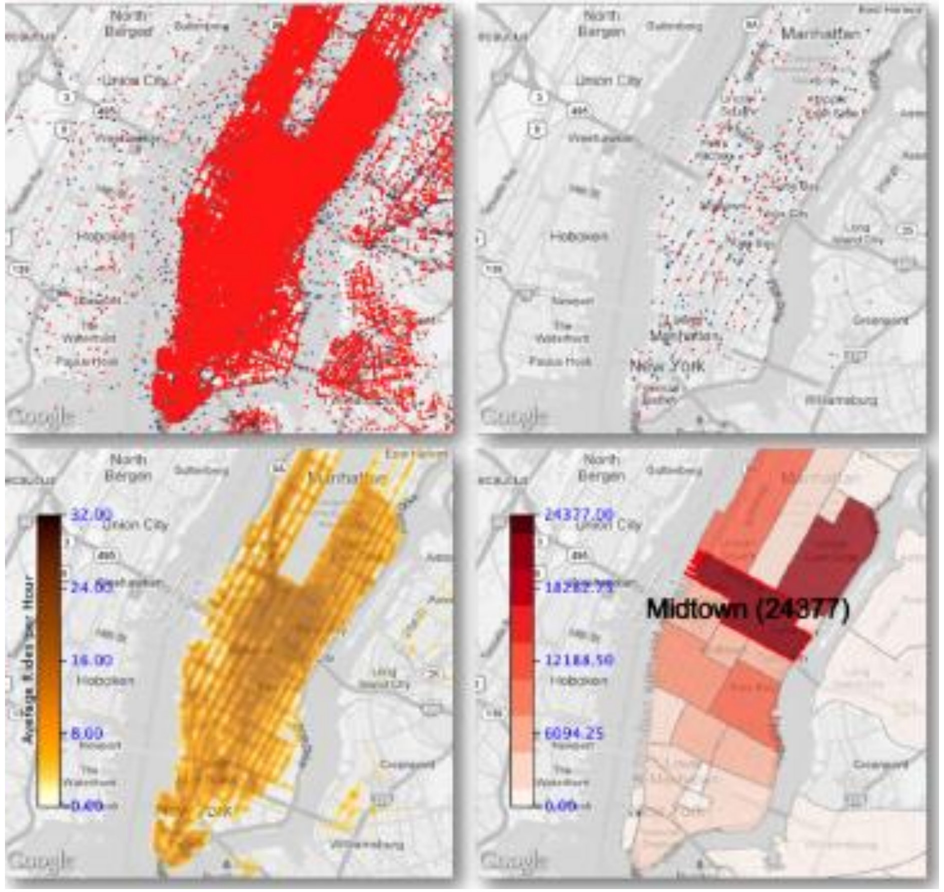


The example shows taxi trips from lower Manhattan area to LaGuardia airport area (upper part of image) and Kennedy airport area (lower part). The volume of trips are given in the lower hourly graphs for Sundays in May 2011 (left) and Monday (right), with blue for LaGuardia and red for Kennedy.

(Source: Ferreira et al., 2013)

New York City Taxi example – further capabilities

- Side-by-side “sensor” maps over time
- Visual queries for pick-up AND drop-off
- Constraints of attributes of taxi id, distance traveled, fare, and tip amount
  - Enables economic analysis
- Complex queries.
  - Use set-theoretic functions on simple queries
- Level-of-detail reduced the number of points shown on the map.
  - Done by hierarchical sampling of point cloud
- Density heat maps
- Different visualizations



(Source: Ferreira et al., 2013)

Analytics

- Analytics refers to a variety of exploratory and statistical methods that can be applied to Big Data and Small Data in order to see patterns, trends, associations, comparisons, and sometimes apply statistical methods to smaller samples.

Among the analytics available in Tableau are:

Mapping
Percentage
Totals
Forecast
Trend Lines
Special Values
Table Layout
Legend
Filters
Highlighters
Parameters
Calculated Fields
Cycle Fields

Spatial Big Data connected with Analytics

- Big Data often contain geo-referencing, by lat.-long. coordinates or by addresses that can be geocoded to lat.-long.
- Spatial big data have the advantage in analysis that maps can be produced, the big data can be organized and data-mined more thoroughly with the spatial dimension added, and spatial analysis methods can be utilized for more thorough analysis on smaller parts of the big data.

Big Data’s Unstructured Data Sources

- The “V” for variety implies a mixture of data types.
- Includes:
  - Sensor Data. Is projected to grow more rapidly than human population.
  - Smartphones
  - Internet of Things (IoT)
  - Lidar
  - Satellite imagery
  - Social media

  Note: all of these are sources for Spatial Big Data.

Since all these data types are allowable with Big Data, the challenge becomes how to organize the unstructured data.

Management of Big Data and Spatial Big Data

- Some of biggest challenges with Big Data (and Spatial Big Data) are the management of it.
- Big Data people do not just fit in with IT, and Spatial Big Data people might not fit in with GIS department, if there is one.
- The responsibility of the Big Data person is not just to conceive, design, and build, but rather to perform continual monitoring of streams of data. There needs to be proactiveness to “sampling, analyzing, and acting on data as necessary.” (Davenport, 2014).
- This shifts away from the traditional view of GIS and IT departments of being project-driven. Once a project is completed, it goes into maintenance for 5 years. With Big Data, there is continual flow requiring unrelenting discovery and agility.

Qualitative Studies of Location and Big Data:

- Case studies are appearing that emphasize the managerial, leadership, and strategic aspects of Location and Big Data.
- In a case study in Big Data at Work by T. Davenport (2014), the big data example of Bathworks is give. Bathworks is a manufacturer of plumbing products in the U.S. Spatial data are applied in several ways.
  - For any Bathworks vehicle, the facilities manager knows locations, distance traveled or one day, average or peak speeds, acceleration/braking patterns. If the patterns are wasteful of energy or risky to the driver, reminder e-mails and text messages are sent.
  - Bathworks also has the potential to apply location analytics to its 23,000 building spaces, which are monitored for humidity, temperature, levels of light and human presence.
- The workshop keynote speaker Dr. Stephanie Woerner and her collaborators have developed case studies of Big Data that have important location elements.
- Several other case study examples will appear in an upcoming special issue of MISQ Executive on Big Data and Analytics.

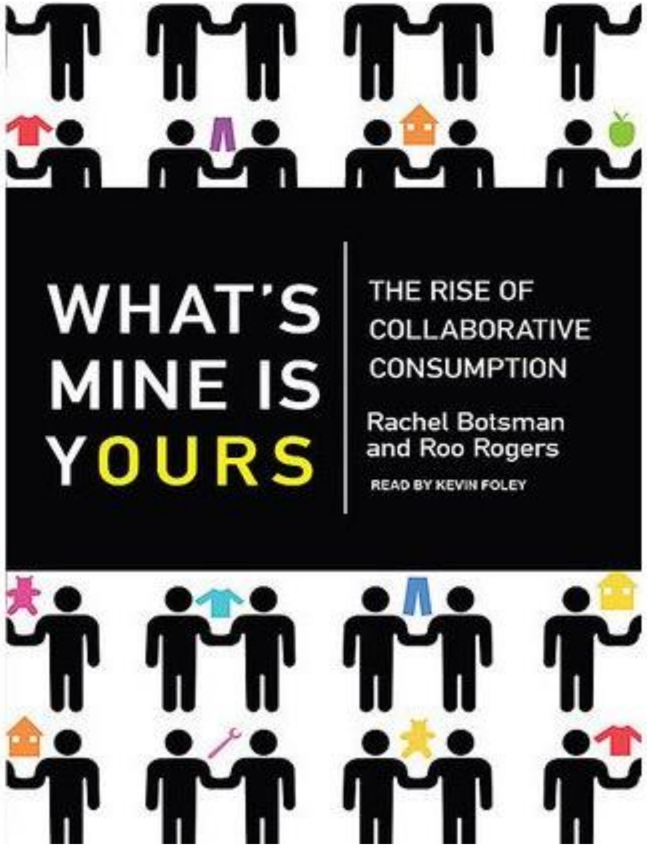
Questions still unanswered with Big Data

- How will Spatial Big Data affect organizational processes.
  - One possible trend is towards centralization of data in the Cloud, after decades of decentralization.
- How will concern be addressed about privacy invasion and targeting from Big Data.
  - The appeal to unsuspecting users can come from it being “clothed” in social media (Foursquare) or retail discounting.
  - A backlash against this privacy and security intrusion is likely
- How will Big Data and Analytics change decision-making.
- To what extent will human managers and decision-makers be sufficiently aware and empowered to override the results of Big Data.

Introduct conomy

Sharing Economy: Definitions

- Botsman and Rogers (2010): Hyper-consumption (20th century) → Collaborative Consumption (21st century).
  - Botsman (2015): an economic system of decentralized networks and marketplaces that unlocks the value of underused assets by matching needs and haves, in ways that bypass traditional middlemen.
- Stephany (2015)
  - The sharing economy is the value in taking underutilized assets and making them accessible online to a community, leading to a reduced need for ownership of those assets.



Sharing Economy: Definitions

- Frenken *et al,* 2015
  - Economy in which consumers grant other consumers temporary access to under-utilized assets possibly for money.
- Maselli et al (2016) expanded this definition.
  - Discarded the temporary access aspect by considering all goods that are shared among consumers in a second-hand economy.
  - Take into account the provision of services from one consumer to another via contests or auctions, instead of only counting the trade in under-utilized assets.

- Collaborative Economy
- Gig Economy
- On-demand Economy
- Peer Economy
- Renting Economy

5 Characteristics of Sharing Economy

(Sundararajan, 2016)

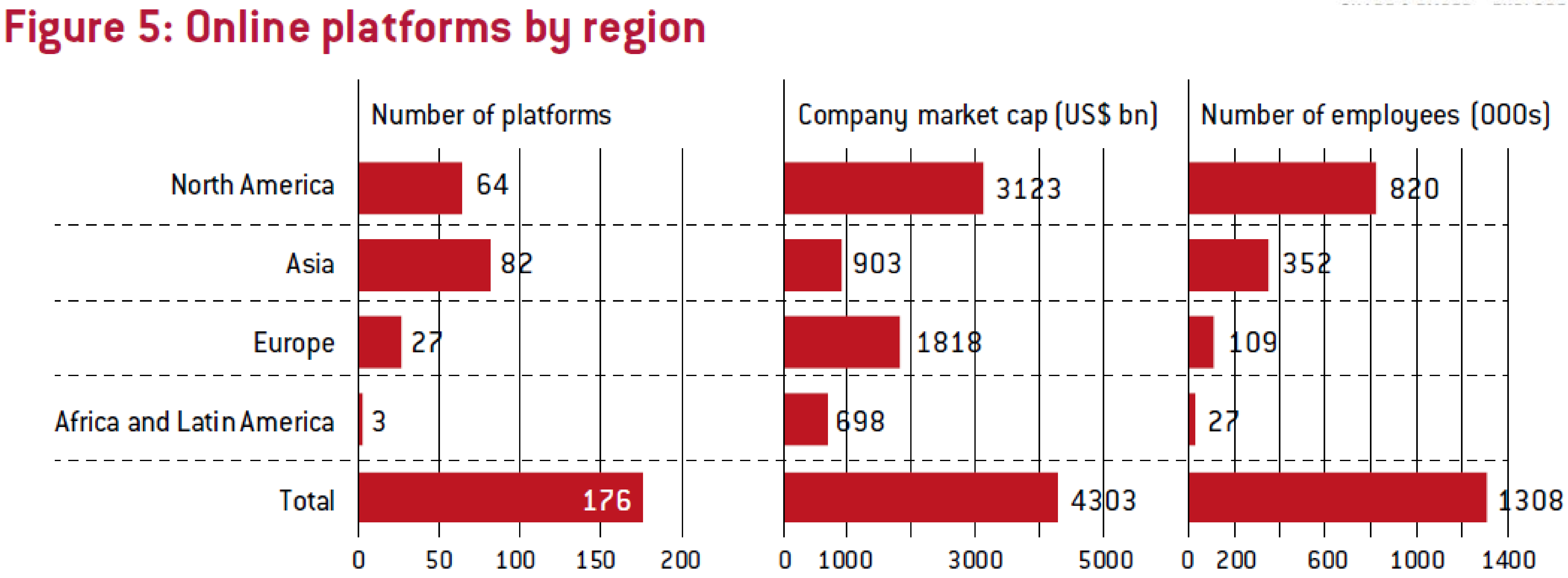
- Largely market-based
- High impact capital
- Crowd-based networks rather than centralized institutions
- Blurring lines between personal and professional
- Blurring lines between fully employed and casual labor, between independent and dependent employment, between work & leisure.

The term generates criticism!

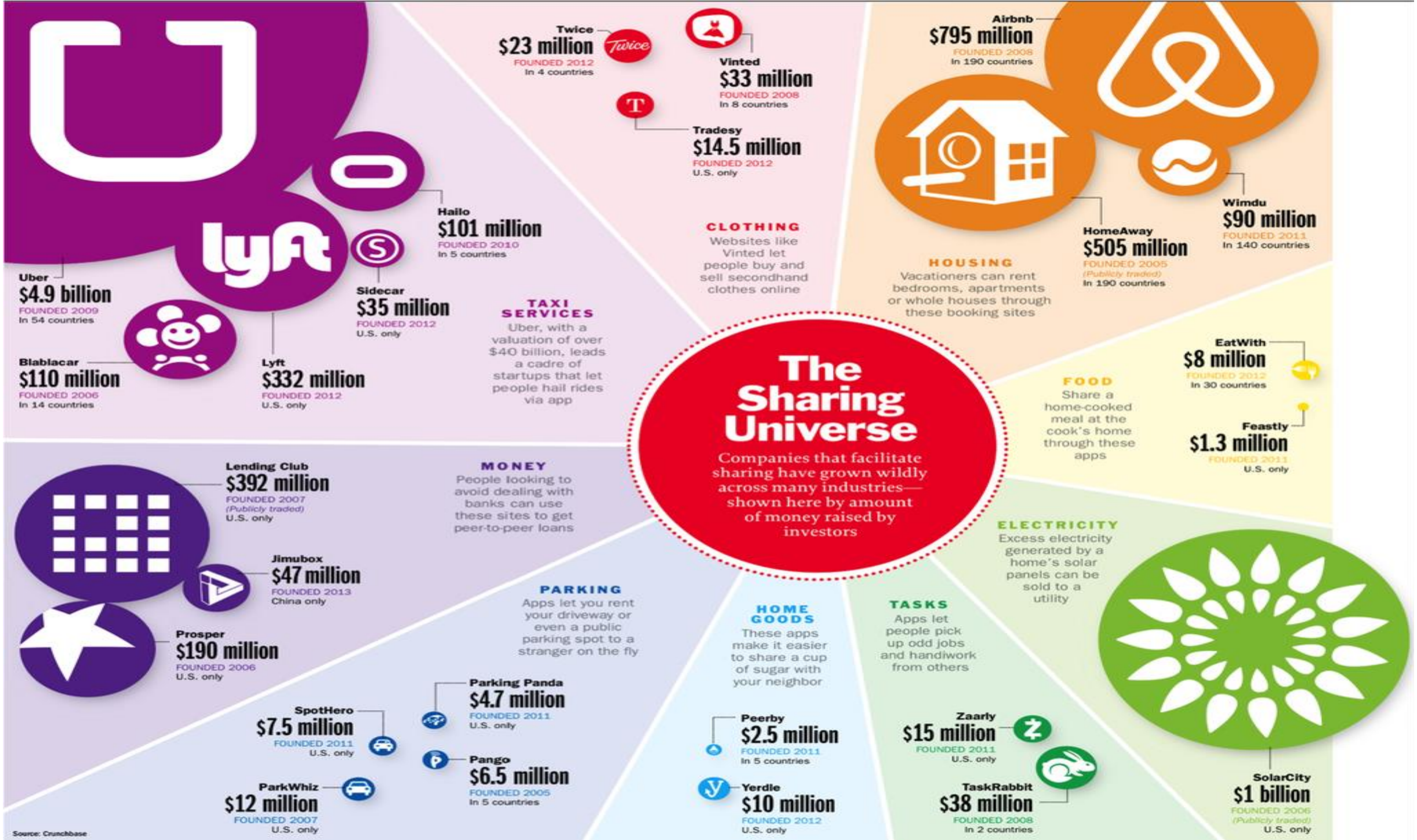
- Some commentators argue that the word “sharing” is a “misnomer” employed to mask the essentially commercial nature of the activity on these platforms.
- The term misleadingly “frames technology-enabled transactions as if they were altruistic or community endeavors” and ….
- “create[s] a halo of positive branding to avoid the discussion of what regulatory structures need to be modernized to deal with these platforms.”
- Terms such as ‘sharing’ and ‘collaborative’ incorrectly imply services being provided for free” although service providers are simply using their assets to earn money.” (FTC Report, Nov 2016).

Worldwide Phenomenon

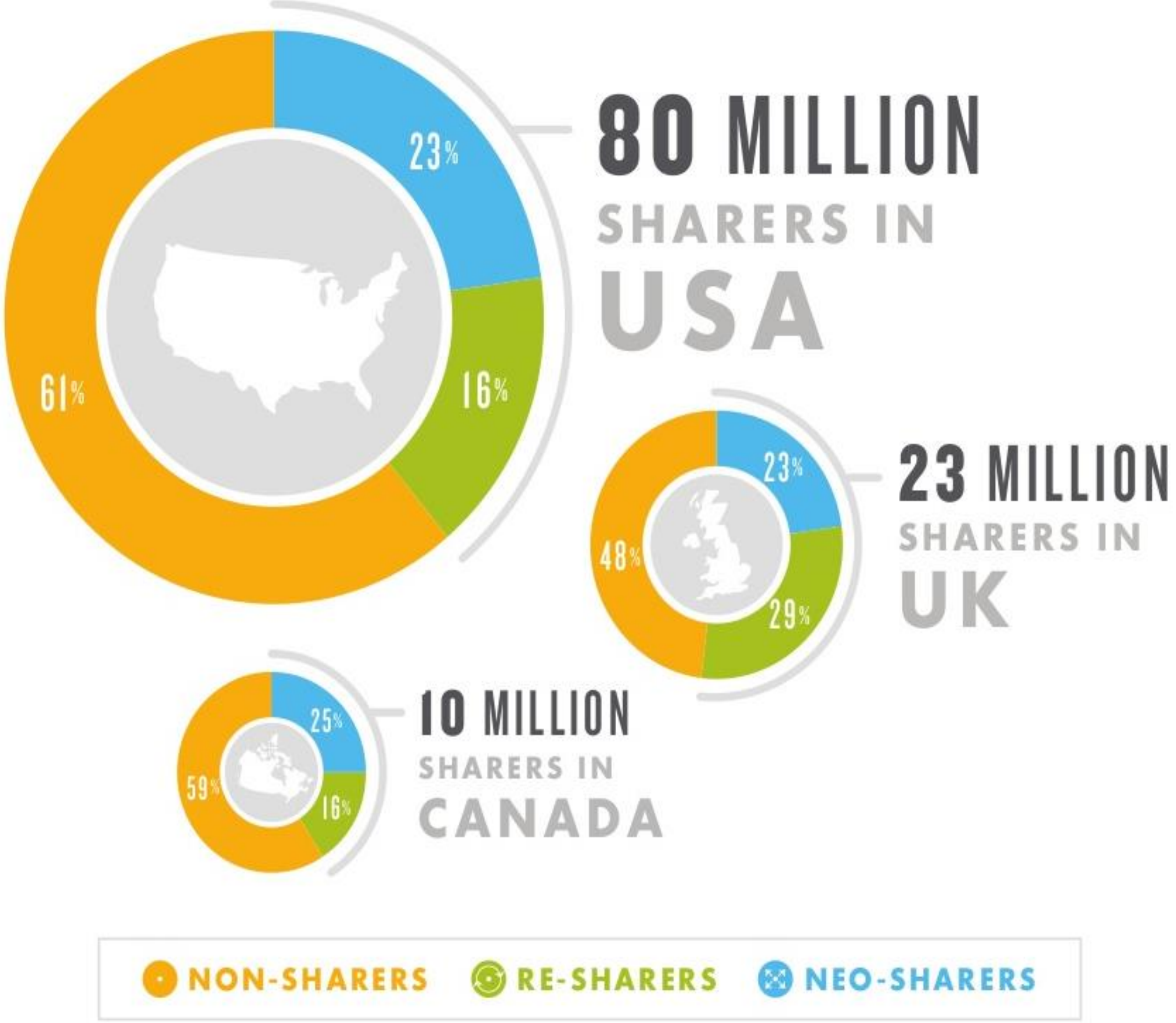
Interactive Map of Sharing Economy Companies, Based on Company's HQ

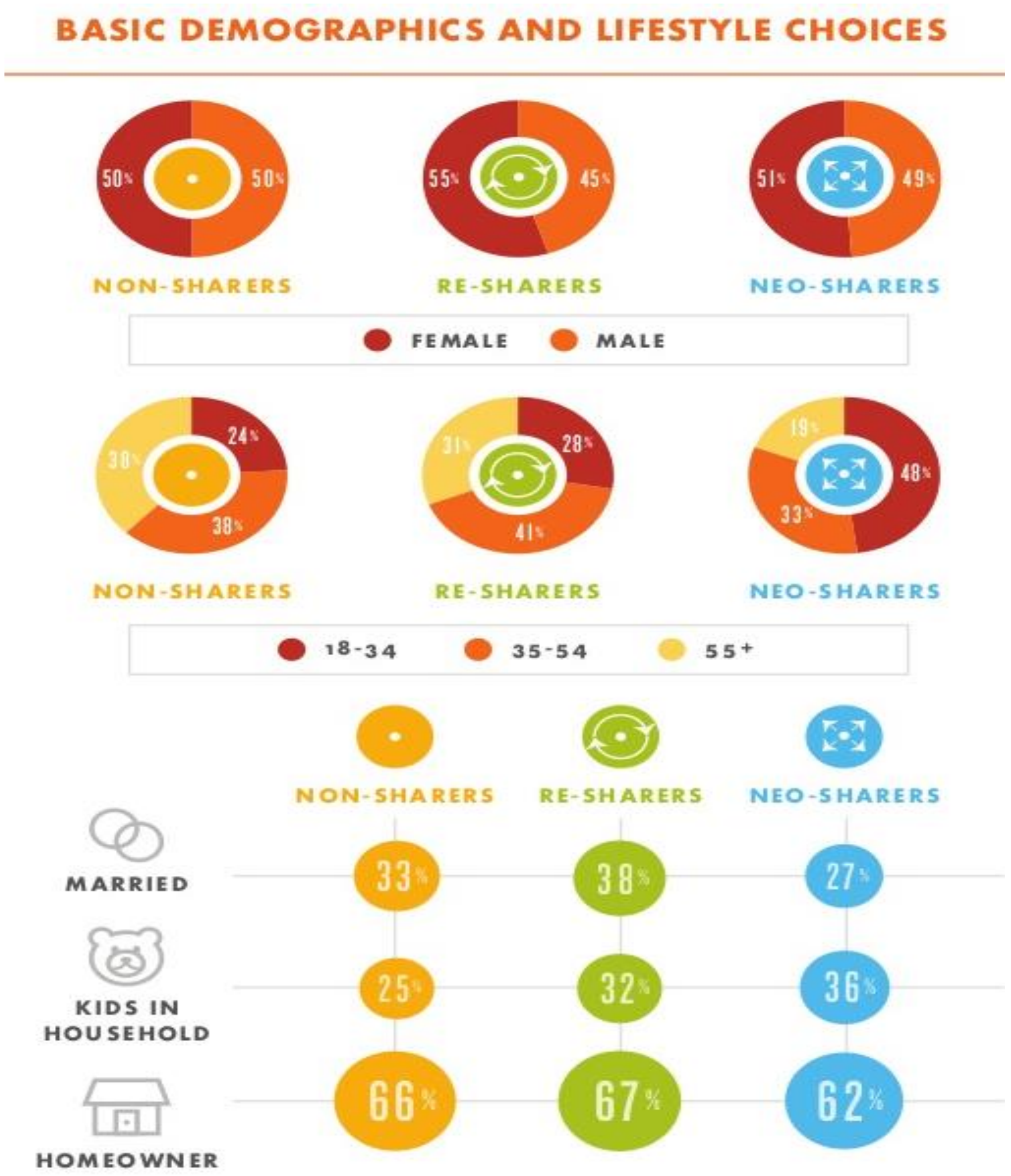


Source: Evans and Gawer (2016).


The Sharing Universe
Companies that facilitate sharing have grown wildly across many industries—shown here by amount of money raised by investors
TAXI SERVICES
Uber, with a valuation of over $40 billion, leads a cadre of startups that let people hail rides via app
Uber
$4.9 billion
FOUNDED 2009
In 54 countries
Blablacar
$110 million
FOUNDED 2006
In 14 countries
Lyft
$332 million
FOUNDED 2012
U.S. only
Sidecar
$35 million
FOUNDED 2012
U.S. only
Hailo
$101 million
FOUNDED 2010
In 5 countries
CLOTHING
Websites like Vinted let people buy and sell secondhand clothes online
Twice
$23 million
FOUNDED 2012
In 4 countries
Vinted
$33 million
FOUNDED 2008
In 8 countries
Tradesy
$14.5 million
FOUNDED 2012
U.S. only
HOUSING
Vacationers can rent bedrooms, apartments or whole houses through these booking sites
Airbnb
$795 million
FOUNDED 2008
In 190 countries
HomeAway
$505 million
FOUNDED 2005
(Publicly traded)
In 190 countries
Wimdu
$90 million
FOUNDED 2011
In 140 countries
FOOD
Share a home-cooked meal at the cook's home through these apps
EatWith
$8 million
FOUNDED 2012
In 30 countries
Feastly
$1.3 million
FOUNDED 2011
U.S. only
ELECTRICITY
Excess electricity generated by a home's solar panels can be sold to a utility
SolarCity
$1 billion
FOUNDED 2006
(Publicly traded)
U.S. only
TASKS
Apps let people pick up odd jobs and handiwork from others
Zaarly
$15 million
FOUNDED 2011
U.S. only
TaskRabbit
$38 million
FOUNDED 2008
In 2 countries
HOME GOODS
These apps make it easier to share a cup of sugar with your neighbor
Peerby
$2.5 million
FOUNDED 2011
In 5 countries
Yerdle
$10 million
FOUNDED 2012
U.S. only
PARKING
Apps let you rent your driveway or even a public parking spot to a stranger on the fly
Parking Panda
$4.7 million
FOUNDED 2011
U.S. only
SpotHero
$7.5 million
FOUNDED 2011
U.S. only
Pango
$6.5 million
FOUNDED 2005
In 5 countries
ParkWhiz
$12 million
FOUNDED 2007
U.S. only
MONEY
People looking to avoid dealing with banks can use these sites to get peer-to-peer loans
Lending Club
$392 million
FOUNDED 2007
(Publicly traded)
U.S. only
Jimubox
$47 million
FOUNDED 2013
China only
Prosper
$190 million
FOUNDED 2006
U.S. only
Source: Crunchbase

Sharing Economy Landscape in U.S. (Pew, Research Center, 2016; Vision Critical, 2016; Petropoulos, 2017)



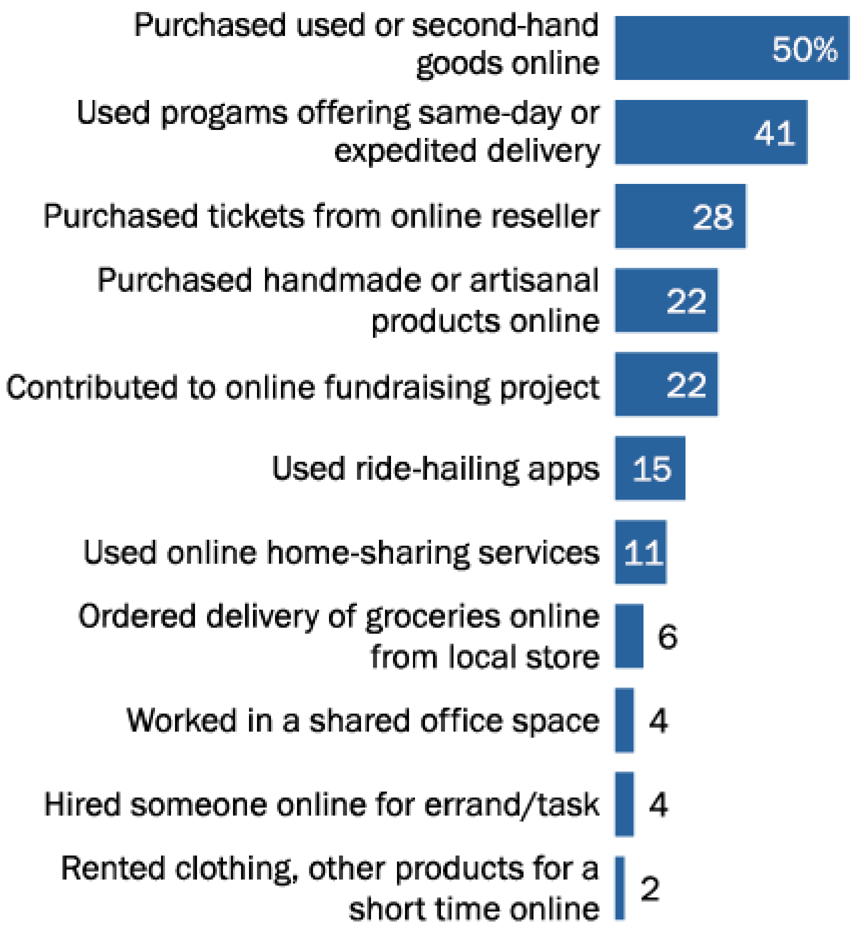


Sharing Economy Landscape in U.S.

(Pew, Research Center, 2016; Petropoulos, 2017)

72% of Americans have used some type of shared or on-demand online service

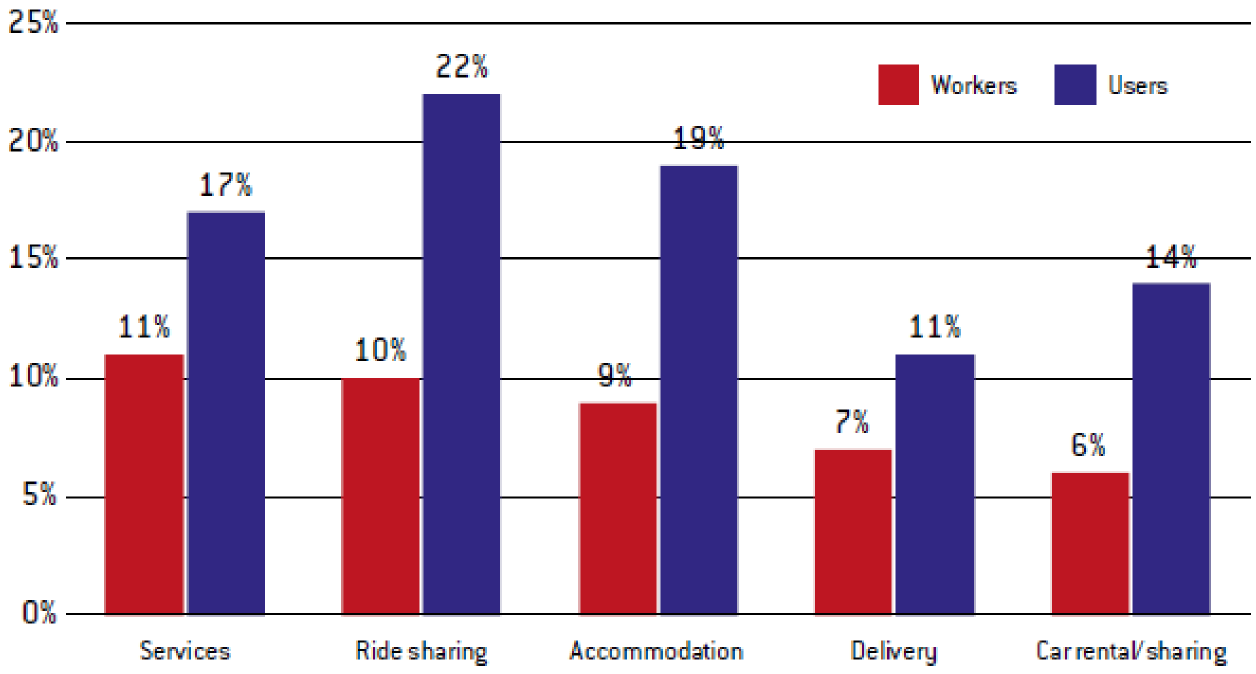
% of adults who have ...



Many are unfamiliar with the vocabulary of the new digital economy

% of adults who have heard of the following terms

Figure 2: US population share that participates in collaborative economy

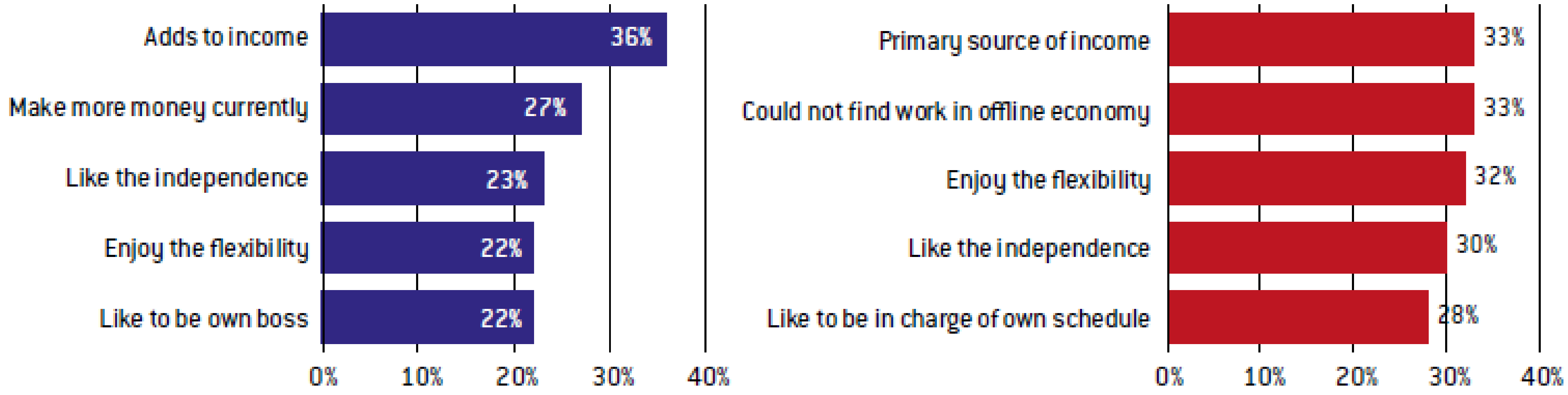


Source: Burston-Marsteller, the Aspen Institute and TIME (2015).

Some motivations to participate….

(Petropoulos, 2017)

Figure 4: Reasons for joining collaborative economy platforms, casual workers (left panel) and regular workers (right panel)

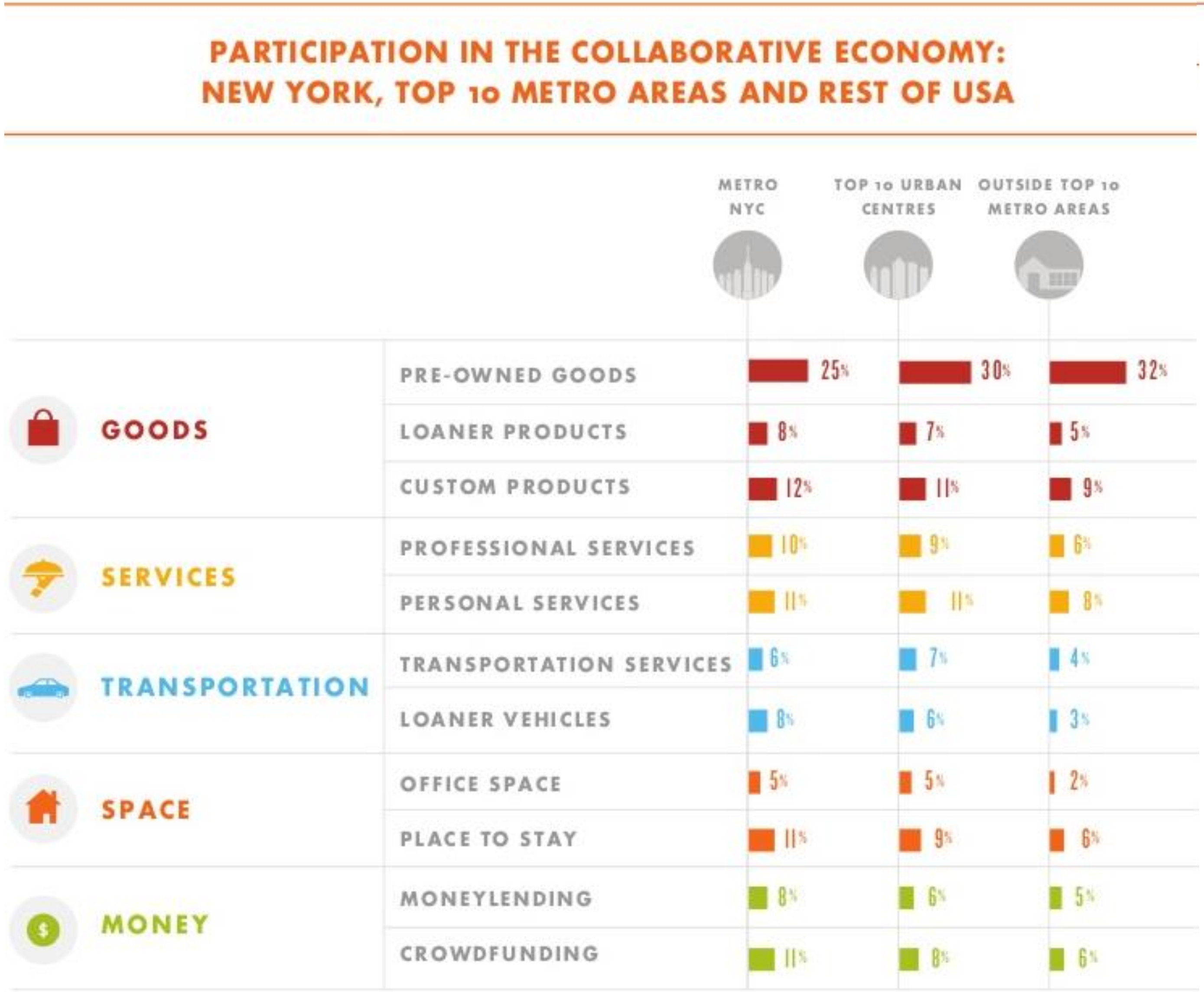


Source: Burston-Marsteller, the Aspen Institute and TIME (2015).

Spatial/Locational Aspects of Sharing Economy

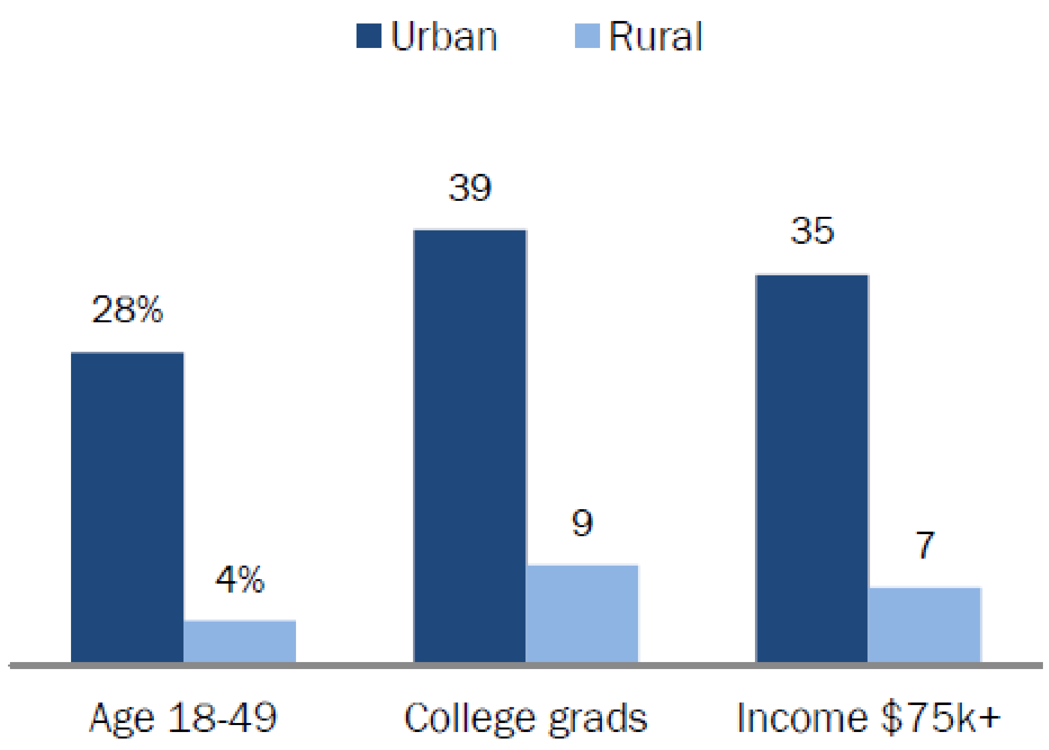
- Where we live says a lot about us.
- Possess unique demographic, economic, social, employment attributes.
- Where one lives impacts one’s attitudes towards trust, waste, and social connectedness.
- Sharing happens someplace.
  - Sharing Economy participants live/work/share someplace.
- Cities are sharing economies (Sundararajan, 2016).

Is sharing an urban phenomenon?



Ride-hailing use is especially high among some urban residents; use by rural dwellers is consistently low

% of urban/rural Americans in each category who have used ride-hailing apps



Study 1: Taxis as sensors of city life (Ferreira et al, 2013)

- Develop new models to visually query very large data warehouses of taxicab trips in NYC.
- Conduct wide range of spatio-temporal queries and origin-destination queries that enable the study of mobility in NYC.
- Case studies motivated by traffic engineers and economists show how model and system enable domain experts to perform tasks that were previously unattainable.
- Public policy implications.

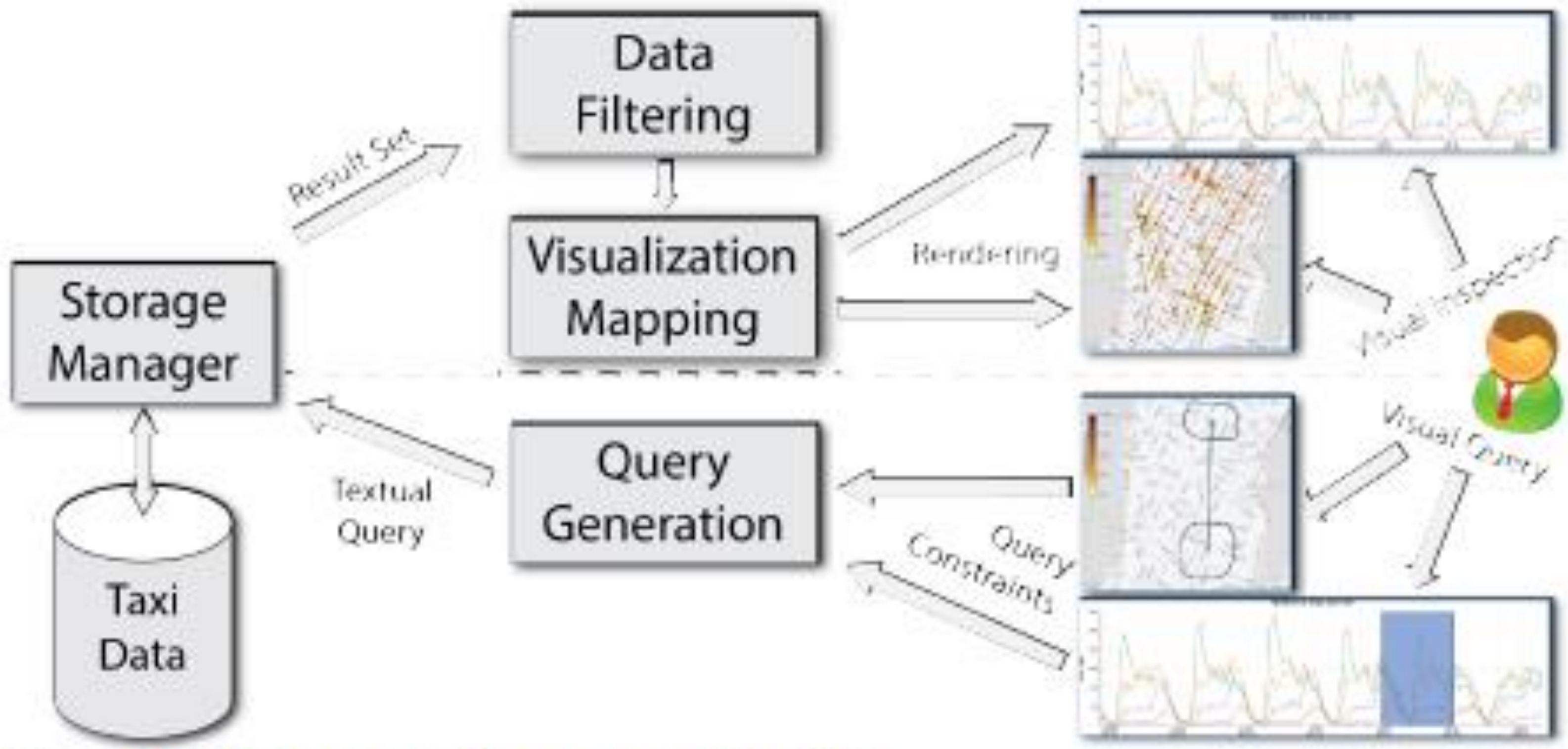
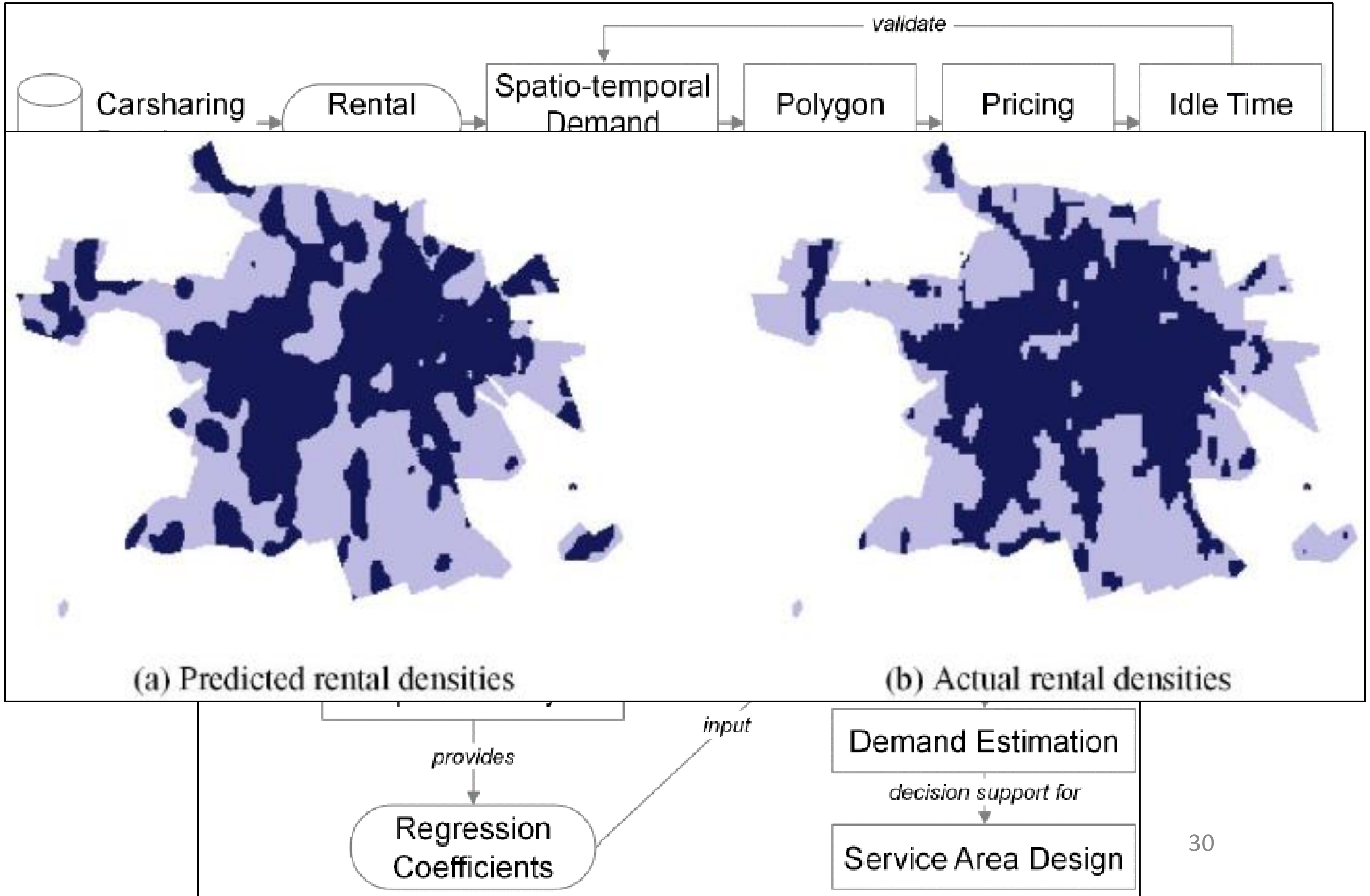


Fig. 3. High-level architecture of TaxiVis.

Study 2: Moving in time and space – Location intelligence for carsharing decision support, Willing et al, (2017)

Decision Support Systems Special Issue on Location Analytics and Decision Support.

- Apply location intelligence and analytics to provide decision support to carsharing operators.
- If a car is dropped off in "dead zone," relocation cost to a "pick-up zone" is much higher.
- Hence, the system enables operators to develop zone-based flexible pricing schemes.
- Spatial and temporal variation in carsharing demand is explained and predicted.
- Prediction quality is validated by analyzing data from Amsterdam and Berlin.



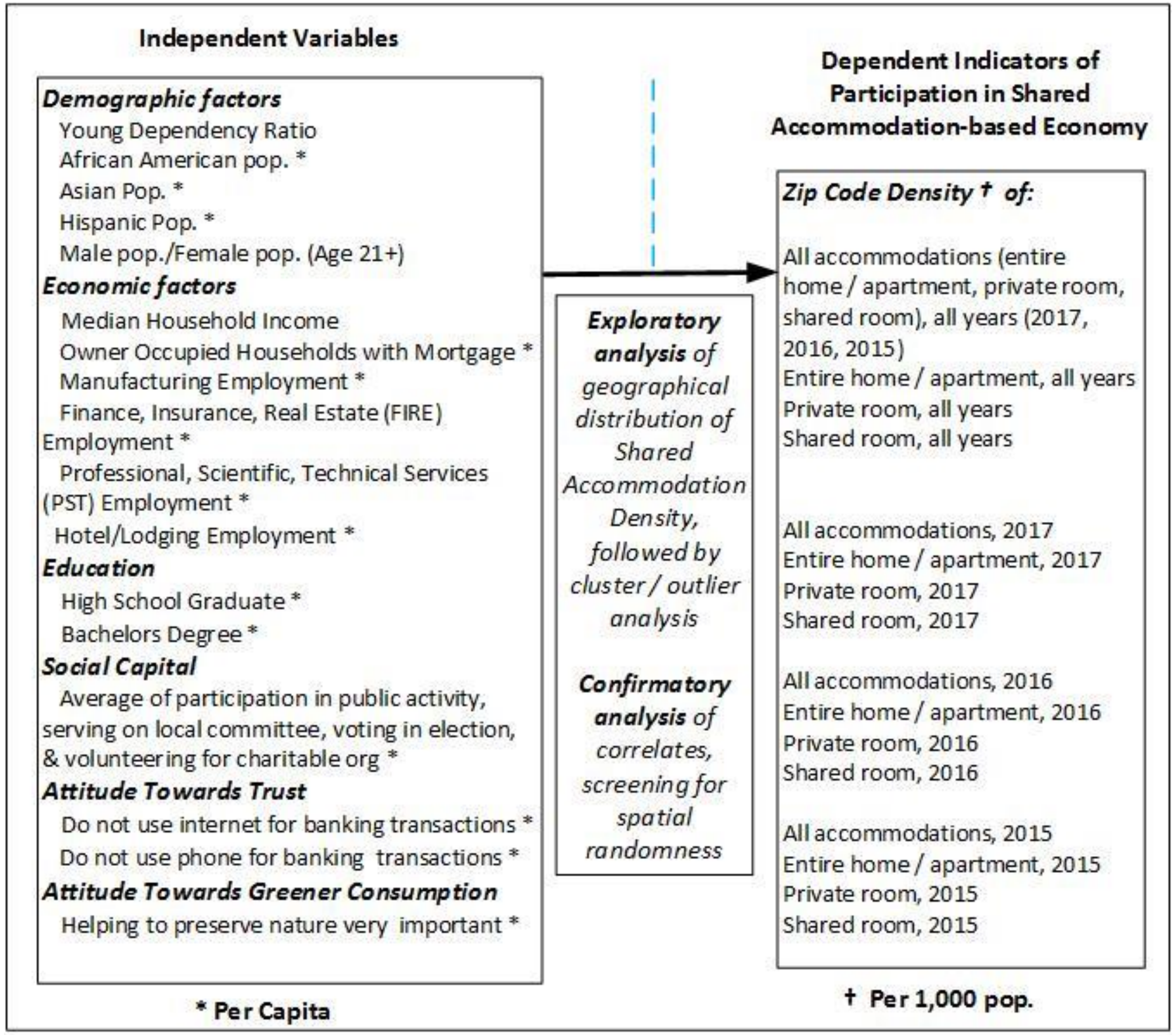
(a) Predicted rental densities

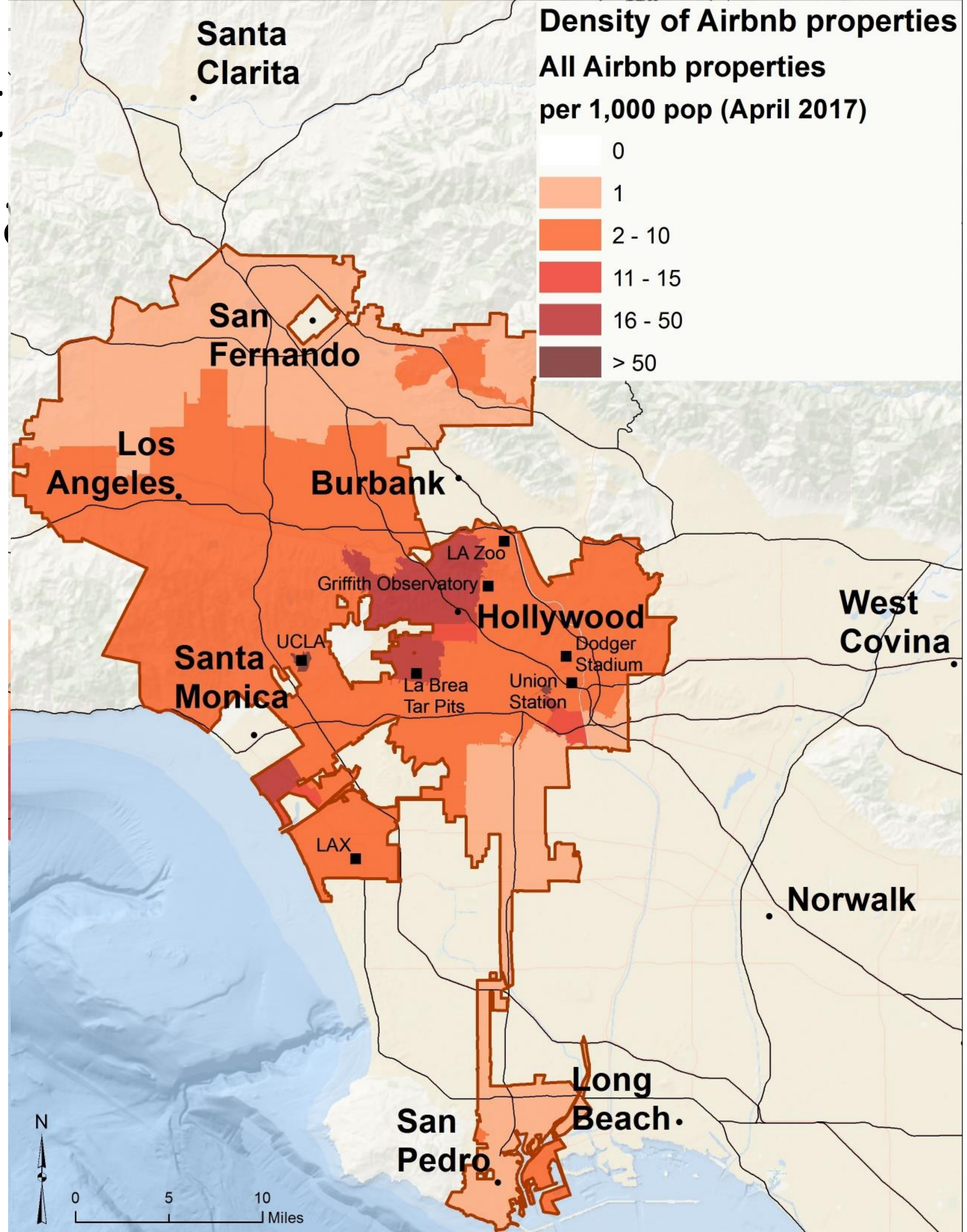
(b) Actual rental densities

Study 3: Spatiotemporal Patterns and Socioeconomic Dimensions of Shared Accommodations: The Case of Airbnb in Los Angeles, California (Sarkar, Koohikamali, & Pick, ISSC 2017)

1. What are spatial and temporal patterns of host participation in the shared accommodation-based economy in the greater Los Angeles area between 2015 and 2017?
2. Are geographic agglomerations of host participation in shared accommodation based economy present in the greater LA area as estimated by cluster and outlier analysis?
3. What are the associations of demographic, socioeconomic, occupational, social capital variables as well as attitudes towards trust and greener consumption with host participation in shared accommodation-based economy in the greater Los Angeles area and the city of Los Angeles?
4. Can a regression model of such associations account for presence of spatial bias in participation?

Conceptual Model of Host
Shared Accommodation-ba






INSIDE THIS WEEK: TECHNOLOGY QUARTERLY
The Economist
MARCH 9TH–15TH 2013
Economist.com
Hugo Chávez's rotten legacy
Is the stockmarket right?
Management tips for the Vatican
Getting Britain to grow
The comet that could hit Mars
The sharing economy
Room $38 a night
Dog $5 a walk
Parking space $20 a day
Loft storage $300 a year
Pickup truck $9 an hour
Bike $18 a day
Hermès Birkin bag $100 a party
Saxophone $55 a month
Surfboard $80 a week
Lawnmower $6 a day
Speedboat $300 a day
On the internet, everything is for hire

Questions (???) , Discussion

Contact info.
James_pick@Redlands.edu
Avijit_Sarkar@Redlands.edu