

Designing Breadth-Oriented Data Exploration for Mitigating Cognitive Biases

Po-Ming Law
Rahul C. Basole



What is Breadth-Oriented Exploration?

Exploring more = Breadth?

Like Talking about A Large Apple...

How large is large?



Why Breadth in Data Exploration?

It is the purpose of exploratory data analysis!

“Exploratory data analysis is detective work.”

John W. Tukey

What Causes the Lack of Breadth?

City	Region	Causes of Lack of Breadth														
		Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education	
Abidjan	Africa	4	4	4	4	2	3	3	2	4	2	2	4	4	5	
Bangkok	Asia	3	3	3	3	2	2	1	4	4	4	4	4	4	1	
Beijing	Asia	2	1	1	3	2	3	1	2	2	1	4	3	3	3	
Copenhagen	Europe	1	1	2	1	1	1	1	1	2	1	1	2	1	1	
Dubai	Middle East	1	1	2	2	3	2	2	2	3	1	1	2	4	3	
Kiev	Europe	3	1	1	3	2	2	2	1	3	3	3	4	2	1	
Tehran	Middle East	2	3	3	3	5	5	3	2	3	4	4	4	2	3	
Zurich	Europe	1	1	1	1	1	1	1	1	2	1	1	2	2	2	
•																

Where should I go after I graduate?

What Causes the Lack of Breadth?

City	Region	Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education
Abidjan	Africa	4	4	4	4	2	3	3	2	4	2	2	4	4	5
Bangkok	Asia	2	3	3	3	1	3	2	1	4	2	1	4	4	1
Beijing	Asia	2	1	1	3	2	3	1	1	2	2	1	4	3	3
Copenhagen	Europe	1	1	2	1	1	1	1	1	2	1	1	2	1	1
Dubai	Middle East	1	1	2	2	3	2	2	2	3	1	1	2	4	3
Kiev	Europe	3	1	1	3	2	2	2	1	3	3	3	4	2	1
Tehran	Middle East	2	3	3	3	5	5	3	2	3	4	4	4	2	3
Zurich	Europe	1	1	1	1	1	1	1	1	2	1	1	2	2	2

•

What Causes the Lack of Breadth?

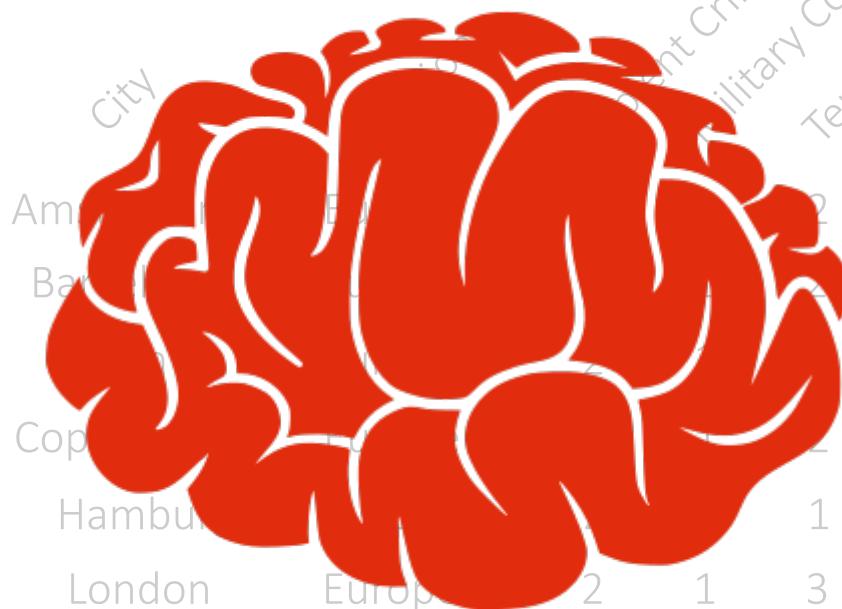
City	Region	Causes of Lack of Breadth													
		Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education
Athens	Europe	1	2	2	2	1	1	2	1	2	3	2	2	3	2
Barcelona	Europe	2	1	2	1	1	1	1	1	2	1	2	2	2	1
Belgrade	Europe	3	2	2	3	2	2	2	3	5	1	4	2	2	2
Hamburg	Europe	2	1	1	1	1	1	1	1	1	1	1	2	2	2
London	Europe	2	1	3	2	1	1	1	2	2	1	1	1	1	1
Oslo	Europe	1	1	1	1	1	1	2	2	1	1	1	2	1	1
Rome	Europe	2	1	1	2	1	1	1	1	2	1	3	2	1	1
Stockholm	Europe	1	1	1	1	1	1	2	1	1	2	1	1	2	1

Europe sounds like an awesome place!

What Causes the Lack of Breadth?

City	Region	Causes of Lack of Breadth														
		Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education	
Athens	Europe	1	2	2	2	1	1	2	1	3	2	2	3	2	2	
Barcelona	Europe	2	1	2	1	1	1	1	1	2	1	2	2	2	1	
Belgrade	Europe	3	2	2	3	2	2	2	2	3	3	1	4	2	2	
Hamburg	Europe	2	1	1	1	1	1	1	1	1	1	1	1	2	2	
London	Europe	2	1	3	2	1	1	1	1	2	2	1	1	1	1	
Oslo	Europe	1	1	1	1	1	1	2	2	1	1	1	1	2	1	
Rome	Europe	2	1	1	2	1	1	1	1	1	2	1	3	2	1	
Stockholm	Europe	1	1	1	1	1	1	2	1	1	2	1	1	2	1	
•																

What Causes the Lack of Breadth?



My subconscious mind

City	America	Bangkok	Copenhagen	Hamburg	London	Paris	Rome	Ant Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity provision	Corruption	Climate Rating	Education
Amsterdam	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Buenos Aires	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Copenhagen	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Hamburg	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
London	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Paris	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Rome	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Amsterdam	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Buenos Aires	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Copenhagen	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Hamburg	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
London	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Paris	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Rome	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

**These cities look familiar to me and they get high scores.
Let me choose them.**

What Causes the Lack of Breadth?

City	Region	Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education
Amsterdam	Europe	2	1	2	1	1	1	1	1	2	1	1	1	1	1
Barcelona	Europe	2	1	2	1	1	1	1	1	2	1	2	2	1	
Berlin	Europe	2	1	2	1	1	1	1	1	1	1	1	2	2	
Copenhagen	Europe	1	1	2	1	1	1	1	1	2	1	1	2	1	
Hamburg	Europe	2	1	1	1	1	1	1	1	1	1	1	2	2	
London	Europe	2	1	3	2	1	1	1	2	2	1	1	1	1	
Paris	Europe	2	1	1	1	1	1	1	2	1	1	2	1	1	
Rome	Europe	2	1	1	2	1	1	1	1	2	1	3	2	1	
					•	•	•								

What Causes the Lack of Breadth?

City	Region	Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education
Budapest	Europe	2	1	1	2	1	1	1	1	2	2	1	3	2	1
Dusseldorf	Europe	2	1	2	1	1	1	1	1	1	1	1	1	2	2
Frankfurt	Europe	2	1	2	1	1	1	1	1	1	1	1	1	2	2
Geneva	Europe	1	1	1	1	1	1	1	1	2	1	1	2	2	2
Helsinki	Europe	1	1	1	1	1	1	2	1	1	2	1	1	2	1
Lisbon	Europe	2	1	1	2	1	1	1	2	2	1	2	1	1	1
Oslo	Europe	1	1	1	1	1	1	2	2	1	1	1	2	1	1
Reykjavik	Europe	1	1	1	1	1	1	2	2	2	2	1	2	1	1
⋮															

Missing lots of incredible cities...

What Causes the Lack of Breadth?

City	Region	Violent Crime	Military Conflict	Terrorism	Healthcare Provision	Religious Restrictions	Censorship	Sporting Facilities	Restaurants Availability	Road Network	Housing	Electricity Provision	Corruption	Climate Rating	Education
Budapest	Europe	2	1	1	2	1	1	1	1	2	2	1	3	2	1
Dusseldorf	Europe	2	1	2	1	1	1	1	1	1	1	1	1	2	2
Frankfurt	Europe	2	1	2	1	1	1	1	1	1	1	1	1	2	2
Geneva	Europe	1	1	1	1	1	1	1	1	1	2	1	1	2	2
Helsinki	Europe	1	1	1	1	1	1	2	1	1	2	1	1	2	1
Lisbon	Europe	2	1	1	2	1	1	1	1	2	2	1	2	1	1
Oslo	Europe	1	1	1	1	1	1	2	2	1	1	1	1	2	1
Reykjavik	Europe	1	1	1	1	1	1	2	2	2	2	2	1	2	1
•															

What Causes the Lack of Breadth?

Familiar names: London Paris Rome

V

Strange names: Oslo Dusseldorf Reykjavik

Familiarity Heuristic

What Causes the Lack of Breadth?

Familiar names: London Paris Rome

Use of **heuristics** is a potential cause
of lack of breadth in data exploration

Strange names: Oslo Dusseldorf Reykjavik

Familiarity Heuristics

Problem

Solution

Information Overload



Heuristics



Lack of Breadth



Miss Insights

Show me the
other cities I
might be
interested in

What is Breadth-Oriented Exploration?

Exploring more = Breadth?

What is Breadth-Oriented Exploration?

Exploring more = Breadth?

A strategy to **overcome lack of breadth** in exploration due to information overload

How do you design a system to overcome the lack of breadth?

Three Design Considerations:

- Unit of Exploration

- User-driven vs System-driven Exploration

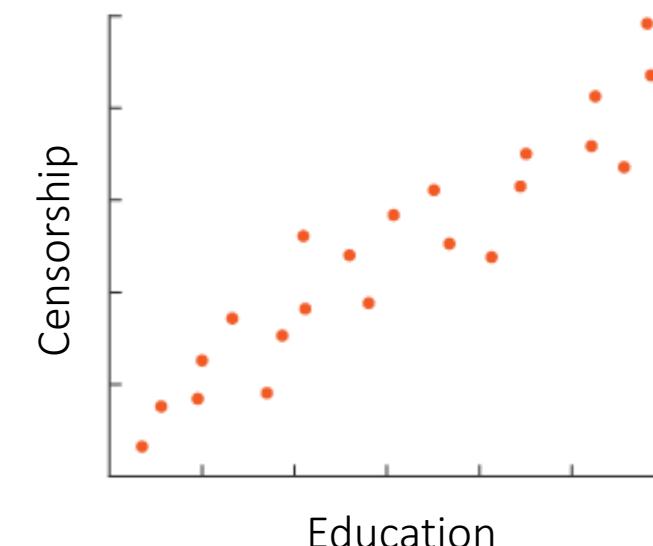
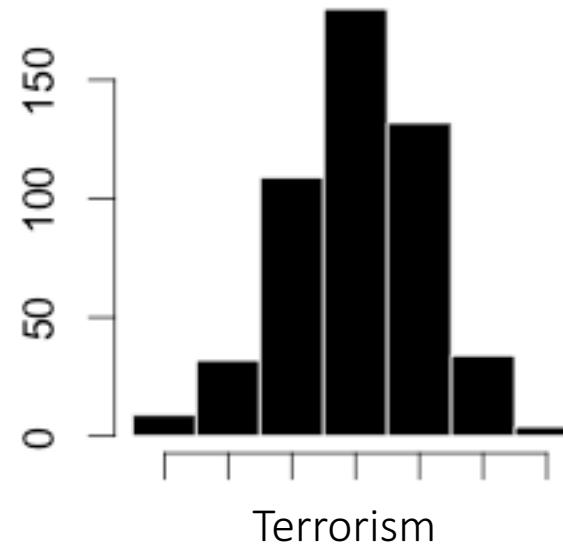
- Related vs Systematic Exploration

First Consideration: Unit of Exploration

What things to which the system should expose users more?

Data case or dimension?

We may want to look at different combinations of dimensions during exploratory data analysis



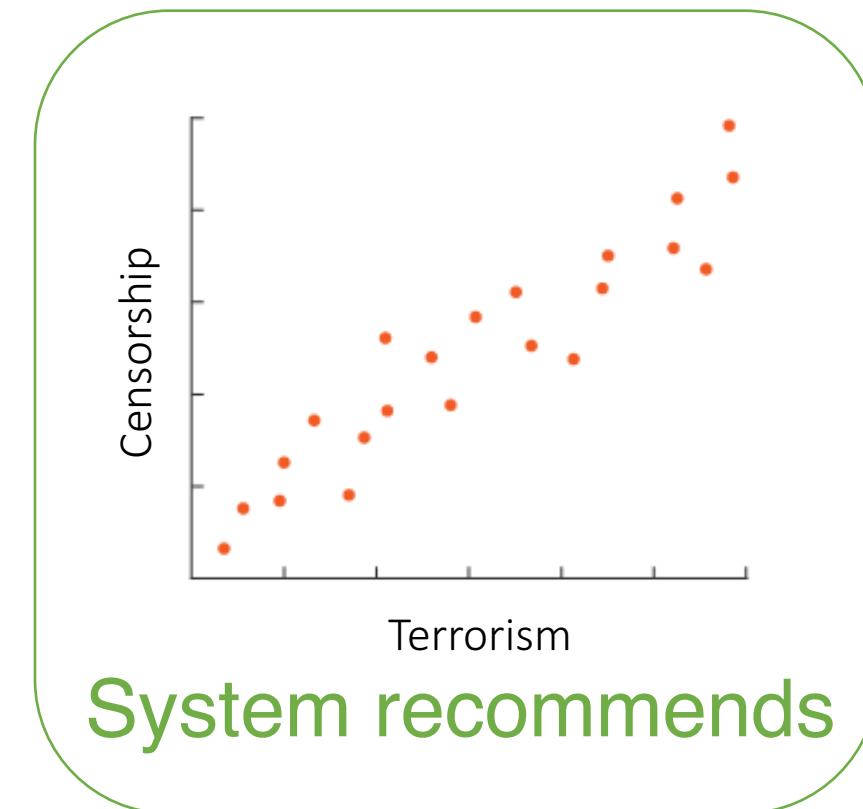
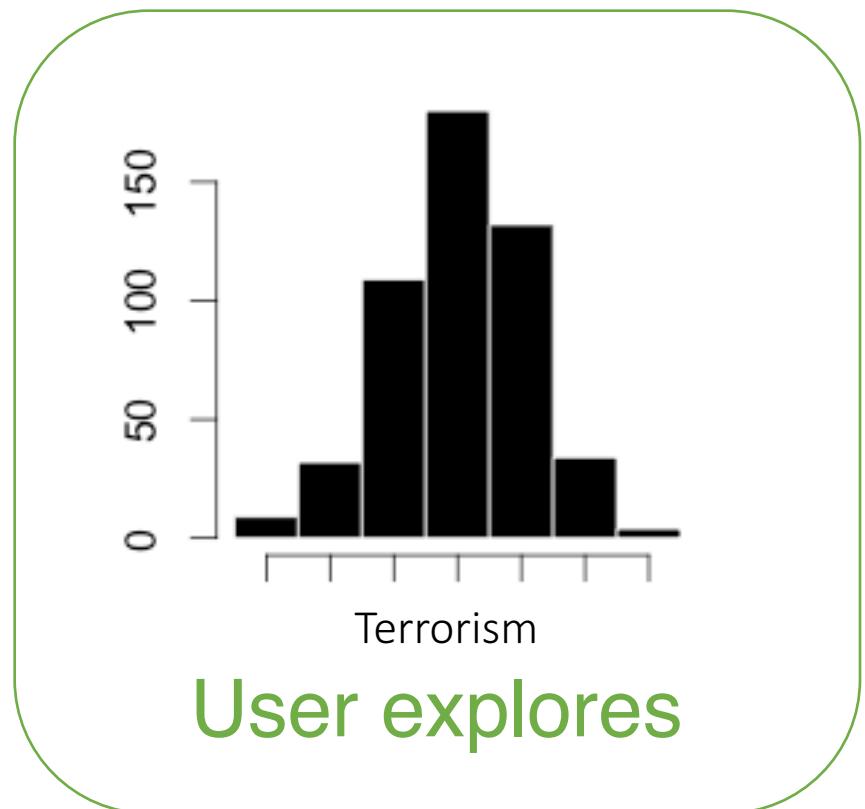
Use of Dimension as a Unit of Exploration

1. Showing dimension coverage information



Use of Dimension as a Unit of Exploration

2. Showing new dimensions as you explore

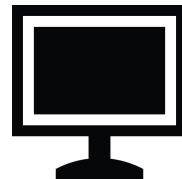


Unit of Exploration

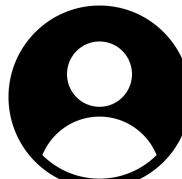
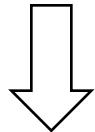
How does it apply to a system for exploring the city dataset? ➔ Data case as a unit



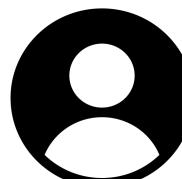
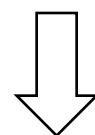
Second Consideration: User-driven vs System-driven Exploration



Information about exploration



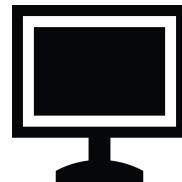
Awareness of exploration



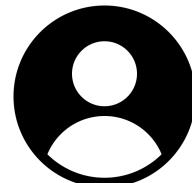
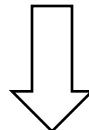
Less-biased exploration

Users are responsible for ensuring breadth

Second Consideration: User-driven vs System-driven Exploration



Select information to show



Less-biased exploration

System is responsible for ensuring breadth

User-driven vs System-driven Exploration

How does it apply to a system for exploring the city dataset? ➔ **System-driven exploration**



Third Consideration: Related vs Systematic Exploration

Only applies to system-driven exploration...

Frequently Bought Together



Price For Both: £364.24

Add both to Basket

These items are dispatched from and sold by different sellers. [Show details](#)

Showing something related but different

Third Consideration: Related vs Systematic Exploration

Only applies to system-driven exploration...

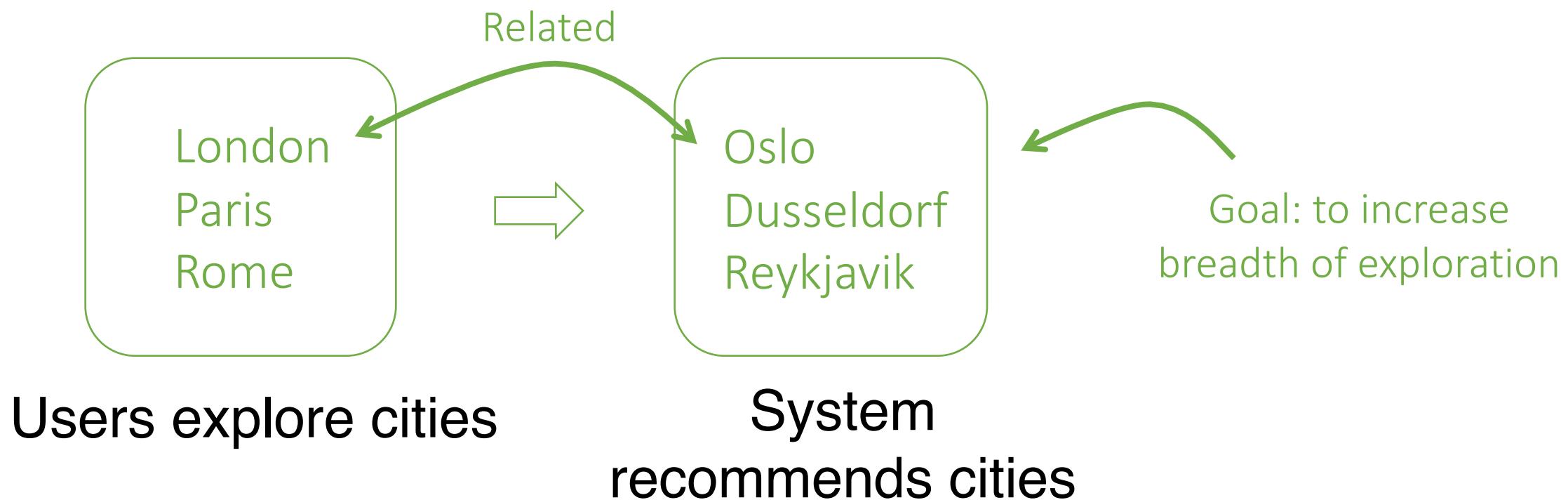
Holiday: Europe → Australia → North America

Work: North America → Asia → Europe

Some steps you should go through in data analysis

Related vs Systematic Exploration

How does it apply to a system for exploring the city dataset? ➔ Related exploration



What is Breadth-Oriented Exploration?

Exploring more = Breadth?

Strategy to **overcome lack of breadth** in exploration due to information overload

by overloading users with more information???

Paradoxical? Or... is it?

Unit of Exploration

What to show



About information selection

User-driven vs System-driven Exploration

Who is responsible to select



About information selection

Related vs Systematic Exploration

Based on what to select



About information selection

What is Breadth-Oriented Exploration?

Exploring more = Breadth?

Strategy to **overcome lack of breadth** in exploration due to information overload by **selecting** potentially useful information to show

by overloading users with more information???

Sounds Challenging...

How to select to ensure relevance of information
and prevent information overload?

Po-Ming Law
pmlaw@gatech.edu

Rahul C. Basole



Designing Breadth-Oriented Data Exploration for Mitigating Cognitive Biases