國立臺灣師範大學
National Taiwan Normal University

Interactive Visual Analytics System for Paleoclimate Causality

Chang, Jung-Chen

Advisor：Wang, Ko-Chih, Ph.D.

VIDALab
NTNUCSIE

Outline

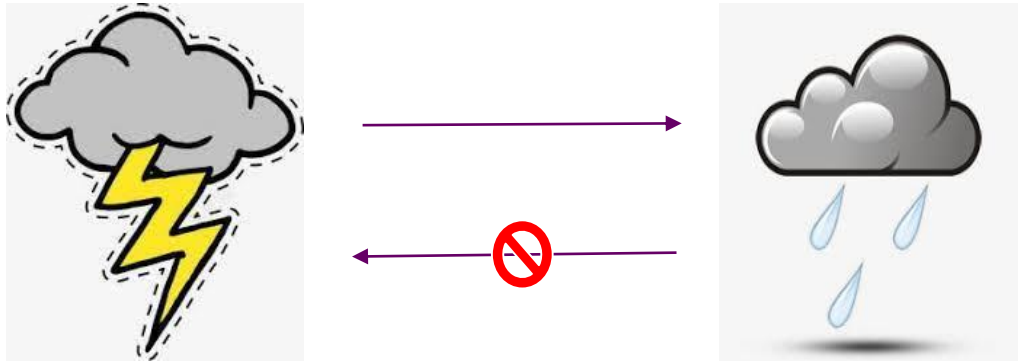
- Introduction
  - Ancient Climate
  - Challenge
- Task and Goal
- Method
  - Dataset
  - Data Preprocessing
  - Association Rule Learning
- Visual Design
- Use Case
  - Discovery of Unexpected Association Rule
  - Spatial Diffusion of Association Rule
- Conclusion and Future Work

Introduction

- Climate is a long-term regional weather pattern influenced by many interacting factors.
- Analyzing ancient climates is an important study.
- Some relations are directional.

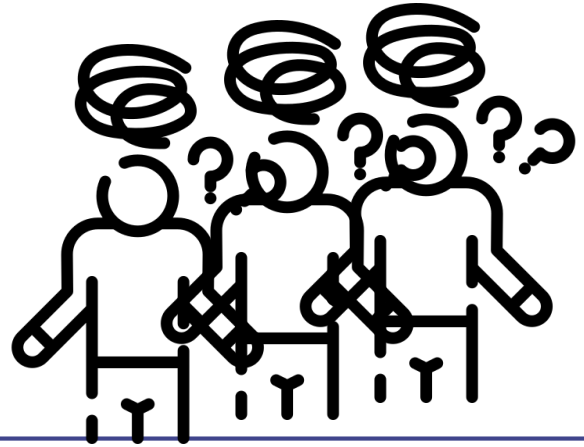
researcher

analyze



Challenge

- No complete and systematic measurement of climate data for the past.
- The search space is so large that it is easy to miss opportunities to discover "surprising" relations.
- Find "asymmetric" relations between events.



Goal and Task

Goals

- G1: Evaluate the strength of relations of events in both event directions.
- G2: Observe the dynamic relation spatial pattern movement over time.
- G3: Explore valuable and special event relations in the large spatiotemporal dataset.

Tasks

- T1: An algorithm to mining relations of events in both event directions.
- T2: Explore spatial regions and time spans with similar or dissimilar association rule sets.
- T3: Compare the detailed difference between association rule sets.
- T4: Visualization to facilitate relation pattern change in the spatiotemporal domain.
- T5: Help experts formulate hypotheses and verify them in a dataset with large spatial and temporal domains.

Dataset - REACHES

REACHES (Reconstructed East Asian Climate History Encoding System)

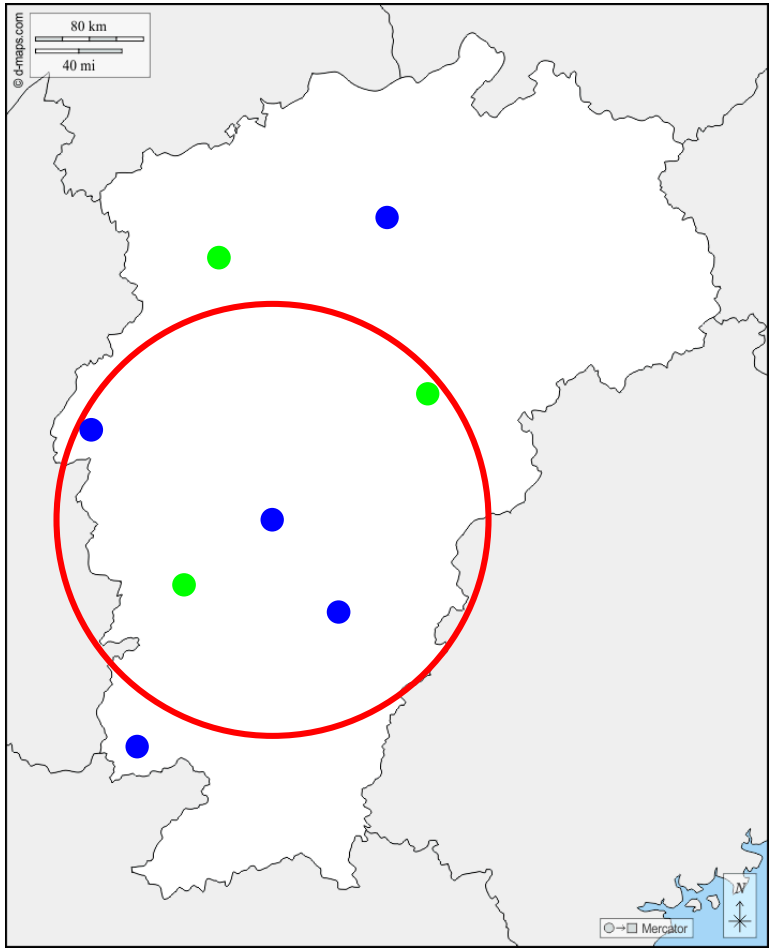
- time range : 1368 - 1911, spatial scope : whole China
- event category : climate event & disaster event

Attributes name	Description
year_lunar_st	Start year of the subrecord/event(s) in lunar calendar
place_provin	Provinces
place_longit	Assigned longitude represented by the location of the city hall
place_latitu	Assigned latitude represented by the location of the city hall
event_code	Event attribution of the record

Event category
Drought
Famine
Flood
Rainfall
Thunder
Wind
Locust
Snow
Uncertain_crop
Crop
Grain
Pestilence
Disaster management

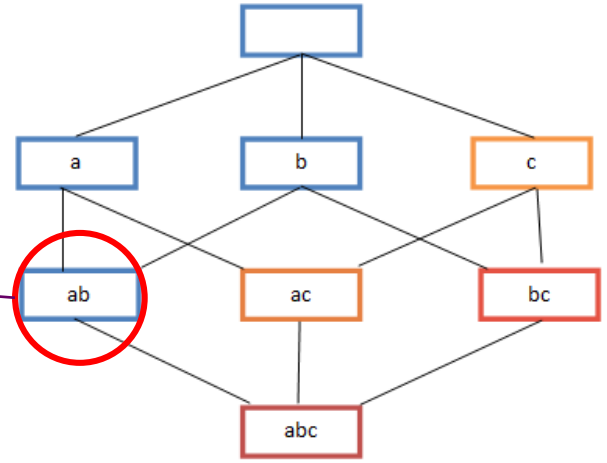
Data Preprocessing

- Frequent pattern mining
  - Usually use on transaction data
  - Only need frequency
  - Calculated relation results are directional
- Transform data into transaction-like pattern
- After discussing with experts, set distance to 100km and time to 1 year



Association Rule Learning

- Find combinations that frequently co-occur in a dataset.
- Rule is defined as an implication of the form : $A \rightarrow B$
- Three standards we use :
  - Support : The frequency or occurrence of an itemset in the data.
    Set the **minimum support** to filter.
  - Confidence : How often the B-event appears in data containing the A-event.
    Set the **minimum Confidence** to filter.
  - Lift : The strength of association between the A-event and B-event of an association rule.



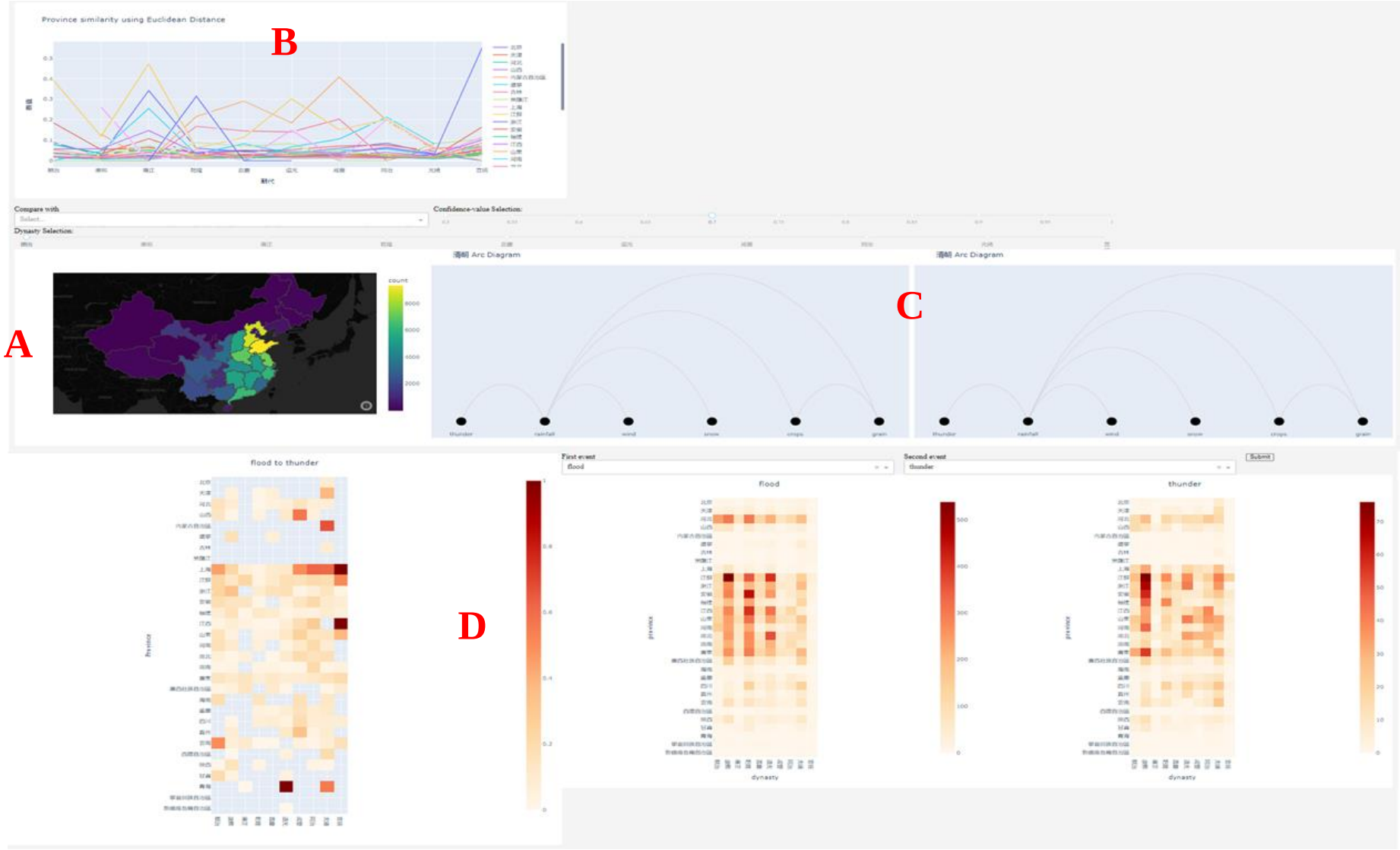
Association Rule Learning

Record ID	Drought	Famine	Rainfall	Snow	Wind	Thunder	Pestilence
1	1	1	0	0	0	0	1
2	0	0	1	0	0	1	1
3	0	0	0	1	1	0	0
4	1	1	1	0	0	1	1
5	0	1	0	0	0	0	0

Item	Support
Drought	2/5 = 0.4
Famine	3/5 = 0.6
Snow	1/5 = 0.2
Thunder	2/5 = 0.4
Pestilence	3/5 = 0.6

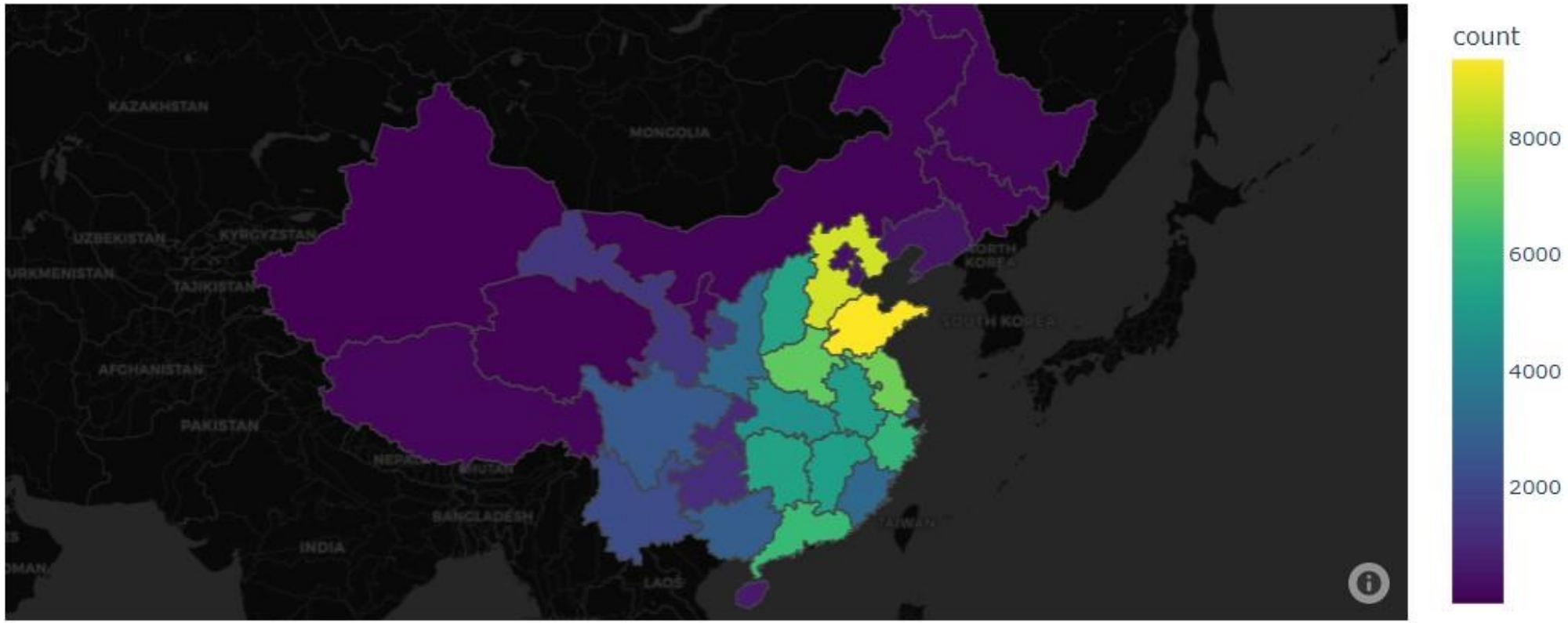
If occur A, then occur B(A → B)	Support	Confidence
(Drought → Famine)	2/5 = 0.4	2/2 = 1
(Drought → Thunder)	1/5 = 0.2	1/2 = 0.5
(Famine → Pestilence)	2/5 = 0.4	2/3 = 0.66
(Pestilence → Thunder)	2/5 = 0.4	2/3 = 0.66

Visual Design – Overview



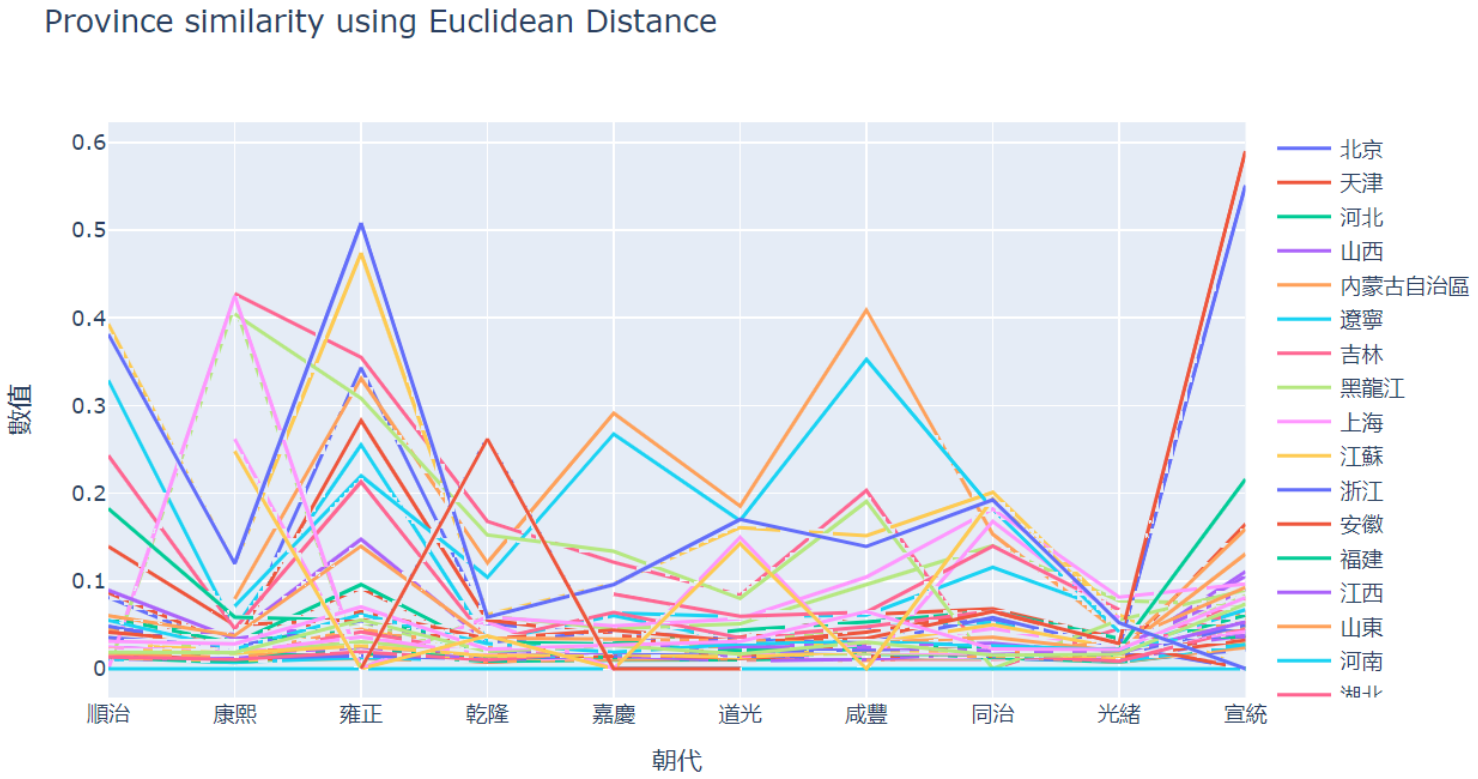
Visual Design – Map

- Easier to understand the spatial variation of association rules.



Visual Design – Line Chart

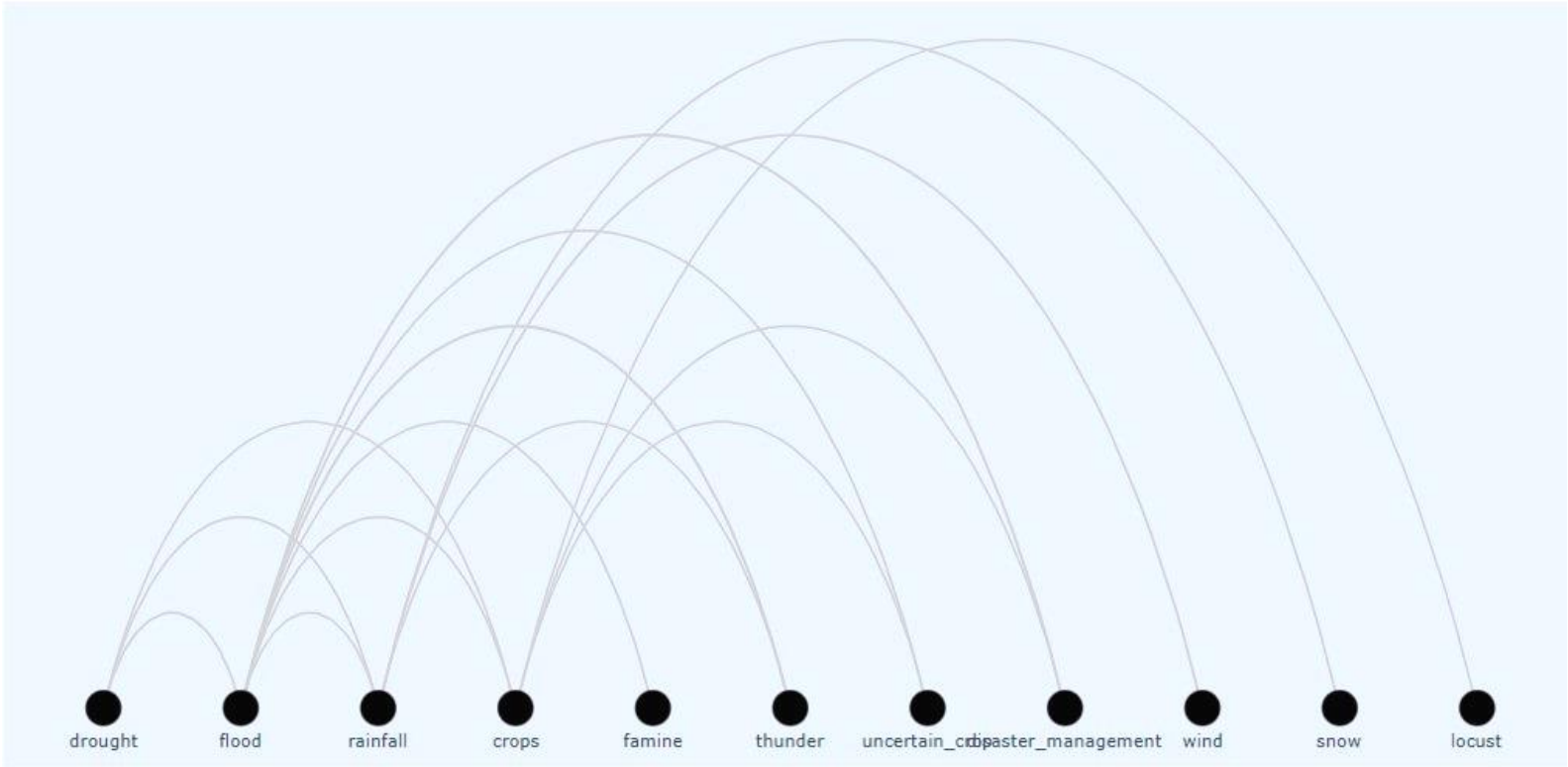
- Compare the association rule set between the base province and other provinces under different emperors.



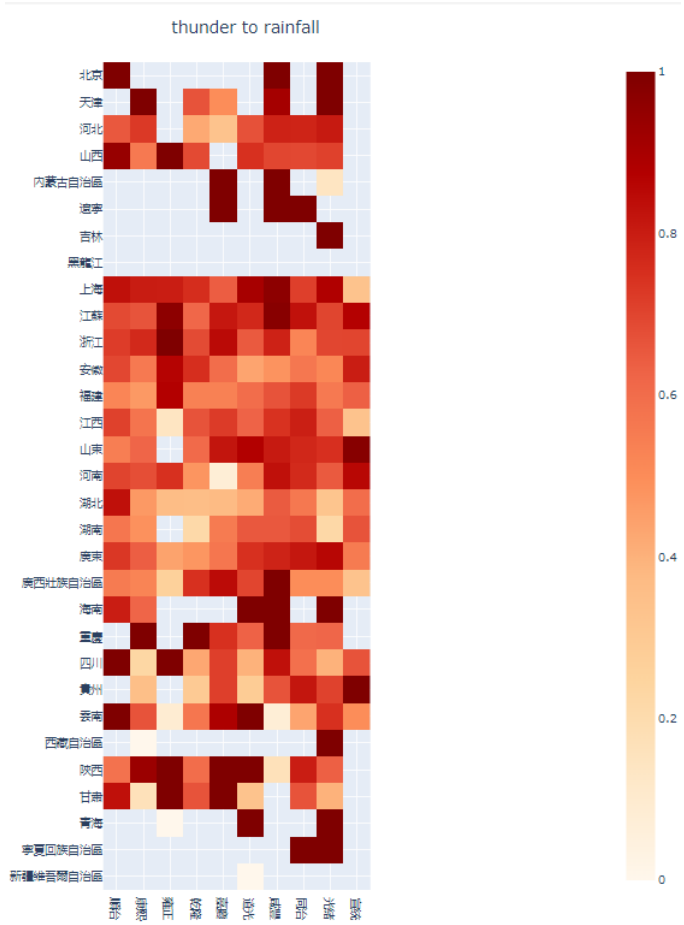
All provinces compared with Qing

	antecedents	consequents	confidence_x	confidence_y
0	(disaster_management)	(flood)	0.555855	0.590909
1	(thunder)	(rainfall)	0.628458	0.678571
2	(wind)	(rainfall)	0.609467	0.566667
3	(crops)	(rainfall)	0.535201	0.551471

Visual Design – Association Rule Graph



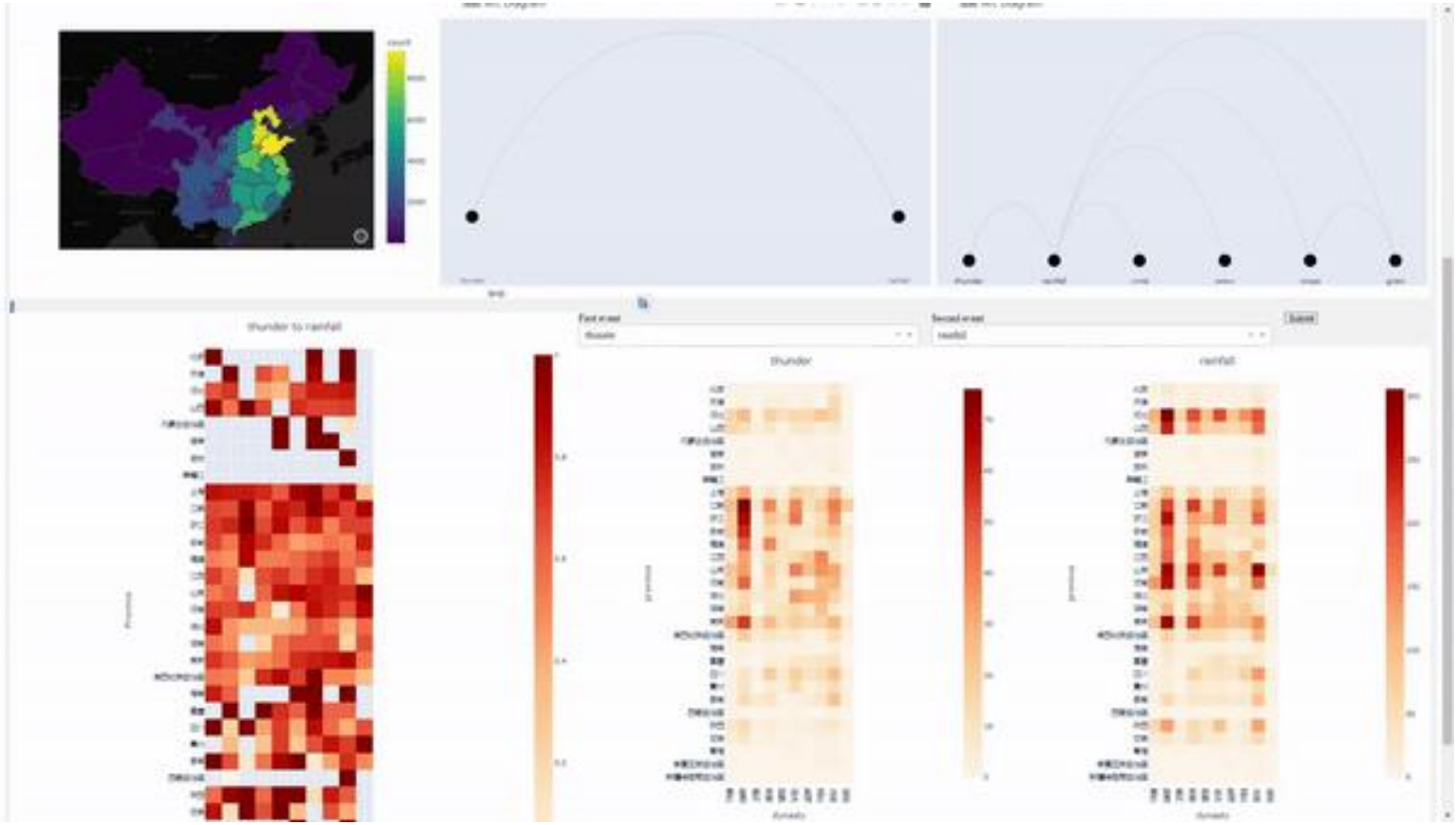
Visual Design – Heatmap

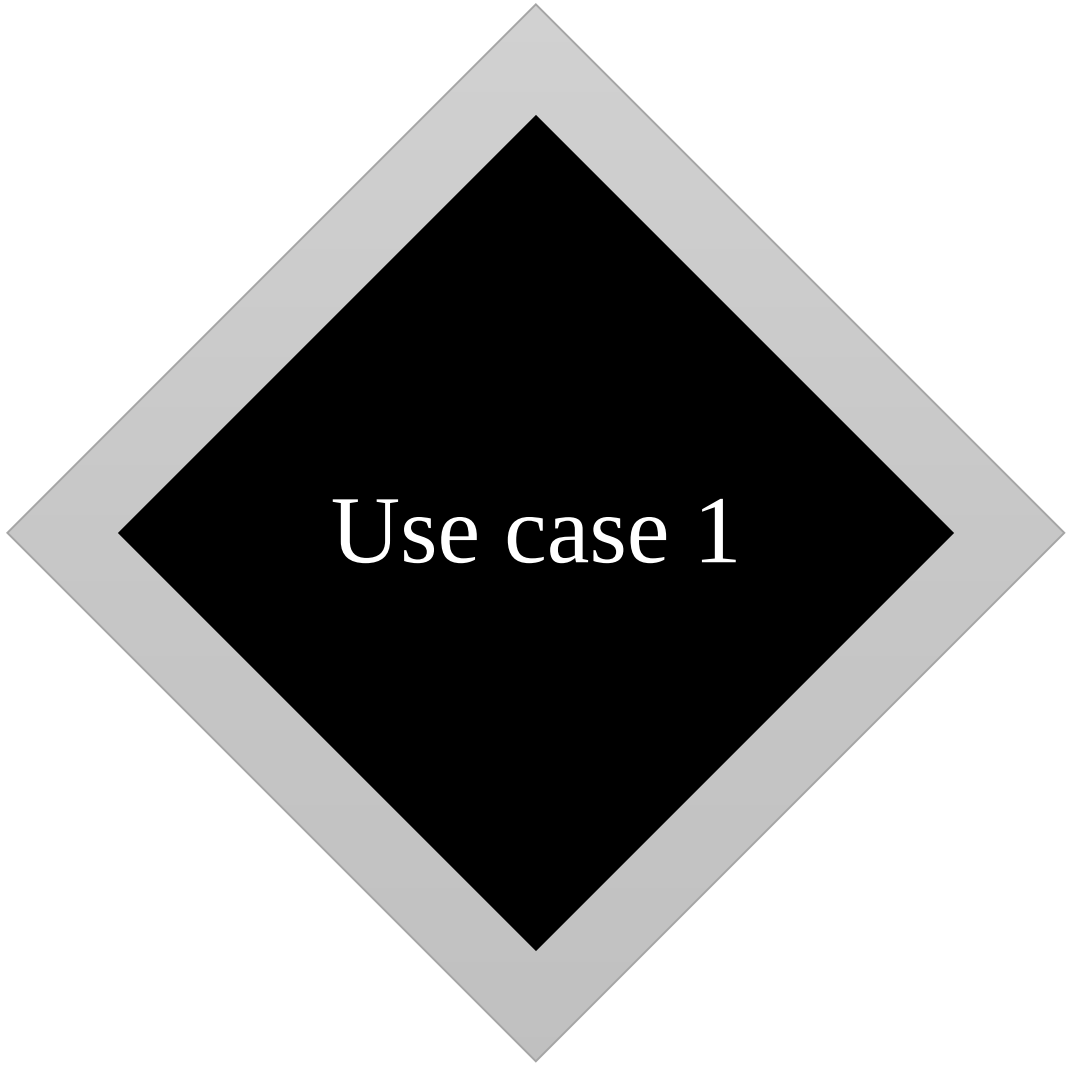


Association rule heatmap

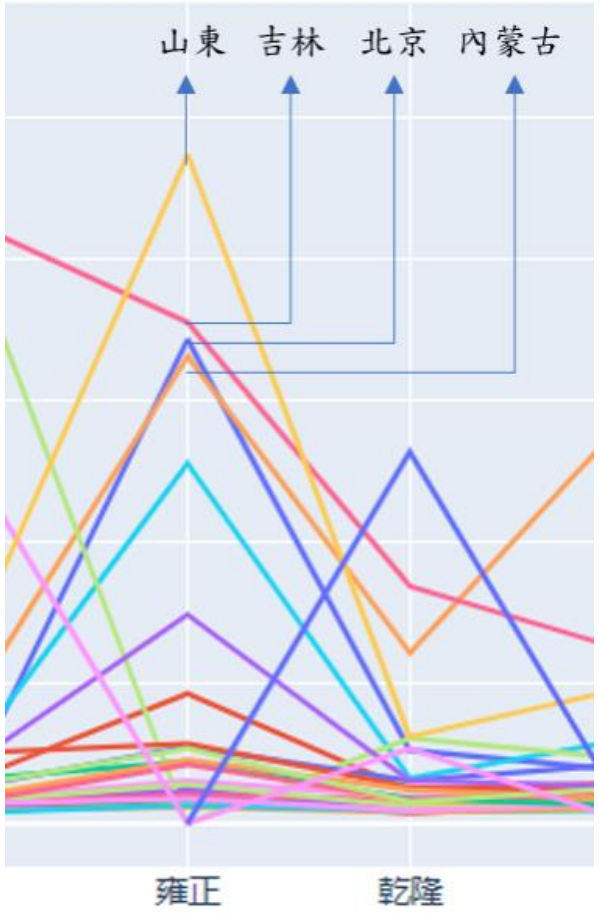
Event frequency heatmap

Visual Design – Interaction

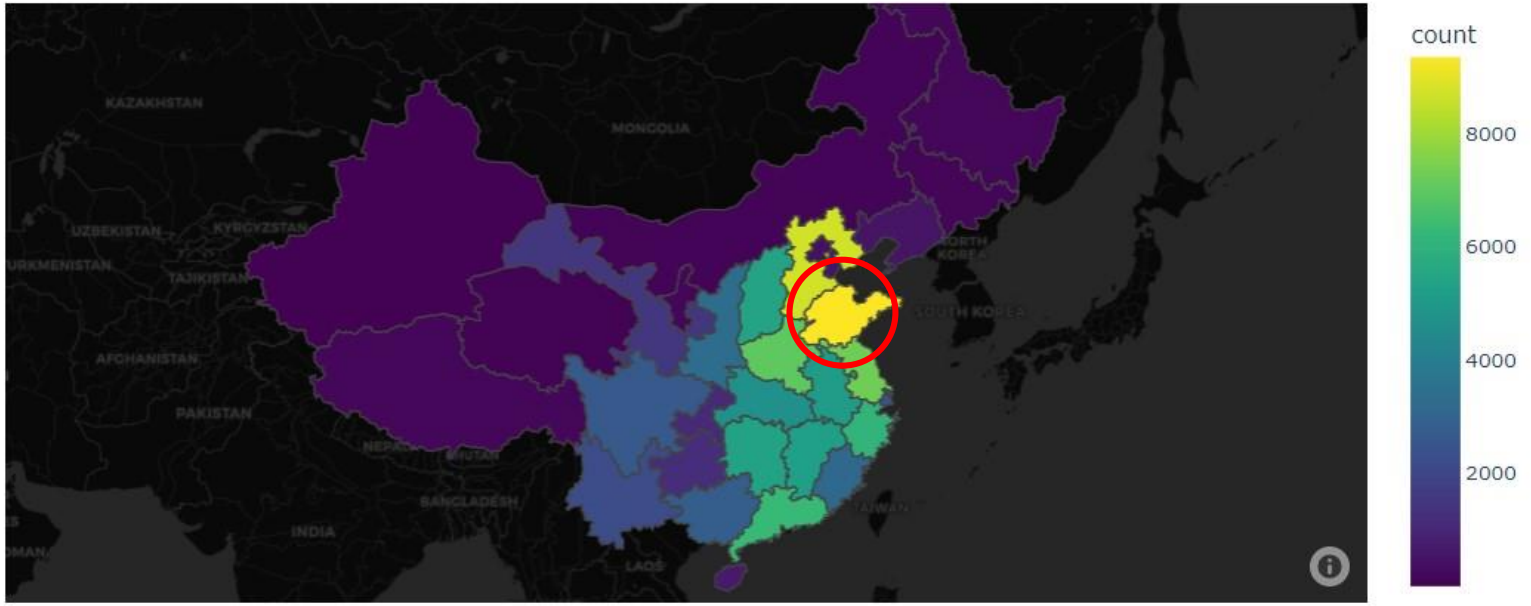



Use case 1

Use Case 1: Discovery of Unexpected Association Rule



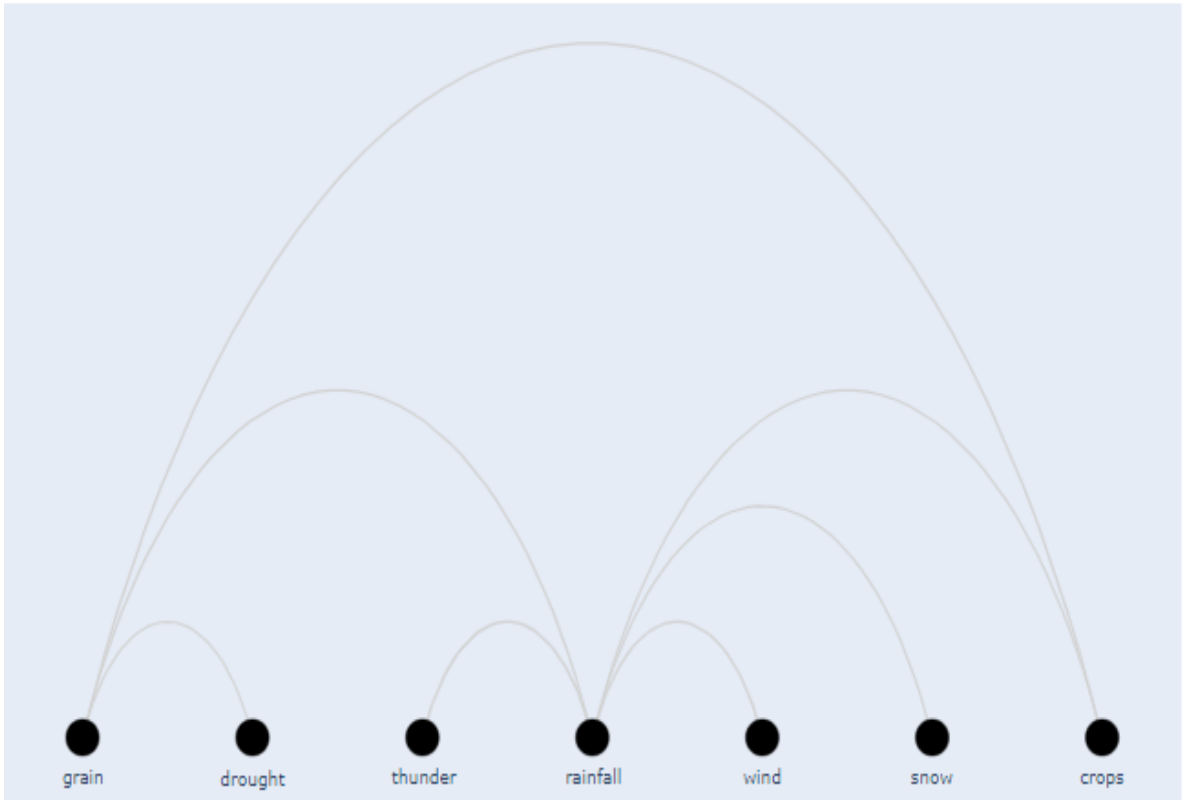
- Some provinces were different from the whole Qing dynasty.
- Check the map and find “山東” special.



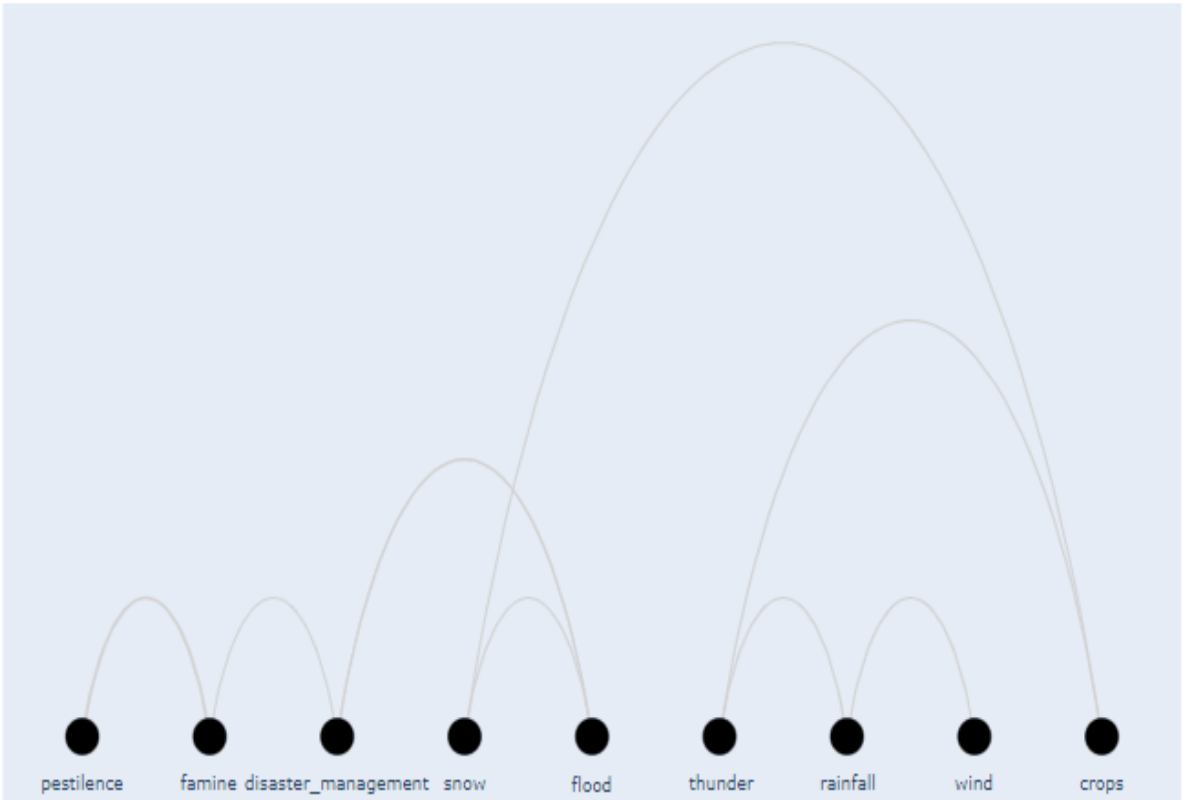
Use Case 1: Discovery of Unexpected Association Rule

Set confidence-value = 0.5, and observe arc diagram graph

清朝 Arc Diagram

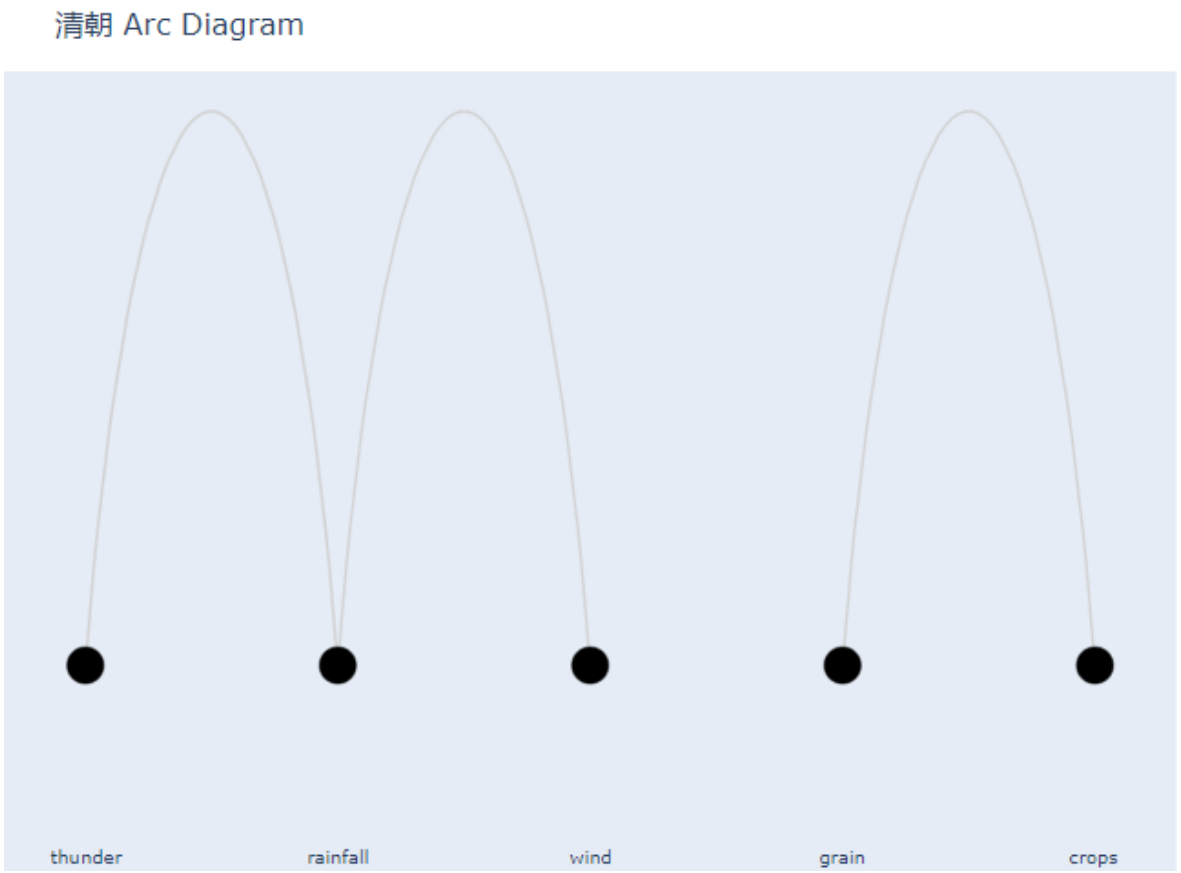


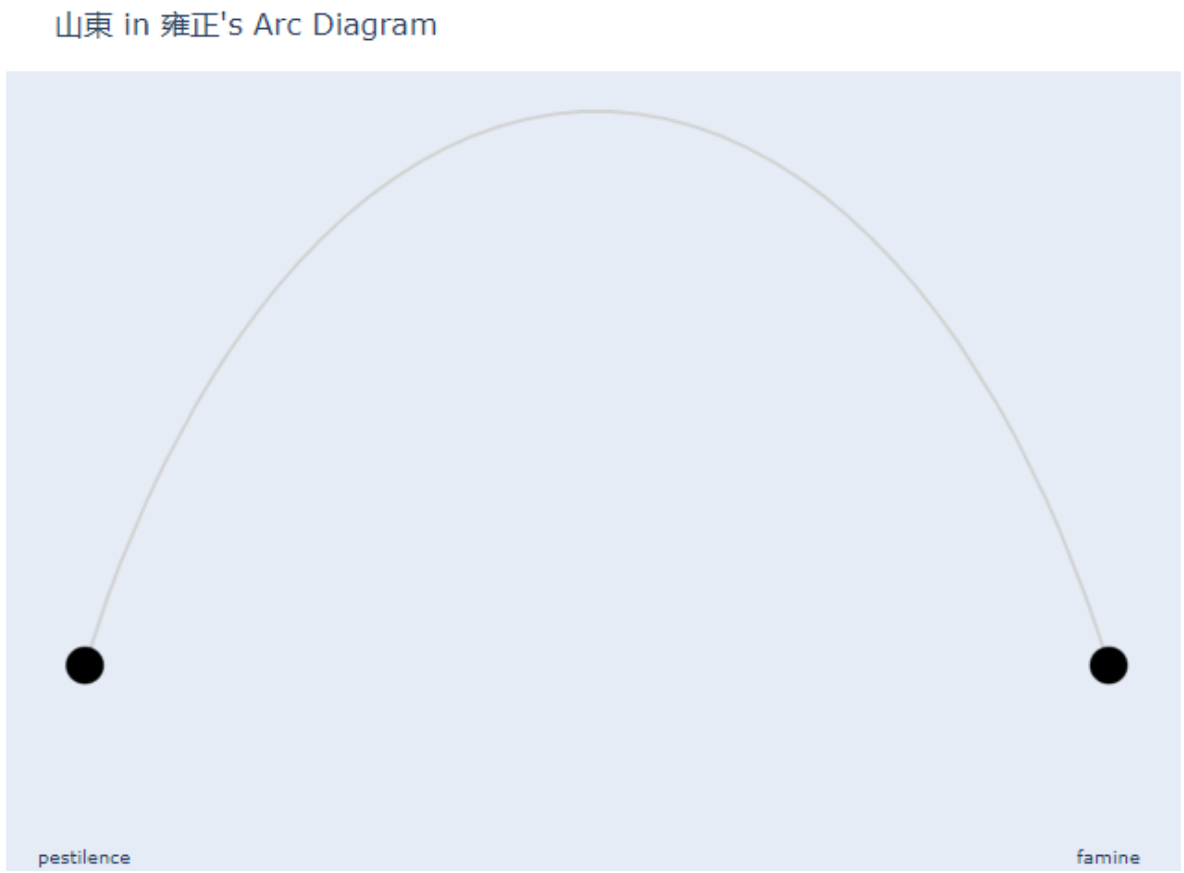
山東 in 雍正's Arc Diagram



Use Case 1: Discovery of Unexpected Association Rule

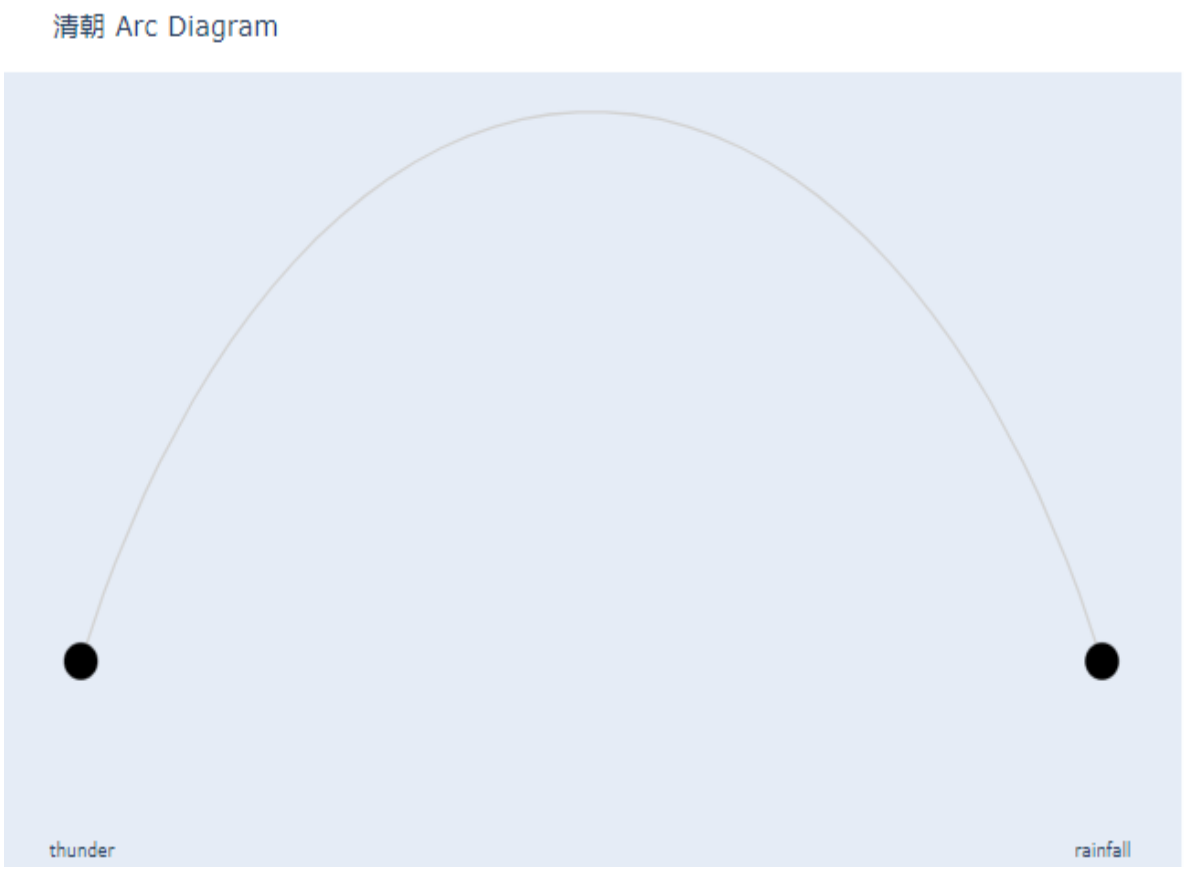
Confidence-value = 0.65

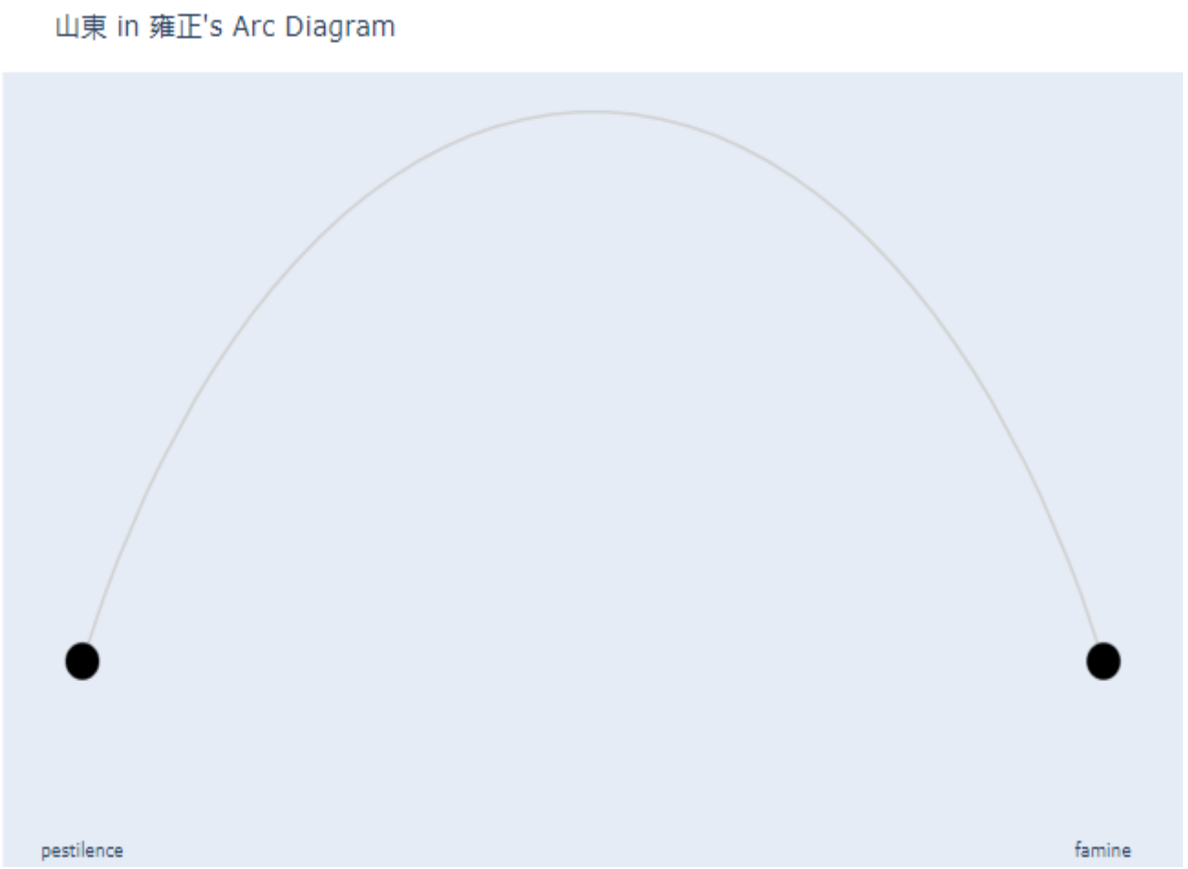




Use Case 1: Discovery of Unexpected Association Rule

Confidence-value = 0.7

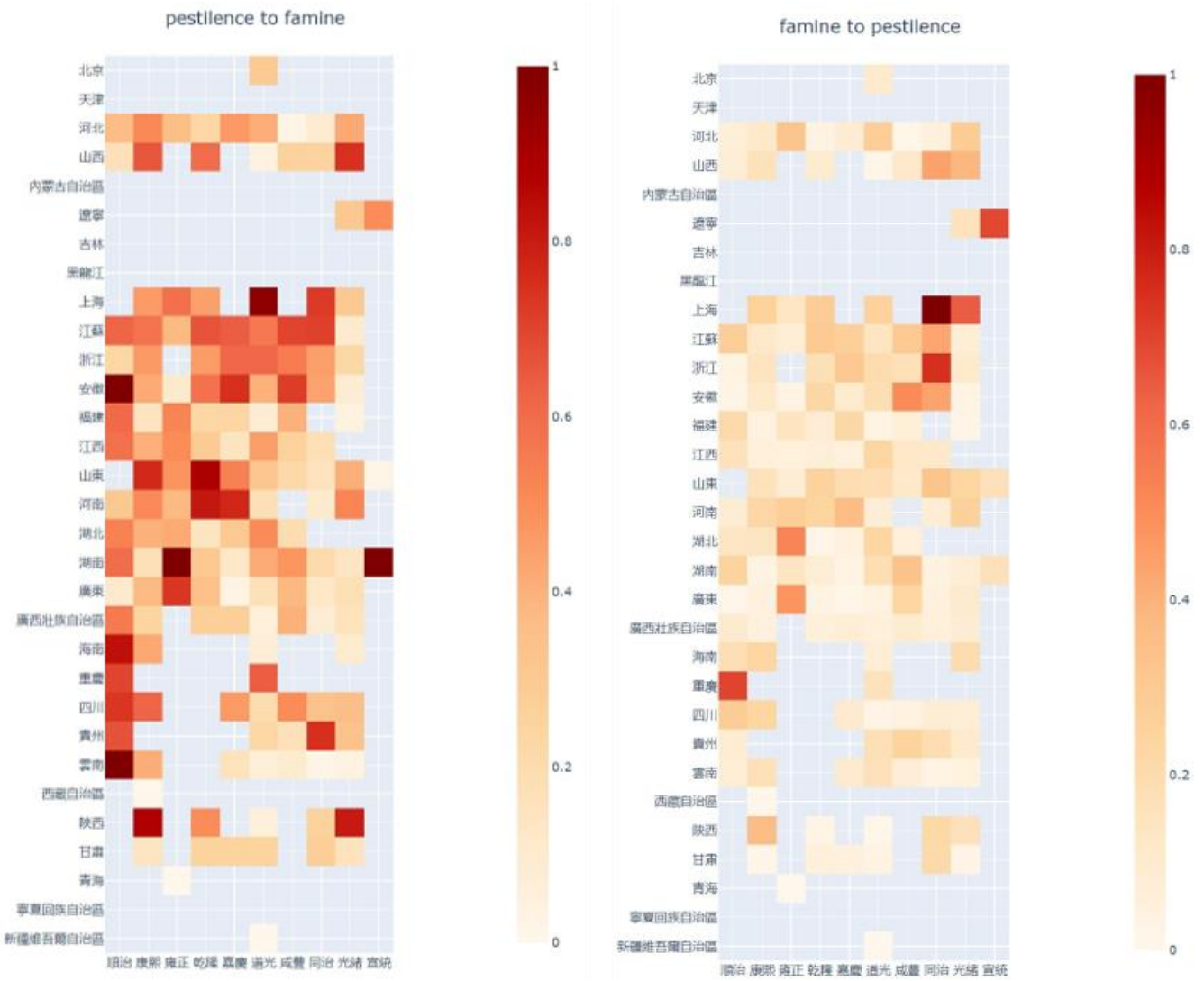




Use Case 1: Discovery of Unexpected Association Rule

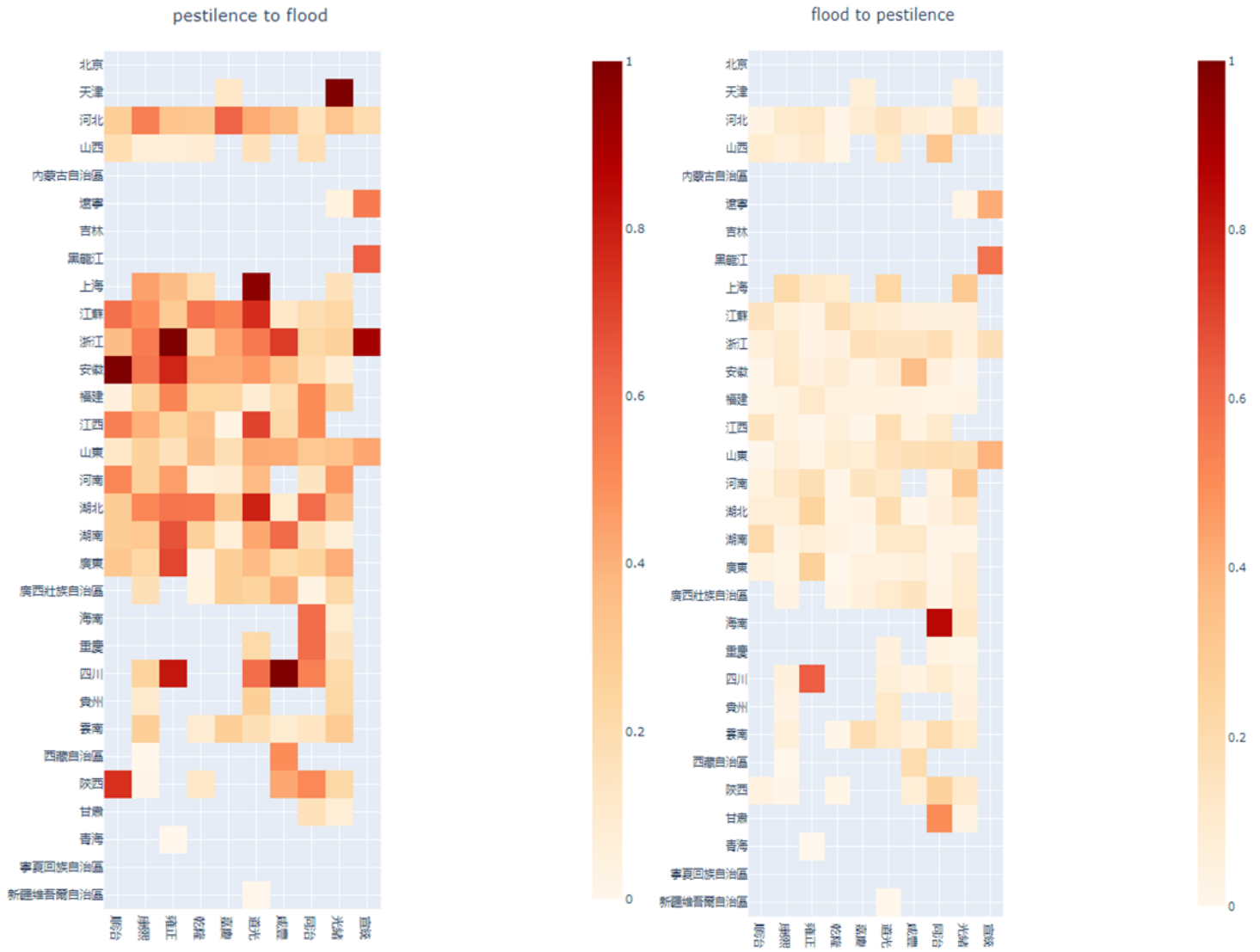
Use heatmap to observe **Pestilence** and **Famine**

- Find that in rule : **Pestilence → Famine** more than 1/3 combination's confidence value are large.
- In rule : **Famine → Pestilence** is opposite.



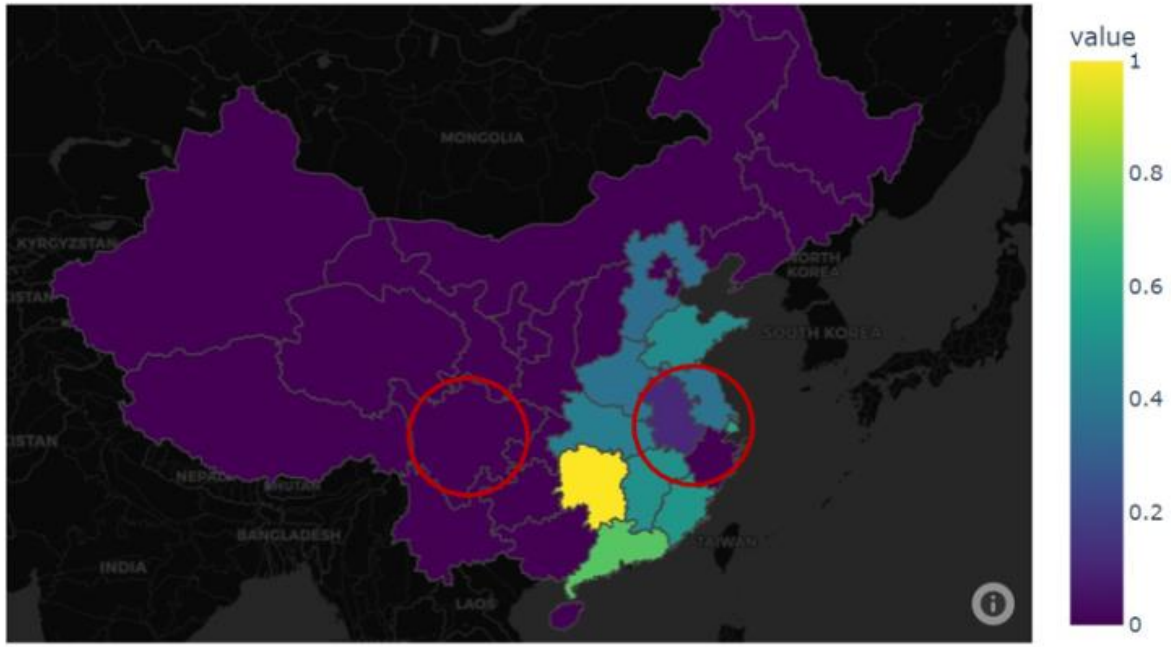
Use Case 1: Discovery of Unexpected Association Rule

- Expert repeats the steps and discovers another association rules : **Pestilence → Flood**

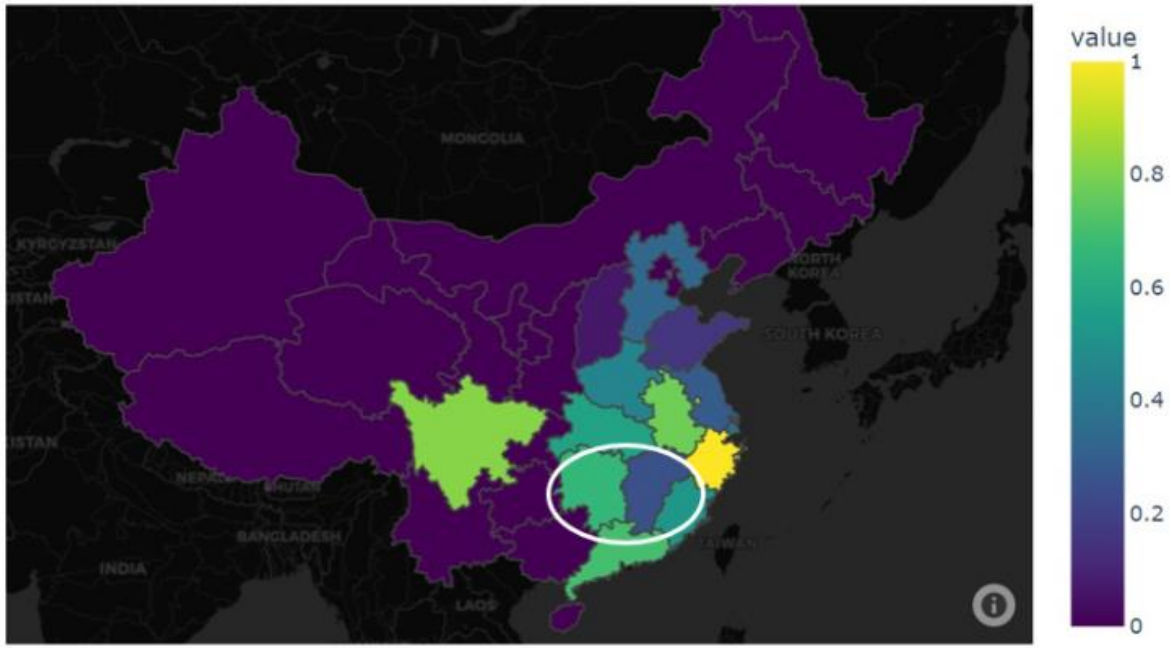


Use Case 1: Discovery of Unexpected Association Rule

- Compare two association rule in same emperor on the map.



Pestilence → Famine



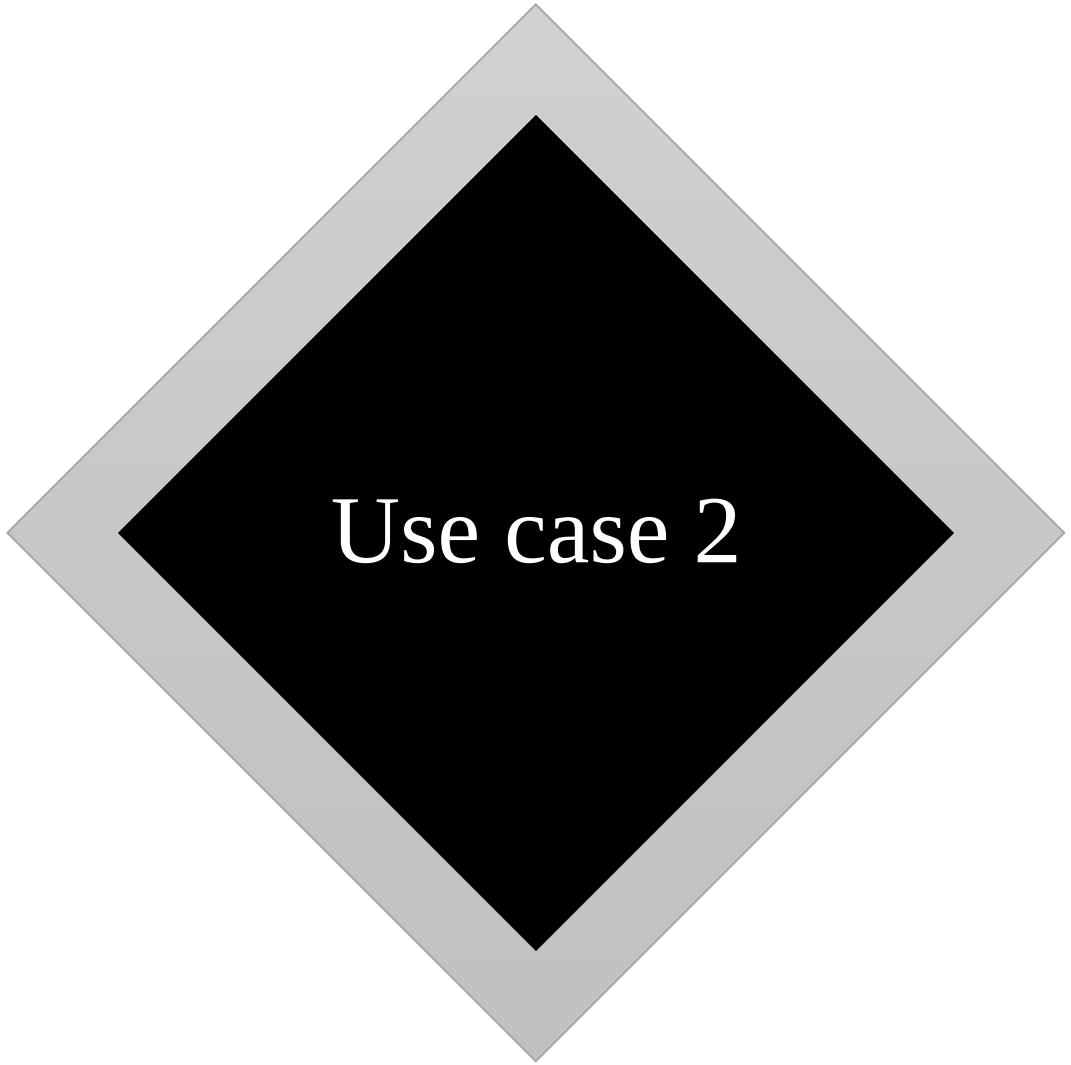
Pestilence → Flood

Use Case 1: Discovery of Unexpected Association Rule

Use case 1 conclusion:

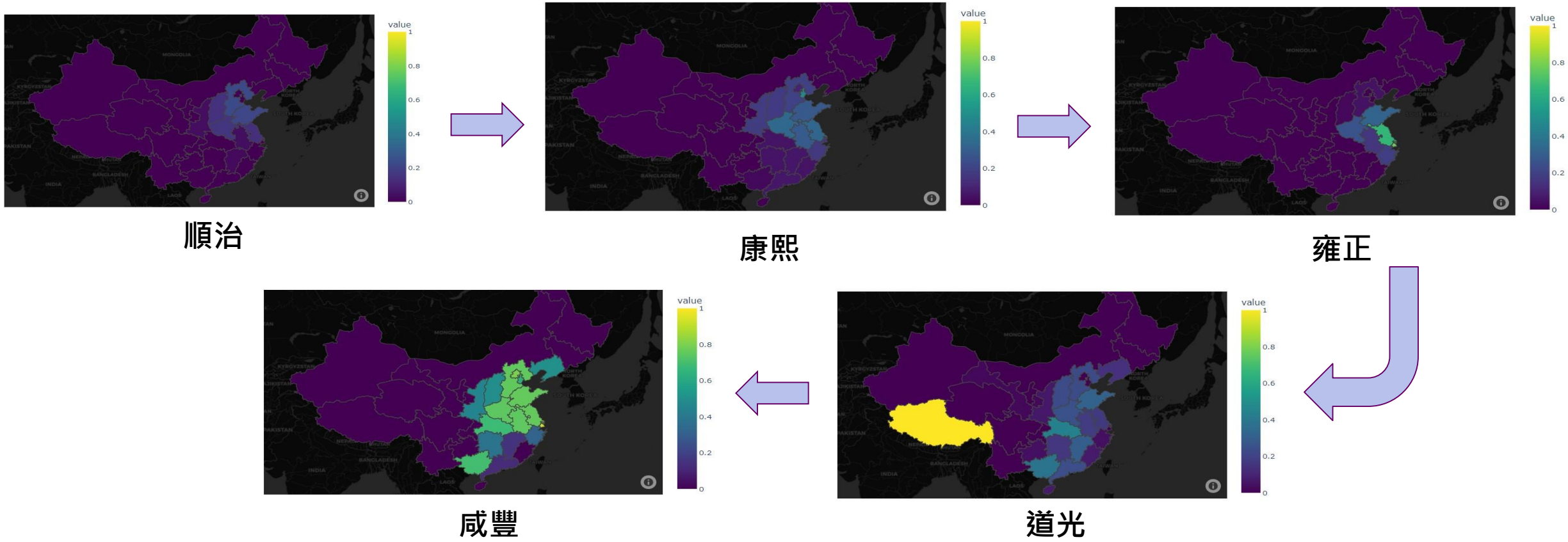
1. Experts find surprising relationships by using our system.

1. Experts find an association rule "Pestilence → Famine" that is directional.


Use case 2

Use Case 2: Spatial Diffusion of Association Rule

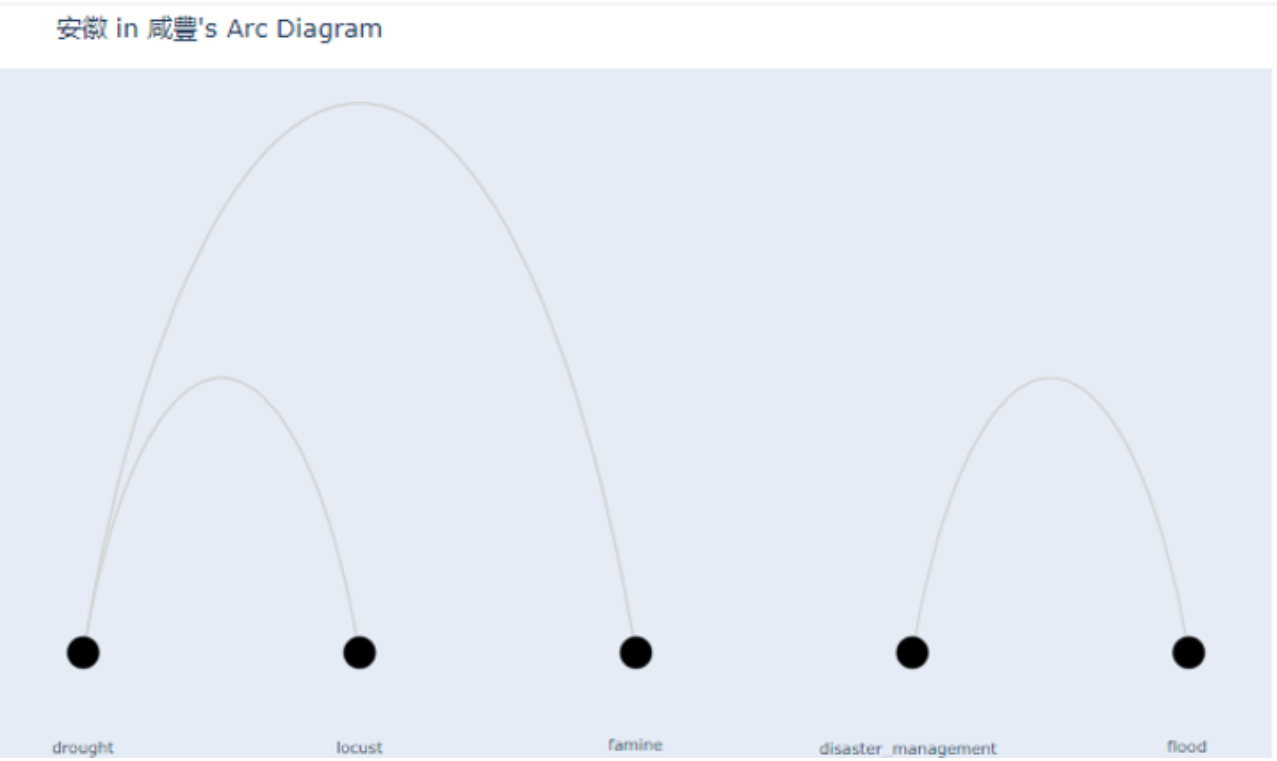
- Expert selected an association rules with diffusion phenomena.

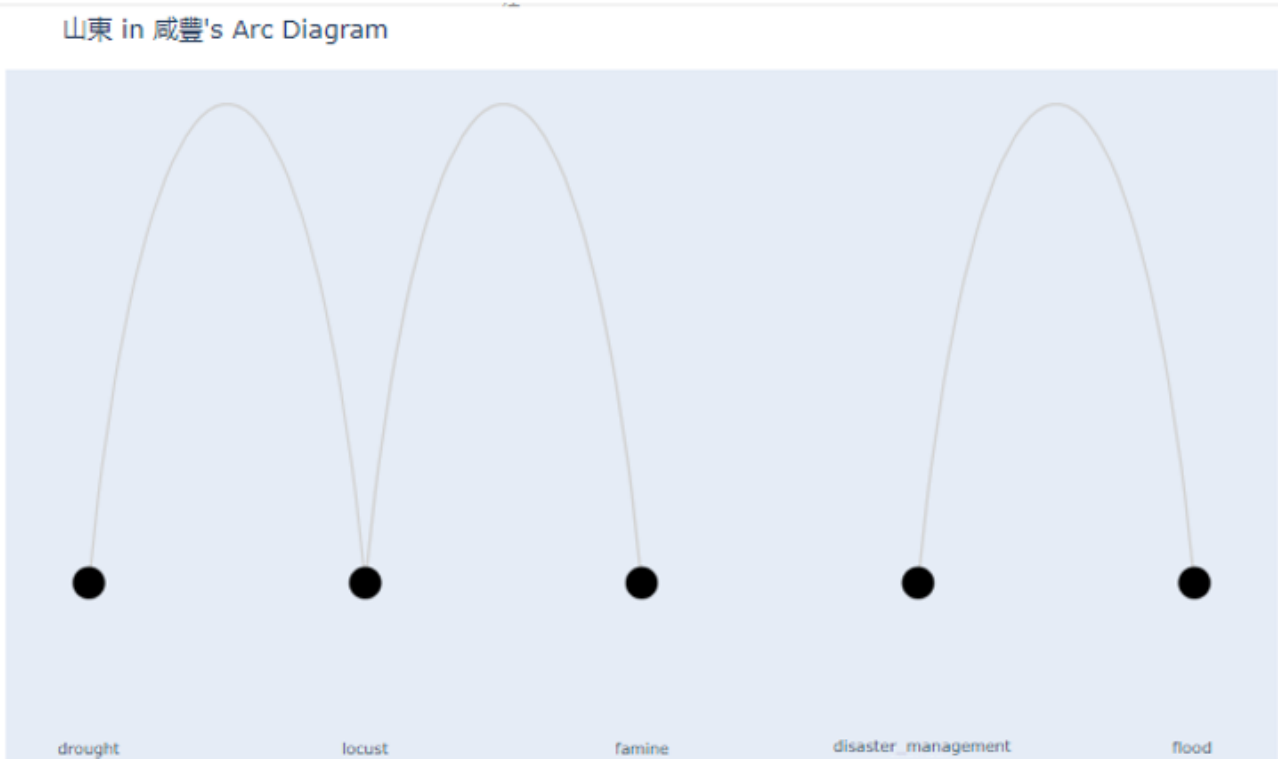


Association rule : Drought → Locust

Use Case 2: Spatial Diffusion of Association Rule

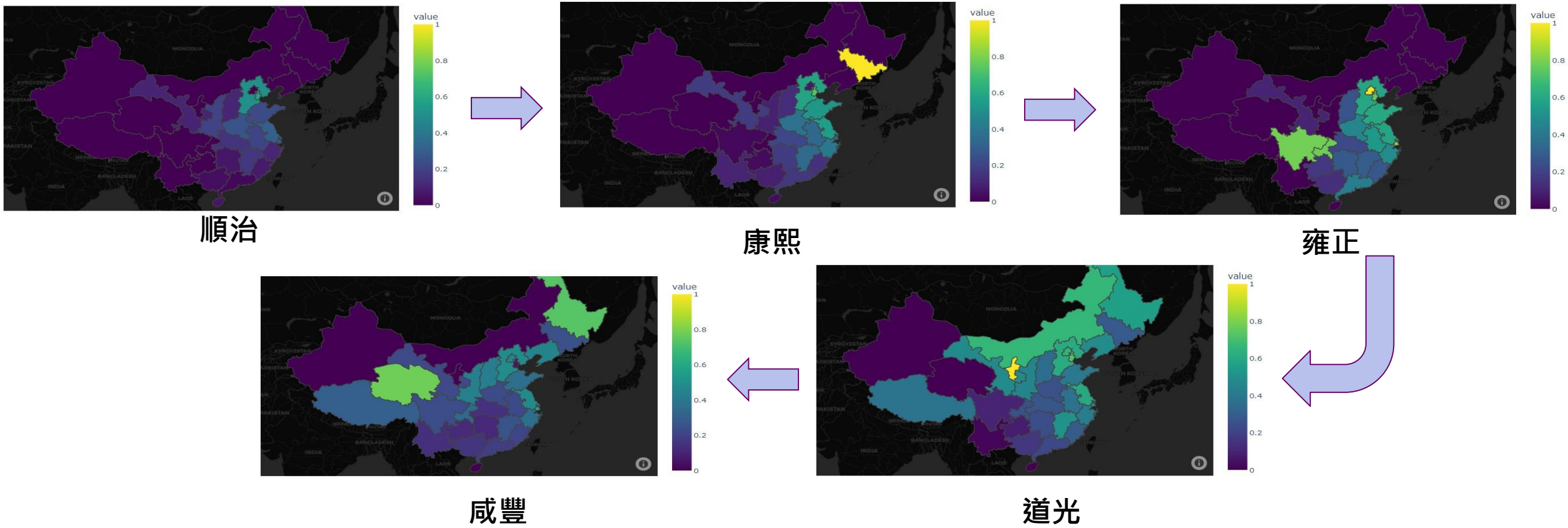
- Experts choose similar provinces to observe and find there are other rules that exist in these two provinces.





Use Case 2: Spatial Diffusion of Association Rule

- Expert selected an association rules with diffusion phenomena.



Association rule : Flood → Disaster management

Use Case 2: Spatial Diffusion of Association Rule

Use case 2 conclusion:

1. Experts find an association rule "Flood → Disaster management" has a similar spatial diffusion with the association rule "Drought → Locust".

1. Experts can decide which association rules they are interested to study.

Conclusion and Future Work

- An association rule learning based algorithm for exploring the unexpected relation.
- An interactive visualization system for analyzing large-scale spatiotemporal climate dataset.
- [Future] Identify more than two events' relationships and their interactions.
- [Future] Organize and consider different spatial and temporal ancient climate datasets.

Thank you for your attention

Q&A