Mapping Invisible Threads: Unveiling Economic Landscapes through Micro-Level, High-Frequency 19th-Century Postal Data

Mengyue Zhao

University of Oxford, Class of 2019

beczhaozmy0@gmail.com

The Official Register of the United States (OR)

- ◈ 12 Volumes of Official Register

- ◈ 191,891 Records of Postmaster Compensation

- ◈ All 47 states

- ◈ Covers 1849 – 1871

- ◈ Biannual

- ◈ Linked to Postal Revenue through a national formula

- ◈ Superior to Alternative Sources:
  - ◈ Census Data (decadal frequency)
  - ◈ Tax Records (inconsistencies across regions)
  - ◈ Local Newspapers (no clear quantitative metric)
  - ◈ …

Table 3.1 A Typical Row from the Official Register

Name of the Post Office	County	State	Postmaster Name	Annual Postmaster Compensation
Clingman	Wikes	North Carolina	J.C. Greene	100.54

Agenda

◈ Context

◈ Machine Vision Techniques

◈ Sensitivity analysis

◈ Hotspot Analysis

From Postmaster Compensation to Town-Level Postal Volume

"For the year 1853, a postmaster takes 50 per cent on an amount not exceeding 100 dollars, 40 per cent on an amount over 100 dollars and not exceeding 400 dollars, 35 percent on an amount over 400 dollars but not exceeding 2400 dollars and 15% on an amount over 2400 dollars."

-- The Record of the Thirty-Second Congress (Act of March 1)

Postmaster Compensation
$$= a\%$$
$$\times \text{Annual Postal Revenue}$$

where a = 50 when Annual Postal Revenue <= 100,

a = 40 when 100 < Annual Postal Revenue <=400,

a = 35 where 400 < Annual Postal Revenue <=2400,

a = 15 where Annual Postal Revenue > 2400

Postmaster Compensation

Postal Volume

Local Economic Activities
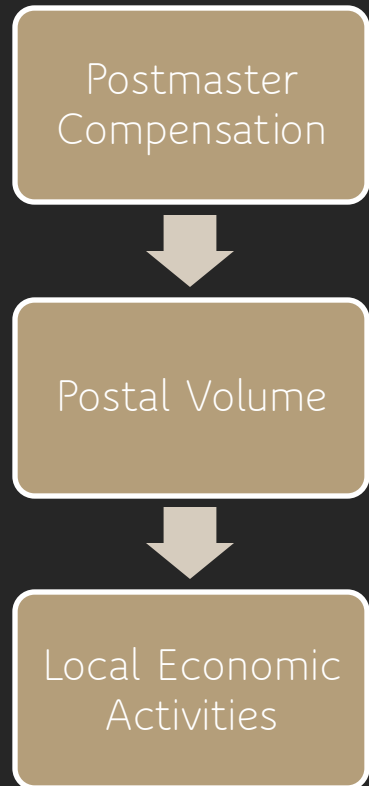
The US Federal Government maintained monopoly status over letter and packet delivery up to the year 1882.

◈ **Primary Revenue Sources**: The main revenue sources were "letter postage, including foreign postage, and stamps sold" and "newspaper and pamphlet postage." The remaining revenue was from non-postal activities. (96% in 1850 and 92% in 1871)

◈ **Monopoly statutes** prohibited private express delivery of letters and packets but allowed delivery of packages containing non-letter items (Johnston 1968, p. 389).

◈ **Package Delivery Exclusion**: The U.S. postal system did not deliver packages before 1913, when Parcel Post was introduced (Kielbowicz 1994).

Methodology

Clean and Link the Datasets Using Matching Algorithms

Link the OR dataset to 19 Century Railroad and Canal Shapefiles

Explore the Sensitivity of the OR Dataset through correlation

Delineate the Extent of Markets with Hotspot Analysis

Page Segmentation Result – My Workflow

Page Segmentation Result -- Omnipage

Page Segmentation Result – Adobe Acrobat

Page Segmentation Result – Tesseract

OCR Output - Customized Tool v.s. Adobe Acrobat

Table 1. OCR Output – Customized Tool

	Post Office	County	Postmaster	Compensation
1	Post Office	County	Postmaster	Compensation
2	Blake's Ferry	Randolph	J J Bradley	$17 78
3	Blocker's	Tuscaloosa	William G Blocker	
4	Blount Spring	Blount	B. H Sapp	
5	Blountsville	Blount	John L. Hopkins	66 38
6			H. A Gillespie	33 81
7	Blowing Spring	Limestone	J Martin	
8	Blue Lick	Franklin	A T Bonds	
9	Blue Mountain.	Calhoun	J W Dendy	
10	Blue Pond	Cherokee	A Poore	20 20
11	Blue Spring	Morgan.	C Ratliff	65
12	Bluffport	Sumter	W H Sanders	

Table 3. OCR output – Adobe Acrobat

P ost office. County. P o~tmaJ1te r.
B!ak 1>]y ... . .• . • . . Bald win .. .. J. II . Stanmyre~
Blake'~ F erry. . Randolph . .. . J. J. Bradley
B locker'~ . . .. Tu6c11!008u .. , Yilliu111 G. B lock<;ir ..
B loun t Spring. Blount. B. H. Supp ....
Bloumsv ill!l --·· · ·-·- - -· Blount .. J ohu L. IIopk ius ..
Blon·in~ S pring ...
B lue Lick . .•. ....
B lue Mooutnin ..
Blue Pond . . . . . .
Blul' Spring . .
Blulfpor t ....
Bluff Spring

Machine Learning Iteration Process

Rebuilding Trust in Historical Records: Leveraging Deep Learning and Generative AI for Enhanced Data Accuracy and Comprehensive Dataset Collection
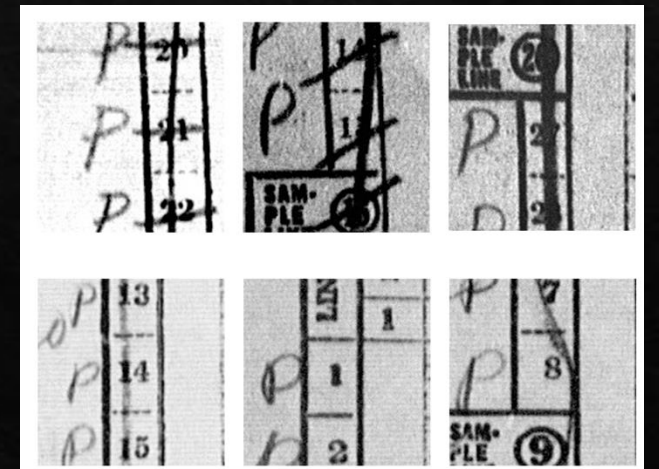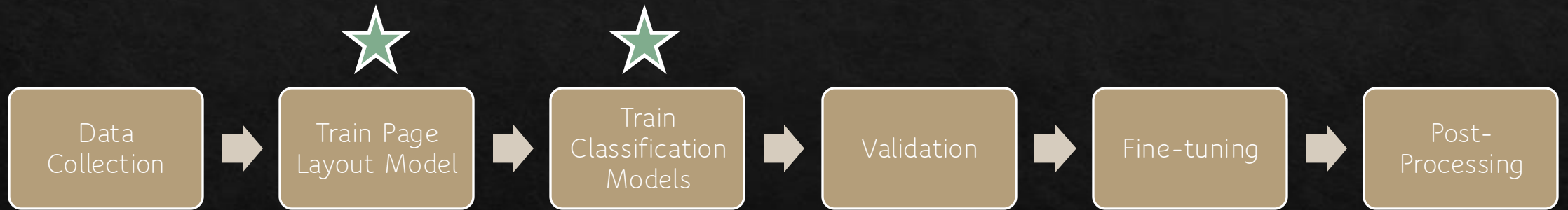
Data View

Table 3.1 A Typical Row from the Official Register

Name of the Post Office	County	State	First and Last Name of the Postmaster	Annual Postmaster Compensation
Clingman	Wikes	North Carolina	J.C. Greene	100.54

Table 3.2 County Level Data

County	State	Postal Revenue	Total Manufacturing Output	Total Farming Output	Population
A	North Carolina	35200.00	356096.00	89703.30	100,100

Table 3.3 Town-Level Data

Name of the Post Office	County	State	Postmaster Name	Postal Revenue	Longitude	Latitude
Clingman	Wikes	North Carolina	J.C. Greene	250	36.1784613	-80.9636877

Correlation Between Postal Revenue and Other Economic Indicators

- **Low R-squared**: Postal revenue reflects a variety of activities beyond manufacturing and farming.
- **Manufacturing and Farming**: Both correlated with postal revenue.
- **Only Manufacturing Significant with Population Control**: Manufacturing shows a distinct, independent impact.

Table 1.1 Regression of Postal Revenue on Manufacturing Output, Farm Output, Population and Total Output Per Capita

	Dependent Variables					
	Revenue					
	(1)	(2)	(3)	(4)	(5)	(6)
Manufacturing Output	0.001*** (0.00004)				0.0002* * (0.0001)	
Farm Output		0.002*** (0.0003)				-0.0003 (0.0003)
Population			0.20*** (0.010)		0.16*** (0.021)	0.204*** (0.011)
Total Output				114.71		

|						
Control For Population	No	No	NO	NO	YES	YES
State Fixed Effect	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES
Observations	70,504	71,741	71,863	70,442	70,502	71,714
R2	0.021	0.018	0.022	0.020	0.021	0.023
Residual Standard Error	89,207.98 (df = 70455)	88,605.83 (df = 71691)	88,370.25 (df = 71813)	89,222.83 (df = 70393)	89,171.66 (df = 70452)	88,408.89 (df = 71690)

*p<0.1; **p<0.05; ***p<0.01

**** Total Output Per Capita = (Manufacturing Output + Farm Output)/ Population
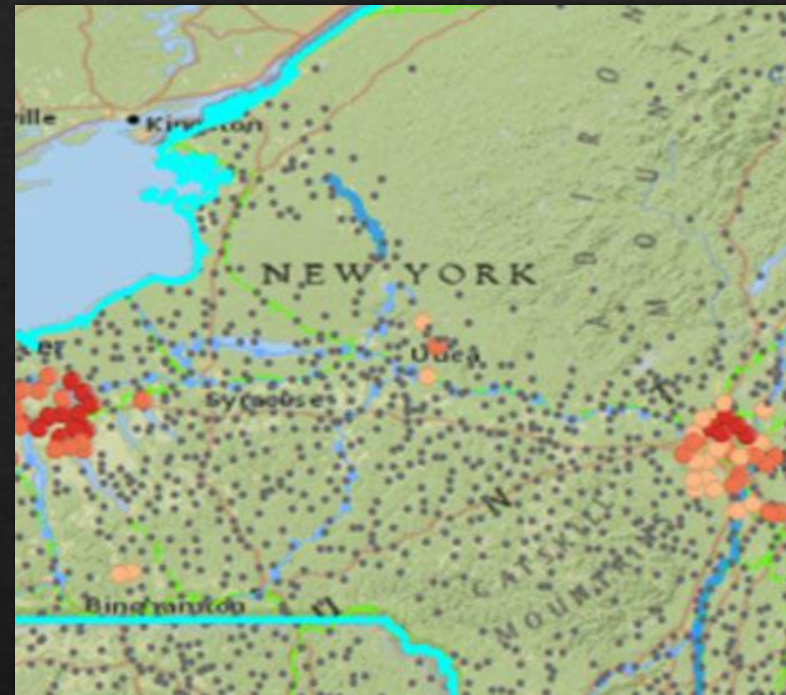
Fixed-Bandwidth Hotspot Analysis

The small neighbors are important.
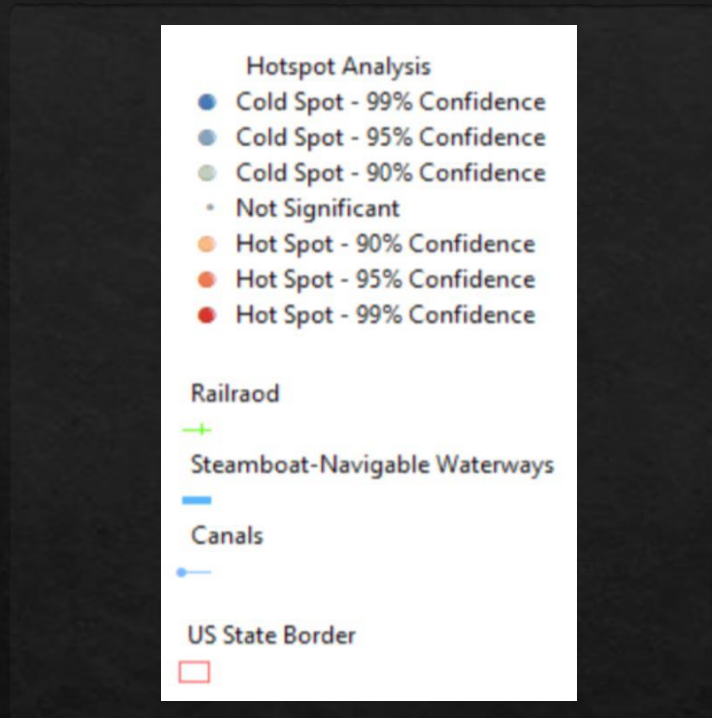
Legend

Chicago (1849 – 1871)

1849

1853

1857

1861

1867

1871

- Illinois and Michigan Canal
- Chicago, Rock Island and Pacific Railroad
- Rock Island Line in 1857
- Joliet
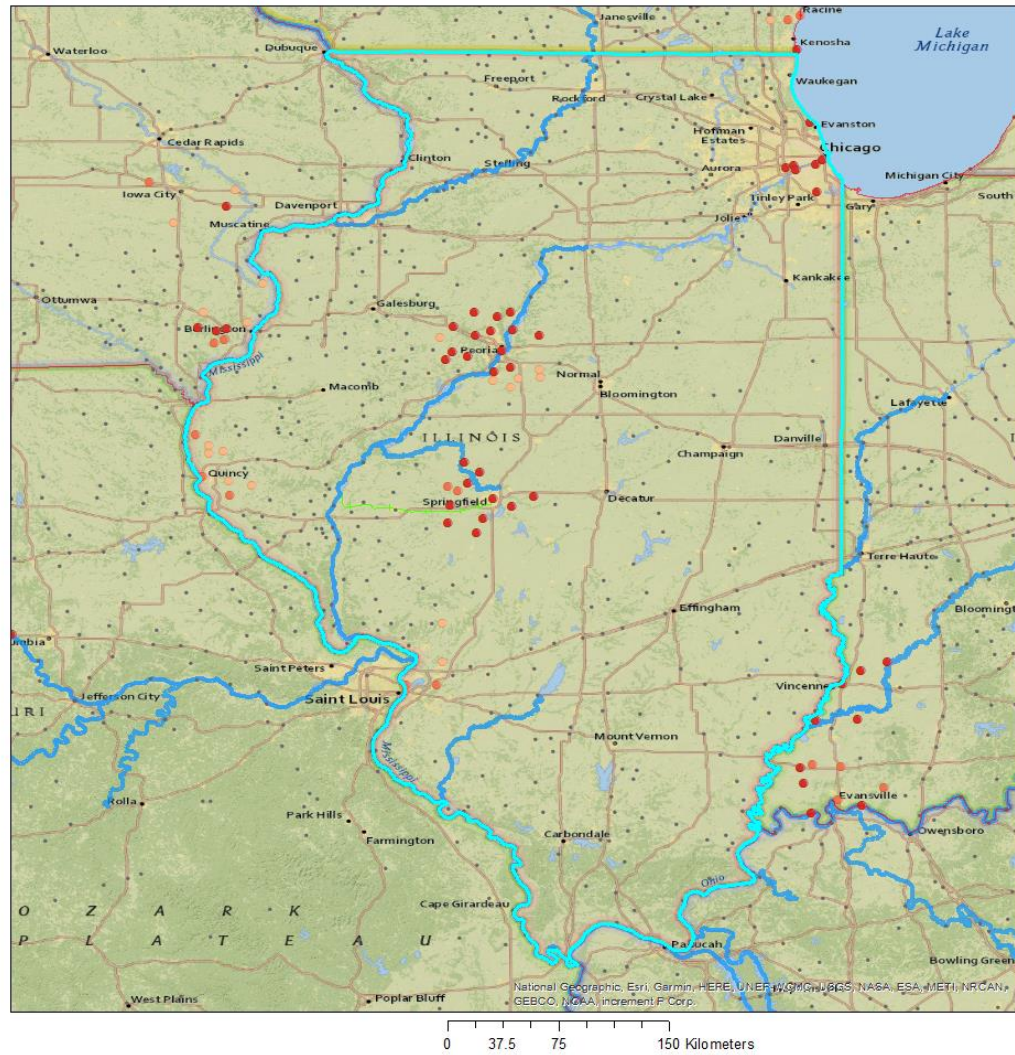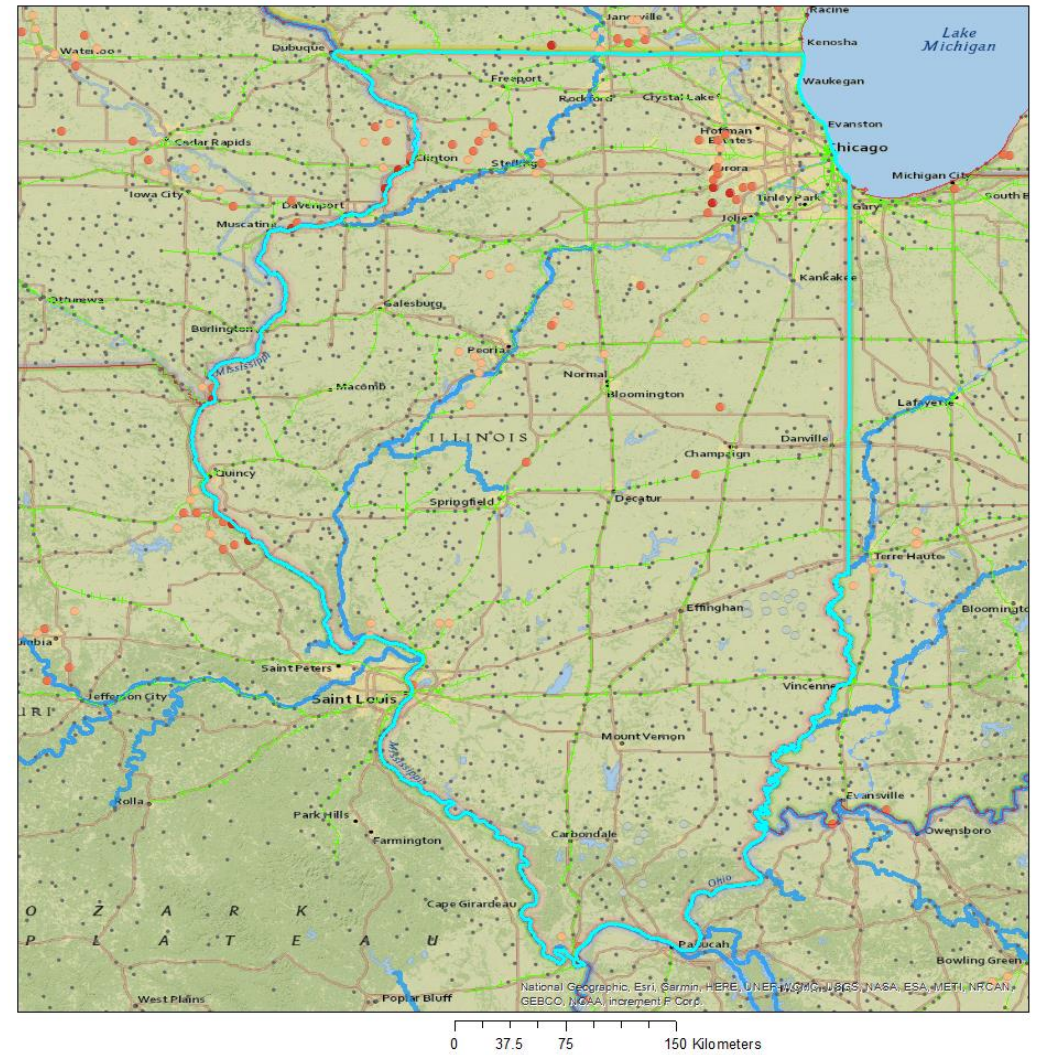
Figure A9.1. Hotspot Analysis for Illinois (1849)

Figure A9.12. Hotspot Analysis for Illinois (1871)

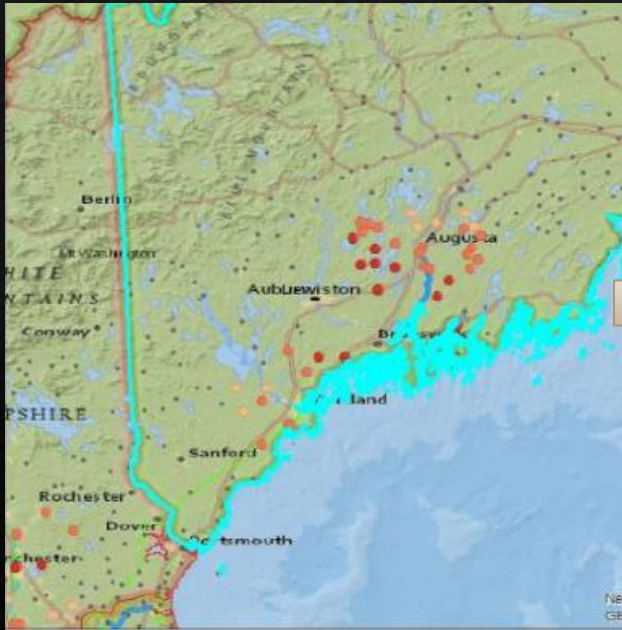Chicago, Springfield and Illinois, St. Louis Border Area

Maine (1847-1871)

Figure A4.12. Hotspot Analysis for Maine (1871)

What Happens When the Railroad Arrive in Town?

Clustering (1849) Diffusion (1855) Merge (1871)
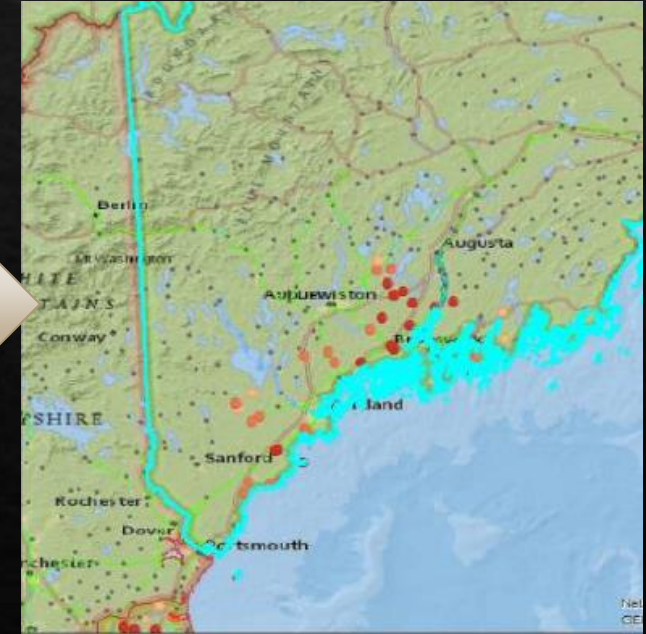
California (1857-1871)

Figure A11.12. Hotspot Analysis for California (1871)
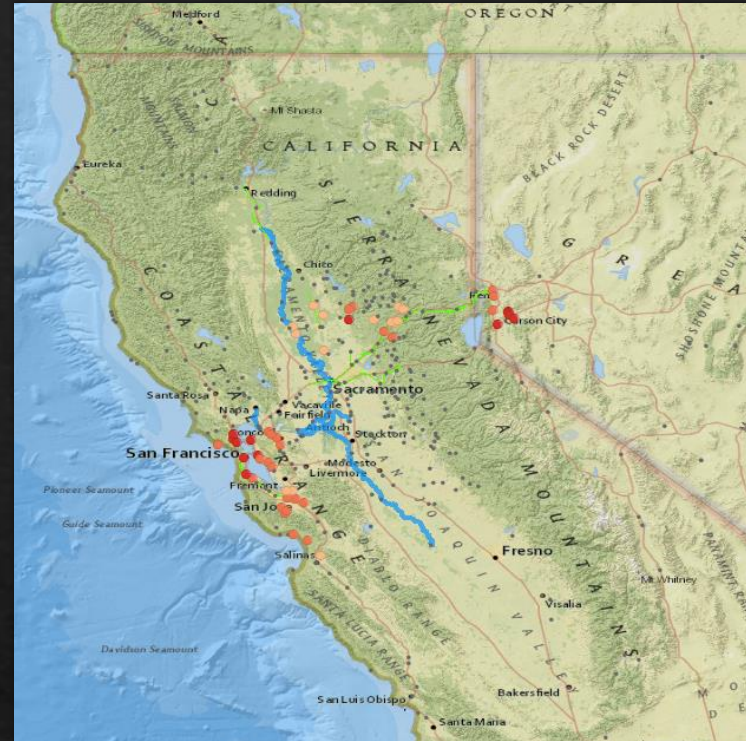
Postal Revenue of California (1857 vs. 1865)

* Bay Area Cluster (war, national market, transportation system)
* Virginia and Truckee Railroad – connects to Carson City silver mine (new transportation corridors)

Postal Revenue of Postal Office Within 30km from San Francisco (1847 v.s. 1871)

1847

Post Office	County	Postal Revenue
San Leandro	Alameda	8,248
Alameda	Alameda	4,342
San Pablo	Contra Costa	1,855
San Rafael	Marin	1,120
Mill Valley Air Force Station	Calaveras	595
Oakland	Alameda	0

1871

Post Office	County	Postal Revenue
Oakland	Alameda	486,384
San Rafael	Marin	24,673
San Leandro	Alameda	22,673
San Mateo	San Mateo	11,340
San Quentin	Marin	5,003
Alameda	Alameda	4,342
San Pablo	Contra Costa	4,102
Millbrae	San Mateo	400

Summary

- **postal revenue data, with its high spatial granularity, is a valuable indicator of economic activity**

- Diffusion Pattern: initially centered in key economic hubs, postal revenue patterns reveal a distinct diffusion as transportation infrastructure, such as railroads and canals, expands.

Table 2.1 Summary Statistics- the Entire Dataset

Number of Records	Mean	Minimum	Maximum	Standard Deviation
191,891	16755.52	0	1146151	77485.72

25th percentile	Median (50th percentile)	75th percentile	90th percentile	99th percentile
424	1266.00	4169	15563.00	66426.80

New York
(1849 – 1871)

Figure A6.12. Hotspot Analysis for New York (1871)
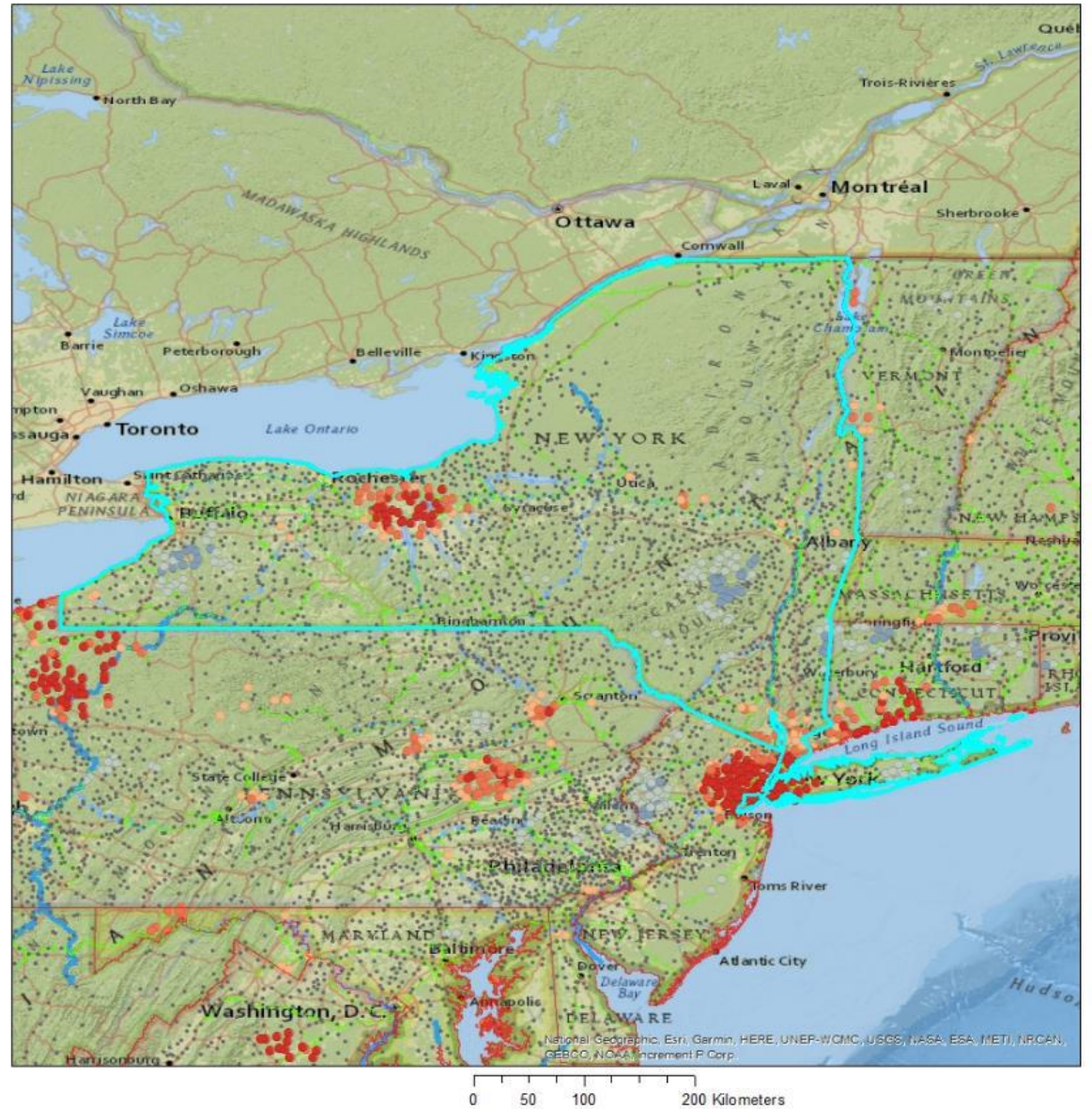
Oregon (1849 – 1871)

Figure A12.12. Hotspot Analysis for Oregon (1871)