

On the Outs: Rethinking Outlier Detection in the Economic Census

Sarah Scott

Supervised by Clayton Knappenberger, Angelina Peck, and Emily Wiley

Economy-Wide Statistics Division

Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau.



The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7504847, subproject P-7514952, Disclosure Review Board (DRB) approval number: CBDRB-FY22-EWD001-003).

1 The Problems

Varied Approaches to Outlier Review

- Decentralized tools and workflow across trade teams
- Existing tools lead to duplication of analysts' efforts

Timing of Micro and Geographic Review

- Prioritization of micro-record review
- Limited resources allocated to state and sub-state review based on existing timelines

2

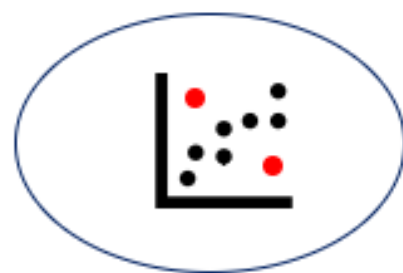
The Solution



Allows analysts to filter by NAICS, state, county, and other important metrics



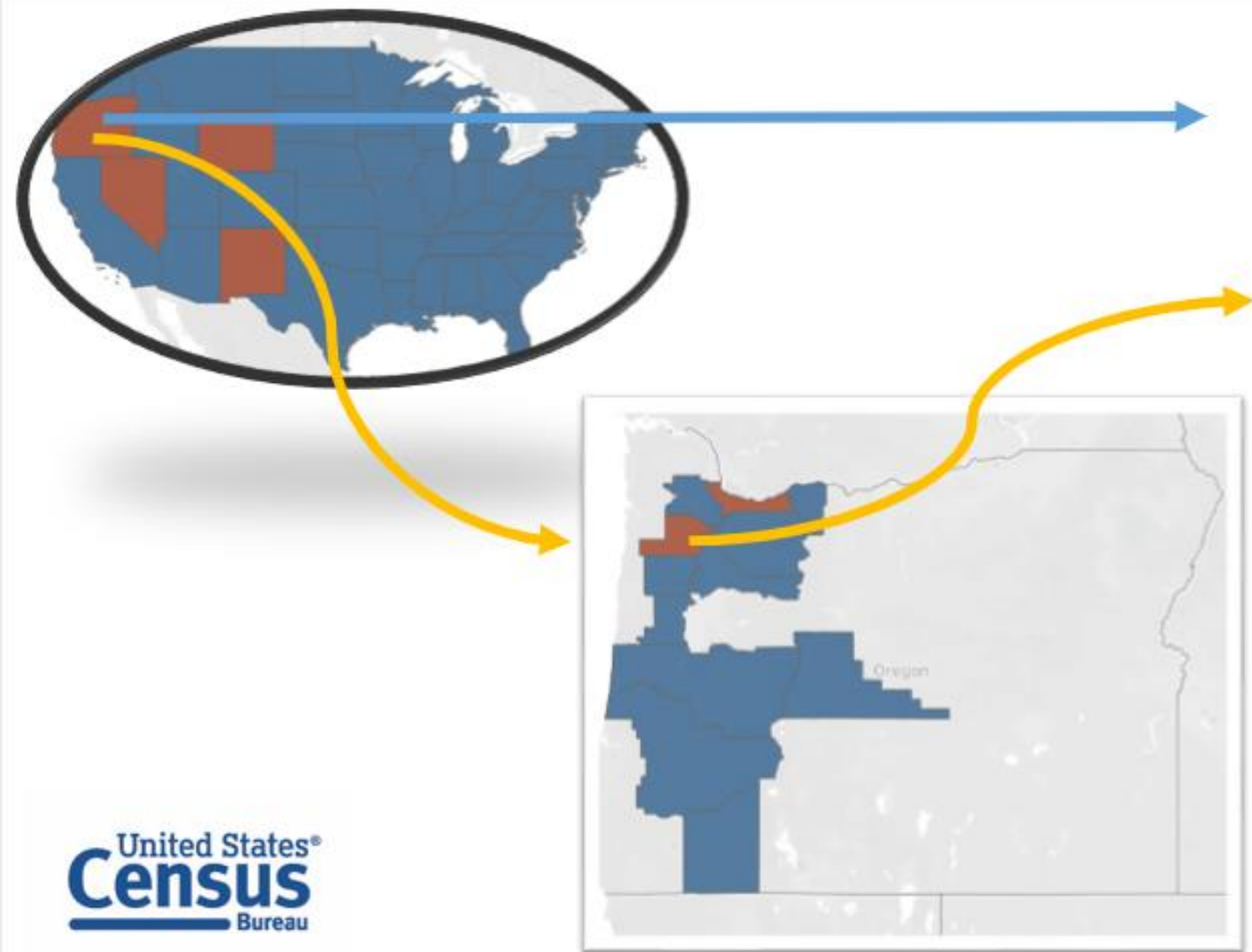
Uses state and county-level maps for a geographic-centered approach



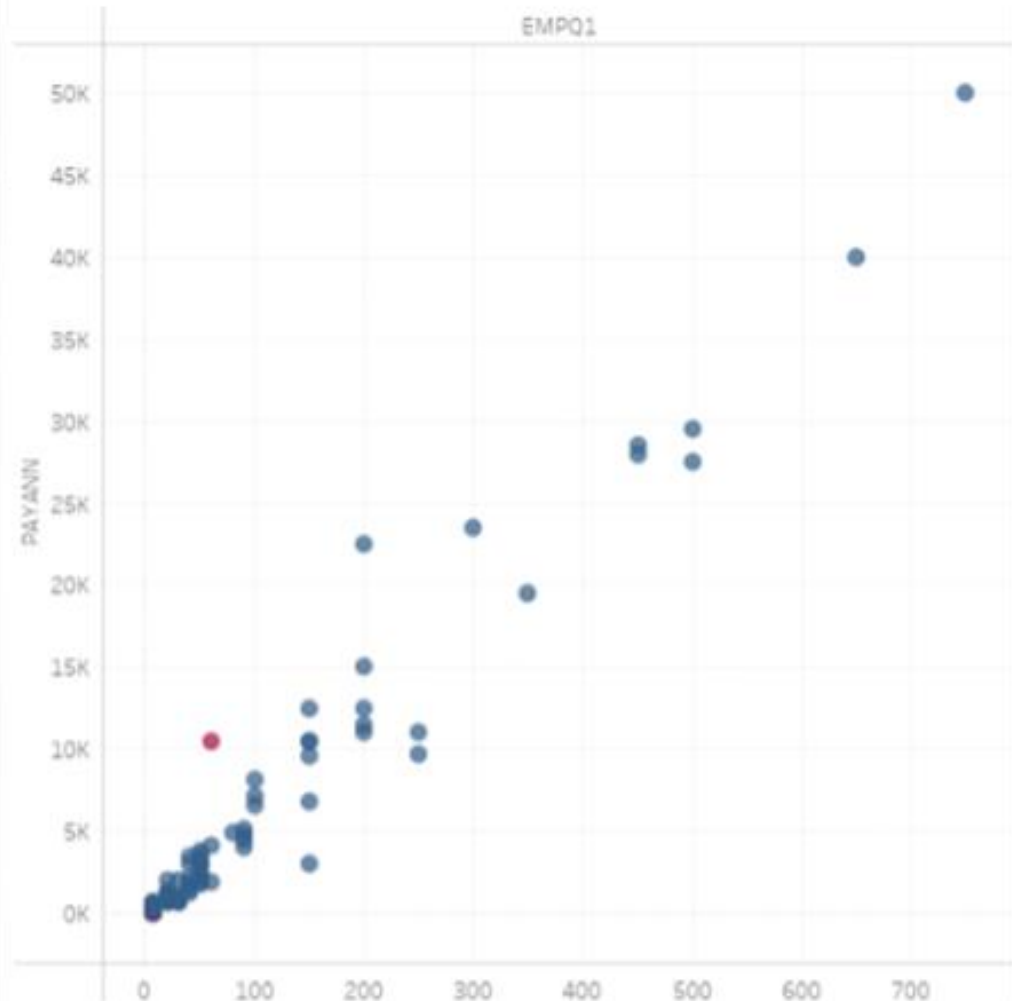
Graphs establishments on scatter plots to visually highlight outliers

3

Dashboard Walkthrough



Establishment Level



3

Dashboard Walkthrough

Estab. Details

Outlier C..	STAT..	COUNTY C..	SURVUID	NAME1	RCPTOT_F															
2	OR	41071	2006636826	Scotty's Hotdog Stand	R															
1	OR	41071	2006458531	Business B	B															
			2006580132	Business C	B															
			2006928642	Business D	R															
0	OR	41001	2023413808	Business E	RB															
		41003	2006609298	Business A	H															

ANPW_EMPAVPW: 0.000

SURVUID (Notes)	Day of Notedate	Noteidty	Userid	
2006636826	December 2, 2019	EST	A1497	replaced RCPTOT with correct number

4

Next Steps

