Raj Shah

School of Communication and Information

Rutgers University–New Brunswick

Information Visualization Section 7 (04:547:321:G7)
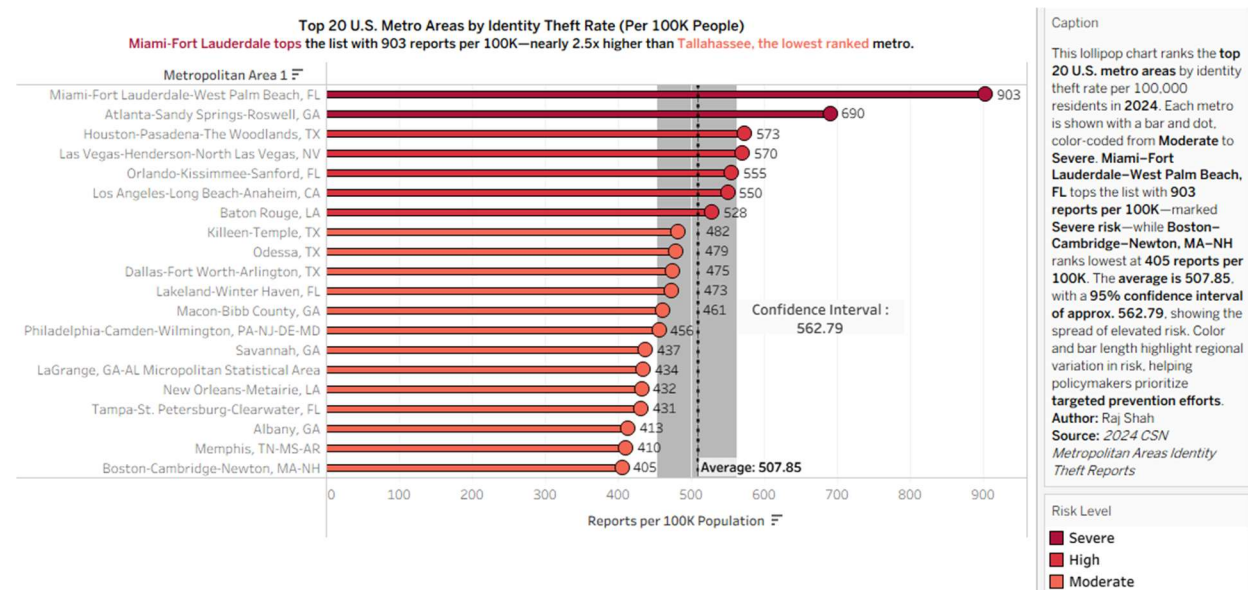
Jennifer Silvia Muller

July 20th, 2025

<u>Visualizing U.S. Identity-Theft Risk and Federal Response</u>

The Consumer Sentinel Network's 2024 metropolitan identity-theft file and the Federal Trade Commission's (FTC) open-government enforcement datasets give complementary views of digital-fraud pressure points in the United States. I joined both sources in Tableau Public to learn where identity theft is most acute, how burdens vary with population size, and whether federal non-merger enforcement is keeping up with complaint growth. Throughout the work I leaned on visualization guidance from Schwabish, Tufte, and Berinato to shape every chart and narrative choice.

The identity-theft file lists 384 metropolitan statistical areas and supplies two crucial measures: raw complaint counts and rates per 100 000 residents. After dropping rows with missing population values, I imported the cleaned CSV and built calculated fields to rank metros and flag risk levels. The FTC tables span 1995-2020 and record each non-merger matter's year, industry, and legal disposition. A simple union produced one tidy table for annual tallies and categorical pivots.

Top 20 U.S. Metro Areas by Identity Theft Rate (Per 100K People)
Miami-Fort Lauderdale tops the list with 903 reports per 100K—nearly 2.5x higher than Tallahassee, the lowest ranked metro.

Caption

This lollipop chart ranks the **top 20 U.S. metro areas** by identity theft rate per 100,000 residents in 2024. Each metro is shown with a bar and dot, color-coded from **Moderate** to **Severe**. **Miami–Fort Lauderdale–West Palm Beach, FL** tops the list with **903 reports per 100K**—marked **Severe risk**—while **Boston–Cambridge–Newton, MA–NH** ranks lowest at **405 reports per 100K**. The average is **507.85**, with a **95% confidence interval of approx. 562.79**, showing the spread of elevated risk. Color and bar length highlight regional variation in risk, helping policymakers prioritize **targeted prevention efforts**. **Author:** Raj Shah **Source:** *2024 CSN Metropolitan Areas Identity Theft Reports*
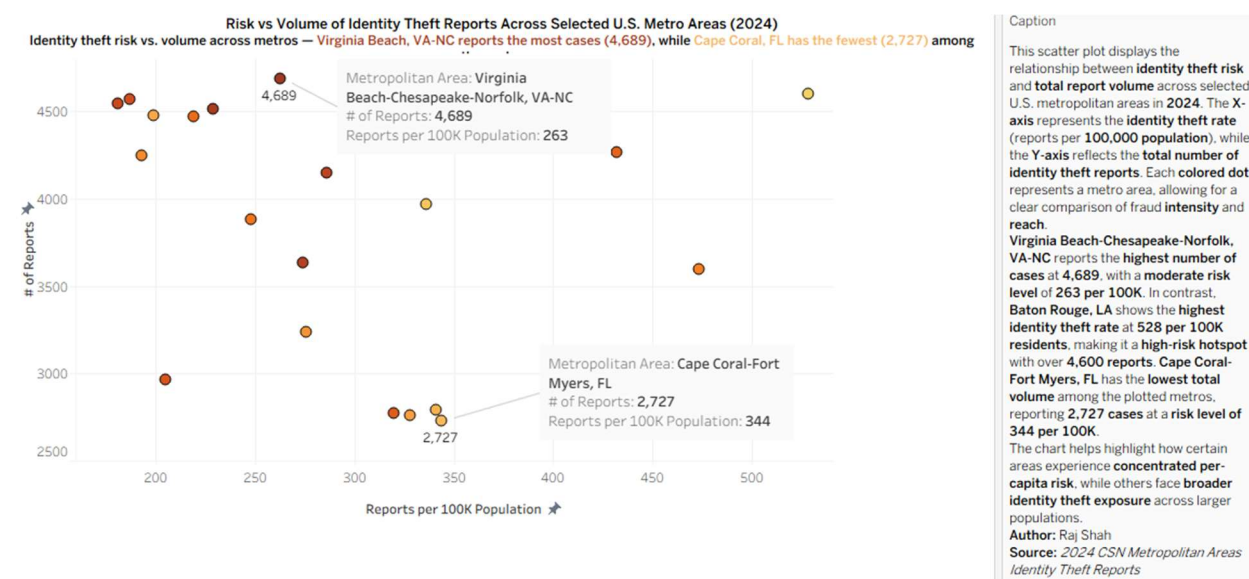
To introduce per-capita variation I designed a lollipop chart that ranks the 20 riskiest metros. The bar-plus-dot format, celebrated by Schwabish (2021) for combining magnitude and pinpoint labels, shows Miami–Fort Lauderdale–West Palm Beach at 903 reports per 100 000, nearly 2.5 times Boston–Cambridge–Newton's 405. A dashed reference line marks the sample mean (507.85), and a translucent confidence band (≈ 453–563) communicates statistical uncertainty. Acting on Tufte's (2001) warning against redundant ink, I removed default gridlines and ticks, relying on subtle reference bands to give context without visual noise.

Raw volume tells a different story. A horizontal bar chart limited to the 16 largest caseloads puts Los Angeles far ahead with 71 624 complaints—about eight times the national metro average of 32 013. Bars appear red or green depending on whether they beat that mean, adopting the "status indicator" aesthetic Berinato (2016) recommends for instant benchmark reading. New York places second in volume but only middling in per-capita risk, confirming that dense populations can conceal large absolute counts behind modest rates.
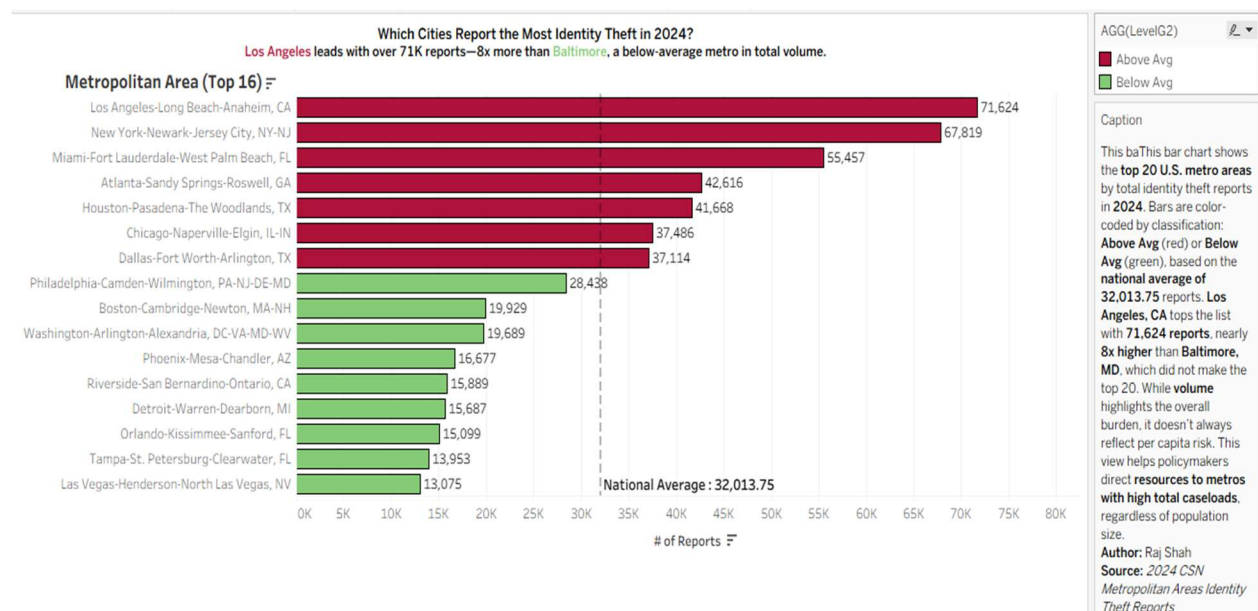
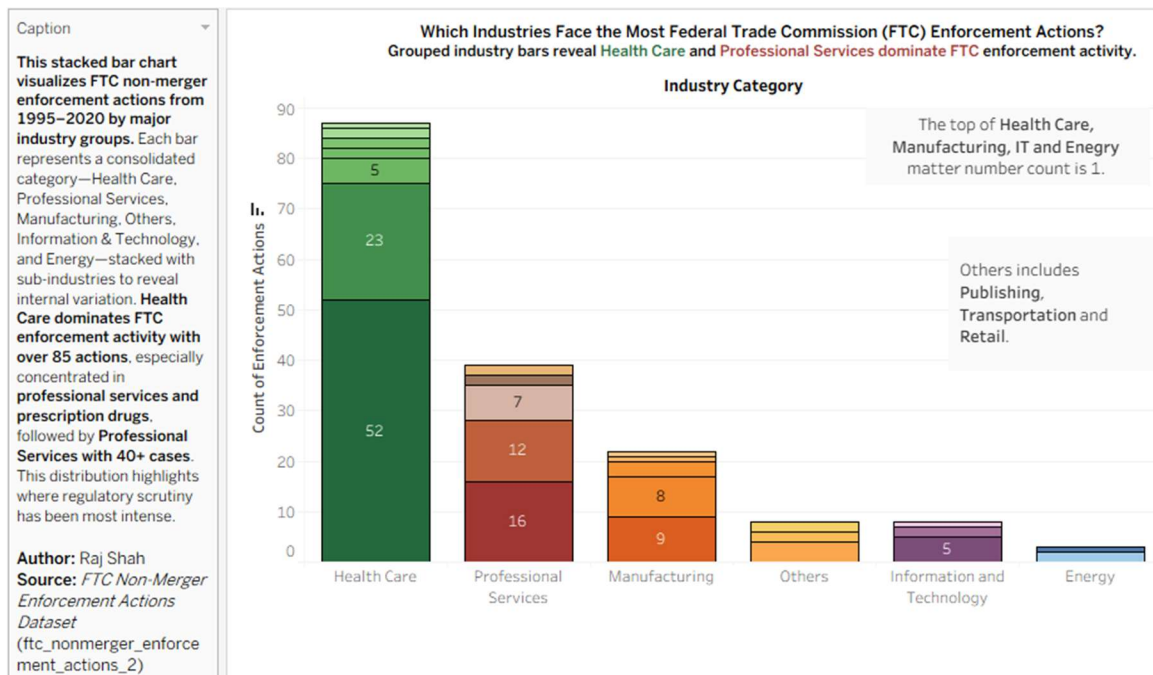I graphed a scatterplot with per-capita risk on the x-axis and total complaints on the y-axis. The national averages—191 reports per 100 000 residents and 791 total complaints— serve as reference lines. **Virginia Beach** rises highest on the y-axis with 4 689 complaints but shows only a moderate risk rate of 263. **Baton Rouge**, in contrast, posts a rate of 528 per 100 000 yet registers a similar complaint count. **Cape Coral–Fort Myers** lands in the middle: 2 727 complaints at a rate of 344. These points show that smaller metros can still concentrate large volumes of fraud. Following Schwabish's (2021) guidance, I toned down the color palette so bright hues highlight only categorical differences rather than every individual dot.

Risk vs Volume of Identity Theft Reports Across Selected U.S. Metro Areas (2024)
Identity theft risk vs. volume across metros — Virginia Beach, VA-NC reports the most cases (4,689), while Cape Coral, FL has the fewest (2,727) among

Metropolitan Area: Virginia Beach-Chesapeake-Norfolk, VA-NC
# of Reports: 4,689
Reports per 100K Population: 263

Metropolitan Area: Cape Coral-Fort Myers, FL
# of Reports: 2,727
Reports per 100K Population: 344

# of Reports

Reports per 100K Population

Caption

This scatter plot displays the relationship between **identity theft risk** and **total report volume** across selected U.S. metropolitan areas in 2024. The X-axis represents the **identity theft rate** (reports per **100,000 population**), while the Y-axis reflects the **total number of identity theft reports**. Each **colored dot** represents a metro area, allowing for a clear comparison of fraud **intensity** and **reach**.
Virginia Beach-Chesapeake-Norfolk, VA-NC reports the **highest number of cases** at **4,689**, with a **moderate risk level** of 263 per 100K. In contrast, Baton Rouge, LA shows the **highest identity theft rate** at 528 per 100K residents, making it a **high-risk hotspot** with over **4,600 reports**. Cape Coral-Fort Myers, FL has the **lowest total volume** among the plotted metros, reporting 2,727 cases at a risk level of 344 per 100K.
The chart helps highlight how certain areas experience **concentrated per-capita risk**, while others face **broader identity theft exposure** across larger populations.
Author: Raj Shah
**Source:** *2024 CSN Metropolitan Areas Identity Theft Reports*

FTC data puts those patterns in policy context. A 25-year line chart shows non-merger actions peaking in 2003 at 23 matters and slumping to 1 in 2011. A downward least-squares trend line and a horizontal mean (6.68) frame the long-run decline. Tufte's (1997) "data-ink ratio" convinced me to strip the original area fill, sharpening each year's variation. Story annotations at the 2003 peak and 2011 trough, a tactic endorsed by Berinato (2016), highlight turning points without a bulky legend. Given these hotspots, the next logical question is: *Has federal enforcement kept pace?*

Federal Trade Commission (FTC) Enforcement Actions Over Time: Yearly Trends & Peaks
A 25-Year Trendline of FTC Non-Merger Enforcement Actions (1996–2020) — High in 2003, Low in 2011

Caption
This line chart shows a 25-year trend in FTC non-merger enforcement actions from 1996 to 2020. The data tracks yearly case counts, peaking in 2003 with 23 actions and dipping to just 1 action in 2011. A dotted trendline reveals a gradual decline in enforcement over time, while the average line (6.68) offers context for long-term activity. The 2003 spike may reflect heightened scrutiny in healthcare or tech, while the 2011 drop hints at shifting agency priorities. This view highlights how federal enforcement intensity varies across decades.
Author: Raj Shah
Source: FTC Non-Merger Enforcement Actions Dataset (1996–2020)

Industry analysis required collapsing 30-plus FTC sector codes into six groups. A stacked bar chart reveals Health Care with over 85 matters, followed by Professional Services with 40+. Consistent segment ordering and in-bar counts aid comparison; Schwabish (2021) advocates direct labeling when color alone may confuse. Palette choices were checked for color-blind accessibility and adjusted once initial reds and greens proved indistinct.

Caption
This stacked bar chart visualizes FTC non-merger enforcement actions from 1995–2020 by major industry groups. Each bar represents a consolidated category—Health Care, Professional Services, Manufacturing, Others, Information & Technology, and Energy—stacked with sub-industries to reveal internal variation. Health Care dominates FTC enforcement activity with over 85 actions, especially concentrated in professional services and prescription drugs, followed by Professional Services with 40+ cases. This distribution highlights where regulatory scrutiny has been most intense.

Author: Raj Shah
Source: FTC Non-Merger Enforcement Actions Dataset (ftc_nonmerger_enforcement_actions_2)

Which Industries Face the Most Federal Trade Commission (FTC) Enforcement Actions?
Grouped industry bars reveal Health Care and Professional Services dominate FTC enforcement activity.

Finally, a donut chart summarizes legal dispositions: Consent Orders Accepted for Comment (73 %), Part III Administrative Complaints (16 %), and Permanent Injunctions (11 %). Tufte (2001) is skeptical of circles, yet with only three slices and direct labels the form avoids angle misreading, and the blank center houses a concise subtitle that reduces the dominant wedge's weight.
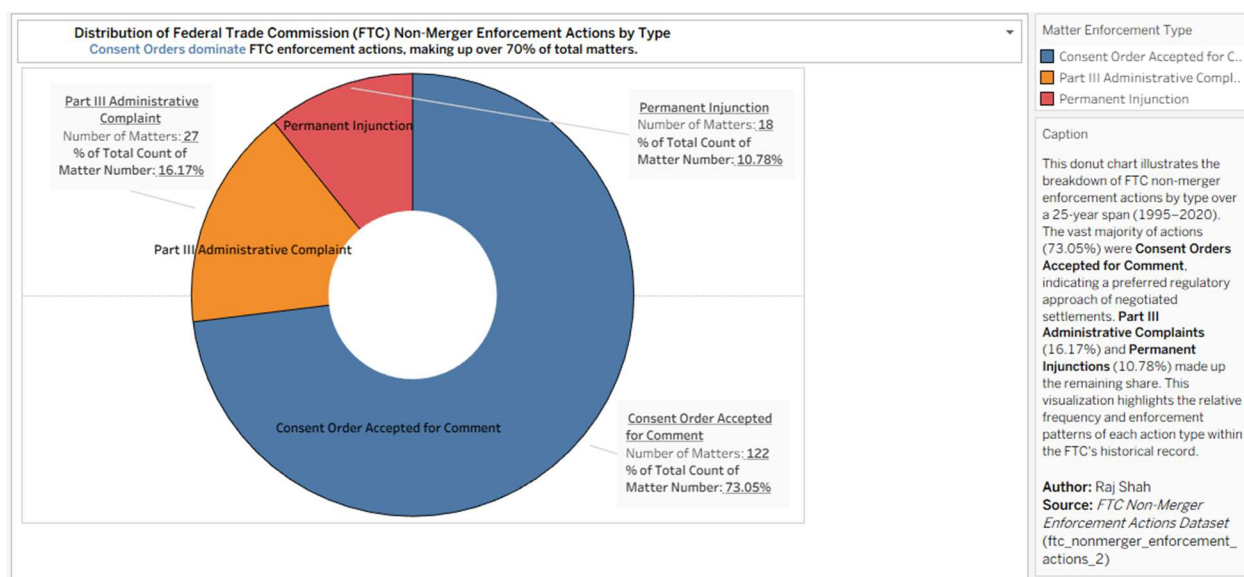
Three technical hurdles stood out. First, the two sources share no geographic key beyond state names, so a direct metro-to-enforcement merge was impossible; I instead present identity-theft and FTC views side by side. Second, initial calculated fields using WINDOW_AVG collapsed every metro into one bar; switching to fixed-level-of-detail expressions restored row-level classification. Third, scatterplot over-plotting in the low-risk quadrant was eased by jittering positions three units and adding tooltips for exact values.

Findings show a sharply skewed burden: a handful of metros dominate both rate and volume. Miami–Fort Lauderdale blends extreme per-capita risk with high counts, making it an obvious target for focused prevention. Los Angeles rules sheer volume but sits just above

average risk, implying that any intervention there must scale to a vast population. The FTC trend line warns that enforcement has not kept pace with the surge in complaints and that actions are concentrated in Health Care, leaving sectors such as retail fraud potentially under-policed.

Visualization scholarship guided every refinement. Schwabish's lollipop format balanced precision and aesthetics; Tufte's minimal-ink rules improved clarity; Berinato's "one-thing" principle framed subtitles like "Los Angeles leads with over 71 K reports—8 × Baltimore," turning statistics into narrative hooks. Iterative redesign through those lenses produced a dashboard capable of informing policymakers and the public about evolving fraud threats and enforcement gaps.

Future research might link identity-theft incidents to demographic or socioeconomic indicators to test vulnerability hypotheses, or slice FTC complaints by subtype (e.g., Social Security fraud) for finer alignment. Even so, the present visuals demonstrate how open data and sound design convert spreadsheets into actionable insight.

References

Berinato, S. (2016). *Good charts: The HBR guide to making smarter, more persuasive data visualizations*. Harvard Business Review Press.

Schwabish, J. (2021). *Better data visualizations: A guide for scholars, researchers, and wonks*. Columbia University Press.

Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Graphics Press.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Graphics Press.