From Spreadsheet to Atlas

How one monthly spreadsheet becomes a publication-grade field report for any division, region, or the nation.

JEFF FRANZEN · JUNE 15, 2026 · BIOMED FIELD-REPORT PIPELINE

The starting point is the HS Sponsor File — a monthly Excel spreadsheet of roughly **232,000 rows** spanning fiscal years 2022–2026. Each row is one blood-drive sponsor for one year: account name, account type, address, number of drives, RBC units collected, and a latitude/longitude. The data is rich. The problem is that nobody can see a geographic pattern in 232,000 rows of a spreadsheet.

The pipeline below turns that file into a clean, branded, multi-page map report. It is fully automated — seven repeatable steps, no hand-drawing and no hand-typed numbers. The toolkit (page 3), a walk-through of how a map is drawn (page 4), and sample pages from the North Carolina Region report (pages 5–6) follow.

The Pipeline

1 READ

Load the spreadsheet

Open the monthly HS Sponsor File and cache it so every later step and every geography runs in seconds rather than re-reading 232,000 rows each time.

2 PLACE ON THE MAP

Join every sponsor to a real county

Each sponsor carries coordinates. We match every sponsor into the authoritative American Red Cross county geography — 3,140 counties carrying Division / Region / Chapter names plus 2023 population and demographics. Now every drive is tied to a real place in the Red Cross hierarchy.

3 REPAIR THE DATA

Catch and fix a coordinate error

A data-quality check found that many FY2025 records carried coordinates that placed sponsors in the wrong location (for example, a sponsor listed in Ithaca, NY plotted into North Carolina). FY2022–24 and FY2026 were clean. Because each sponsor recurs across years with correct coordinates, the FY2025 locations were rebuilt from each sponsor's known-good years. The underlying file issue has been documented separately for the team that produces the file.

4 AGGREGATE

Roll up by county and fiscal year

Group the cleaned data by county and fiscal year: total units, total drives, and sponsor counts. This is the numeric backbone every map and table draws from.

5 COMPUTE THE LENSES

Five maps, each answering one question

Collections — where is blood collected? · **Growth** — where is it rising or falling, FY22 → FY25? · **Yield** — where do drives work hardest (units per drive)? · **Opportunity** — where is the untapped market (units per 1,000 residents; low coverage = target)? · **Mix** — who hosts the drives (leading sponsor type)?

6 RENDER — EXPLAINED SIMPLY

A printer that draws maps from data

Think of this step as a printer that *draws* instead of types. Free, widely used mapping software (called GeoPandas and matplotlib) reads two things — the outline shape of every county, and the numbers we computed in the steps above. It then paints each county the correct shade, drops a labeled dot for each sponsor, adds the trend chart, the donut, the legends, and the county labels, and writes out finished, brand-styled pages as a single PDF. Nobody draws a map or types a number by hand. Change the spreadsheet and the maps redraw themselves — exactly the same way, every time.

7 SCALE

Any geography, refreshed monthly

The same engine runs for the nation, any of the 6 divisions, or any of the 47 regions — 54 reports — on demand, and refreshes automatically each month when the new spreadsheet arrives.

Why It Can Be Trusted

The software does the counting and the placing; the maps only present what the software computed. Nothing is hand-drawn or hand-typed, so every number is **reproducible and auditable**, the branding is always consistent, and each map answers exactly one operational question. Re-running next month's file reproduces the entire binder in minutes — and the same automated check that caught the FY2025 coordinate problem runs every time.

> *A spreadsheet nobody could read becomes a boardroom-ready field report — produced the same way, every month, in minutes.*
>
> THE ONE-SENTENCE VERSION

The Toolkit — Python · GeoPandas · matplotlib

Three free, open-source tools do all the work — none is Red-Cross-specific; they are the standard toolkit for data and mapping. One Python script chains them together, so re-running it on next month's file reproduces every page identically.

Py

PYTHON — THE LANGUAGE

The engine everything runs in

A free, widely-used programming language — the standard for data work. One Python script opens the spreadsheet, runs the math, and writes the finished PDF. No spreadsheet macros, no manual steps: run the script, get the binder — and the same script next month gives the same result, every time.

gpd

GEOPANDAS — THE GEOGRAPHY

A spreadsheet that understands maps

GeoPandas adds location awareness. It takes each sponsor's latitude/longitude and the county shapes and answers the spatial questions: *which county does this drive fall in? · sum the units for every county · keep only the sponsors inside the selected region.* This is the step that turns 232,000 address rows into county-level totals tied to the Red Cross hierarchy.

plt

MATPLOTLIB — THE DRAWING

The pen that paints the page

Given the county shapes and the numbers GeoPandas computed, matplotlib draws everything: it shades each county, places the labeled circles, draws the trend line, donut, and legends, applies the Red Cross styling, and exports a print-ready PDF. It's what makes the output look like a publication instead of a chart dump.

*The flow, in one line: the spreadsheet → **Python** opens it → **GeoPandas** places & counts every sponsor by geography → **matplotlib** draws the branded pages → a PDF. Free tools, no licenses, runs on any laptop, identical output every month.*

HOW THE MAPS ARE MADE

Inside Step 6 — How One Map Gets Drawn

This is the step that surprises people: the map is not designed in Photoshop or placed by hand. Code draws every element in a fixed order, the same way every time. Here is exactly what happens to turn the numbers into one finished page.

1 START WITH TWO INGREDIENTS

The county shapes and the numbers

The code begins with the outline of every county (the shapes) and the table GeoPandas produced — for example, RBC units for each county. Nothing else. Everything on the page is built from these two things.

2 COLOR EACH COUNTY — THE CHOROPLETH

Paint-by-numbers, the data picks the shade

The code sets up a color scale (pale cream → deep red). Each county is filled by where its number lands on that scale: more units, darker red. A person never chooses a county's color — the value does, so the picture can't be fudged.

3 ADD THE SPONSOR MARKERS

A circle for each sponsor, sized by drives

For every sponsor the code drops a circle at its latitude/longitude, with the circle's size set by how many drives it ran — big circle, big sponsor. They stack on top of the colored counties.

4 LABEL THE LEADERS

Name only the top counties, with a halo

Labeling all 3,140 counties would be a wall of text, so the code writes the names of just the top counties, centered in each, with a soft cream halo behind the letters so they stay readable over any color.

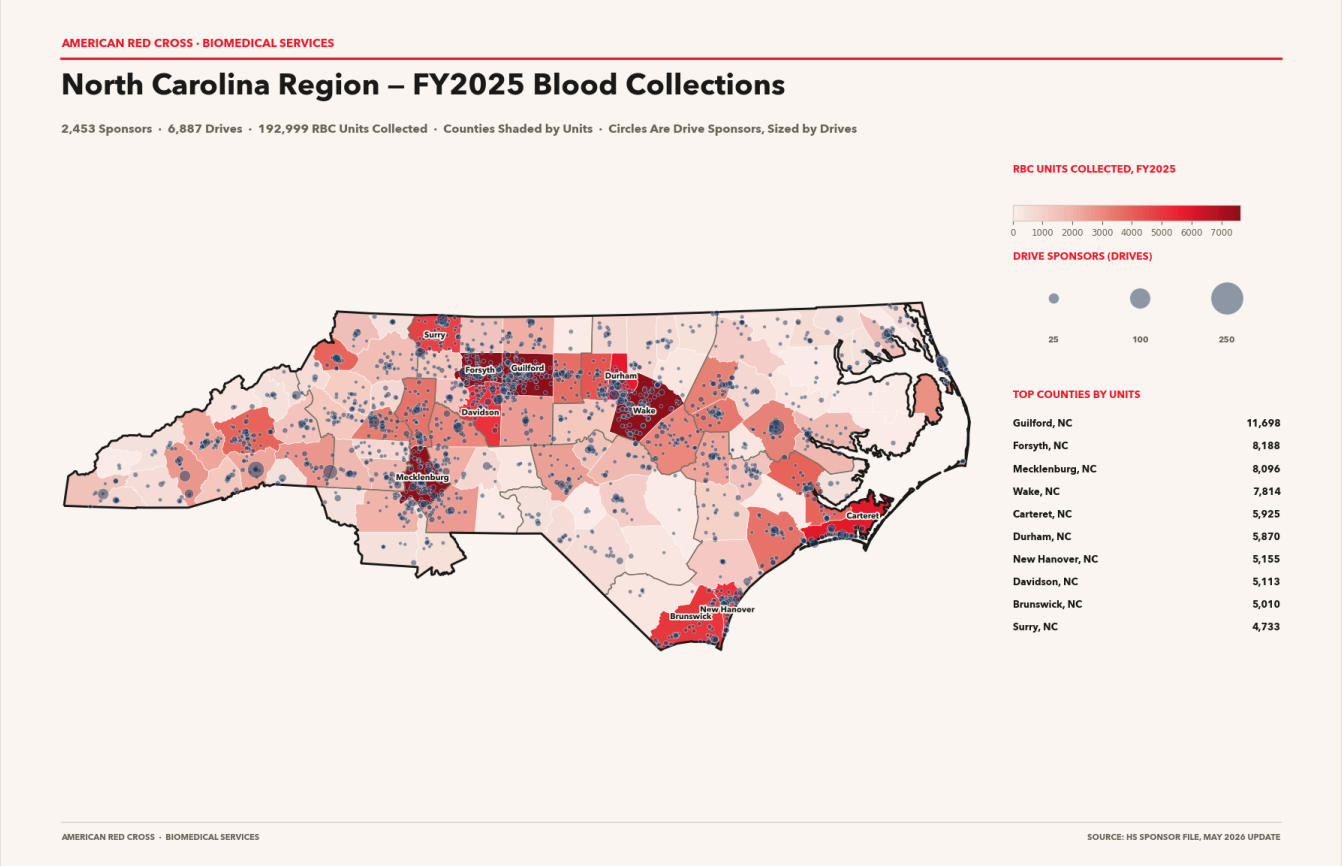
5 ADD THE CHROME & EXPORT

Title, legend, charts, frame — then save to PDF

Finally the code draws the title, the color key (legend), the trend line and donut, the cream page frame and the ARC-red accents — then saves the whole canvas straight to a PDF page. No screenshot, no touch-up. Change next month's numbers, re-run, and the identical page redraws itself in seconds.

What It Produces — Sample Page

From the North Carolina Region report. Every geography gets this same layout, generated automatically.



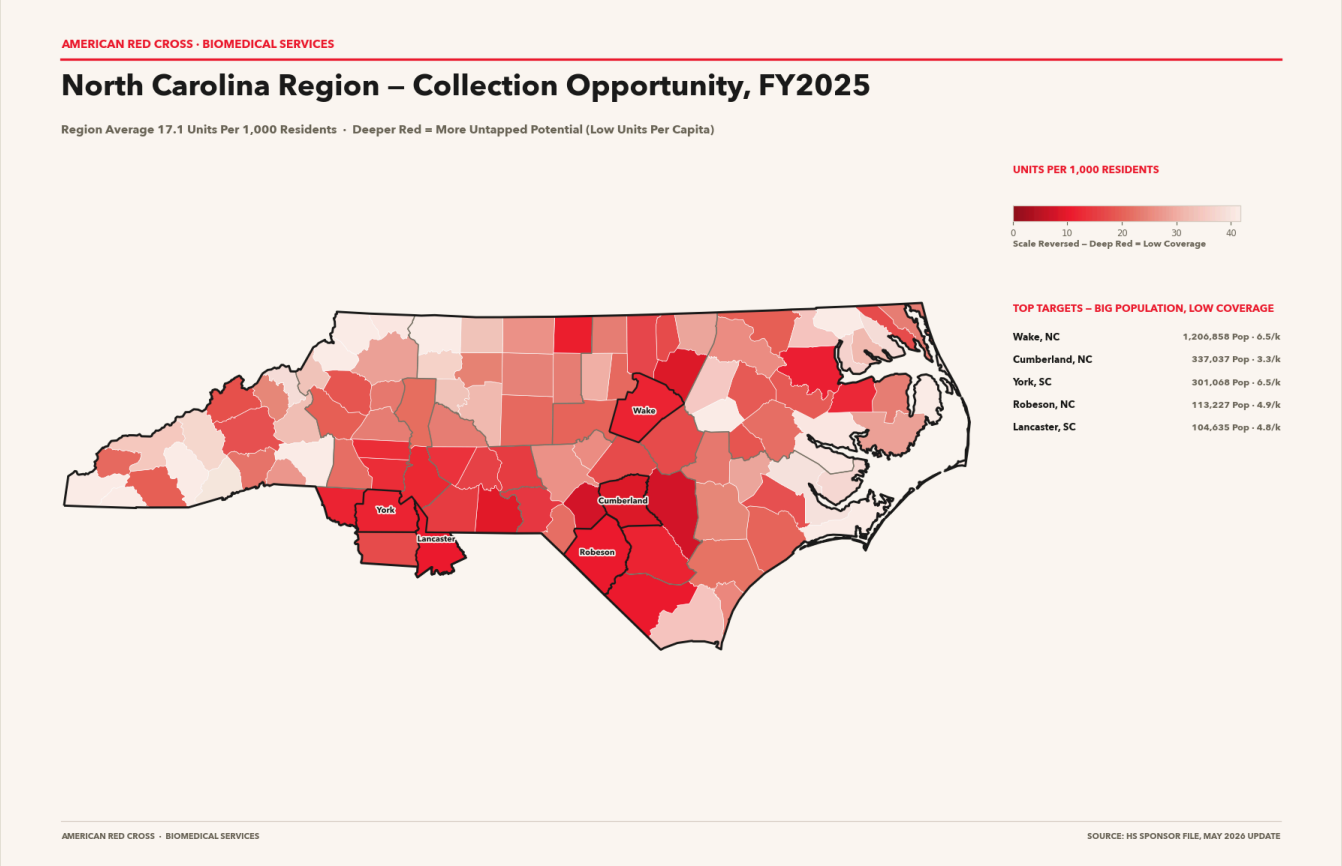
COLLECTIONS. Counties shaded by RBC units collected; a circle for every sponsor, sized by number of drives; the leading counties labeled; a ranked list and color key on the right. Answers **where blood is collected** — here, the Piedmont (Guilford, Forsyth, Mecklenburg, Wake) leads.

How to Read Any Map in the Report

The maps share one visual grammar, so a manager learns it once. **The shade of a county** encodes the headline measure for that page — darker means more (or, on the Opportunity map, less coverage). **Circles** are individual sponsors, sized by how many drives they ran. **Labels** name only the leading counties so the map stays readable. **The right-hand rail** always carries the color key and a ranked top-ten list. Same grammar on every page and every geography.

Sample output page. The full report is eight pages: cover, performance summary, five analytical maps, and a sponsor roster.

The Decision Map — Where to Add Drives



OPPORTUNITY. Counties shaded by units collected per 1,000 residents on a **reversed** scale — deeper red means **lower** coverage, i.e. more untapped potential. Big-population, low-coverage counties are outlined and named as targets (Wake, Cumberland). Answers **where to add drives**.

The Payoff

Beyond the polish, the reports point to action. The Opportunity map above flags where to add drives. The **Growth** map shows where collection is slipping before it surfaces in annual totals. And the repair step proved the pipeline does not just present data — it validates it, and caught a real problem in the source file along the way.