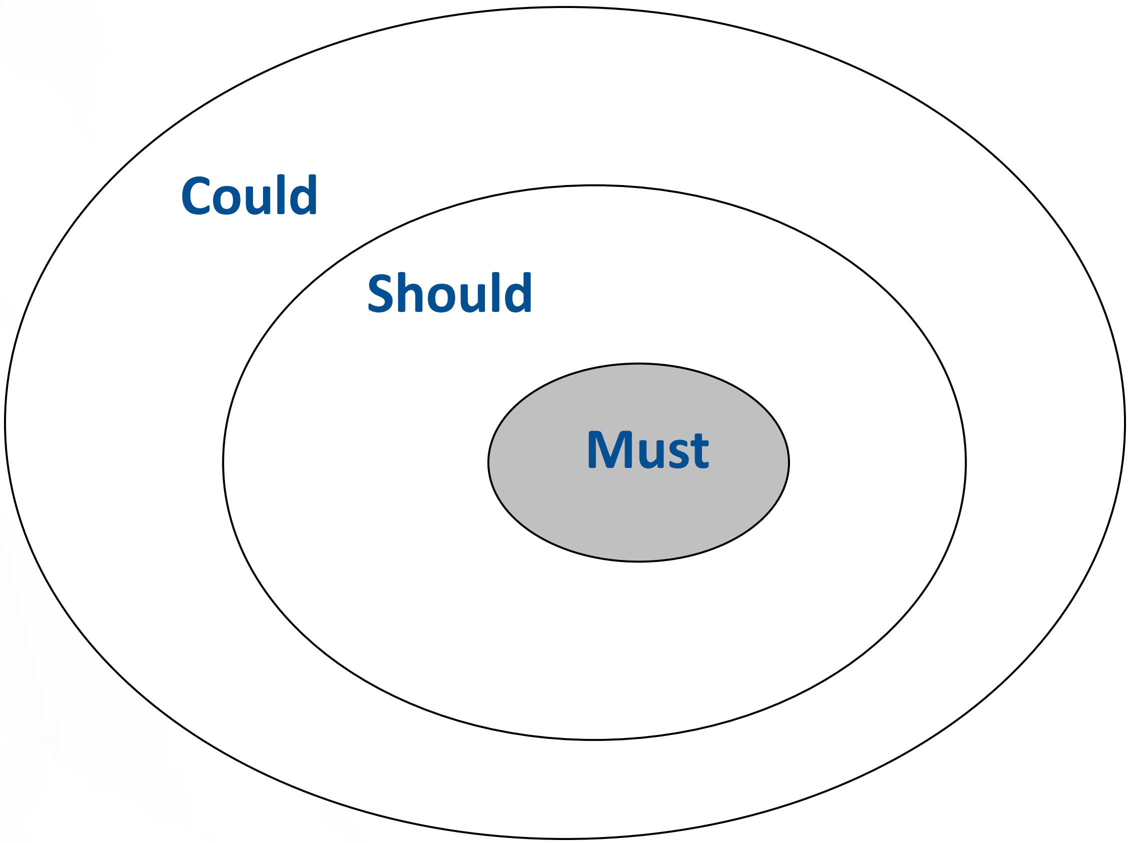
How to Turn 300 Pages of Data into a 30-page Report

Kate van Gelder

Pacific Northwest Intergovernmental Audit Forum, Olympia
September 5, 2013

The Goal



Today's Talk

- Understand your readers
- Summarize your data
- Simplify your message
- Write concisely

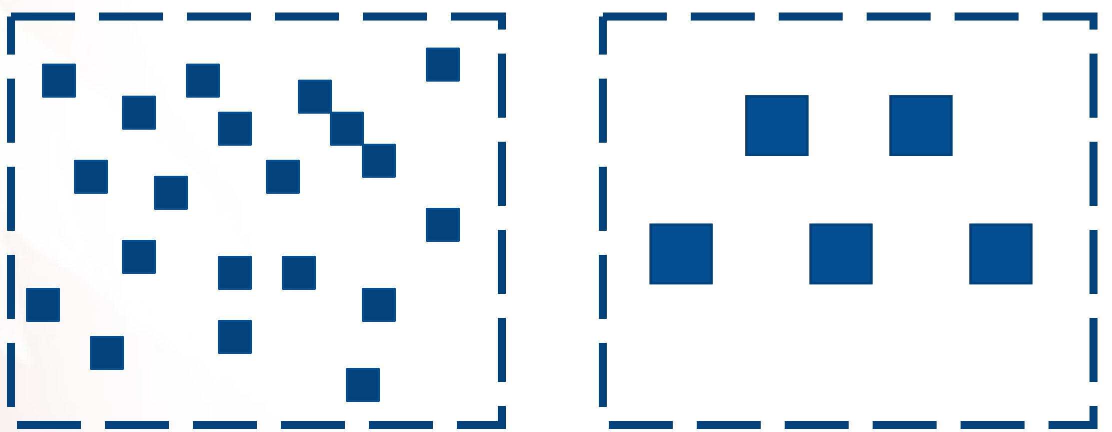
Accommodate Your Readers

- Understand distractions
- Chunk information
- Provide context
- Inform and educate

Understand Distractions

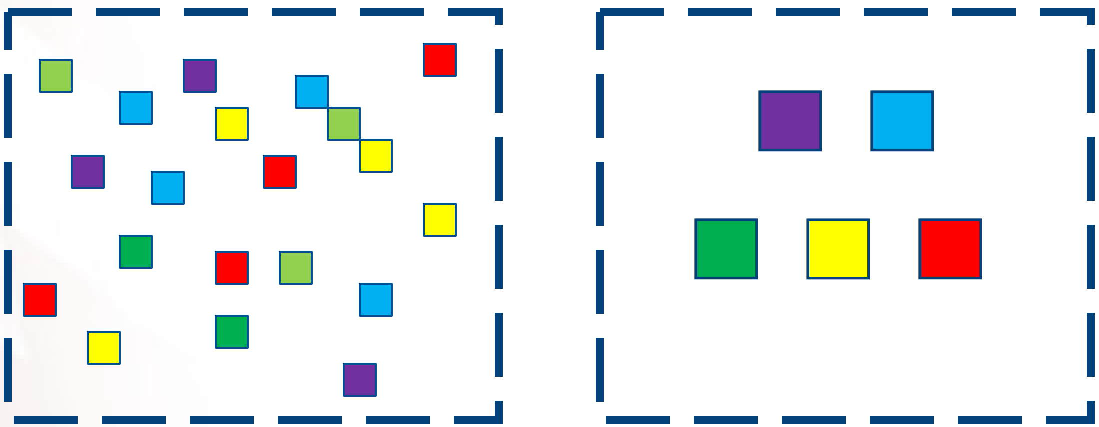
- Information overload
- Short attention span
- New media
  - E-mail
  - Facebook
  - Twitter

Chunk Information



but limit the number of chunks

Chunk Information



but limit the number of chunks

GAO
Accountability * Integrity * Reliability

Provide Context

What might the reader be familiar with that you can link new information to, so that it makes sense?

- Time
- Place
- Program
- Event

Inform and Educate

- *Inform*

  Most of VA's elderly outreach is web-based.

- *Educate*

  Most of VA's elderly outreach is web-based, but two-thirds of older adults don't use the Internet.

Summarize Your Data

- Interpret for your reader
- Show rather than tell

Interpret For Your Reader

- Was there an increase / decrease?
- Is the trend upward / downward?
- Are most / few states implementing the policy?
- Did funds increase while participants decreased?
- Is participation greater for Whites than Hispanics?
- Are women more at risk than men?

Show Rather Than Tell

	Defined Benefit (DB)	Defined Contribution (DC)
What is it?	A monthly payment during retirement based on the employee's salary and tenure	A pool of cash, enhanced by financial investment, that can be drawn upon during retirement [i.e. 401(k)]
Who finances the plan?	Employer, who bears all costs	Worker, often with a limited employer match
Who manages the plan?	Employer, who selects provider and may select investments	Employer selects provider, but most plans allow worker to select investments
How do the contribution costs compare?	Costs born entirely by employer Costs can fluctuate, but are generally smoothed by amortization	Costs shared by worker and employer Less fluctuation in costs, but payments must be pre-funded or made immediately
Advantages for workers	Dependable lifetime income Benefit grows with length of service	Portable Investment gains can increase available cash
Disadvantages for workers	Not portable Workers who change jobs earn a smaller benefit	Investment losses can decrease available cash No reward for length of service
Advantages for employers	Can serve a worker retention tool	Workers bear all risk of investments
Disadvantages for employers	Employer solely responsible for funding monthly payments and bears all risk of investments	Portable, so less valuable as a retention tool Higher administrative costs

Why Graphics Are Effective?

- **For writers, a graphic**
  - organizes the information
  - helps with analysis

- **For readers, a graphic**
  - supports and supplements analysis of the information
  - translates complex data to a visual to aid understanding
  - makes the presentation of the message more understandable, accessible, and memorable

GAO
Accountability * Integrity * Reliability

Two Types of Graphics

Table

A comparison—in columns and rows—of information or categories of information

Examples: Amounts of money or details of programs

Figure

A pictorial or diagrammatic representation of information

Examples: Bar charts, pie charts, line charts, time lines, maps, photographs, organizational charts, and flow charts

Streamline Text with a Table

	Defined Benefit (DB)	**Defined Contribution** (DC)
What is it?	A monthly payment during retirement based on the employee’s salary and tenure	A pool of cash, enhanced by financial investment, that can be drawn upon during retirement [i.e. 401(k)]
Who finances the plan?	Employer, who bears all costs	Worker, often with a limited employer match
Who manages the plan?	Employer, who selects provider and may select investments	Employer selects provider, but most plans allow worker to select investments
How do the contribution costs compare?	Costs born entirely by employer Costs can fluctuate, but are generally smoothed by amortization	Costs shared by worker and employer Less fluctuation in costs, but payments must be pre-funded or made immediately
Advantages for workers	Dependable lifetime income Benefit grows with length of service	Portable Investment gains can increase available cash
Disadvantages for workers	Not portable Workers who change jobs earn a smaller benefit	Investment losses can decrease available cash No reward for length of service
Advantages for employers	Can serve a worker retention tool	Workers bear all risk of investments
Disadvantages for employers	Employer solely responsible for funding monthly payments and bears all risk of investments	Portable, so less valuable as a retention tool Higher administrative costs

Amounts: Table

Table 2: Comparison of the Number of Firms Supporting Iran Oil and Gas Development Activities, during the Period of 2005 through 2009 and 2010 through May 2011

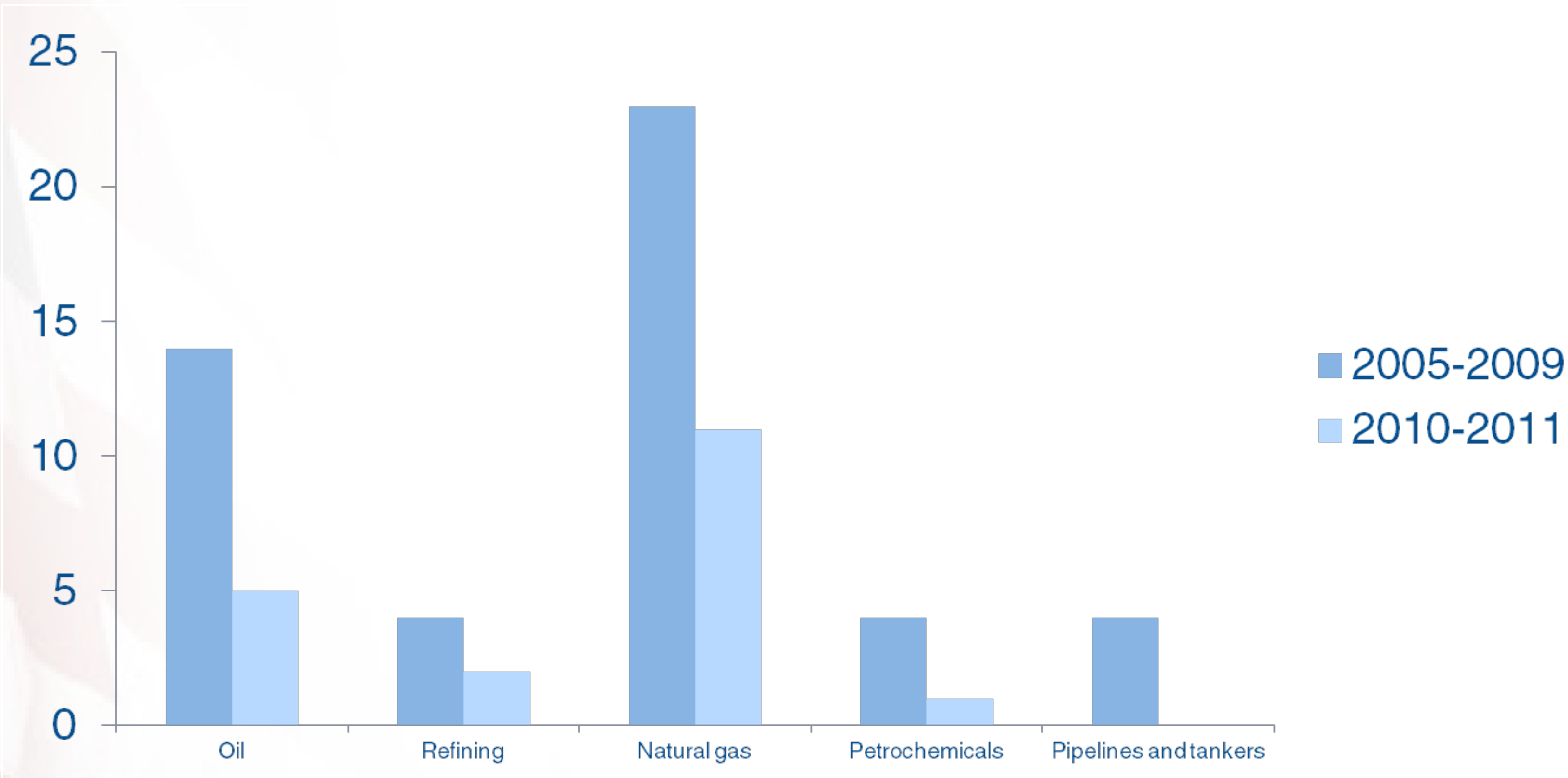
Activity type	Number of firms reported as having commercial activity between 2005 and 2009[a]	Number of firms reported as having commercial activity between 2010 and May 2011[a]
Oil exploration and production	14	5
Refining capacity	4	2
Natural gas	23	11
Petrochemicals	4	1
Pipelines and oil tankers	4	0

Source: GAO analysis of open source information.

[a]Multiple companies reported having commercial activity in more than one oil and gas development activity.

Similar information arranged in columns and rows, so that a reader can easily scan the information.

Amounts: Clarity of a Bar Chart



Streamline Text with a Figure

The Corrosion Office Used a Rigorous Multistep Process to Select Projects for Funding

For the fiscal year 2011 project review and selection, we observed that the Corrosion Office used a rigorous multistep process to determine if proposed projects were acceptable for funding.

- **Step 1:** In mid-June 2010, the military departments submitted 81 CPC project proposals to the Corrosion Office, as shown in table 1 above. At this point, Corrosion Office support staff assembled the project plans into binders for review by the project selection panel convened by the Director of the Corrosion Office. The fiscal year 2011 panel had five members: the Director, Corrosion Office (chair); Associate Director, Materials and Structures, Office of the Director, Defense Research & Engineering (vice-chair); and an official from each of the following organizations within the Office of the Under Secretary of Defense (Acquisition, Technology and Logistics): Defense Acquisition University; Installations and Environment; and Logistics and Materiel Readiness, Maintenance Policy and Programs.[18]

- **Step 2:** In mid-July 2010, 2 weeks after project information was provided to the panel, the panel members assembled for their preliminary evaluation of the proposals. This preliminary evaluation, which we observed, was conducted at a meeting immediately prior to the annual DOD Corrosion Forum and resulted in projects being designated as either a "go" (meaning that the projects are deemed acceptable in their current form) or a "no go" (meaning that the projects require additional information or changes in scope to be acceptable to the panel). We observed that the panel used criteria for this preliminary evaluation that are not made available to the submitters of project proposals and are different from those used for final project selection later in the process.[19]

- **Step 3:** Following the preliminary evaluation and during the Corrosion Forum, the panel held individual feedback sessions with project managers from the military commands, such as Naval Air Systems Command, Army Aviation and Missile Command, and Air Force Civil Engineer Support Agency, so feedback could be done in person. The panel provided feedback on each project, regardless of whether it was designated as a "go" or "no go." A panel member told us that the panel provided feedback on all projects so that project managers could address—if they choose to do so—any perceived weaknesses in their "go" projects and improve their ranking in the final evaluation, as well as revise the "no go" project submissions. Following the feedback, the project managers had three options: prepare and submit information addressing the feedback provided by the panel, re-submit project proposals in their original form, or remove projects from consideration for that year's funding process. Project managers told us that they sometimes decide to remove their "no-go" projects from consideration and that the military departments may implement such projects using other funding. A project selection panel member told us that if a project manager decided to modify a project proposal to address the panel's feedback, this modified proposal was due to the Corrosion Office no later than 2 weeks after the feedback session. Upon receipt of any revised proposals, the panel conducted another review of all proposals (original and resubmitted), which involved each panel member independently scoring the projects on judgmental criteria and providing written comments.[20]

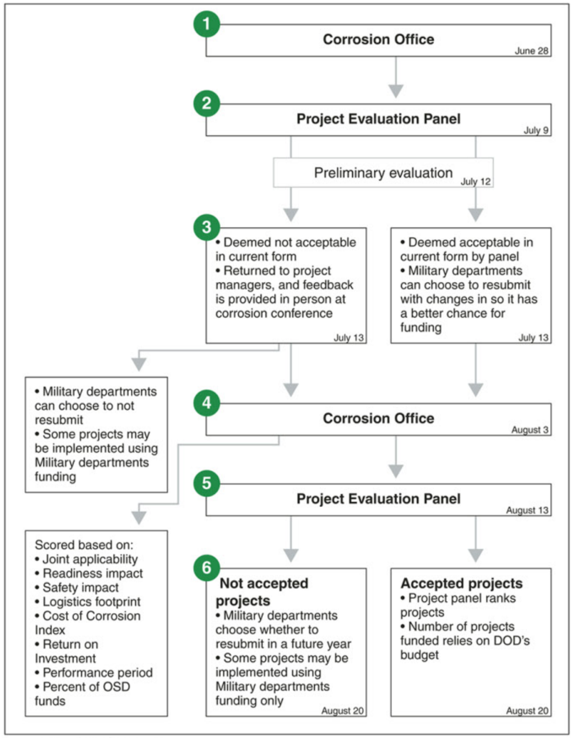
- **Step 4:** In mid-August 2010, Corrosion Office support staff used an analytical tool to rank the projects based on the average of the scores recorded by each panel member for eight criteria: the five judgmental criteria above and three quantitative criteria—ROI, Corrosion Office funding as a percentage of total project cost, and the project performance, or implementation, period.

- **Step 5:** Following the ranking of projects using the analytical tool, the selection panel reconvened for a final evaluation of the projects. The panel arranged the ranked list that resulted from the analytical tool described above into four categories: best, acceptable–prioritized for funding, acceptable–not prioritized, and not acceptable. According to the staff, the "best" projects would likely all be funded, the "acceptable–prioritized for funding" projects would be funded by priority until the Corrosion Office funding is exhausted. Corrosion Office support staff informed the panel that, based on historical funding levels, they anticipated having $7 million in available funding

[18] The panel member from Logistics and Materiel Readiness, Maintenance Policy and Programs did not participate in the project selection meetings we observed.

[19] The criteria used for the preliminary evaluation include whether the proposed project requires greater than $500,000 of Corrosion Office funds to complete, uses similar technology to a previously approved project, or is anticipated to take more than 2 years to complete. The preliminary evaluation did not consider the joint applicability of the project, but this was a criterion in the final project evaluation.

[20] The judgmental criteria are: joint applicability, readiness impact, safety impact, logistics benefits, and anticipated contribution of the project to reducing the cost of corrosion. Corrosion Office officials told us that they believe the criteria to be clearly identified in the DOD *Corrosion Prevention and Mitigation Strategic Plan.*

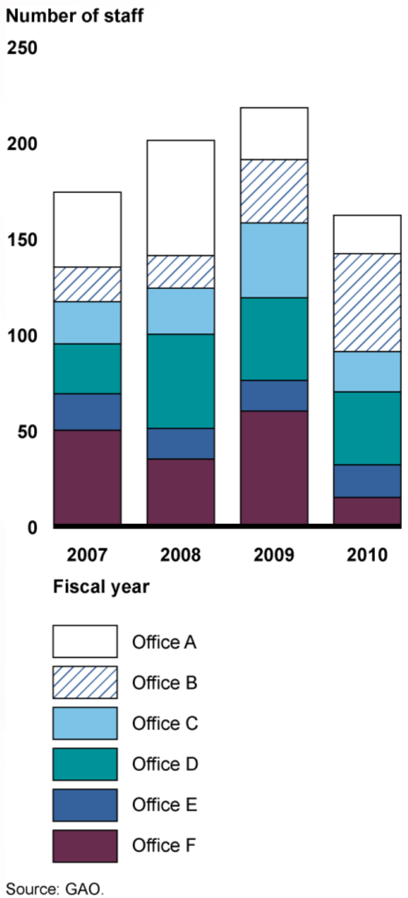


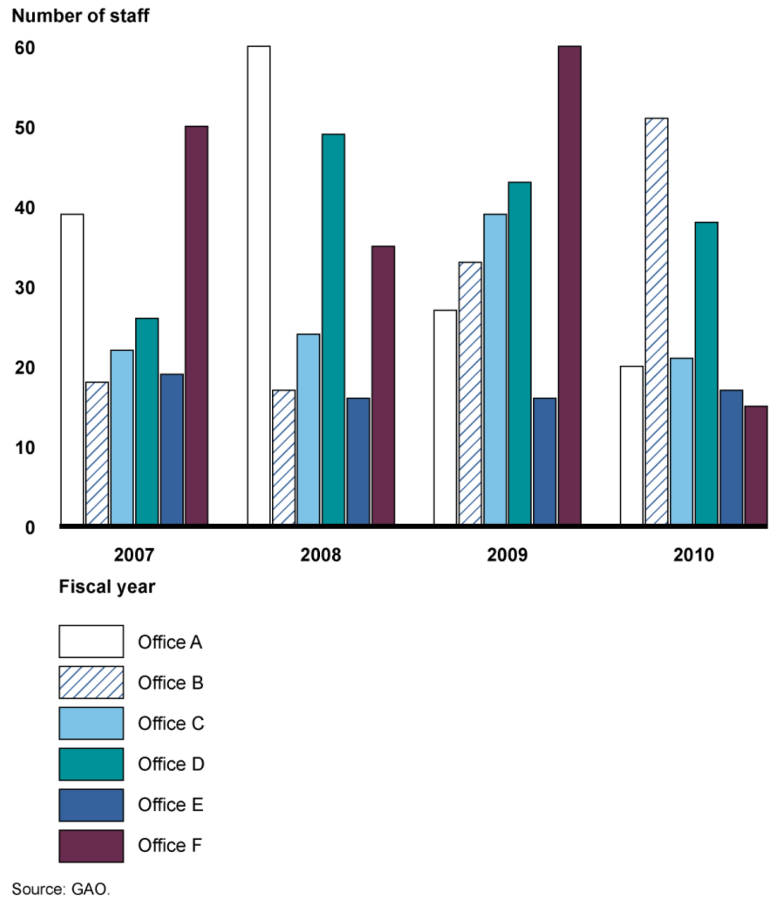
Source: GAO analysis of OSD selection process.

Six Broad Categories: Graphic Options

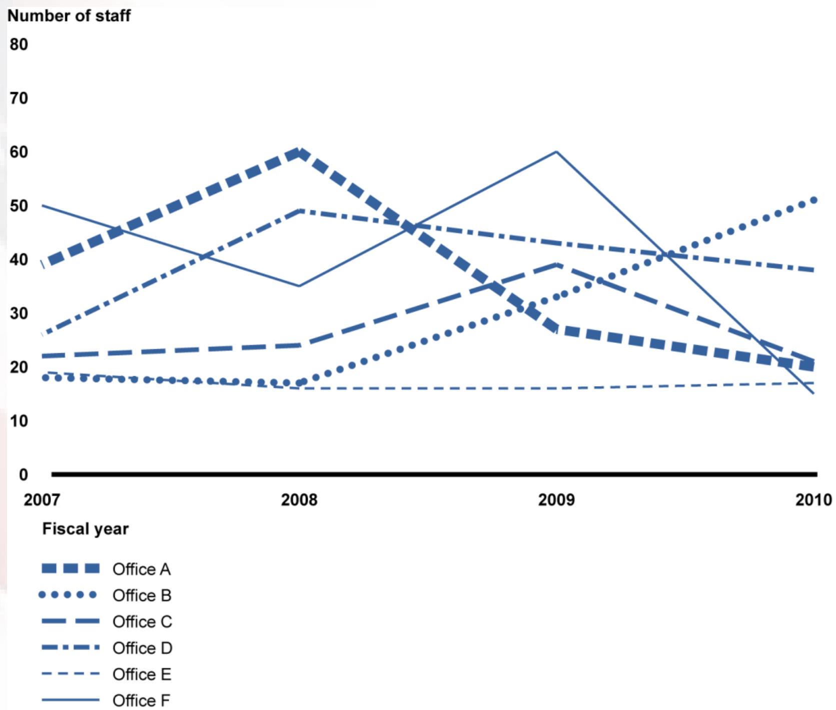
Categories	Graphic Options
1. Amounts	Tables and charts (for example, bar, line, area, range, and pie)
2. Geography	Map
3. Explanatory	Organizational chart, illustration, photograph, and matrix
4. When	Time line
5. Cause and effect	Flow chart
6. Analysis, deduction, and prediction	Any combination of the above

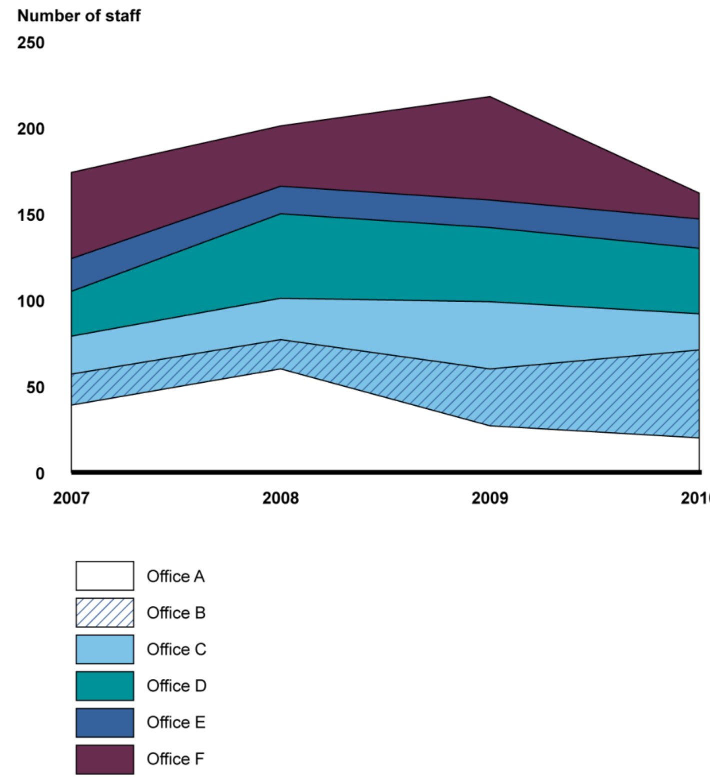
Amounts: Stacked & Cluster Bar Charts





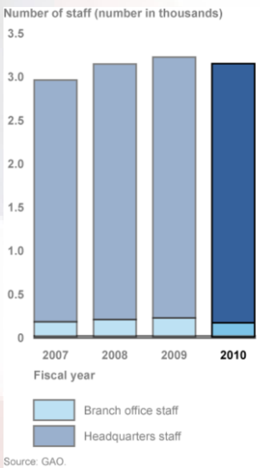
Amounts—Trends: Line & Mountain Charts

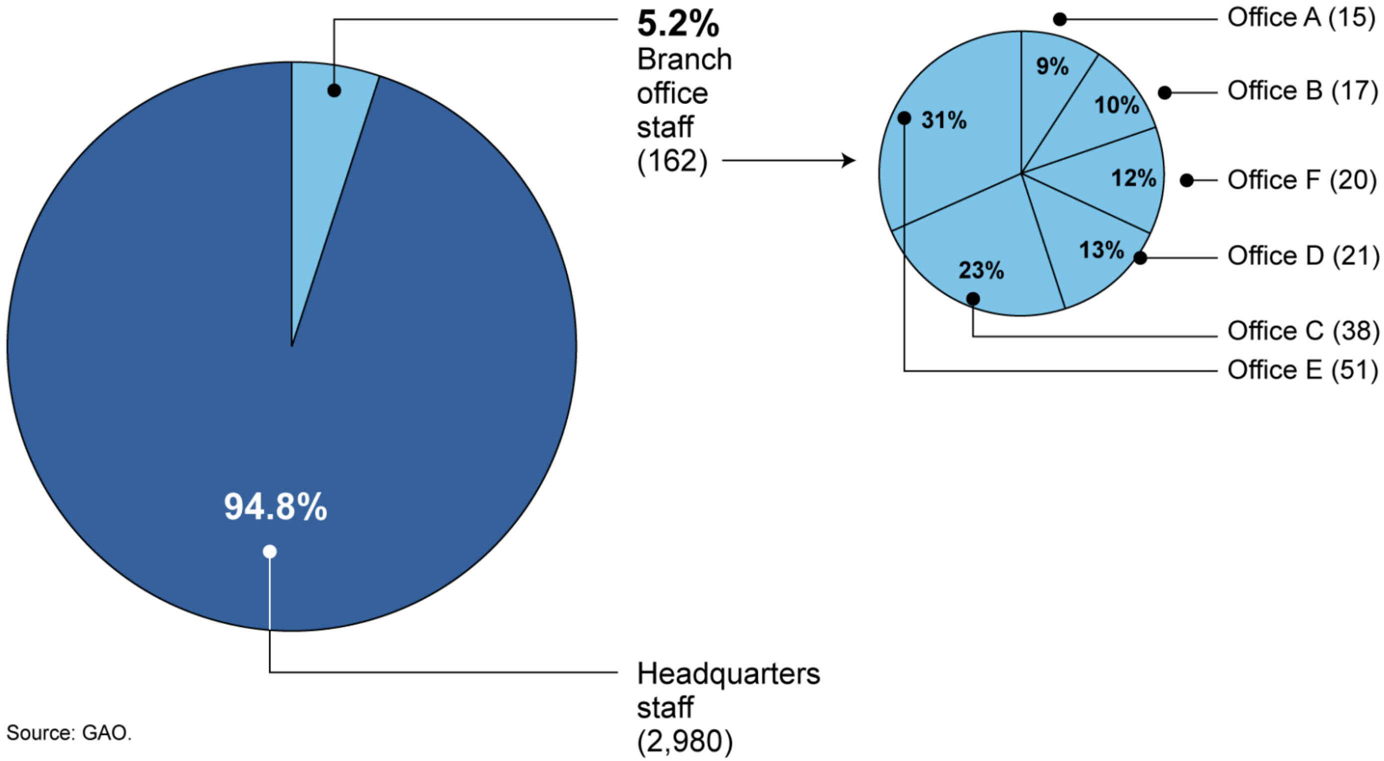




Source: GAO.

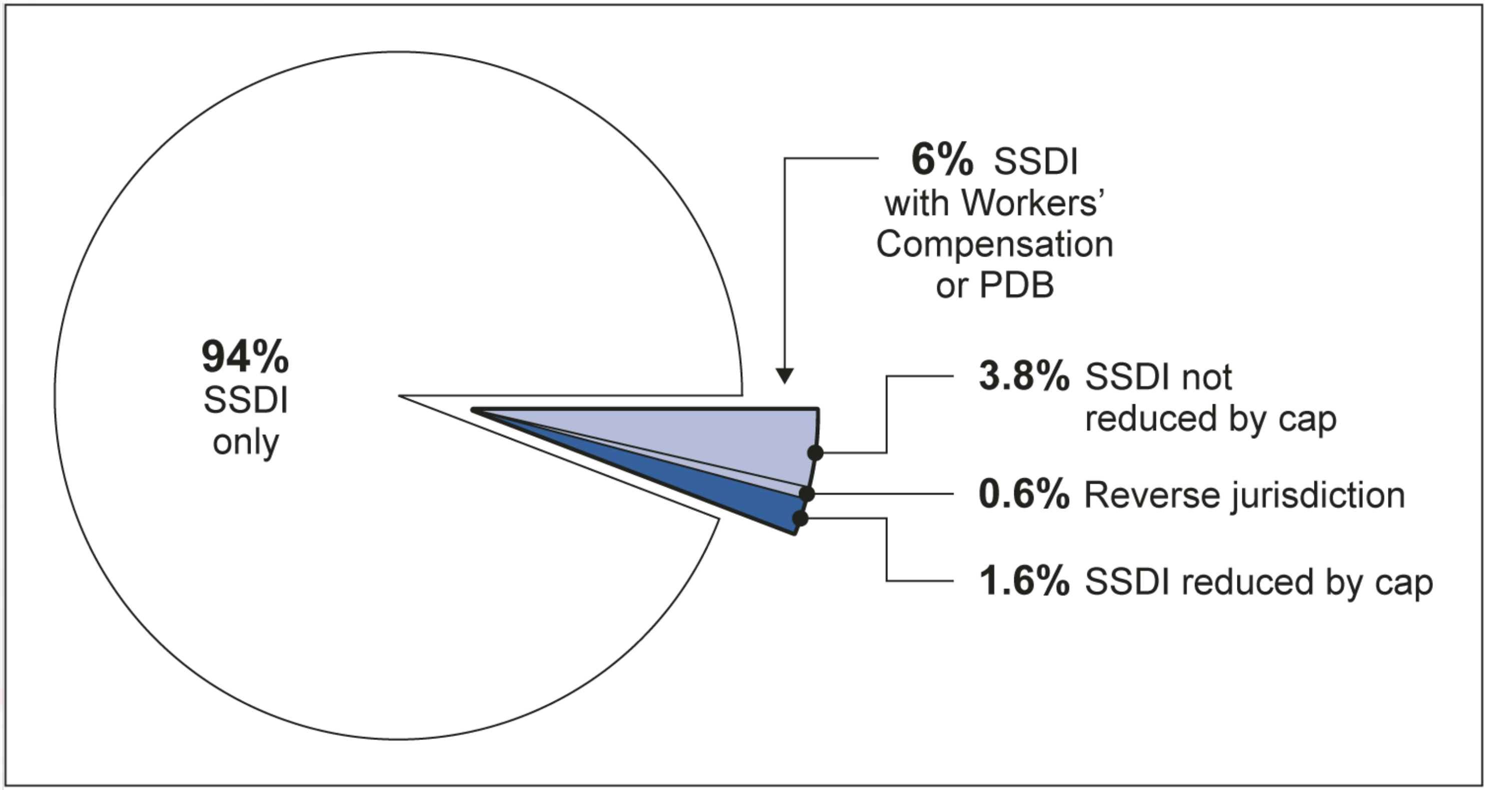
Amounts—Parts of a Whole: Pie Chart





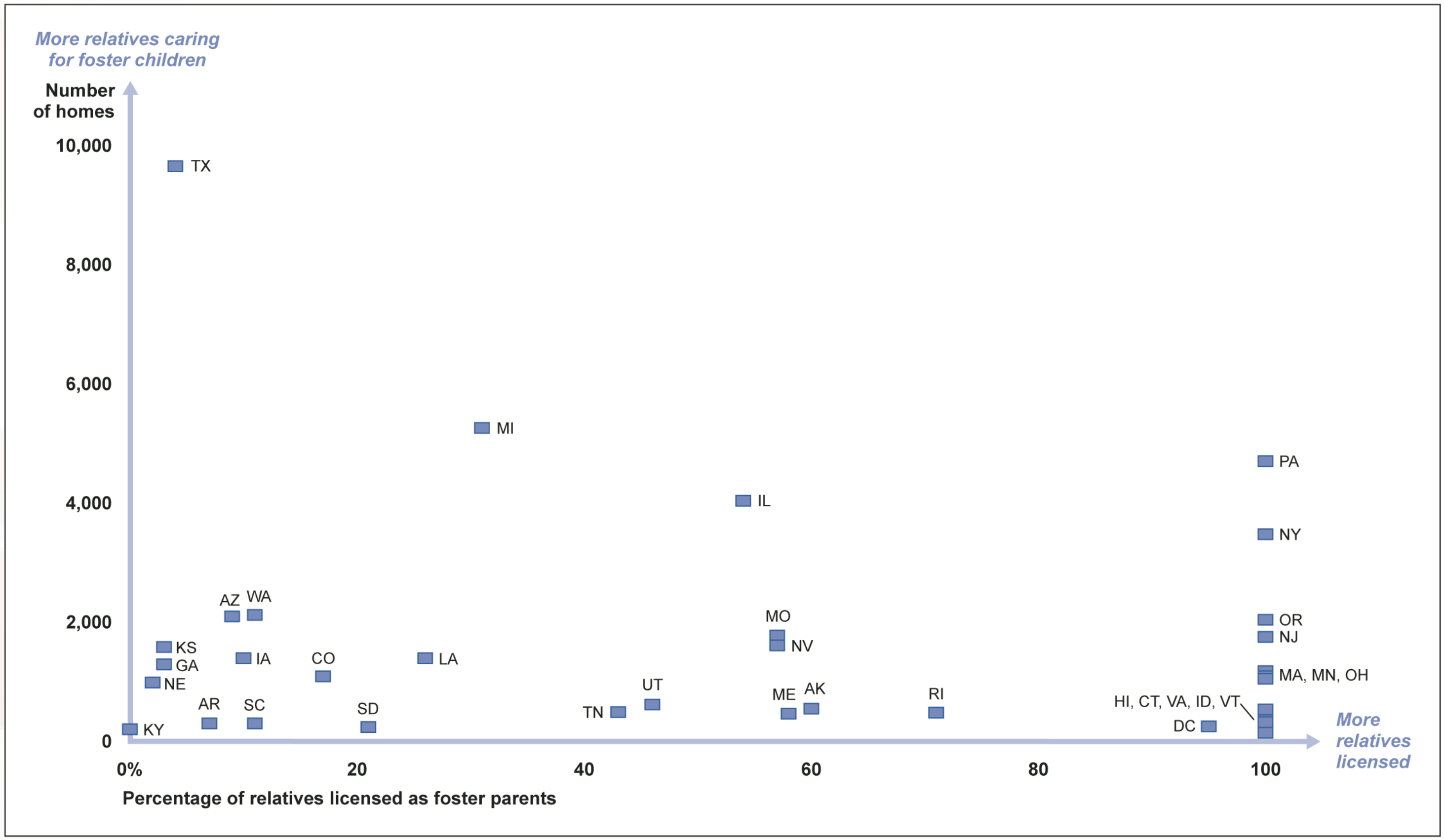
Source: GAO.

Pie Example: Exploding Out One Slice



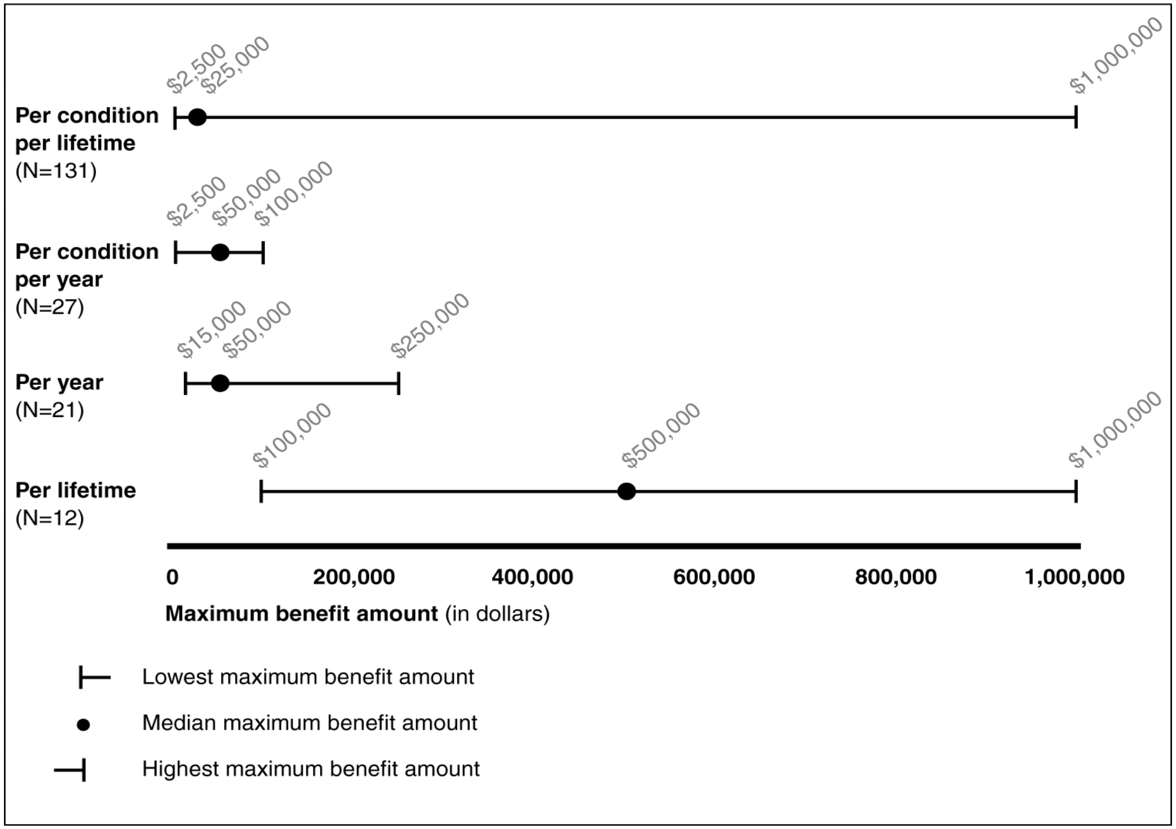
Source: Analysis of SSA data.

Scatter Example: Funding for Foster Families



Source: GAO survey of state child welfare administrators.

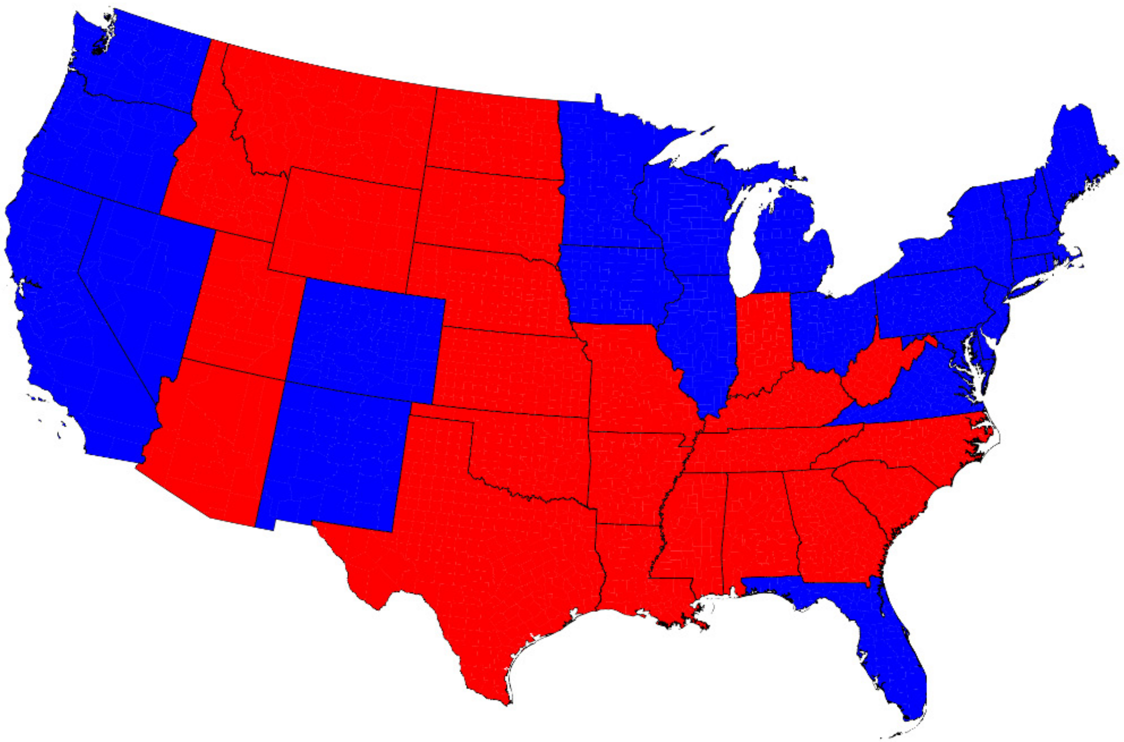
Amounts: Range Chart



Source: GAO random sample of 340 colleges.

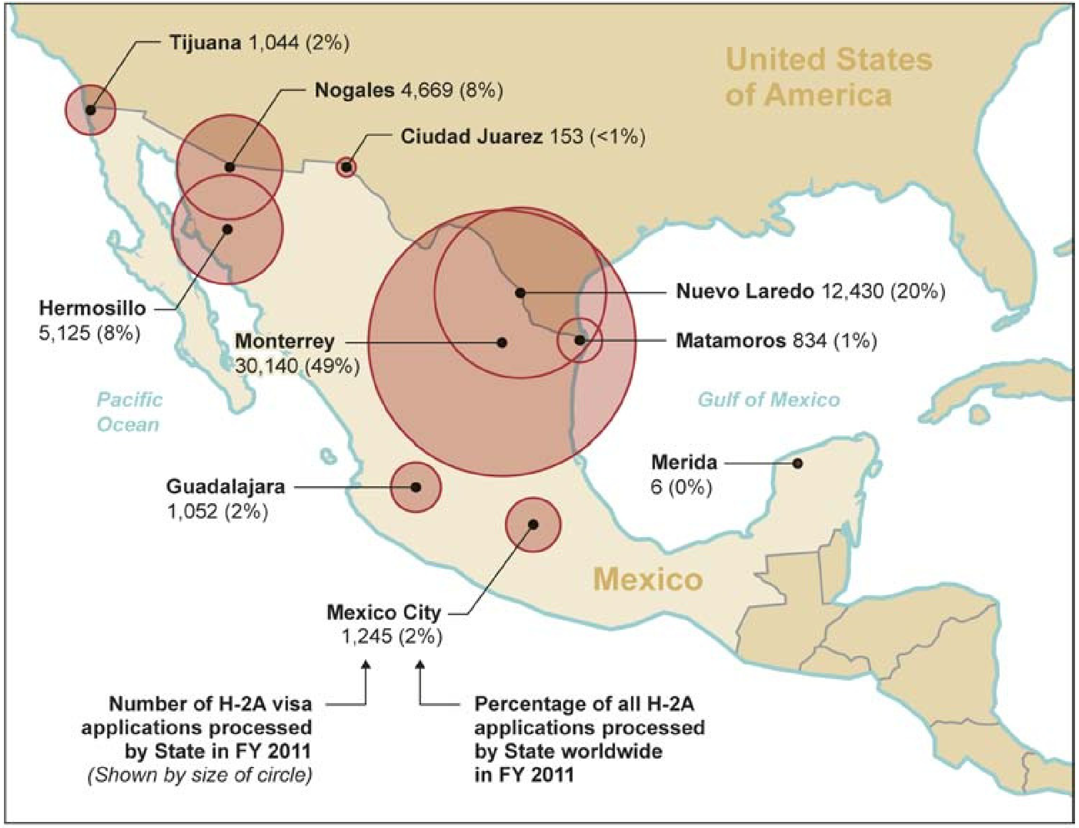
Geography: Maps

2012 Presidential results by state



Geography: Maps

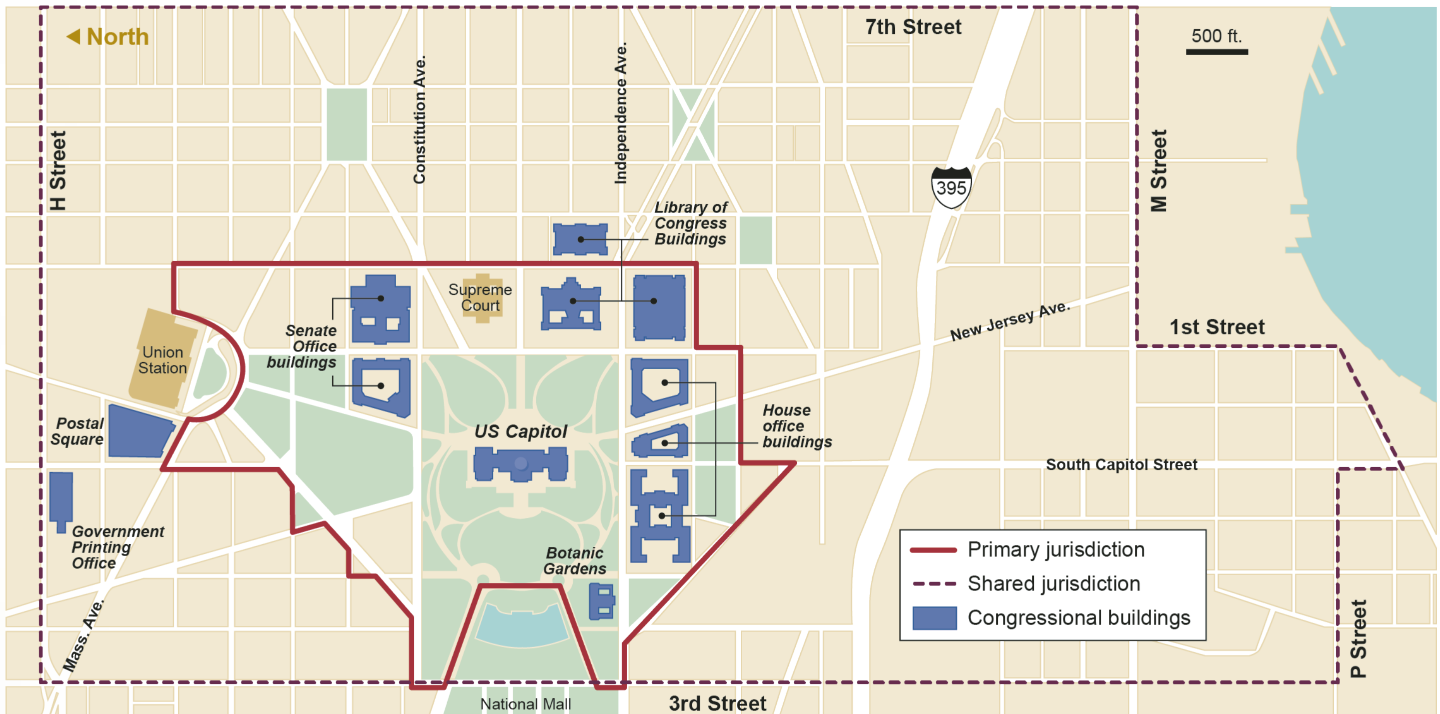
Maps combined with a chart "locate" the numbers



Sources: GAO analysis of State via issuance data; National Atlas of the United States (base map).

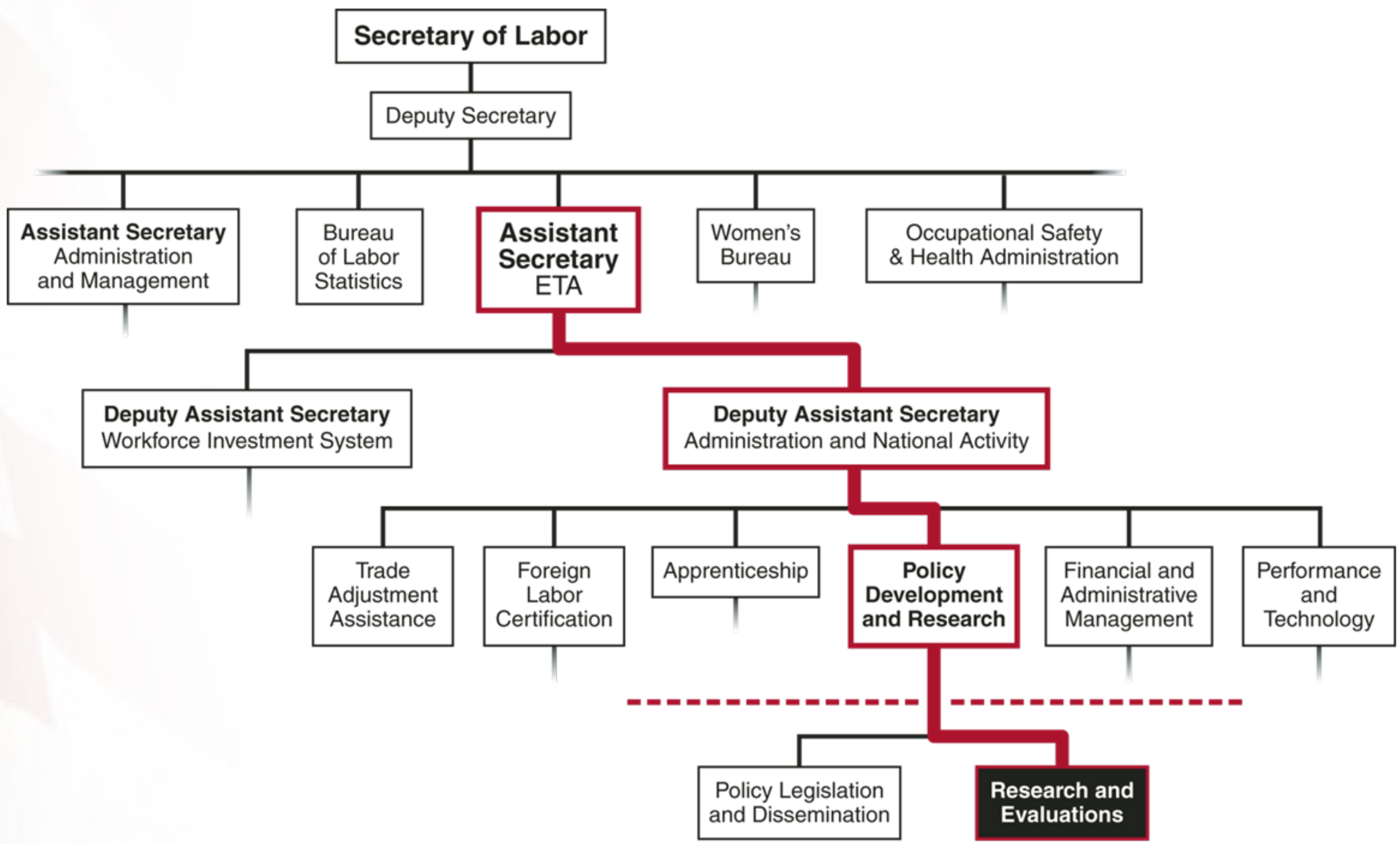
Geography: Maps

Maps show boundaries and physical relationships

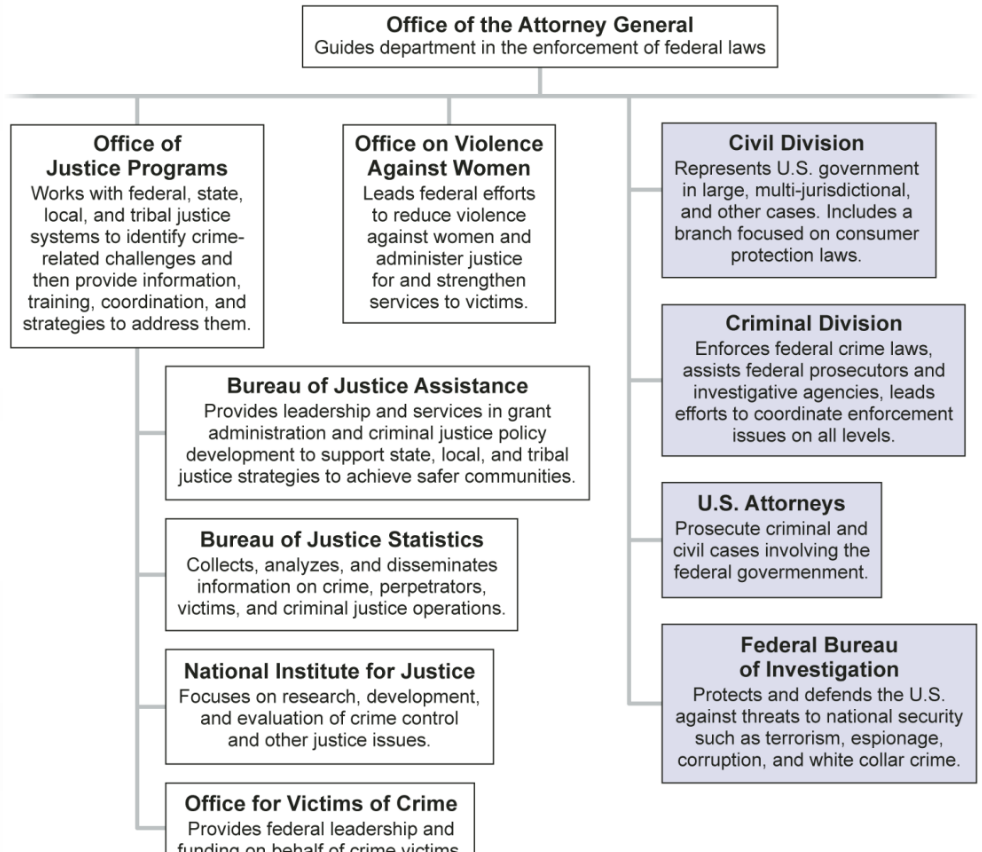


Source: USCP.

Explanatory: Organizational Chart

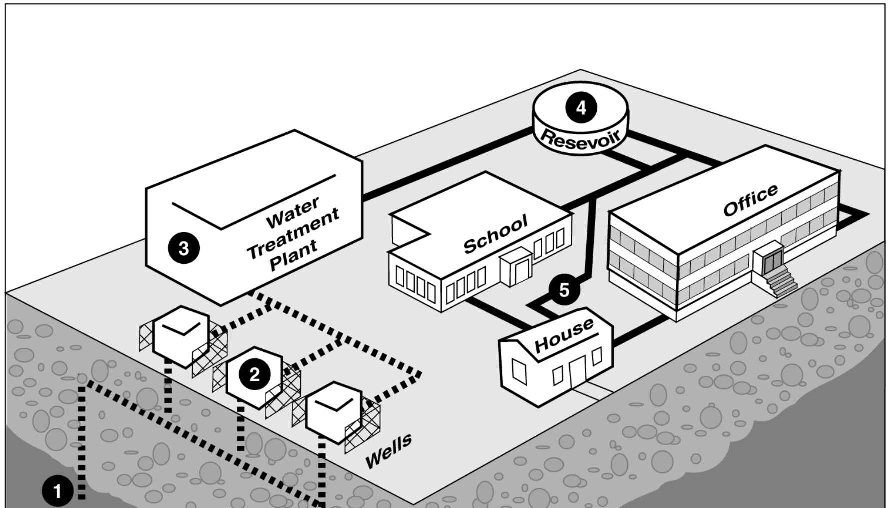


Org Charts: Highlight the Point for Readers



Source: Justice.

Explanatory: Illustration



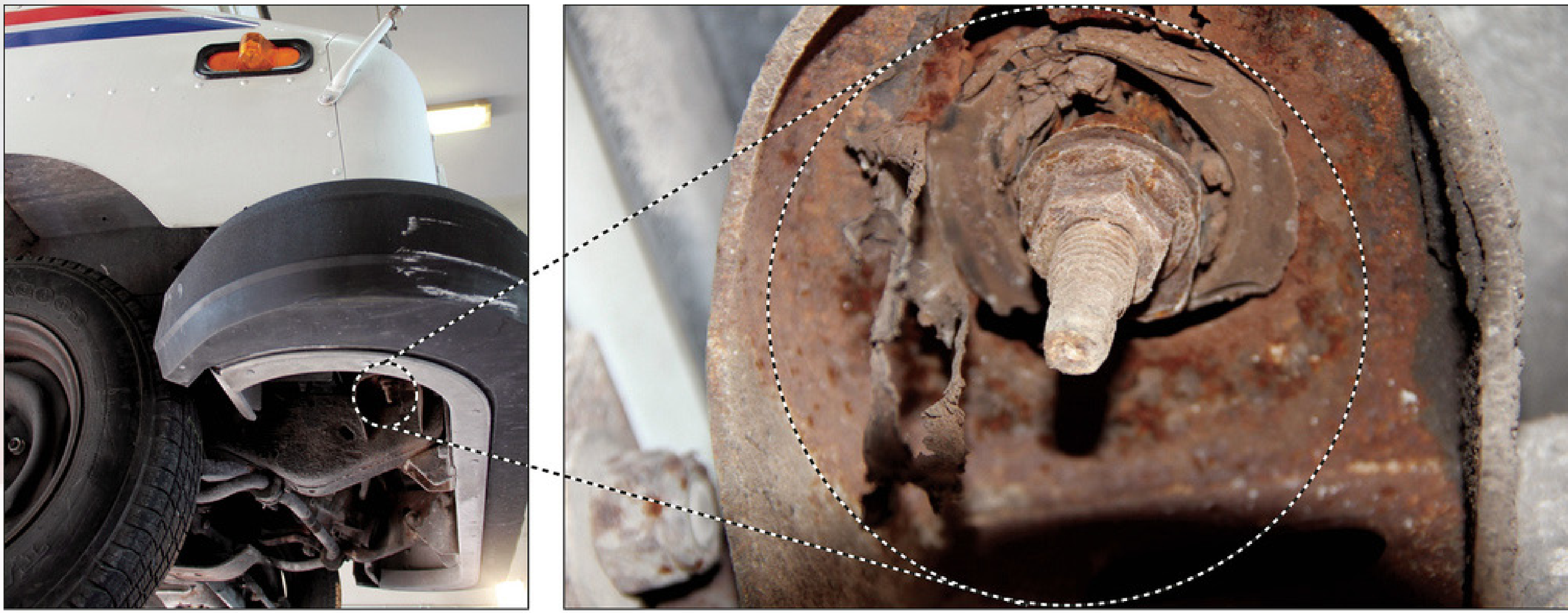
❶ The drinking water at Camp Lejeune is obtained from groundwater pumped from a freshwater aquifer located approximately 180 feet below the ground. ❷ Groundwater is pumped through wells located near the water treatment plant. ❸ In the water treatment plant, the untreated water is mixed and treated through several processes: removal of minerals to soften the water, filtration through layers of sand and carbon to remove particles, chlorination to protect against microbial contamination, and fluoride addition to help prevent tooth decay. ❹ After the water is treated, it is stored in reservoirs and elevated storage tanks. ❺ When needed, treated water is pumped from the reservoirs and tanks to facilities such as offices, schools, or houses on the base.

▪▪▪▪▪▪ Untreated water

▬▬ Treated water

Source: GAO, Art Explosion, and Marine Corps Base Camp Lejeune.

Explanatory: Photograph



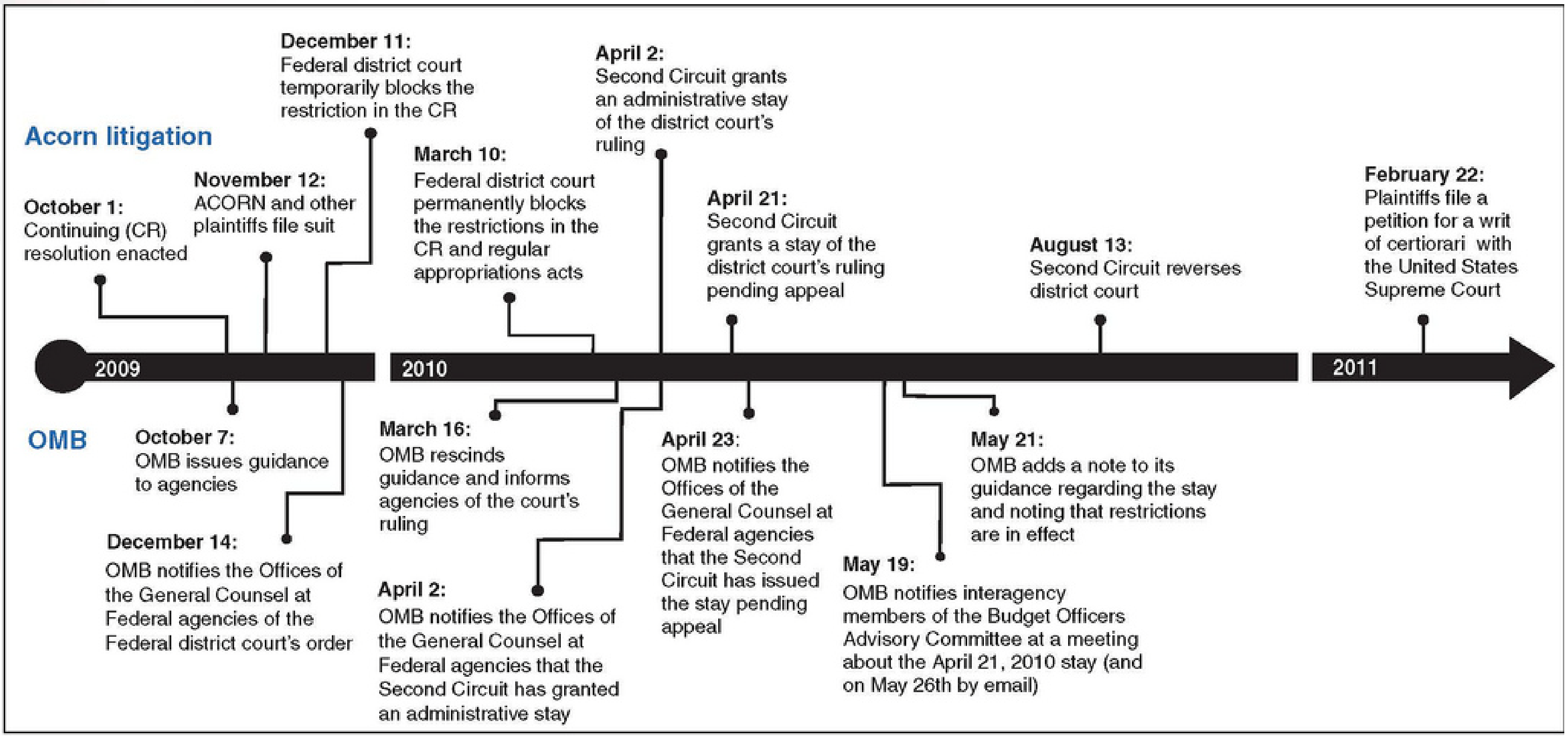
Source: GAO.

Explanatory: Matrix

Delivery options	Australia	Canada	Finland	Germany	Sweden	Switzerland
Traditional post office	✓	✓	✓	✓	✓	✓
Home delivery	✓	✓	✓	✓	✓	✓
Cluster boxes		✓	✓		✓	✓
24-hour automated parcel locker				✓		
Pick up at nonpost-owned retail counter	✓	✓	✓	✓	✓	✓
Digital and hybrid mail	✓	✓	✓	✓	✓	✓

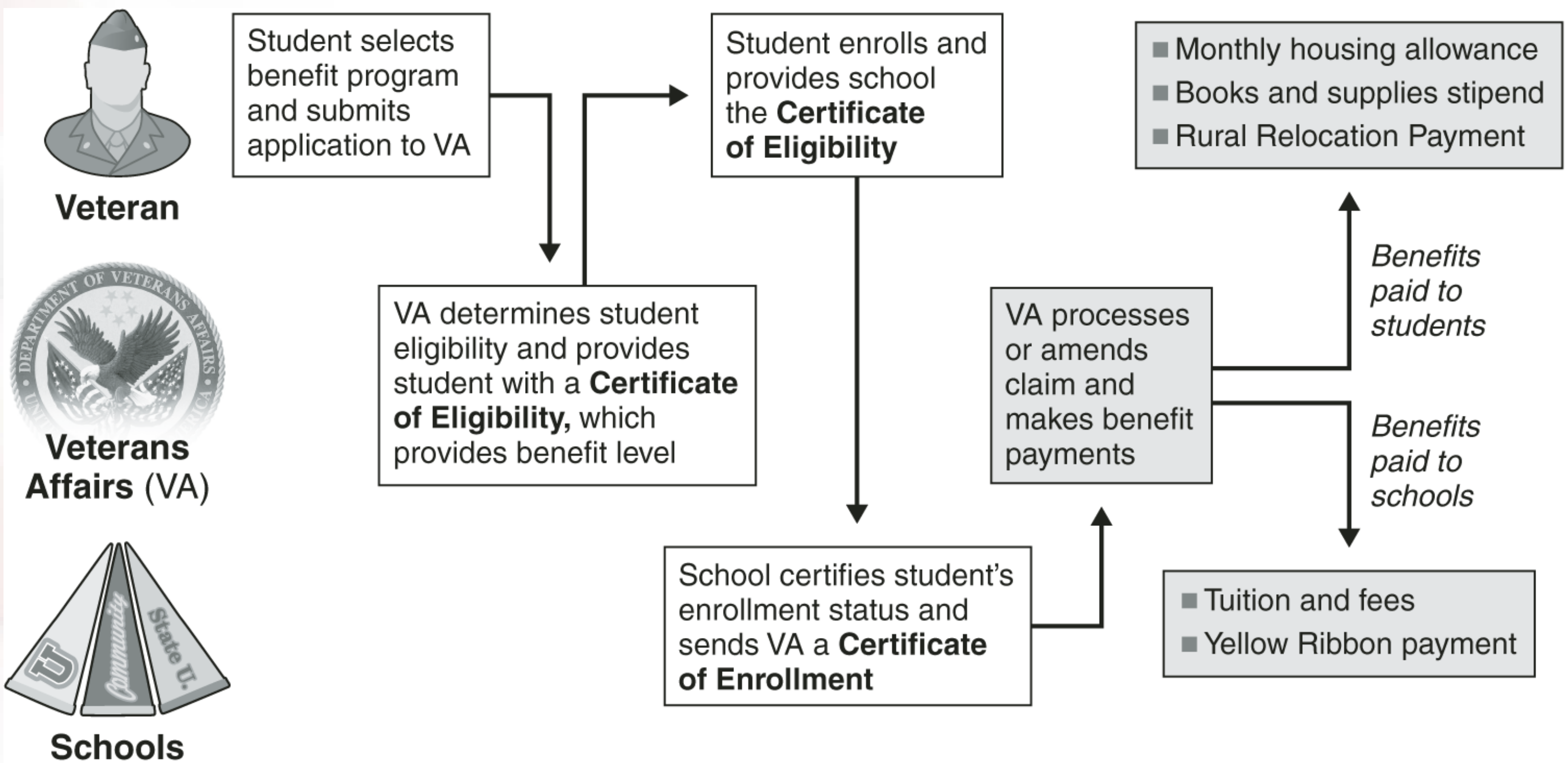
Sources: GAO (analysis); logos used with permission from Australia Post, Canada Post, Itella, DP DHL, Posten AB, and Swiss Post.

When: Timeline



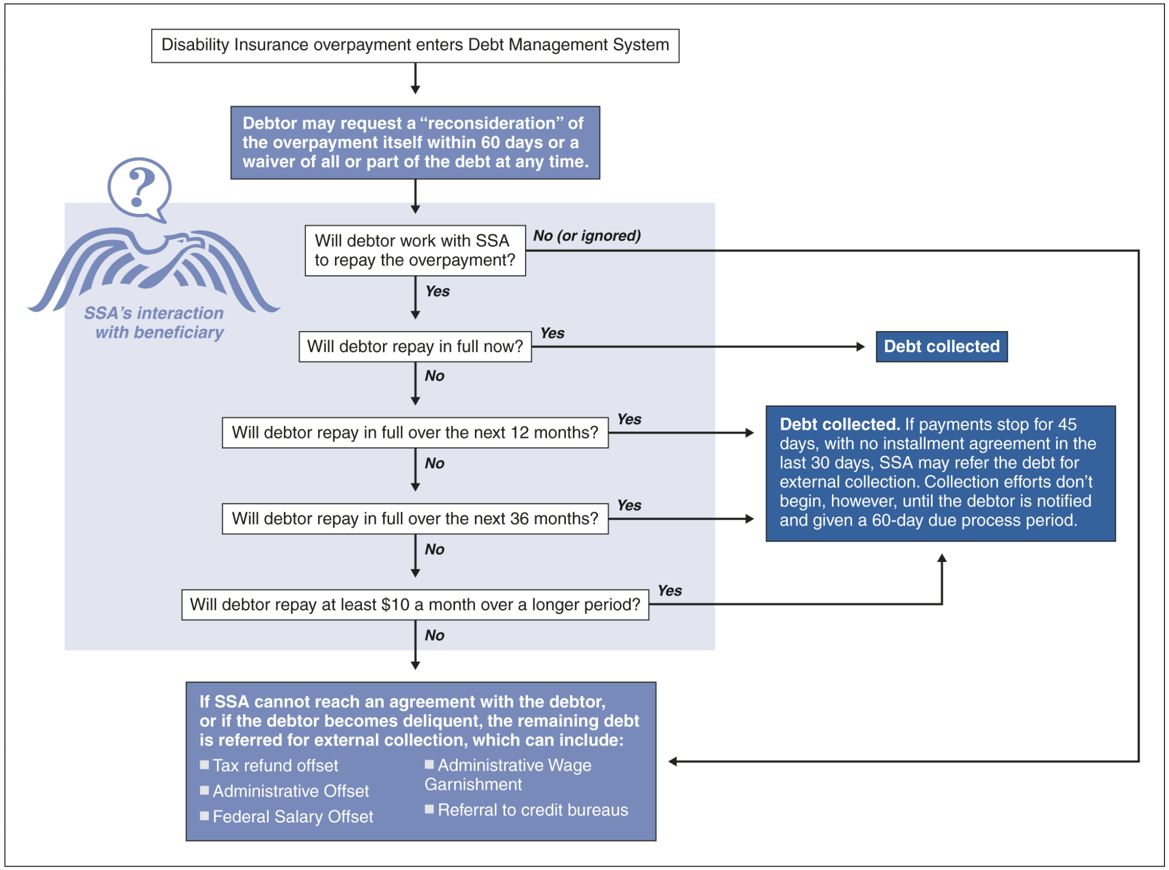
Source: GAO analysis of court and OMB documents.

Flow Chart: VA Payments (showing actors)



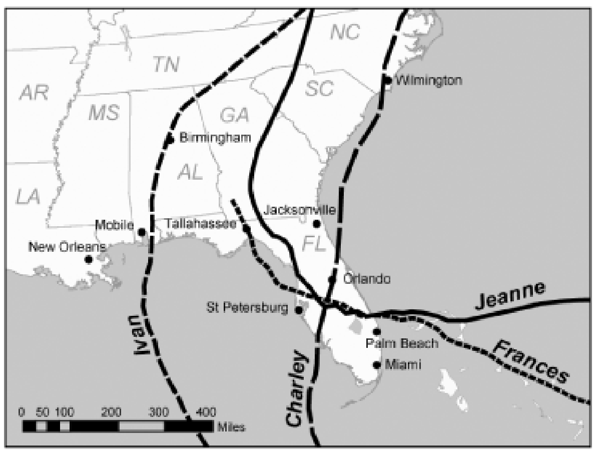
Source: GAO analysis of VA documents; GAO (images).

Cause and Effect: Flow Chart



Source: GAO analysis of SSA policies.

Analysis, Deduction, and Prediction: Combination



Four hurricanes made landfall in the United States within a 6-week period in 2004. The storms caused an estimated $27 billion in insured losses in the United States and the Caribbean. As a whole, the 2004 hurricane season is the second most expensive insurance event in U.S. history, behind the September 11, 2001, terrorist attacks. The chart below details the impact of the four storms on Florida.

	Date of landfall	Saffir-Simpson Hurricane Scale	Estimated insured losses (in billions)	Number of claims reported (in thousands)
Charley	13 Aug. Sept.	1 2 3 **4** 5	$7.885	449
Frances	5 Aug. Sept.	1 **2** 3 4 5	4.801	504
Ivan	16 Aug. Sept.	1 2 **3** 4 5	3.914	197
Jeanne	25 Aug. Sept.	1 2 **3** 4 5	4.153	393

Sources: Benfield Group Limited, Florida Insurance Council, Florida Office of Insurance Regulation, Guy Carpenter, Swiss Re, and Risk Management Solutions (map).

Identifying Graphic Opportunities

1. Scan for key words (for example: “data,” “location,” “how,” “process,” “why,” “what,” and “frequency.”
2. Jot down initial ideas for graphics (for example: what type of graphic or combination of graphics, such as a map with a bar chart).
3. Think about next steps or questions (for example, requesting information or sources from an agency, finding existing models or samples, and defining the range of values or the time period that needs to be covered).

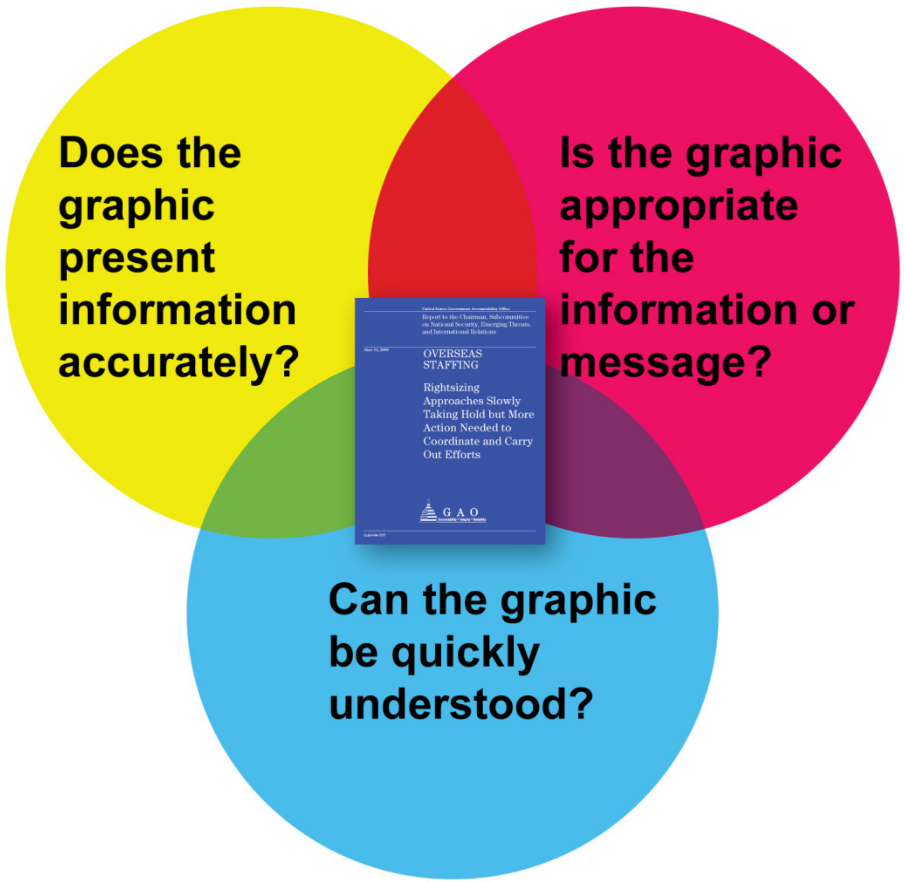
FOCUS HERE:
(first two columns of design matrix)

Researchable Question(s)	Information Required and Source(s)
1. What is known about the nature and prevalence of mortgage foreclosure rescue and loan modification schemes? a. What are the federal laws prohibiting these schemes? b. What is known about the nature of these schemes? c. What data exist, if any, on the incidence and geographic location of these schemes?	1a. Key federal laws and regulations on prohibited practices, such as • FTC Act provisions concerning unfair and deceptive practices and • prohibitions on charging fees in advance for related services. (*Sources*: Statutes and regulations identified by key federal agencies and “selected” states.) 1b. Descriptions of typical foreclosure rescue and loan modification scams: • how they work • perpetrators • potential victims (*Sources*: Enforcement cases--filed by the FTC, DOJ, and selected U.S. Attorneys and State AGs; Subject matter experts at federal agencies—FTC, HUD, DOJ, Treasury-FinCEN, selected U.S. Attorneys and State AGs, and national non-profits, including HPF and NW, Lawyer’s Committee, as well as reports, letters, and testimonies that all of the above may have published.) 1c1. Identify available data, such as: • complaint data (*Source*: FTC’s Consumer Sentinel, selected state agencies, and national nonprofits—HPF & Lawyer’s Committee). • enforcement data (*Source*: DOJ, FTC, and U.S. Attorneys & State AGs from selected states) • investigative data (*Source*: Federal investigative data—at FinCEN, FBI, and SIGTARP) 1c2. Expert opinion on what the data may say about prevalence. (*Source*: Data providers identified in 1c2.) descriptions in 1c2) 1c3. **Criteria** for assessing if available data are reliable.

Example

1. **Key words:**
“geographic location”
2. **Initial ideas:**
A list, table, or map
3. **Next steps or questions:**
If we do a map, are Federal Trade Commission data at state or county level? Does FTC have data? Are the data current? Can we show multiple types of data (complaints, enforcements, investigations)–is any one type more important than the others? Can we show “incidence and geographic location” together?

Key Questions



Simplify Your Message

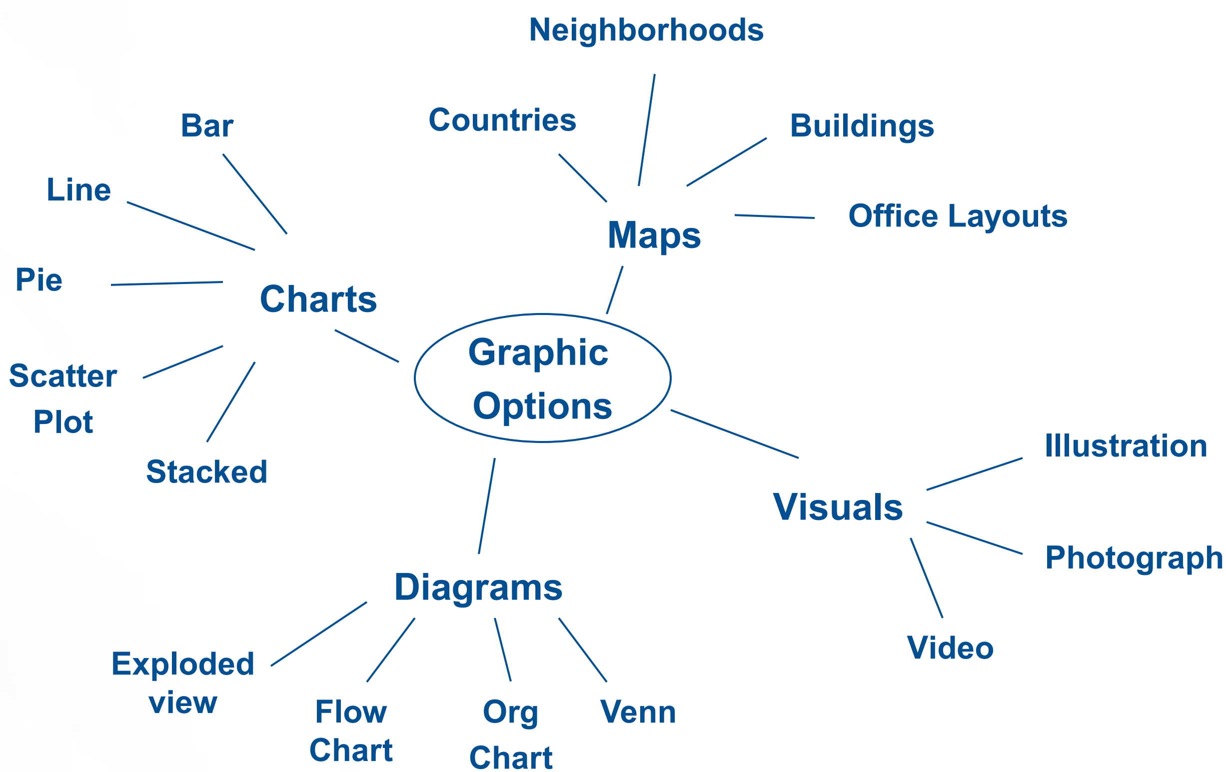
- Outline
- Mind map
- Write on walls

Outline

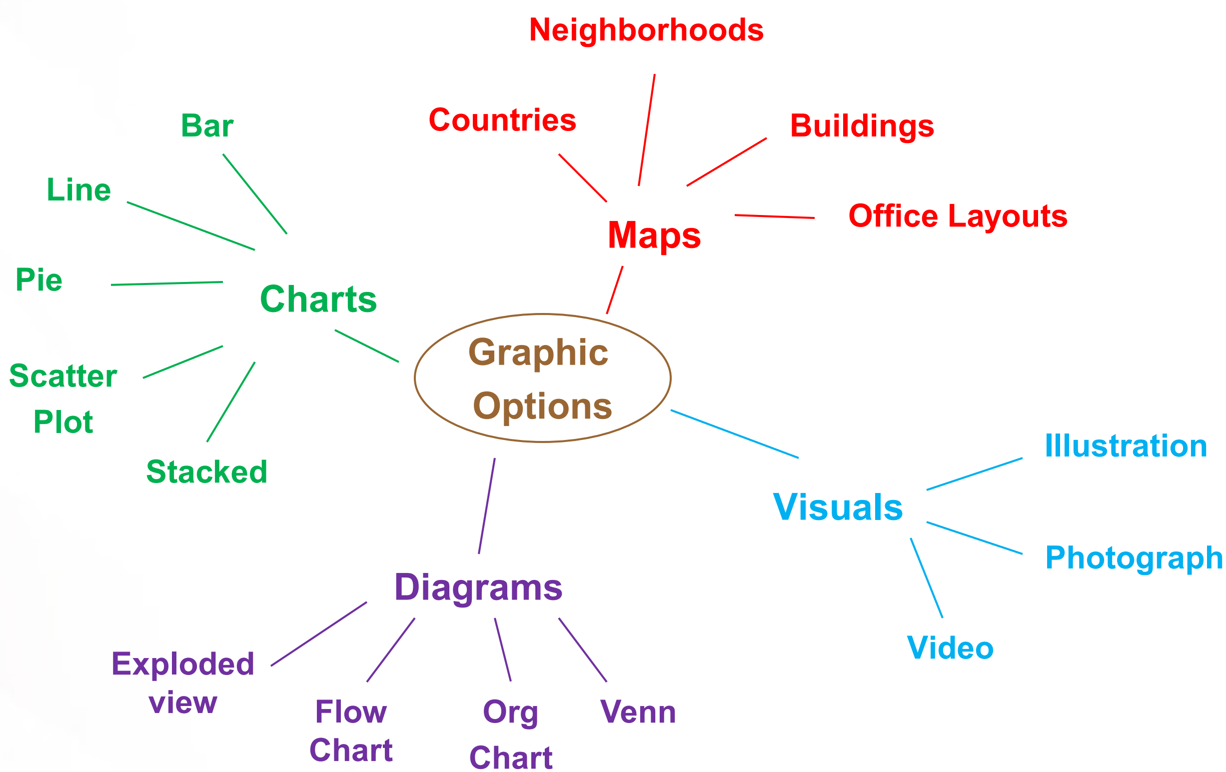
Objective 1

- **First most important finding**
  - Support or evidence for this finding
  - More support or evidence
  - More support or evidence
- **Second most important finding**
  - Support or evidence for this finding
  - More support or evidence
  - More support or evidence
- **Third most important finding**
  - Support or evidence for this finding
  - More support or evidence
  - More support or evidence

Mind Map


Neighborhoods
Bar
Countries
Buildings
Line
Maps
Office Layouts
Pie
Charts
Graphic Options
Scatter Plot
Stacked
Illustration
Visuals
Photograph
Diagrams
Video
Exploded view
Flow Chart
Org Chart
Venn

Mind Map

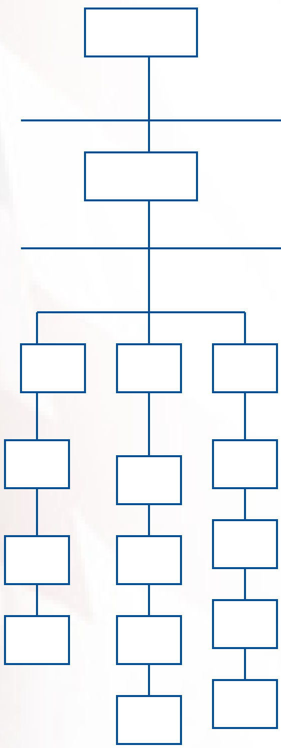

Neighborhoods
Countries
Buildings
Office Layouts
Maps
Bar
Line
Pie
Charts
Scatter Plot
Stacked
Graphic Options
Visuals
Illustration
Photograph
Video
Diagrams
Exploded view
Flow Chart
Org Chart
Venn

Write on Walls

•Objective 1	•Objective 2	•Objective 3

•Message statements

•Key points supporting message statements



•Supporting detail for each key point

These divisions allow us to organize visually

Write Concisely

- Use plain language
- Eliminate
    - wordiness
    - Redundancy
    - Acronyms
- Write strong topic sentences

Plain Language

Instead of	*Use*
Purchase	Buy
Reside	Live
Endeavor	Try

www.plainlanguage.gov

Eliminate Wordiness

Before

In order to ensure that we can draw a conclusion about the extent to which the lack of data point to problems in overseeing the effectiveness of the program, we should endeavor to assess the sufficiency of the database, notwithstanding the fact that the agency has in the not too distant past issued assurances of its quality. (57 words)

After

To adequately determine whether program oversight is possible with existing data, we should assess the database quality. (17 words)

Eliminate Redundancy

Before

Child Protective Services agencies collect information about children who are **maltreated** and the circumstances surrounding **maltreatment**—including children who died from it—to aid efforts to prevent **maltreatment**. (28 words)

After

Child Protective Services collect information about children who are harmed by and die from **maltreatment**. This information, such as the circumstances surrounding **maltreatment**, is used for prevention efforts. (28 words)

Eliminate Acronyms

Before

Training for VSRs, RVSRs, and DROs in the VBA has been a recurring problem according to VA’s OIG.

After

Providing adequate training to veterans service representatives—including those who rate claims—and decision review officers, has been a recurring problem according to VA’s inspector general.

Write Strong Topic Sentences

Weak

The cost estimates for the four projects lacked credibility.

Strong

The cost estimates for the four projects we reviewed did not exemplify the four characteristics of high-quality cost estimates as established by best practices—credible, well-documented, accurate, and comprehensive.

Resources

- The Back of the Napkin: Solving Problems and Selling Ideas with Pictures
  Dan Roam

- The Elements of Style
  William Strunk & E.B. White

- The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts and Figures
  Dona M. Wong

- Kate van Gelder, (206) 287-4815, vangelderk@gao.gov