BENCHMARKING

A practical guide to help governments compare themselves to other cities.

What Works Cities | Bloomberg Philanthropies
# Table of Contents

- Introduction 1.1
- What is Benchmarking? 1.2
  - Internal Benchmarking 1.2.1
  - External Benchmarking 1.2.2
  - The Benchmarking Mirage 1.2.2.1
- What is a Proxy Measure? 1.3
- Useful Links 1.4
Benchmarking

Benchmarking is the process of continuously comparing and measuring an organization against its own past performance or comparing its performance to that of another comparable organization. A benchmarking practice provides information that will help an organization take action to improve its own performance. Governments of all levels of capability can leverage benchmarking in their performance management practice. This short guide explains benchmarking using practical examples and provides quick links to helpful benchmarking resources. The guide also describes the related practice of using "proxy measures," which are another useful tool in the performance management toolbox.

The Center for Government Excellence at Johns Hopkins University (GovEx), a partner in Bloomberg Philanthropies’ What Works Cities initiative, created this practical guide to help governments advance performance management practices by leveraging benchmarking and proxy measures. The guide includes: definitions of commonly used terms; real examples of benchmarking and proxy measures in practice; and a list of useful resources.
What is Benchmarking?

Benchmarking is a practice that attempts to answer one of life's most ubiquitous questions: *what does success look like?* As individuals, we benchmark every day. We look around at our friends and colleagues and say things like this to ourselves:

- I should be making more money because my coworker is making more money and we are the same age.
- I should own a house by now, because all my friends own houses already.

As individuals, this isn't a very healthy practice. But as a government organization, benchmarking can prove very constructive because there are certain goals that are universally accepted as in the public's best interest. For example, governments want to:

- help individuals, communities and businesses to thrive and prosper
- ensure people are healthy, safe and secure
- educate children with knowledge and skills that will help them contribute and succeed
- be good stewards of limited shared resources
- ensure people with access and functional needs are safe, included and empowered

These are just a few examples of some universally accepted government interests, and it is okay for governments to compare themselves to one another to assess how they are performing, relative to everyone else. This is called **External Benchmarking**. But comparing one organization to another is not the only way to benchmark. Simply comparing current performance to past performance is a good place to start. This is called **Internal Benchmarking**.
Internal Benchmarking

Internal Benchmarking is simply comparing yourself . . . to yourself. If an organization wants to know what success looks like, they should first look to data on what they have been capable of in the past. The most iconic example of internal benchmarking is COMPSTAT.

COMPSTAT is the New York City Police Department's (NYPD) accountability process. It is both a management philosophy and a tool for police department managers to identify spikes in crime using comparative statistics (or internal benchmarking). One of the main features of COMPSTAT is its relentless commitment to benchmarking current performance against the past. In Figure 1 below, which is taken directly from the NYPD website, the department compares crime this week to the same week last year. It also calculates percent changes over the past two, five and 22 years. For additional context, the NYPD provides a historical perspective, comparing current performance to performance in 1990, 1993, 1998, and 2001.

Figure 1
Internal Benchmarking is a very common practice among police departments. SeaStat is the Seattle Police Department's (SPD) accountability process. Like COMPSTAT, it is both a management philosophy and a tool for police department managers to identify spikes in crime using internal benchmarking. One of the main features of SeaStat is its commitment to benchmarking current performance against the past. In Figure 2 below, which is taken
directly from the SeaStat website, the department compares the number of shots fired year-to-date in 2015 to the number of shots fired in previous years. It also makes distinctions between fatal and non-fatal shots fired across multiple years.

Figure 2

![SHOT FIRED Diagram]

Public Safety agencies are not the only ones to leveraging internal benchmarking. It is also a very common public health practice. Figure 3 is pulled from Kansas City, Missouri's KCSTAT. It represents an example of a local public health department benchmarking current life expectancy of black and white males and females against their life expectancy in the past. In all cases, life expectancy has increased in Kansas City.

Figure 3
Life expectancy, KCMO 1999-2003 vs. 2009-2013

<table>
<thead>
<tr>
<th></th>
<th>1999-2003</th>
<th>2009-2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>White female</td>
<td>79.1</td>
<td>80.3</td>
</tr>
<tr>
<td>Black female</td>
<td>74.6</td>
<td>77.1</td>
</tr>
<tr>
<td>White male</td>
<td>73.8</td>
<td>75.2</td>
</tr>
<tr>
<td>Black male</td>
<td>65.7</td>
<td>68.7</td>
</tr>
</tbody>
</table>

*Non-Hispanic white male and female and non-Hispanic black male and female.*
External Benchmarking

External Benchmarking is simply comparing one organization to another. Many government organizations want to know what other comparable jurisdictions are doing, because they want to see their performance in context. After all, if crime rates remain constant in one city, it may not raise alarm bells. But if crime rates are going down in every other city, then there might be a problem.

The first step in external benchmarking is finding a comparable jurisdiction, and there are several tools to help:

- The US Census Bureau's American Fact Finder enables users to select states, counties and metropolitan statistical areas and compare them across multiple dimensions (e.g. population, income, employment, etc).
- The Bureau of Labor Statistics has labor force, wages, and price data for multiple geographic regions
- Wikipedia has several lists of cities in the United States ranked by population

However, finding an appropriate benchmarking comparison can be difficult, and perfect matches do not exist. Governments should think about the benchmarking mirage before putting significant resources behind a benchmarking effort.

Once an organization identifies a comparable cohort, external benchmarking can begin. The key is finding commonly measured services (and metrics) between and among the jurisdictions being compared. For example, if two government organizations collect trash AND measure pounds of trash collected, then they can compare themselves to one another. Here are some common services that most jurisdictions have in common:

- Refuse Collection & Recycling
- Police, Fire and Emergency Services
- Street Maintenance
- Building Inspection & Code Enforcement
- Human Resources
- Water & Sewer Services
- Parks & Recreation
- Public Education

Figure 4 below is an example from the Arizona Valley Benchmarking Initiative and compares the fire department response times across multiple jurisdictions in the same geographic area.
The Community Foundation, sponsored by Boulder County, Colorado, publishes a Boulder County Trends Report on Key Indicators which leverages external benchmarking to provide context to the performance in Boulder County. Figure 5 is a snapshot of one page of the report, which emphasizes the county's desire to become more diverse. The diversity data is benchmarked across multiple jurisdictions in the County as well as compared to the State of Colorado and the U.S. Population.

Figure 5
The same publication even uses trend lines to visually compare performance in one jurisdiction against another, as you can see in Figure 6.

Figure 6
If naysayers say benchmarking is impossible because your government is so unique, remind them that the District of Columbia Government, which is probably the most unique form of government in the United States (because it functions simultaneously as a city, county and state with no federal voting representation in Congress), manages to benchmark its performance against other comparable jurisdictions.
The Benchmarking Mirage

Unless you live in Minneapolis or St. Paul, there is no such thing as a "twin city." Many governments are searching for a meaningful comparison, but it can be a frustrating path. The issues below highlight the reasons why finding an exact "twin" is a mirage - as soon as you find a good comparison, you realize there is more complexity to take into account. Instead of trying to find an exact match, find a cousin with a strong relationship to one specific service area.

Thinking about Population

Population is a simple word with a lot of embedded complexity. Population volume, the number of people who live a jurisdiction, is important. However, comparing one jurisdiction to another based solely on the population volume is not enough. Population density, a measure of population volume per unit area, is also critical. For example, Chicago is America's third most populous city, but is the 14th in population density - so comparing Chicago to New York and Los Angeles based on population volume alone can disregard important differences in the distribution of people across geographic areas. Population demographics, quantifiable characteristics of a given population, also matter tremendously. For example, Phoenix and Philadelphia have similar total populations (between 4 and 6 million depending on what geographic boundaries you include). However, Philadelphia is four times as dense as Phoenix, and Phoenix has almost three times as many Hispanic residents as Philadelphia. Population fluctuations, changes in population over time, also matter. Some jurisdictions see huge increases in population during the workday, while others stay roughly the same. Some jurisdictions are losing residents while others are gaining them. Taking all of these population considerations into account makes it nearly impossible to find an exact jurisdictional comparison. So consider narrowing the scope of what you are trying to benchmark against. Instead, consider using other factors, which correlate to population, to find comparable jurisdictions. For example, find a jurisdiction with the same weather or geographic landscape. Look for cities with the same type of transportation infrastructure (i.e., highways vs. subways), or cities with comparable budget expenditures. These characteristics might be a more direct route to finding a comparable jurisdiction.

Thinking about Mobility

Mobility is a key driver of a jurisdiction's economy and population. Cities that are accessible through robust public transportation networks develop differently than those connected by motorways. The advent of car sharing and services like Uber are also impacting the way people move around a place. But just because two jurisdictions have a bus system, doesn't mean they are worth comparing. To find meaningful comparisons across cities with rail and
bus systems, consider ridership differences, the number of lines and stations, miles of service, and fare differences. When thinking about highway-connected jurisdictions, think about its relationship to geopolitical boundaries, congestion, and urban sprawl. After all, benchmarking DMV services in a city where most people rely on cars should likely differ from cities where public transportation is the norm.

Thinking about Infrastructure

The United States is a young country, but the infrastructure varies widely and is aging at varying rates. Housing stock in the Northeast faces a different set of issues than housing stock on the west coast. Transportation networks were built at different times, with varying levels of reach into suburbs and across geopolitical boundaries. Technology is unevenly deployed across jurisdictions based on resources and proximity to academia and the tech sector. Building codes on the west coast focus on earthquake resiliency more than building codes on the east coast. Some jurisdictions have combined sewer/stormwater systems and others keep them separate. These infrastructure differences create important differences in the way services get delivered and should be taken into account when comparing one place to another.

Thinking about Politics and Governance

All politics is local, and that has a huge impact on external benchmarking. Jurisdictions have varying governance structures. Some have Strong Mayors and others have Strong Council/Managers. Some jurisdictions have Mayoral control of schools, others have an elected school board. Some cities deliver health and human services, while many cities reserve those services for the county governments. Some services are delivered based on political boundaries, others are delivered based purely on physical boundaries. For example, Chicago previously conducted trash collection by Ward while most cities pick up trash based on geography.
What is a Proxy Measure? A proxy is an indirect measure of the desired outcome which is itself strongly correlated to that outcome. It is commonly used when direct measures of the outcome are unobservable and/or unavailable. An organization should use a proxy measure when there is little or no data available about the program being implemented, but the outcome the program is designed to influence has an existing and commonly accepted proxy.

<table>
<thead>
<tr>
<th>This Measure</th>
<th>Is a Proxy for</th>
<th>And can be found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross Domestic Product (GDP)</td>
<td>Quality of Life, Standard of Living</td>
<td>Department of Commerce: Bureau of Economic Analysis has GDP by state &amp; metropolitan statistical area</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>Economy, Labor Market</td>
<td>Department of Labor's Bureau of Labor Statistics publishes regular updates to employment, unemployment and wage data.</td>
</tr>
<tr>
<td>Air Quality Index (AQI)</td>
<td>Environmental Quality</td>
<td>The Environmental Protection Agency (EPA) publishes Air Quality data by city and county.</td>
</tr>
<tr>
<td>Infant Deaths (Mortality)</td>
<td>Maternal, Infant &amp; Child Health</td>
<td>The CDC publishes Linked Birth and Infant Death Data with links to local public health agencies</td>
</tr>
<tr>
<td>Obesity Rates</td>
<td>Health &amp; Well-being of a community</td>
<td>The CDC publishes Nutrition, Physical Activity and Obesity: Data, Trends and Maps which can be viewed at the state and city level for many cities</td>
</tr>
<tr>
<td>Homicide Rates</td>
<td>Public Safety</td>
<td>Local jurisdictions are the best source for the most current crime data, but the FBI publishes annual reports.</td>
</tr>
<tr>
<td>Reading &amp; Math Proficiency Levels</td>
<td>Education Quality</td>
<td>The National Center for Education Statistics publishes proficiency scores on standardized tests for many participating jurisdictions.</td>
</tr>
<tr>
<td>General Fund Personnel and Operating Expenses per Capita, FTEs per 1,000 population</td>
<td>Government Operations &amp; Management</td>
<td>Personnel Costs, Operating Costs, and FTE counts are usually available in government budget documents, and population is available through the US Census Bureau</td>
</tr>
</tbody>
</table>
What is a Proxy Measure?
External Benchmarking Resources

External Benchmarking is the process of continuously comparing and measuring one organization against another to gain information that will help the organization take action to improve its performance. The list below provides quick links to helpful benchmarking resources.

National Resources

The Community Indicators Consortium seeks to bridge the gap between community indicators and performance measurement, providing ways for community groups and governments to coordinate efforts and jointly enhance knowledge about the use of indicators to leverage positive change. The CIC provides two resources which may be helpful as a benchmarking tool:

- The **Indicator Projects Database** is a searchable selection of indicator projects managed at the community level. The CIC does not endorse any projects and the information is generated by the projects themselves. The CIC keeps the list current by regularly asking projects to update their page.
- The **Indicator Resources Database** is a wide range of knowledge resources designed to advance the work of their members and others in the field. It includes publications, presentations, reports generated by indicators projects, web sites, webinar archives, conference presentations, and many other items relevant to the field.
For those interested in the relative size of government, the US Census Bureau conducts an annual survey of federal, state and local governments and publishes the payroll costs and number of full/part time employees working at each level of government. The data is a snapshot of the March pay period for each jurisdiction and breaks down into programs (i.e. housing, finance, parks, education, etc).

The Open Data Network connects publishers of high-value data to the businesses, developers, and analysts that use, reuse, and enrich data. The site has easy links to data published by members of its network, including almost 200 city, county, state and federal agencies publishing data about infrastructure, housing & development, finance, politics, education, public safety, health and transportation. This is a great resources for governments to use data published by other governments to establish performance benchmarks, but is limited to governments who use a Socrata platform.

ICMA Insights is, among other things, a national comparative performance database. Participating jurisdictions can look at the data from other jurisdictions, see how their own results compare, share successful approaches to service delivery, and set appropriate benchmarks and standards for their own performance on selected measures.

The Brookings Institution publishes Data Resources for State and Metropolitan Areas which provides interactive demographic, social and economic data and in-depth profiles for the U.S.’s 100 largest metropolitan areas.

The Urban Institute produces and/or supports multiple resources that are valuable as benchmarking tools.

- The Urban Institute National Data Repository provides free, well-documented versions of national data files with meaningful indicators of community well-being. This includes home Mortgage data, zip/County business patterns, education statistics, and American Community Survey data.
- Performwell, which provides measurement tools and practical knowledge that human services professionals can use to manage their programs’ day-to-day performance.
- The National Neighborhood Indicators Project (NNIP) is coordinated by the Urban Institute and leverages NNIP partners to build information on neighborhood conditions in partner cities and communities. Their list of data sources has indicators that cover topics such as births, deaths, crime, health status, educational performance, public assistance, and property conditions.

State-Specific Resources
ClearGov was launched in June of 2015 with the goal of helping average citizens better understand how their tax dollars are spent and how their local government is performing. It is basically a benchmarking tool for financial data. It is currently limited to three states (California, Massachusetts, and New York), but more states are coming soon (according to their website).

The North Carolina Benchmarking Project is produced by the UNC School of Government and provides a comparative basis for local governments to assess service delivery and costs. It allows municipalities to compare themselves with other participating units and with their own internal operations over time. The benchmarking process includes compiling service and cost information, cleaning the data for accuracy, calculating the selected performance measures, and comparing the results.

- One outgrowth of the North Carolina Benchmarking Project was the creation of the North Carolina Fiscal Benchmarking Tool for Counties and Municipalities, which is a web-based dashboard that helps local governments in North Carolina analyze and communicate their financial condition.

The Arizona State University Center for Urban Innovation does a number of research projects aimed at accelerating and transforming local government, including publishing a Valley Benchmarking Cities Report which highlights benchmarking data for 11 participating jurisdictions in the Phoenix “Valley of the Sun.”

The Florida Benchmarking Consortium is an intra-state local government benchmarking consortium with over 40 member local governments. Each member local government participates across 18 local government performance management-focused service areas, potentially using a combined total of over 670 performance measures that have been dynamically crafted, massaged, managed and improved over time by service area experts from the many industries of Florida's local governments.

The Michigan Local Government Benchmarking Consortium brings together cities, townships, and counties from across Michigan and produce meaningful and relevant performance measures for the purpose of benchmarking and multijurisdictional performance comparisons.

International Resources
The World Council on City Data has Indicators for City Services and Quality of Life, the first international standard on city data, was published in May 2014 by the International Organization for Standardization (ISO). This new international standard was developed using the framework of the Global City Indicators Facility. This is a demand-led standard, driven and created by cities, for cities. ISO 37120 defines and establishes definitions and methodologies for a set of indicators to steer and measure the performance of city services and quality of life. The standard includes a comprehensive set of 100 indicators — of which 46 are core — that measures a city’s social, economic, and environmental performance.

The Scottish Local Government Benchmarking site brings together a wide range of information about how all Scottish councils perform in delivering better services to local communities, including the cost of services and how satisfied citizens are with them. Bear in mind that Scottish councils are democratically elected bodies which represent very diverse communities in terms of geography, population, deprivation levels and community needs. The information presented here should be read and understood in that context, as some of the variation highlighted by the data is significantly affected by such differences between communities.

The Ontario Municipal Benchmarking Initiative (OMBI) fosters and supports a culture of service excellence in municipal government by creating new ways to measure, share and compare performance data and operational practices. OMBI acts as a source of credible information to assist Council, Senior Management, Staff and Citizens to understand how their municipality is performing over time and in relation to others.