

Open Data - How Much Does It Cost?

(Note: originally published in 2016, and prices may have changed.)

Everyone wants to know the costs and benefits of their actions. This is especially important for local governments, which often face resource constraints. And it certainly applies to open data. Open data advocates, practitioners, and providers are quick to point out its benefits: that it promotes transparency, improves efficiency and effectiveness of government services, fosters innovation, creates new business opportunities, etc. But when asked, “how much does it cost?” the answers may be “no information is available,” “nobody can tell,” or “it depends.” This post is intended to highlight some of the common cost elements and how much cities are investing in open data initiatives.

What Are the Common Cost Elements of Opening Data?

One of the important first steps in undertaking any program involving significant cash expense is cost estimation. Because of the perennial reality of limited resources, cost estimates for projects must be accurate, transparent, and reliable. Across all levels of governments, open data is becoming a key innovative tool for delivering results to residents. The amount cities are paying for developing open data programs depends on several factors, including but not limited to the type of portal functionality, data publishing services, available datasets, city size, type of datasets, add-on requirements, and city objectives. But what does this amount to? The Open Data Institute [highlighted](#) the following common costs elements for consideration when developing open data program:

Set Up and Technical Cost

- Open data portal development costs
- Cost of storage including cloud storage
- Cost of creating APIs
- Hardware and other overhead costs

Administration/Governance

- Getting legislation or policy written and approved
- Maintaining and enforcing the policy or legal requirements
- Legal costs to comply with open data legislation
- Managing requests or questions from the ecosystem
- Digitization of paper records or upgrading of data systems
- Cost of changing the way data is collected
- Creating a coordination unit within government

Skills Development and Community Engagement

- Capacity building of developers
- Awareness raising activities to promote use
- Capacity building for the use of data within government
- Coordination unit human resource costs

Sustainability Cost

- System maintenance cost
- Data inventorying and publishing
- Person hours for updating and administering the portal
- Liability costs in case of publication of nonpublic information
- Impact and analytics cost
- Staff compensation and training costs
- Portal contract renewal
- Cost of expanding portal functionalities and capabilities such as add-ons

How Much Are Local Governments Paying for Opening Data?

The desire of city leaders to use data and evidence to deliver results to their residents is growing rapidly. Though demand for public data continues to increase, only a few cities have the resources and structures to regularly release data. A recent [report](#) published by the Bloomberg Philanthropies' [What Works Cities Initiative](#) found that city leaders nationwide are eager to address the challenges facing their residents by using data but are constrained by limited resources and expertise. The study revealed that of the 70% of cities whose leadership one way or the other have shown dedication to publishing data only 18% have established processes for releasing data publicly. Although, inadequate resources present a barrier to opening data, the lack of cost information poses potential threats to quenching the energies that have characterized open initiatives.

Typically, it should not be difficult to get costs information by consulting available price lists from providers. Interestingly, open initiatives are supposed to increase transparency, but vendors tend to apply different transparency rules when it comes to pricing of their services. Not only do they shy away from being transparent about their pricing, they also require clients not to disclose cost information through non-disclosure and confidentiality agreements. Open data portals vary widely across cities. An [analysis](#) of contracts by the KPCC data team from selected cities with few miles apart from one another mostly within the state of California revealed interesting cost disparities. For instance; Pasadena, Anaheim, and Santa Clarita maintain open data portals with Junar, an open data service platform vendor but Pasadena's annual contract is in excess of \$20,666.67. On average, contracts with Socrata are relatively higher at least for the few selected contracts from cities that are available. Of the contracts examined, [Chattanooga's contract](#) amount of \$49,578 appears to be the highest. Figure 1. is a

summary of the annual cost for the selected cities from three different vendors; Junar, Opengov, and Socrata.

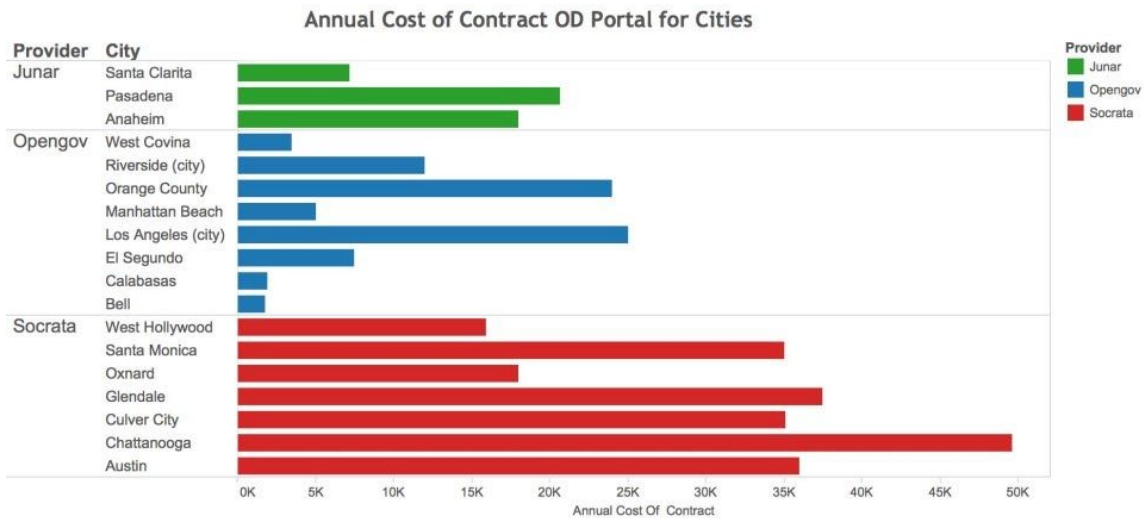


Figure 1. Annual Cost of Contracts

According to some service providers, pricing of data portals takes into account a city's size, but there is little evidence to support that assertion. Figure 2. represents the cost of contracts by city with their corresponding population. The city of Los Angeles has a population of 3.88 million and maintains a lower contract price than a number of the cities with fewer thousands of population.

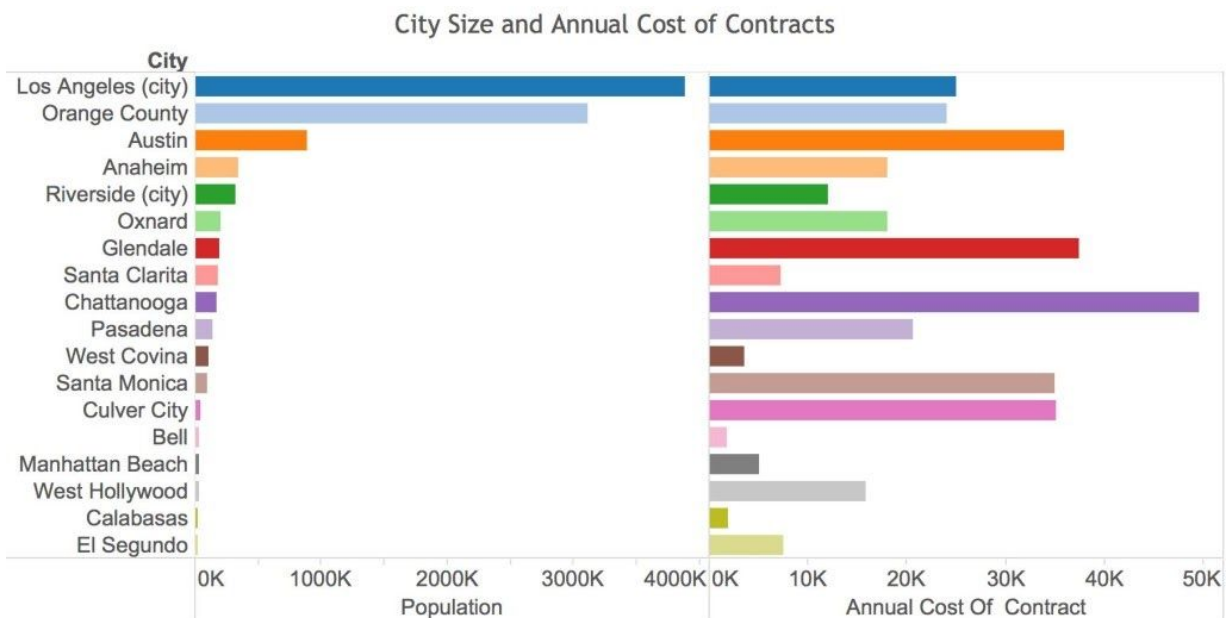


Figure 2. Contract and Population

So What Determines the Cost?

Open data is garnering considerable momentum and significance in today's governance practices. Thanks to a number of open data advocacy organizations, standards have been developed to advance open data practices, unleashing all of its benefits to society. However, pricing standards have lacked adequate attention necessary to measure the value or cost-effectiveness of government data portals. Besides setup and the technical and administrative costs of opening data, city leaders are always eager to know what other details determine the eventual cost. As it turns out, there is little information to satisfy such genuine curiosity. However, Socrata's pricing structure in [this spreadsheet](#) (link downloads .xlsx) could be useful to understand some of the key cost determinants. The price list is primarily determined by the number of datasets, views and visualizations capabilities, desired number of users, storage, bandwidth, and APIs.