Publication of County Data

Objective:	To establish County government policy and procedure that reflects our belief in transparency and ease of public access to data and information by encouraging and guiding public disclosure and publication of data by Tompkins County.	**Policy/Procedure Number:**	*01-55*
		Effective Date:	*November 21, 2023*
Reference: *(All applicable federal, state, and local laws)*		**Responsible Department:**	*County Administration*
		Modified Date (s):	
Legislative Policy Statement:	It is the policy of Tompkins County government that the public should generally have free and easy access to County government data when it is required, appropriate, and/or reasonably requested. This policy affirms that basic to our democratic society is the people's right to examine data related to the (1) County's programs and performance and (2) information managed by the County regarding the community. By publishing data that includes demographic and performance information, Tompkins County government furthers its commitment to providing equitable provision of government services to all.	**Resolution No.:**	*2023-248*
		Next Scheduled Review:	*November 2028*

General Information:
This policy addresses the publication of aggregate data in a user-friendly format when it is relevant to the wider community. This policy does not address specific records requests, which are managed through *Administrative Policy 01-06: Public Access to Records*, as well as public portals that allow the public to search for and access records and data.

The purposes of this Policy are to:

- Provide guidance for Tompkins County departments when preparing data for public presentation;
- Ensure a consistent user experience for the public accessing data from Tompkins County;
- Increase the public benefit of available data by providing legible and easily accessible data;
- Make the best use of software and tools procured by Tompkins County for the purposes of data analysis and publication; and
- Help identify data appropriate for public presentation.

I. Definitions:

Aggregated Data - Multiple data sources combined into one set to create a larger picture of a particular issue.

Chart - A graphical representation of data.

Complete Dataset - All relevant and necessary information for a given purpose. A complete dataset should not have any missing, duplicated, or irrelevant values that could affect analysis.

Data - A collection of facts, observations, or statistics collected for reference or analysis.

De-identified Data - Also known as anonymized data, refers to information that has been processed in a way that removes or alters certain identifying elements to protect the privacy of individuals while still maintaining the data's utility for analysis

and research purposes and includes, but is not limited to, Personally Identifiable Information (PII), Health Insurance Portability and Accountability Act (HIPAA) and Criminal Justice Information (CJI). The goal of de-identifying data is to reduce the risk of re-identifying specific individuals from the dataset.

Disaggregated Data - Separation of a dataset into smaller units to show underlying trends and patterns.

FOIL - The Freedom of Information Law (FOIL) gives the public the right to access, with certain exceptions, documents and information about the functions, procedures, policies, decisions and operations of government department and agencies.

Graph - Diagram representing the connectedness or two (2) or more data points.

Managed Data - Third party data accessed and/or used by Tompkins County, may be data aggregated into a larger dataset.

Open Source - Data made freely available for public use and sharing. Data used from an open-source database should be attributed to its source.

Record - Any information kept, held, filed, produced, or reproduced by, with, or for an official County agency, board, or committee, and/or for the County Legislature, in any physical form whatsoever. Under the law, all records are accessible, except records or portions of records that fall within one of defined categories of deniable records as stated in §87(2) of the Public Officer's Law (also see item B of the Procedure section of *Administrative Policy 01-06: Public Access to Records*). Access refers to existing records, and therefore no record need be created in response to a request.

II. Policy:

A. Tompkins County may publish data when it is owned by the County or its officers or when it is owned by a third party and managed or accessible by the County or its officers.

B. When required, appropriate, and/or reasonably requested by a third-party, Tompkins County's departments should make data owned by the County available to the public.
- Publication is deemed required when federal, state, or local law states such a requirement, or when it is subject to a request made under the Freedom of Information Act (Tompkins County Public Access to Records Policy, 01-06).
- Publication is deemed appropriate when data owned by the County is of interest to the public, publication of the data supports a department or county-led initiative, or publication is not onerous based on available resources and staff capacity within a department or the organization.
- Publication is deemed reasonably requested when a community member or organization requests the data be published and the publication is deemed appropriate by the above considerations.

C. Publication of data in accordance with this policy does not require an action of the Legislature. The Tompkins County Legislature may request publication of certain County data by resolution.

D. Training will be made available on an annual basis to inform staff and Department Heads on how best to analyze and publish data in accordance with this policy.

E. County Administration will maintain a listing of all published data made available by departments.

F. The County Administrator will name a Data Officer who will be responsible for maintenance of data analytics and publication tools, open-source data publication tools, management of data publication requests, and offering guidance and assistance to department implementing data publication efforts.

G. County Administration shall procure and maintain tools for data analysis and publication for all departments, as needed.

H. County Administration will offer a central form to receive data publication requests from the public and will distribute requests to the appropriate departments.

III. Procedure:

A. The County may publish performance data on the services or activities carried out by departments. Performance-related data can help the public to understand how well the County is operating and meeting its goals, as well as help to further the organization's commitment to providing equitable government services to all.

B. When appropriate, departments will use the tools made available by County Administration for the analysis and publication of data and will consider migrating existing data from other platforms or tools to the centralized tools.

C. Departments in concert with County Administration will help determine whether reasonably requested data is appropriate for publication. In some instances, someone who requests data be published may be redirected to the County FOIL process or other public portals or information made available by the County or third-party organizations.

D. Guidelines

1. Publication of data should be completed in a reasonable time frame after the request or requirement. Time frames should be determined based on the complexity of the data and the County's capacity to prepare it for publication.

2. Data published by the County should include demographic information when appropriate in alignment with the organization's commitment to the equitable provision of government services to all.

Data published by the County should adhere to the principles of ethical governance and conscious design (adapted from the Federal Data Strategy Action Plan).

Ethical Governance

- Uphold Ethics: Monitor and assess the implications of local data practices for the public. Design checks and balances to protect and serve the public good.

- Exercise Responsibility: Practice effective data stewardship and governance. Employ sound data security practices, protect individual privacy, maintain promised confidentiality, and ensure appropriate access and use.
- Promote Transparency: Articulate the purposes and uses of data to engender public trust. Comprehensively document processes and products to inform data providers and users.

Conscious Design

- Ensure Relevance: Protect the quality and integrity of the data. Validate that data are appropriate, accurate, objective, accessible, useful, understandable, and timely.
- Harness Existing Data: Identify data needs to inform priority research and policy questions; reuse data if possible and acquire additional data if needed.
- Anticipate Future Uses: Create data thoughtfully, considering fitness for use by others; plan for reuse and build in interoperability from the start.
- Demonstrate Responsiveness: Improve data collection, analysis, and dissemination with ongoing input from users and stakeholders. The feedback process is cyclical; establish a baseline, gain support, collaborate, and refine continuously.

4. Data should be available aggregated and disaggregated as appropriate. Data should be viewable by variables that show trends and patterns. Charts and graphs should be prepared, when possible, to show data. Aggregate data may pull information from multiple sources.

5. Data published by the County should be from a complete dataset. Complete datasets should be available in addition to any charts and graphs published by the County. Complete datasets should be published in an open-source format.

6. Datasets must be de-identified before publication.

E. When possible, departments will collect detailed demographic information on clients/users so that the equitable provision of government services to all can be examined by staff and/or the public. Data may be published in a static format or updated on an ongoing basis. Departments will determine whether data will be published one time or on an ongoing basis, and if ongoing, how frequently it will be updated. Information on how often data will be updated should be made available in conjunction with its publication.