Data Management Committee

Data Usage Guideline Document

Document: #1025

**Data Usage Guideline - Usage of Departmental / Unit / Local Data**

**on UW Tableau Server**

**Purpose**

The purpose of this document is to formalize agreement for publishing institutional data from departmental, unit or local data[[1]](#footnote-1) sources to the Tableau Server installation at the University of Washington.

Academic and administrative units across UW create and use visualizations in Tableau to aid in data analysis, exploration and decision-making. For this, units may use data from systems of record such as administrative business systems, from the Enterprise Data Warehouse (EDW), or from institutional data from their own departmental, unit or local data repositories. This agreement is to set a guideline that would govern the usage, storing and sharing of institutional departmental, unit or local data to build and share visualizations using the UW Tableau Server.

Departmental, unit and local data are subject to the same policies and procedures which govern the use of institutional data[[2]](#footnote-2). This includes the classification of information according to Administrative Policy Statement (APS) 2.2 Sections 4 and 5, and APS 2.4[[3]](#footnote-3).

**Overview / Problem Statement**

Tableau Business Intelligence (BI) products hold great promise for the UW in facilitating decision-making with intuitive end user features and advanced data visualization capabilities. Once visualizations are published to Tableau Server, the data that is used to build the visualizations are deployed to the server as extracts. These extracts are self-contained data sources that can be not only used to build visualizations, but that can also be shared with and used by other members of the same group owning the extract to join with other data sources, and/or be downloaded (with correct server permissions) locally. Access and security controls for local data may differ from those applied in administrative business systems or the EDW.

Currently, the use of Tableau Server is limited to the use of EDW data. This prevents groups which rely on institutional departmental local data from leveraging UW’s investment in Tableau. To support the use of departmental, unit or local data for legitimate business needs, the following proposes measures and controls for hosting of local data sources on Tableau Server in compliance with the UW Institutional Data Management Standards.

The benefits of allowing departmental, unit or local data repositories to be used by Tableau users include:

* To encourage and expand data driven culture at UW
* To enable building and sharing of visualizations within and across units
* To provide transparency about high-value departmental, unit or local data repositories and its usage, to help inform the future expansion of EDW

**Observations, Recommendations and Agreements**

UW-Information Technology (UW-IT) proposes to create a program that will host departmental, unit or local data repositories for use in Tableau based on the following.

**Security**

* Establish and manage a secured environment to host data extracts on Tableau Server
* Leverage Tableau Server permissions to authorize server users to publish/access/share data extracts
* Appoint a single-person on the business side to be a point-person and “owner” of the data extract
* Ensure that data extracts are accessed and used by authorized members of the Tableau Server community with agreement and support of the extract owner

**Publishing**

* Create and maintain a list of Tableau Publishers that are trusted by data custodians to publish data extracts
* Assign point-person to work with UW-IT to support maintenance and availability of data extracts
* Allow individual groups to publish visualizations and data extracts that support their local departmental analytics initiatives only
* Discourage use of data extracts beyond the scope of unit’s own specific needs

**Data**

* **Size:** Limit the size of extracts to the minimum amount of data required to support the visualization**.** Data extracts are not replacement for EDW, as such, their size should be limited to provide only data necessary to satisfy the requirements of the dashboard (or suite of dashboards) they support
* **Type:** Support a multitude of data extracts created from (but not limited to): MSSQL, Excel, Access, MySQL, Text Files, Google Analytics, Oracle
* **Storage:** Manage utilization of server storage resources.Data extracts are created by selecting data from the source database and stored as physical files on Tableau Server using Tableau’s proprietary file format. Storing extracts on Tableau server will consume machine resources, which will be subject to maintenance policies established by administrators of the server
* **Retention:** Data extractsprovide only the data necessary to satisfy the requirements of the dashboard they are supporting, therefore, any requirements to store more data beyond the needs of visualizations that are being created will be discouraged. Unused and outdated data extracts will be deleted from the server
* **Refresh Rate and Processing:** Each extract will be scheduled to refresh its content on a predefined basis. Scheduling and monitoring of extract refresh cycle will be performed by the server administrator. All extracts will be refreshed during off-peak hours (evenings/nights/weekends) to maximize computing resources and to minimize degraded user experience and response time to process and render visualizations. The Tableau administrator will work with the business side to queue the refresh of extracts to meet the business needs. All extracts will support the following refresh rate options: once per day (nightly), weekly, monthly, quarterly, annually. Refresh errors attributed to connectivity issues, such as denied access to the database/file, invalid connection string, will need to be resolved by the business. Such extracts will be flagged and the owner of the data extract will be notified.

**Concurrence / Responsibility**

The undersigned concur with the observations and recommendations herein, and are in agreement with this guideline.

**Signatures**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Gary Quarfoth  Associate Vice Provost & DMC Co-Chair  Office of Planning & Budgeting |  | Date |
|  |  |  |
|  |  |  |
|  |  |  |
| Steve Majeski  Associate Dean for Research Administration & Infrastructure & DMC Co-Chair  College of Arts & Sciences |  | Date |

1. See the UW Institutional Data Management Standard for a definition of terms <http://www.washington.edu/uwit/im/dmc/docs/UWDataManagementStandards.pdf>. [↑](#footnote-ref-1)
2. See UW Institutional Data Management Standard pg. 2 [↑](#footnote-ref-2)
3. See <http://www.washington.edu/admin/rules/policies/APS/02.02.html> [↑](#footnote-ref-3)