# DAT 325 Project One Template

# Data Quality Plan

### Purpose Statement:

We need high quality data to be able to make good business decisions, ensure customer satisfaction, reduce risks, save costs, and stay in compliance. Trying to analyze or make business decisions using poor data will cause more harm than good. Using good data would result in good outcomes for the business since it will be relevant, accurate, and authentic. Having good customer satisfaction would result in getting more business and it may even help save the cost. For example, Netflix used high quality data to improve their algorithm to improve personalization and recommendation, which resulted in saving more than $1 billion dollars per year. That is because of low cancellation rates, they do not have to spend more money to acquire customers to replace the cancelled members. As for reducing risks, instead of gathering data for it to be remain unused (because it is still an asset, it poses as a risk for misuse and lose customers’ trust). Companies use data management plan to filter out the good and bad information to limit their liability. It also ensures a company will comply to the standards and not risk any violations. (Reddy, 2019)

### Organizational Goals:

1. Join data with Wayne Enterprises to develop requirements, standards, and databases to be able to conduct accurate and reliable analyses. To do so, obtain their data, then clean and transform it so the quality is good. Both companies should have different experts working together (data management, data security, and more).
2. Obtain high quality data based on industry standards. Using DAMA guidelines, it contains each dimension, and it explains the purpose of each one regarding to high quality data. Then have data analysts clean and transform the data before using.
3. Maintain data quality and integrity. Have analysts constantly check and ensure for data quality. Constantly conduct audits, making sure there are different departments doing their job, and ensuring the company and its employees are complying. Data quality is usually used for decision making. Data integrity ensures data quality and the reliability of the information. (Johnson, 2022)

### Data Quality Characteristics and Procedures:

* Accuracy: Ensures reliable and appropriate information is being used for good business decision making, accurate data means it is without error and is authentic. (*What is Data Accuracy? (Why is Data Accuracy Important)*)
* Completeness: Appropriate data is in its assigned value in a data set, all necessary data is available. (Technics Publications, LLC, p. 296)
* Validity: Data is in its appropriate format. (Technics Publications, LLC, p. 297)
* Uniqueness: There are no data duplication. (Technics Publications, LLC, p. 297)
* Consistency: Ensuring that the data is consistent with the data in the data set, making sure it makes sense. For example, having a data set on application users and it will be considered inconsistent if the number of active users is greater than the number of registered users. (Johnson, 2022)
* Timeliness: The data is up-to-date and is available to access. (Technics Publications, LLC, p. 297)

### Security and Personnel Responsibility Plan:

Data security and data quality are two different critical aspects when it comes to working with data. To ensure both data security and data quality, it can become an issue. Trying to ensure data quality can compromise security. For example, implementing a data quality management system requires flexible read and write access, it can create security problems because the data system can exchange data with other data system or be manipulated by other people who do not have the same access rights (Talha, 2019). Therefore, strengthening data security can improve data quality. One of the dimensions of data quality is privacy. Enabling data security means the users/clients will have privacy and their information will be protected.

References:

Johnson, J., & Raza, M. (2022, February 15). *Data Integrity vs Data Quality: An Introduction*. BMC Blogs. Retrieved March 20, 2022, from <https://www.bmc.com/blogs/data-integrity-vs-data-quality/>

Reddy, R. (2019, November 22). *The Importance And Benefits Of Ensuring Data Quality*. Acuvate. Retrieved March 20, 2022, from <https://acuvate.com/blog/the-importance-and-benefits-of-ensuring-data-quality/>

Talha, M., Abou El Kalam, A., & Elmarzouqi, N. (2019). Big Data: Trade-off between Data Quality and Data Security. Retrieved March 20, 2022, from <https://www.sciencedirect.com/science/article/pii/S1877050919305915>.

Technics Publications, LLC. The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide). Retrieved March 20, 2022, from <https://learn.snhu.edu/d2l/lor/viewer/view.d2l?ou=6606&loIdentId=24443>

*What is Data Accuracy? (Why is Data Accuracy Important)*. ERP Information. Retrieved March 20, 2022, from <https://www.erp-information.com/data-accuracy.html>