# DAT 325 Project One Template

# Data Quality Plan

### Purpose Statement:

High-quality data is a vital business requirement for any organization in any industry, but especially here at Bruce, Inc. Good business requirements are clear and concise, which in turn allows an organization to make key decisions to benefit the success of the organization. Therefore, high-quality data is a vital business requirement because there are data quality characteristics and procedures that outline concisely how data should be collected and stored to ensure that data is cleaned to produce accurate information, ensure it is readily available, and ensure security from outside threats, which are three of our organizational goals at Bruce, Inc.

### Organizational Goals:

There are three main goals that Bruce, Inc has concerning the data we use. The first goal is to be able to use and trust our data to make informed business decisions. This means that any new data should be cleaned and validated before entering our databases or uploaded to the cloud. Our second goal is that our data should be readily available and accessible. This means that data should be stored in an organized manner that allows it to be used quickly. Finally, the third goal is to keep our data safe and secure from outside organizations, people, etc. This means having proper security measures in place to ensure the integrity of our data is not compromised. Data uploaded to the cloud should be encrypted using the 256-bit AES algorithm.

### Data Quality Characteristics and Procedures:

The six dimensions that define data quality that we adhere to at Bruce Inc are Completeness, Validity, Uniqueness, Timeliness, Accuracy, and Consistency. Completeness in our industry means all required data fields must be complete. If no entry exists in the data from a new dataset, then research needs to be done to find an entry. The second dimension is Validity, which means new data needs to be checked to ensure it is consistent with the domain/requirements of the other data. A reference table will be available. The third dimension is Uniqueness which outlines that all data entries must be unique from one another. If duplicated values are found then one must be deleted or they must be merged into one data value. The next dimension is Timeliness, which directly relates to one of our organizational goals. Data needs to be accessed and available promptly. The fifth and perhaps most important dimension is Accuracy which states that all data needs to be correct according to its real-life representation. Any data in a new dataset that is incorrect should be corrected so that it reflects real-life representation. Finally, the last dimension is Consistency, which is a focal point with the data coming from new sources. All new data must be consistent with the data that we already have stored here at Bruce Inc.

### Security and Personnel Responsibility Plan:

The security of the data that we collect and use here at Bruce Inc. is the shared responsibility of everyone working here. Data producers, consumers, brokers, and stakeholders of all kinds need to be concerned with the security of our data. This means ensuring data is only seen by people who can assess it. Sharing data with any outside organization or people is against our data privacy policy. If there are issues concerning data security it is the responsibility of anyone who might have come across it. If the issue is beyond the scope of your abilities then it is then your responsibility to find somebody inside the organization to help remediate the security issue. If data is being uploaded to our cloud provider, the data needs to be encrypted properly according to company policy. The security of the data inside the cloud is still the responsibility of Bruce Inc.

**References**

Y. Sharma, H. Gupta, and S. K. Khatri, "A Security Model for the Enhancement of Data Privacy in Cloud Computing," *2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019, pp. 898-902, DOI: 10.1109/AICAI.2019.8701398.