Executive Summary of Responses to Data Steward Kick-Off Meeting Exercise
EXECUTIVE SUMMARY

As part of the Data Steward Kick-Off Meeting held in August 2014, a participatory exercise was conducted where attendees were asked, individually and as part of a joint response from their table, to indicate both the data-related problems they were having with NYU administrative data systems and ideas for how to resolve those problems.

Though there were many varied responses, certain themes did clearly emerge. The following highlights follow the structure of Data Governance as laid out during the Kick-Off, where the initiative can be broken down into Policies; Processes; Tools and Technology; and Accountability.

**Policies**
- There are different values, tables, data fields in use in each of NYU's administrative systems. The coding structures should be harmonized for ease of use.
- There should be greater integration among all of NYU's administrative systems, including at the School of Medicine and the global sites.

**Processes**
- Processes need to be developed to (a) prevent duplicate records from entering our data systems, and (b) clean up any duplicates found.
- Incomplete records hamper efficient use of data; processes need to be in place to ensure that all necessary information is captured as early in the data gathering process as possible.
- There needs to be enhanced training, covering such topics as: what information is available to report on; what information is available through which reporting environment; and understanding the various business processes which create data.

**Tools and Technology**
- There needs to be a data dictionary/business glossary so that there is a common understanding of the meaning of terms. [overwhelmingly the most frequent response to the exercise]
- To assist in understanding data, the lineage of data elements (provenance) should be available.
- There should be a stream-lined mechanism for requesting and approving access to data, across all transactional and reporting systems.
- Data fields should be named consistently across systems.
- Data profiling should be introduced, with standard exception reports, so that problems are identified and data can be cleaned.
- Data integrity constraints should be added (mandatory fields, valid values) to transactional systems to capture complete and accurate data as early in a process as possible.

**Accountability**
- Data should be used for the purpose intended (e.g. housing billing dates should not be used to control door access).
- Communication needs to be enhanced so that "down stream" users of data are kept apprised of changes to processes, data values, etc. which will affect them.
- Simple questions can have many answers, and before providing an answer to a data question, the intent of the question needs to be explored and a clearly defined answer supplied.