Does secondary analysis of a data set gathered for another purpose require a new research project for review?

Yes - IF THE DATA IS IDENTIFIABLE.

Projects that use an existing data set which includes identifiable data gathered in earlier research projects may require a new IRB protocol for review. Secondary analysis of existing data may include the review of medical records, student records, data collected from previous studies, audio/video recordings, etc. that were initially collected for another purpose. In order to be existing, the information must be "on the shelf" (i.e., it has already been collected) at the time that the current research is proposed.

Though such projects do not involve interactions or interventions with humans, they may still require IRB review, since the definition of "human subject" at 45 CFR 46.102(f) **includes *living* individuals about whom an investigator *obtains identifiable private information for research purposes.***

In addition to being identifiable, the existing data must include "private information" in order to constitute research involving human subjects. Private information is defined as information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (e.g., a medical or school record). Information that contains identifiers and can be accessed freely by the public (without special permission or application) is not "private" and the research therefore does not therefore involve human subjects. For example, a study involving only analysis of the published salaries and benefits of public university presidents would not need IRB review since this information is not private.

Data analysis activities that meet the definition of research with human subjects may qualify for an exemption or require expedited or even full committee review. Any such project must receive IRB approval or a determination of exemption *before* the investigator accesses the data.

When does the secondary use of existing data not require review?
In general, the secondary analysis of existing data does not require IRB review when it does not fall within the regulatory definition of research involving human subjects, as referenced above.

Note: Although the definition of a human subject includes only living individuals, thereby excluding decedents, there are cases in which the health information of the deceased and death data files may require IRB review.

Public data: Public use data sets (such as portions of U.S. Census data, data from the National Center for Educational Statistics, National Center for Health Statistics, etc.) are data sets prepared with the intent of making them available for the public. The data available to the public are not individually identifiable and therefore their analysis would not involve human subjects.

De-identified data: If the dataset has been stripped of all identifying information and there is no way that it could be linked back to the subjects from whom it was originally collected (through a key to a coding system or by any other means), its subsequent use by the lead researcher or another investigator would not constitute human subjects research, since it is no longer identifiable. Identifiable means the identity of the subject is known or may be readily ascertained by the investigator or associated with the information. In general, information is considered to be identifiable when it can be linked to specific individuals by the investigator(s) either directly or indirectly through coding systems, or when characteristics of the information obtained are such that by their nature a reasonably knowledgeable person could ascertain the identities of individuals. Therefore, even though a dataset may have been stripped of direct identifiers (names, addresses, student ID

numbers, etc.), it may still be possible to identify an individual through a combination of other characteristics (e.g., age, gender, ethnicity, and place of employment).

Example: Many student research projects involve secondary analysis of data that belongs to, or was collected by, their faculty advisor or another investigator. If the student is provided with a *de-identified, non-coded data set*, the use of the data does not constitute research with human subjects because there is no interaction with any individual and no identifiable private information will be used. The project does not therefore require IRB review.

Coded data: Secondary analysis of coded private information is not considered to be research involving human subjects and would not require IRB review if the investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information pertains as a result of one of the following circumstances:

1. The investigators and the holder of the key have entered into an agreement prohibiting the release of the key to the investigators under any circumstances, until the individuals are deceased (DHHS regulations for humans subjects research do not require the IRB to review and approve this agreement);
2. There are IRB-approved written policies and operating procedures for a repository or data management center that prohibit the release of the key to the investigator under any circumstances, until the individuals are deceased; or
3. There are other legal requirements prohibiting the release of the key to the investigators, until the individuals are deceased.

Note: *If a student is analyzing coded data from a faculty advisor/sponsor who retains a key, this would be human subjects research, because the faculty sponsor is considered an investigator on the student's protocol, and can readily ascertain the identity of the subjects since he/she holds the key to the coded data. If the student's work fits within the scope of the initial protocol from which the dataset originates, the faculty sponsor (or investigator who holds the dataset) may wish to consider adding the student and his/her work to the original protocol by means of a modification request rather than having the student submit a new application for review.*

Example: Researcher A plans to examine the relationships between attention deficit hyperactivity disorder (ADHD), oppositional defiance disorder, and teen drug abuse using data collected by Agencies I, II, and III that work with "at risk" youth. The data will be coded and the agencies have entered into an agreement prohibiting release of the key to the researcher that could connect the data with identifiers. The use of the data would not constitute research with human subjects and does not require IRB review.

Will I be required to reconsent participants?

Consent: Researchers using data previously collected under another study should consider whether the currently proposed research is a "compatible use" with what subjects agreed to in the original consent form. For non-exempt projects, a consent process description or justification for a waiver must be included in the research protocol. The UCM IRB may require that informed consent for secondary analysis is obtained from subjects whose data will be accessed.

What is Restricted Use Date?

"Restricted Use Data": Certain agencies and research organizations release files to researchers with specific restrictions regarding their use and storage. The records frequently contain identifiers or extensive variables that combined might enable identification, even though this is not the intent of the researcher. Research using these data sets most often requires expedited or full committee review.