Consider the following quote from Stephen B. Vardeman and Max D. Morris’s paper titled “Statistics and Ethics: Some Advice for Young Statisticians” (*Source: The American Statistician, February 2003, Vol. 57, No 1*).

*“… this discipline provides tools, patterns of thought, and habits of heart that will allow you to deal with data with integrity. At its core statistics is not about cleverness and technique, but rather about honesty. Its real contribution to society is primarily moral, not technical. It is about doing the right thing when interpreting empirical information. Statisticians are not the world’s best computer scientists, mathematicians, or scientific subject matter specialists. We are (potentially, at least) the best at the principled collection, summarization, and analysis of data. Our subject provides a framework for dealing transparently and consistently with empirical information from all fields; means of seeing and portraying what is true; ways of avoiding being fooled by both the ill intent (or ignorance) of others and our own incorrect predispositions.”*

In this handout, we will discuss ethical issues frequently encountered by statisticians or data scientists. We will begin with the ethical principles that have been established for research involving human subjects. As a student studying statistics, you may be tempted to believe that this is something that only your client from another discipline (e.g., a biologist or a psychologist) should have to consider. Note, however, that the “[Ethical Guidelines for Statistical Practice](http://www.amstat.org/about/ethicalguidelines.cfm)” endorsed by the American Statistical Association specifically addresses our responsibilities to research subjects as statisticians.



**Examples of Unethical Research Involving Human Subjects**

In the previous class period, we discussed four examples of unethical studies that were conducted in the past:

* Milgram’s Behavioral Obedience Study (1963)
* Zimbardo’s Stanford Prison Experiment (1971)
* Willowbrook Hepatitis Study (1963 – 1966)
* Tuskegee Syphilis Experiment (1932 – 1972)

 **Responses to Unethical Research Involving Human Subjects**
The following codes were established in response to these (and other) examples of unethical research.

* [Nuremberg Code](http://www.hhs.gov/ohrp/archive/nurcode.html) (1949)
* [Belmont Report](http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html) (1979)
* [Declaration of Helsinki](http://www.wma.net/en/30publications/10policies/b3/) (1964 – amendments added in later years)

These documents vary somewhat, but they all typically emphasize the following principles.

|  |
| --- |
| **Ethical Principles for Research with Human Subjects** |
| **Respect for the Dignity and Autonomy of Persons** | Researchers must respect research subjects as human beings whose participation in any given study is a result of their own free choice. This involves:(1)(2)(3)Persons with limited autonomy (e.g., children, the sick, prisoners) must be protected from coercion by others and from activities that harm them. |
| **Beneficence and Nonmaleficence** | Beneficence refers to contributing to the well-being of others; nonmaleficence involves refraining from actions that might harm others. Researchers must strive to maximize potential benefits while minimizing risks. Risks that are identified in advance must be communicated to prospective study participants. |
| **Justice** | The benefits and the burdens of the research must be distributed equally for all persons. In other words, all persons should be granted access to the benefits of the research, and the risks for harm should not be disproportionately greater for a particular group of people. |
| **Fidelity, Responsibility, and Trust** | Researchers must establish and maintain a relationship of trust with research participants. Researchers must be upfront in describing what the research study will entail, and they must honor all promises that are made as part of the agreement to participate. If full disclosure is not made prior to obtaining informed consent because it might compromise the validity of the study, safeguards must be implemented to protect the welfare and dignity of the human subjects. When unintended negative consequences occur in a study, researchers must remove and/or correct these consequences immediately. |
| **Integrity** | Researchers are obligated to “do good science.” They must truthfully report their results, take reasonable steps to correct errors, and take responsibility and credit for only work that is their own. |

Questions:

1. How do the four aforementioned “unethical” studies comply with these principles?
2. Which of these principles were violated in these four studies, and how so?

**Institutional Review Boards**

The Nuremberg Code influenced, in part, federal regulations set forth by the U.S. Congress in the National Research Act of 1974. This legislation required the formation of an institutional (or internal) review board (IRB) by every university or organization receiving federal funds for research.

The purpose of the IRB is to review research proposals in order to prevent potentially unethical studies from being conducted and to protect the rights and welfare of individual research subjects. For any proposed study, the IRB assures that the following requirements are satisfied:

1. Risks to subjects are minimized,
2. Risks to subjects are reasonable in relation to anticipated benefits,
3. Subjects are selected in a just manner,
4. Informed consent is sought from each subject or his/her legally authorized representative,
5. Informed consent is appropriately documented,
6. When appropriate, the research plan makes provisions for monitoring data collection,
7. Privacy and confidentiality of research subjects are appropriately protected, and
8. When some or all of the subjects are likely to be vulnerable to coercion or undue influence, additional safeguards have been included.

The IRB categorizes the risk associated with research into the following classifications.

|  |
| --- |
| **IRB Risk Classifications of Research Proposals** |
| **Exempt** | The study proposes no risk to the human research subject, the researcher, or the environment. Examples of such research include anonymous questionnaires, standardized tests used in education, and research involving the collection or study of existing data, documents, or pathological specimens (as long as the confidentiality of any personally identifiable information is obtained). |
| **Minimal Risk** | Although safeguards must be present, these studies typically present no more risk than the subject would face in everyday life. Examples include research on perception, cognition, or behavior that involves no manipulation of the subjects. Many survey studies fall into this classification. |
| **Greater than Minimal Risk** | Studies classified in this category can potentially cause stress, pain, injury, or even death. Examples include research involving vulnerable populations, research involving drugs or invasive procedures, and research involving deception. Such a study requires approval by an IRB panel composed of members qualified to review research in that field. |

**Other Ethical Issues Addressed by IRB Panels**

1. An IRB panel must be convinced that researchers have taken every precaution to ensure that participants are not subjected to undue harm or stress.
2. Voluntary informed consent is essential for research involving human subjects. This informed consent should include:

	* An honest, complete description of the nature of the research,
	* A statement that the research is voluntary and that participants can withdraw at any time,
	* An identification of the risks and benefits of the research,
	* A description of how confidentiality will be protected,
	* A description of compensation,
	* A description of what information researchers will share with participants,
	* Contact information for the primary investigator responsible for the research.

1. All information collected in a research project should remain confidential. For example, researchers should ensure that

	* Data are “locked away” in a secure setting
	* Electronic databases are protected

Note that confidentiality is different from anonymity (results are anonymous when the researcher cannot link a given response to a give participant).

1. In an unquestionably ethical study, a researcher should not deceive their study participants. At times, however, researchers may choose to hide the true nature of the study from study participants in order to protect the validity of the study. An IRB may allow researchers to use deception under certain conditions:
* Participants must be provided with as much information as possible about the study to consent voluntarily
* Researchers must convince the IRB that deception is necessary to collect data and that it will cause little or no harm
* Researchers must arrange to fully inform the patients of the true nature of the study in a timely manner at the conclusion of the study. When subjects are deceived, a thorough debriefing of study participants at the conclusion of the experiment may be required.

**Ethical Guidelines for Statistical Practice**

In 2016, the American Statistical Association published “[Ethical Guidelines for Statistical Practice](http://www.amstat.org/about/ethicalguidelines.cfm).” The purpose of these guidelines is as follows.



These ethical guidelines address eight general topic areas:

1. Professional Integrity and Accountability
2. Integrity of Data and Methods
3. Responsibilities to Science/Public/Funder/Client
4. Responsibilities to Research Subjects
5. Responsibilities to Research Team Colleagues
6. Responsibilities to Other Statisticians or Statistical Practitioners
7. Responsibilities Regarding Allegations of Misconduct
8. Responsibilities of Employers, Including Organization, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners

Take some time to look through the document, and note some of the specific ethical considerations listed under each of these eight topics.

**Ethical Guidelines for Data Science**

* Data Science Ethical Framework, UK Government
* Code of Conduct



Six Guiding Principles of the Data Science Ethical Framework

1. Start with clear user need and public benefit
2. Use data and tools wich have minimum intrusion necessary
3. Create robust data science models
4. Be alert to public perceptions
5. Be as open and accountable as possible
6. Keep data secure