

SDS 237 - Field Journal

Anny(Jiaqi) Li

Last updated on 2022-03-01

Instructions: For each journal entry, first fill out the metadata for that entry by entering your name, the date you observed a data environment, the date you wrote up your entry, and the location of the environment you observed in between the quotation marks of the associated code snippet. Be sure to enter the date as month/day/year. Below this code snippet, write your journal entry. Once your done, knit the document and push your changes to GitHub.

Journal Entry 1

`## Warning: All formats failed to parse. No formats found.`

This entry documents a data environment Anny(Jiaqi) Li observed on NA in Chase House 1st floor. The observations were written up on 2022-03-01.

My friend M and me were having a conversation about the limitations of Big Data.

While doing her CS homework in the living room of Chase House, M started talking about how strongly she believed in the predictive and calculating power of Big Data. “I believe that everything in this material world interact with each other and all the phenomenon, whether observed or not by humans, is a result of such interaction” M said with an excitement on her face. She continued: “Since Big Data is capable of discovering the pattern and interaction hiding behind different data, if we could gather enough data that captures everything in the physical world, then we can even predict the future.” Her argument reminded me of the classical determinist idea — Laplace’s demon, which states that if we know how every atom in the universe work, then we could plug this information into a formula and predict the future. It is indeed very exciting to imagine the scenario where humans have full control of everything in the world.

However, technology we have today is not advanced enough to measure and collect everything and I highly doubt that we would ever be able to do that. The Uncertainty Principle in Quantum Physics has already proposed the idea that the position and velocity of atoms cannot be both measured exactly in theory, and the closest measurement we could reach is probability cloud, which only provides us a probability of detecting a given particle in a certain space. Though this principle is not necessarily true, it still makes me to think that there are limitations to the capability of Big Data. Besides the limitation of its capability, Big Data employs mostly **inductive reasoning** because it collects data based on observations, come up with a model to best explain the relationship between data, which could be called generalization, and finally arrive at its conclusion. Unlike **deductive reasoning**, starting from a paradigm, inductive reasoning could not guarantee the conclusion is 100% true even with extremely strong evidence supporting it.

An **interpretivist** could argue that no matter how strong the association between two variables are, it could be just an illusion.

After hearing my view about Big Data, M refuted: “If everyone is as nihilistic as you, the society would achieve nothing, it would be a total failure”. The discussion of **epistemology** of big data, which concerns about what makes data and what counts as valuable data and so on, must continue because it determines the direction of research in a fundamental way.

Journal Entry 2

This entry documents a data environment FILL YOUR NAME observed on 1990-01-01 in FILL LOCATION.
The observations were written up on 1990-01-01.

[Enter your write-up here.]

Journal Entry 3

This entry documents a data environment FILL YOUR NAME observed on 1990-01-01 in FILL LOCATION.
The observations were written up on 1990-01-01.

[Enter your write-up here.]

Journal Entry 4

This entry documents a data environment FILL YOUR NAME observed on 1990-01-01 in FILL LOCATION.
The observations were written up on 1990-01-01.

[Enter your write-up here.]

Journal Entry 5

This entry documents a data environment FILL YOUR NAME observed on 1990-01-01 in FILL LOCATION.
The observations were written up on 1990-01-01.

[Enter your write-up here.]