Epistemology – the science of knowing.
Methodology – a subfield of epistemology, the science of finding out.
How do we know what we know?
How much is direct experience?
How much is passed on?

Do you know the dark side of the moon is cold?
Do you know they speak Portuguese in Guinea Bissau?

Agreement Reality. The things we ‘know’ as part and parcel of the culture we share with those around us.

In what ways does this help?
In what ways does it hinder?

Role of Tradition in establishing agreement reality; don’t necessarily have to experience to have knowledge.

Role of Authority in establishing agreement reality; experts stating assessments about what is a fact.

Agreement reality is not the same as proven reality, which is part of what we try to test in social science inquiry.
Problems in Social Science inquiry.

Inaccurate observation / lack of attention to specific detail. We need to be deliberately paying attention to particular facts and recording them correctly. Problem of recall of specific item if we were not paying attention to this item.

Overgeneralization. We see one case and we think it applies to all cases.

Selective Observation. We see what we expect to see. We dismiss contrary evidence to confirm the version of reality we choose to see rather than allow the contrary evidence to refute our favored interpretation.

Illogical reasoning. The exception proves the rule. Dismiss evidence that refutes one interpretation as exceptions.

The foundation of social science.

Theory, the logical foundation that provides systematic explanations.

Data collection is the observational aspect.

Social theory has to do with what is, not what should be (p. 10).

That is the theme of the Friedman article; normative compared to positive approaches.

Distinct from what should be, or values.

What should be is drawing on your Management class: what is the objective of our project? What are the indicators?

To contrast things, we have to first agree on measurable criteria with which we will evaluate. That is the monitoring indicators approach.
Looking for patterns of regularity; average overall behavior that may have exceptions, but describes the overall tendency. That is our research part of the overall approach.

Overall tendency is where you draw on the statistics training.

Then you take the findings and go back to the management part. How can we take these findings from the research treated by the statistics to do a better job in realizing our objectives?

In doing analysis and understanding behavior we need to understand the context. What are the formal rules and regulations? Where can you park your car and on which side of the street today? Can you eat food in this classroom?

Informal norms and customs. Which side of the stairs is for going up and which side is for going down? Which side of the sidewalk for east to west crossing of the quad?

A critique of social science is that when we identify a regularity, that regularity is trivial. “Female headed households in rural Senegal are poorer than male headed households”.

However, what makes things interesting is that which is understood as true might not hold up to investigation. “Actually, they are not statistically poorer due to the role of male labor outmigration and remittances.”

That which seems trivial ex post was not understood ex ante. “Fulani female headed households are poorer but Wolof are not due to differing land inheritance customs that vary across ethnic groups.”

Figuring out which is the exception and which is the general pattern is not always easy, especially if not treated with some systematic design. Adopting the formal structure of probability, and the law of large
numbers. What is the general tendency and what is an outlier (and what does the outlier tell us).

Table 2: Average herd size and income measures by sub-group

<table>
<thead>
<tr>
<th>Group</th>
<th>Herd Size TLU</th>
<th>Total Income per capita per day</th>
<th>Cash Income as % of Total Income</th>
<th>Total Income variability (cv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Left out</td>
<td>7.3</td>
<td>$0.20</td>
<td>29%</td>
<td>1.32</td>
</tr>
<tr>
<td>2) Moving From</td>
<td>7.2</td>
<td>$0.27</td>
<td>46%</td>
<td>0.90</td>
</tr>
<tr>
<td>3) Staying With</td>
<td>23.7</td>
<td>$0.34</td>
<td>21%</td>
<td>0.82</td>
</tr>
<tr>
<td>4) Combining</td>
<td>26.0</td>
<td>$0.46</td>
<td>35%</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Significant difference in means by groups, t-statistics

<table>
<thead>
<tr>
<th>Group</th>
<th>t12</th>
<th>t12 ***</th>
<th>t12 ***</th>
<th>t12 **</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t13 ***</td>
<td>t13 ***</td>
<td>t13 ***</td>
<td>t13 ***</td>
</tr>
<tr>
<td></td>
<td>t14 ***</td>
<td>t14 **</td>
<td>t14 *</td>
<td>t14 ***</td>
</tr>
<tr>
<td></td>
<td>t23 ***</td>
<td>t23 ***</td>
<td>t23 ***</td>
<td>t23 ***</td>
</tr>
<tr>
<td></td>
<td>t24 ***</td>
<td>t24 ***</td>
<td>t24 ***</td>
<td>t24 ***</td>
</tr>
<tr>
<td></td>
<td>t34</td>
<td>t34 ***</td>
<td>t34 ***</td>
<td>t34 ***</td>
</tr>
</tbody>
</table>

Note that observation has challenges.

The knowledge by people we are observing that they are being observed may lead to changed behavior, making the observed behavior not the ‘natural’ behavior but the response to being observed. They might be performing for us so we are recording their construct, not their natural behavior.

The “Hawthorne Effect”. Observing workers in a plant in Illinois in the 20s and 30s. They changed working hours, break times, lighting in the plant. Worker productivity went up when these changes were made. Eventually they figured out the productivity went up because people were aware they were being observed, not due to the changes.

Hard to observe behavior if behavior changes in response to observation.
People perform.

Note challenge to informed consent. There is an inherent conflict.
What you see is also filtered against what you are prepared to see:

http://www.jstor.org/stable/665280?seq=1#page_scan_tab_contents

Unidentified participant: Mr. Faulkner—

William Faulkner: Yes, sir.

Unidentified participant: In *Absalom, Absalom!* is any one of the people who talks about Sutpen have the right view, or is it more or less a case of thirteen ways of looking at a blackbird with none of them [getting it] right?

William Faulkner: That's it exactly. I think that—that no one individual can—can look at truth. It—it it blinds you. You look at it, and—and you—you see one phase of it. Someone else looks at it and sees a slightly awry phase of it, but taken all together, the truth is—is in what they saw, though nobody saw the truth intact. So—so these are—are true as far as—as Miss Rosa and as Quentin saw it. Quentin's father saw what—what he believed was truth. That was all he saw. But the old man was—was himself a little too big for—for people no greater in stature than Quentin and Miss Rosa and Mr. Compson to see all at once. It would've taken, probably, a wiser or more tolerant or more sensitive or more thoughtful person to see him as he was. It was, as you say, thirteen ways looking at a blackbird. But the truth, I would like to think, comes out, that when the reader has read all these thirteen different ways of looking at the blackbird, the reader has his own fourteenth image of that blackbird, which I would like to think is the true one.

http://faulkner.lib.virginia.edu/display/wfaudio29_1
Social science is study of the aggregate, rather than the individual. We are looking for collective regularities. Individual decisions, but in aggregate, they create aggregate facts.

Total Fertility Rate for example, that changes over time but is not a collective decision, but a characteristic of individual decisions aggregated to the collective level. Not an outcome of social planning (though may reflect social programs and policies). TFR is the average number of children a woman will give birth to in her reproductive years conditional upon her birth cohort.

The goal is not to understand an individual person, but ‘to understand the systems in which people operate, the systems that explain why people do what they do’ (p. 14).

Variable. Sets of attributes that can have different values.

Attributes. Characteristics of peoples or things.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Number of years since birth</td>
</tr>
<tr>
<td>Gender</td>
<td>Female, Male</td>
</tr>
<tr>
<td>Occupation</td>
<td>Professor, Lawyer, Candymaker</td>
</tr>
<tr>
<td>Social Class</td>
<td>Upper, Middle, Lower, Lumpenproletariat</td>
</tr>
<tr>
<td>Height</td>
<td>Value in centimeters</td>
</tr>
</tbody>
</table>

Attributes are agreed upon concepts that sort complicated reality into conceptual ‘bins’. The bins are constructs, choices. We have to figure out what to do with a transgender lawyer who makes candy on the weekends from a wealthy family who self identifies as working class.
So sorting cases into bins is part of the art, or our first step from complicated reality to our artificial constructed version of reality used for analysis.

When looking at two variables, we might see some kind of pattern in correlation. The variables move together in some sense.

Rho=Expected value $$\frac{((x – \text{mean } x)*(y-\text{mean } y))}{(\text{standard deviation } x*\text{standard deviation } y)}$$

That computation is mechanical based on statistical theory.

Do things move together positively or negatively?

Total Fertility rate and wealth?
Height and lifetime expected income?

To put structure on the correlation, we can reach for the idea of causation. That moves beyond statistical manipulation and forces us to reach for a behavioral theory.

Much of what we do in social science research is a search through correlations to try to identify causation.

Why does Total Fertility Rate decrease as expected income increases?
Why do taller people have higher expected lifetime earnings?

We are trying to identify independent variables, the things on the right hand side, the predictors, from the dependent variables, the things on the left hand side, the outcomes.

In classic $$y=f(x)$$ notation, the $$y$$ is the outcome, the dependent variable, the $$x$$ is the independent variable(s).
We spend countless hours arguing about what is the outcome and what is the input. What is the dependent variable? What is jointly determined?

Higher income countries have better governance.

Countries with better governance have higher incomes.

Higher wages go to people who are less likely to be absent for health problems.

People with higher incomes eat more healthy diets.

\[ Q=286-20*p \] is the demand

\[ Q = 88+40*p \] is the supply

Also have to worry about the outcome being related to the relationship in question.

Students are highly satisfied with the IR curriculum revision.

Students who have higher satisfaction complete the survey.

Students think the revision is a disaster.

Students who are unhappy are more likely to fill out the survey.

People who went through the treatment program are employed and no longer homeless.

We did not find the ones who went through and are still homeless.
Purposes for conducting Social science research:

1) Exploratory. What do people around here mean by the word ‘household’? What is a working definition of this concept that we can use to pose questions. Herd owning, common cooking pot, joint labor in fields, wives and co-wives, labor migration concepts....

2) Descriptive. How many households in this area are female headed? What is the average number of children? What is their experience with formal education? What is the mid upper arm circumference for the kids?

3) Explanatory. Do the kids residing in female headed households have higher or lower educational attainment than kids from non-female headed households? How about nutrition as measured by MUAC?

Different methodologies and skills are called for when trying to address each purpose.

In research teams, we tend to have comparative advantages, but we need to do all kinds at some point.
Ideographic explanations. The full listing of all the individual influences that lead to a given outcome. Explain by exhaustive listing the multiple reasons a given unique outcome was realized. Fully explaining the single case / event in question. What are all the factors that led to the Iranian nuclear arms deal? What are all the factors that explain the Brexit vote? What are all the reasons and influences that brought you to be in this seat in this classroom right now?

Nomothetic – identify a few causal factors that lead to general classes of outcomes. A general explanation of what leads to a class of conditions and outcomes. What variables lead to success or failure of a sample of 30 nuclear arms deals over time? What economic, political, and historical factors have influenced referendum votes on political and economic integration in the post war era? What are the most important variables in the characteristics of MA students in Maxwell that predict they will be enrolled in this class?

Induction; moving from the particular to the general. Take observations of individual cases and make statements about general patterns. Begins with whether (I wonder whether this factor matters for that outcome) and moves to why (Hmm, it seems to matter, I am going to try to explain why it matters).

Individual household wealth seems to influence whether children attend secondary school. How to interpret answers on the survey form in terms of overall patterns.
Deduction. Moving from the general to the specific. Begins with why and moves to whether. In other studies, we have found household wealth influences school enrollment rates. Is this the case in southern Ethiopia? How to write questions on the survey form to see if patterns seen elsewhere hold up here as well.
Determinism versus agency; \( y = f(x) \) or \( y = f(x) + u \). By setting up structure of average behavior we have not described individual decision making or motivation, but a description of the aggregate.

Quantification. Putting numbers on things. What year were you born? But why years? Months? Days? A choice is made.

Qualitative data. Are you young or old? What is the threshold that divides? Again, a choice is made.
We also have quantification of qualitative concepts if we agree to measures and constructs.

Qualitative tends to work better for ideographic and quantitative for nomothetic explanations though this does not have to be the case.

Qualitative looks at the multiple causes of a case, quantitative looks at the overall patterns in multiple cases leading to certain categories of outcomes.
Overview of Research Proposals (but also of research projects):
Abstract.
Introduction.
Review of the Literature.
Specify the Problem / Question / Topic noting explicitly why this study is needed and what is being added to knowledge.
Research Design.
   Data gathering methods
   Sampling frame and population to whom this research is applicable
Research Ethics
Plan for analysis of data, noting methods and standards of evidence
Discussion and conclusion; what will we know if you do this research that we don’t know now (reiterate now that it is all detailed as described above).

Tell them what you are going to tell them, tell them, then tell them what you told them. Aristotle.