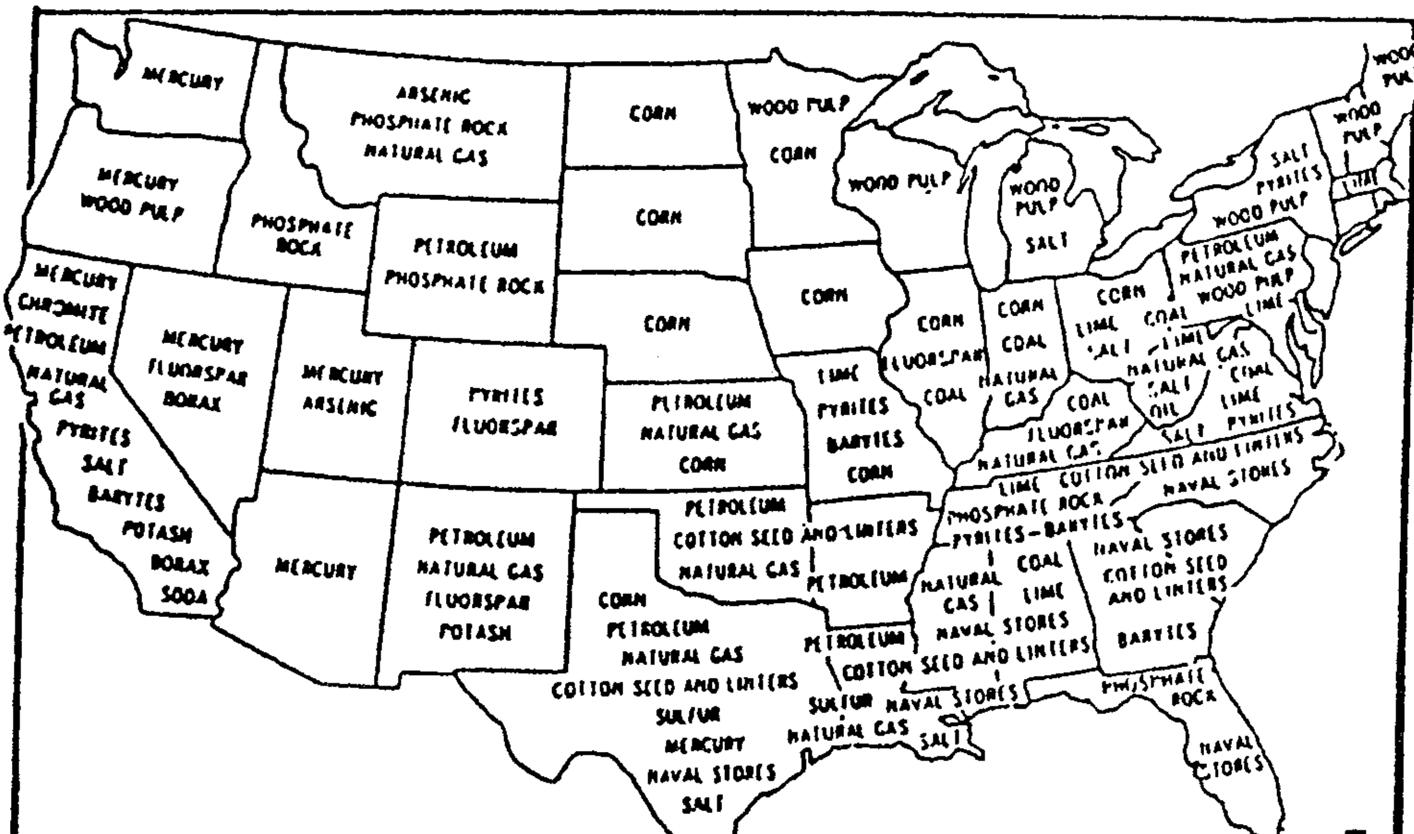
Lecture 06

More on graphical statistics:

Recall that the purpose of a graphical descriptive statistic is to facilitate insight with respect to the dataset. Although there are various standard graphical statistics (e.g. histograms, boxplots, scatterplots), sometimes data with a non-standard structure may benefit from a special-purpose graphical display.

The only limit in developing graphical displays is your imagination. Keep in mind however that the goal is to learn from the display. Therefore, simplicity and clarity are important considerations. On the following pages, we give an example of a non-standard dataset and special purpose graphical displays that aid in addressing various questions.

Raw Materials for Chemical Industry



GR 158 **Superposition de n caractères** **1**

Question 1: Which states produce particular raw materials?

Question 2: Which raw materials do states produce?

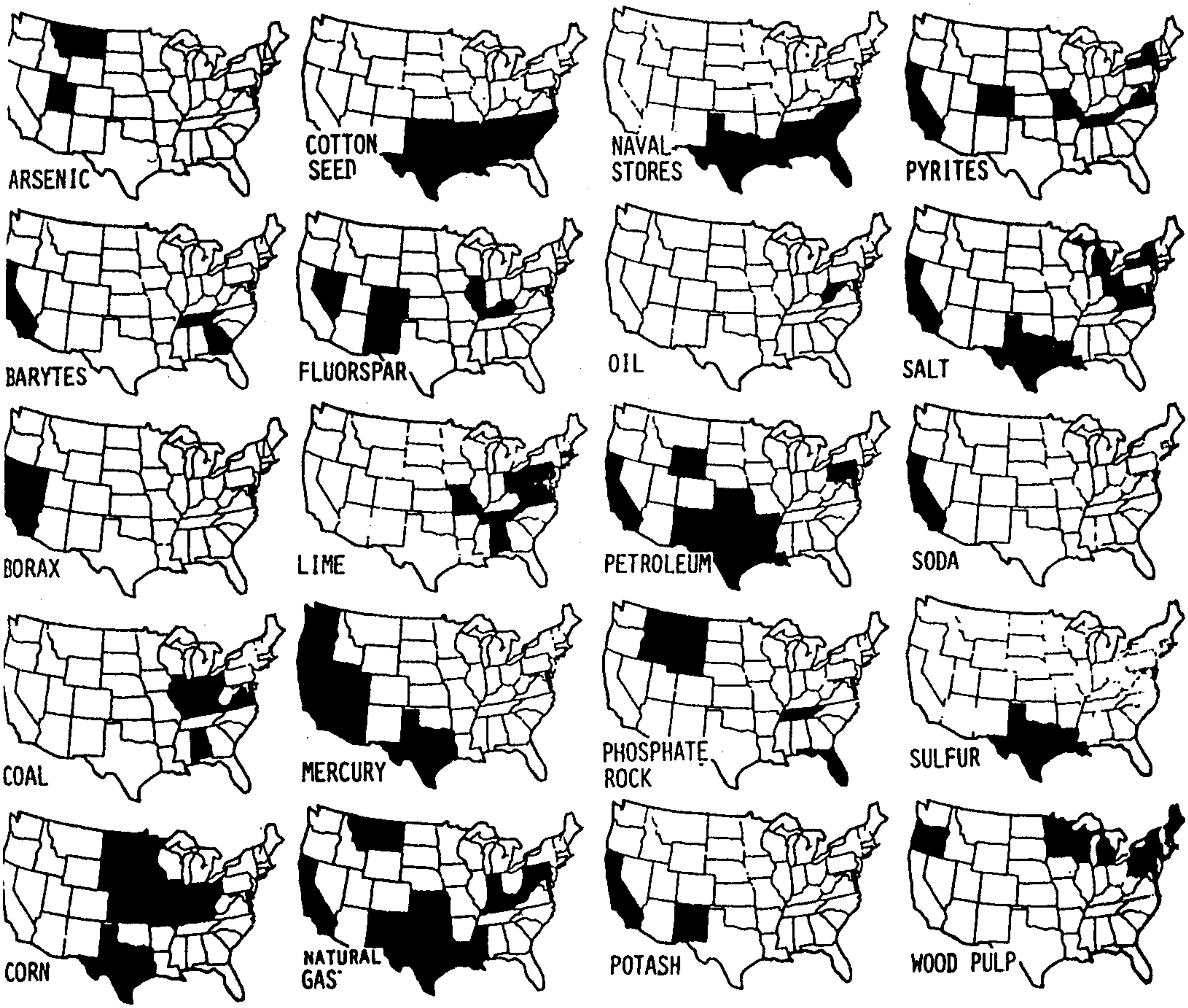


Figure 3. A collection of the one-variable maps arranged alphabetically.

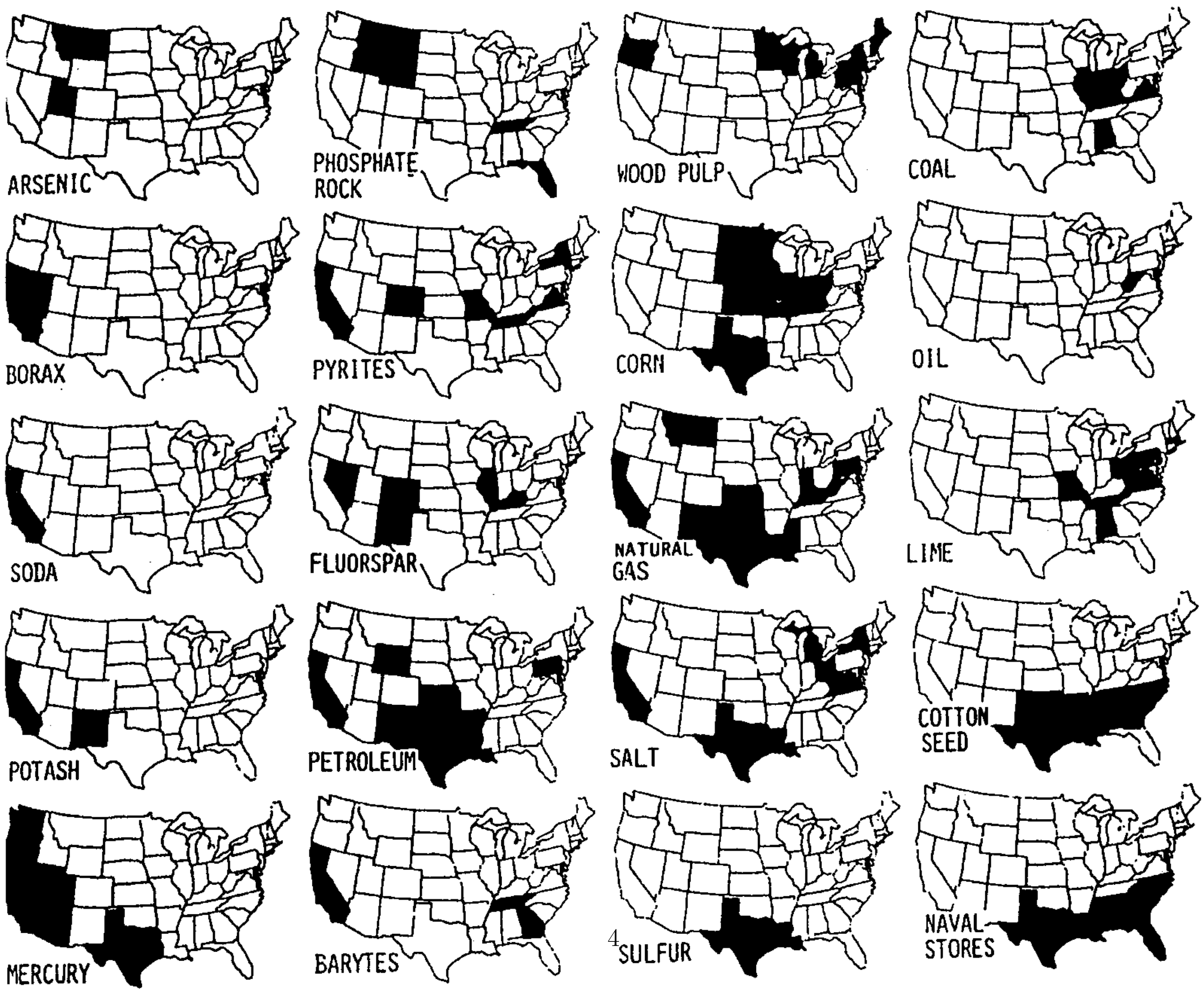


Figure 4. A collection of the one-variable maps arranged geographically.

Introduction to probability:

We think of an *experiment* as any action that produces data.

The *sample space* is the set of all possible outcomes of the experiment.

An *event* is a subset of the sample space.

Example 1: flipping a coin three times.

Example 2: total auto accidents in BC in a year

Example 3: lifespan in hours of 2 components

Problem: Write down the sample space wrt the experiment where you roll a die until an even number occurs.

Set theory for events using Venn diagrams:

- **"A union B"** $\equiv$ "A **or** B**"** $\equiv A \cup B$

- **"A intersect B"** $\equiv$ "A **and** B**"** $\equiv A \cap B \equiv AB$

- **"A complement"** $\equiv \bar{A} \equiv A' \equiv A^c$

Definition: A **and** B **are *mutually exclusive* (*disjoint*) if** $A \cap B = \phi$**.**

DeMorgan's Law: $\overline{A \cup B} = \bar{A} \cap \bar{B}$

Something to think about: Probability is used in everyday language yet it is not well defined. What is meant by the statement "the probability of rain today is 0.7"?

Oxford English Dictionary definition of probability: extent to which an event is likely to occur, measured by the ratio of favourable cases to the whole number of cases possible

Kolmogorov (1933) provided the following definition of probability:

A probability measure P satisfies three axioms

1. **For any event A,** $P(A) \geq 0$

2. $P(S) = 1$ **where S is the sample space**

3. **If $A_1, A_2, \ldots,$ are disjoint,** $P(\cup A_i) = \Sigma P(A_i)$

Discussion points:

Useful derivations from the Kolmogorov defn:

Example: $P(\bar{A}) = 1 - P(A)$

Example: $P(\phi) = 0$

Example: If $A \subseteq B$**,** $P(A) \leq P(B)$

Example: $P(A \cup B) = P(A) + P(B) - P(AB)$

Example:

$$\begin{aligned} P(A \cup B \cup C) = \; & P(A) + P(B) + P(C) \\ & -P(AB) - P(AC) - P(BC) \\ & +P(ABC) \end{aligned}$$