STA113: Probability and Statistics in Engineering

Feng Liang and Robert Wolpert

Institute of Statistics and Decision Sciences

Duke University

- Statistics : the science of data. This involves collecting, classifying, summarizing, organizing, analyzing and interpreting data.

- Two Branches of Statistics

- Descriptive statistics: summarization and description of data set

- Inferential statistics: using sample data to make an inference about a population

- A population is a date set that is the target of our interest.

- A sample is a subset of data selected from a population

Qualitative Data

- Visual plots:
  - table,
  - bar graph,
  - pie chart ...
- Numerical measure:
  - frequency (counts),
  - relative frequency

Types of Data

- Quantitative data: quantity or amount of something
  e.g. stock price, temperature, ...
- Qualitative (or categorical) data: no quantitative
  interpretation, i.e., they can only be classified into categories
  e.g. gender, country names, ...
  e.g. ranks?
  - Nominal: can NOT be meaningfully ordered
  - Ordinal: can be ranked from high to low

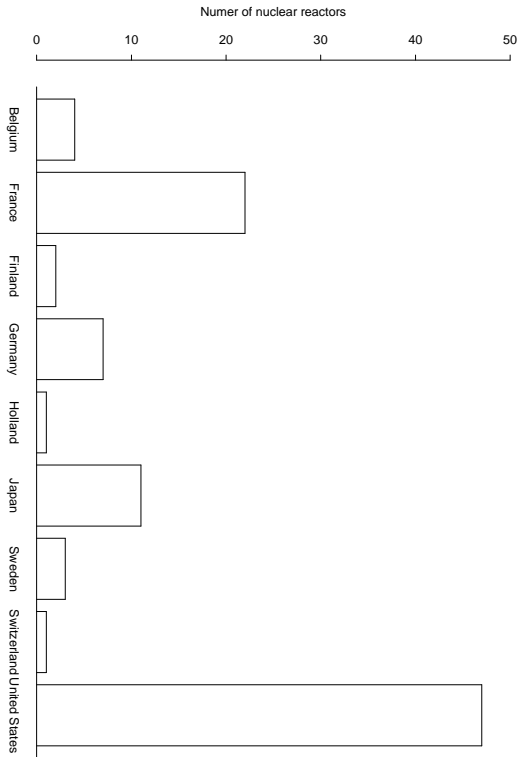Figure 1: Vertical Bar Graph showing the number of large nuclear reactors for nine countries

---

An Example

Table 2.1 Summary Frequency Table for Largest Nuclear Reactors Worldwide

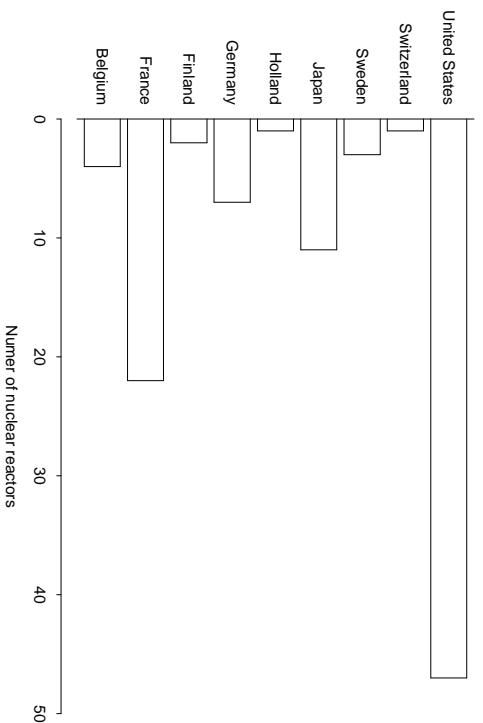| Category	Frequency	Relative Frequency
Country	Number of Nuclear Reactors	Proportion
Belgium	4	0.0408
France	22	0.2245
Finland	2	0.0204
Germany	7	0.0714
Holland	1	0.0102
Japan	11	0.1123
Sweden	3	0.0306
Switzerland	1	0.0102
United States	47	0.4796
TOTALS	98	1.0000

Figure 2: Horizontal Bar Graph showing the relative frequency of large nuclear reactors for nine countries
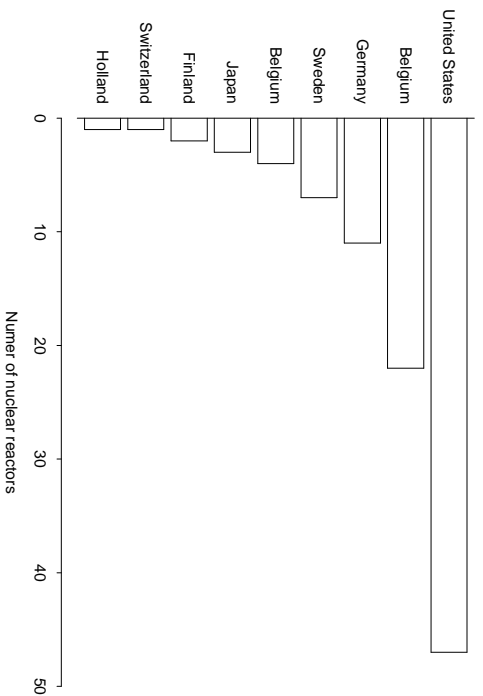
Figure 3: Pareto diagram for the nuclear reactor data