Katari, Manpreet Singh; **Tyagi, Sudarshini**; **Shasha, Dennis**
Statistics is easy. Case studies on real scientific datasets. (English) Zbl 1469.62004
Synthesis Lectures on Mathematics and Statistics 39. San Rafael, CA: Morgan & Claypool Publishers
(ISBN 978-1-63639-091-8/hbk; 978-1-63639-089-5/pbk; 978-1-63639-090-1/ebook). xi, 63 p. (2021).

Publisher's description: Computational analysis of natural science experiments often confronts noisy data
due to natural variability in environment or measurement. Drawing conclusions in the face of such noise
entails a statistical analysis.

Parametric statistical methods assume that the data is a sample from a population that can be charac-
terized by a specific distribution (e.g., a normal distribution). When the assumption is true, parametric
approaches can lead to high confidence predictions. However, in many cases particular distribution as-
sumptions do not hold. In that case, assuming a distribution may yield false conclusions.

The companion book [*D. Shasha* and *M. Wilson*, Statistics is easy! 2nd ed. San Rafael, CA: Morgan &
Claypool Publishers (2010; Zbl 1457.62010)] gave a (nearly) equation-free introduction to nonparametric
(i.e., no distribution assumption) statistical methods. The present book applies data preparation, machine
learning, and nonparametric statistics to three quite different life science datasets. We provide the code
as applied to each dataset in both R and Python 3. We also include exercises for self-study or classroom
use.

MSC:

62-01	Introductory exposition (textbooks, tutorial papers, etc.) pertaining to statistics
62E15	Exact distribution theory in statistics
62D05	Sampling theory, sample surveys
62-08	Computational methods for problems pertaining to statistics
97K70	Foundations and methodology of statistics (educational aspects)
00A06	Mathematics for nonmathematicians (engineering, social sciences, etc.)

Keywords:

basic concepts of resampling; standard statistical measures

Software:

R; Python

Full Text: DOI