◉ Dataset

∧ | 722

Medical Cost Personal Datasets

Insurance Forecast by using Linear Regression

Miri Choi • updated 2 years ago (Version 1)

Data   Tasks (2)   Kernels (237)   Discussion (7)   Activity   Metadata

🧰 **Usability** 8.8   ⚖️ **License** Database: Open Database, Contents: Database Contents   🏷 **Tags** education, health, finance, healthcare, insurance

---

Description

Context

Machine Learning with R by Brett Lantz is a book that provides an introduction to machine learning using R. As far as I can tell, Packt Publishing does not make its datasets available online unless you buy the book and create a user account which can be a problem if you are checking the book out from the library or borrowing the book from a friend. All of these datasets are in the public domain but simply needed some cleaning up and recoding to match the format in the book.

Content

Columns

- age: age of primary beneficiary
- sex: insurance contractor gender, female, male
- bmi: Body mass index, providing an understanding of body, weights that are relatively high or low relative to height,
  objective index of body weight (kg / m ^ 2) using the ratio of height to weight, ideally 18.5 to 24.9
- children: Number of children covered by health insurance / Number of dependents
- smoker: Smoking
- region: the beneficiary's residential area in the US, northeast, southeast, southwest, northwest.
- charges: Individual medical costs billed by health insurance