MEPS tutorials

Mark Bounthavong

2023-07-22

Contents

About

This is a collection of tutorials that use data from the Agency for Healthcare Research and Quality (AHRQ) Medical Expenditure Panel Survey (MEPS).

These tutorials use R and RStudio for data loading, manipulation, analysis, and presentation.

Chapter 1 provides an introduction on how to load and import the data into R.

Chapter 1

Introduction

The Agency for Healthcare Research and Quality (AHRQ) Medical Expenditure Panel Survey (MEPS) is a set of data on U.S. households about their healthcare expenditures. It includes data on the individual / household demographics, socioeconomic status, insurance coverage, and healthcare expenditures. Healthcare expenditures include data on health-related spending, medical conditions, prescriptions, and utilization (e.g., number of office-based visits). MEPS draws upon a nationally representative subsample from the National Health Interview Survey, which is conducted by the National Center for Health Statistics. Hence, MEPS provides researchers with the ability to generate estimates for the representative U.S. population.

1.1 MEPS Data

MEPS data are located on their website in their data files page. You can find data from 1996 to the most recent available year (during the writing of this tutorial, 2020 was the latest release).

The MEPS data files include the Full-Year Consolidated Data files, which is the calendar-year summary of the different longitudinal panels. The Full-Year Consolidated Data files contain information on the annual healthcare expenditures by the type of care; it contains data on spending, insurance coverage, health status, patient satisfaction, and several health conditions. The Full-Year Consolidated Data files also contains information from several surveys (e.g,. Diabetes Care Survey).

1.2 Load MEPS data into R

MEPS data can be downloaded onto your local storage and read into a statistical software program such as Stata or R. But you can also communicate directly

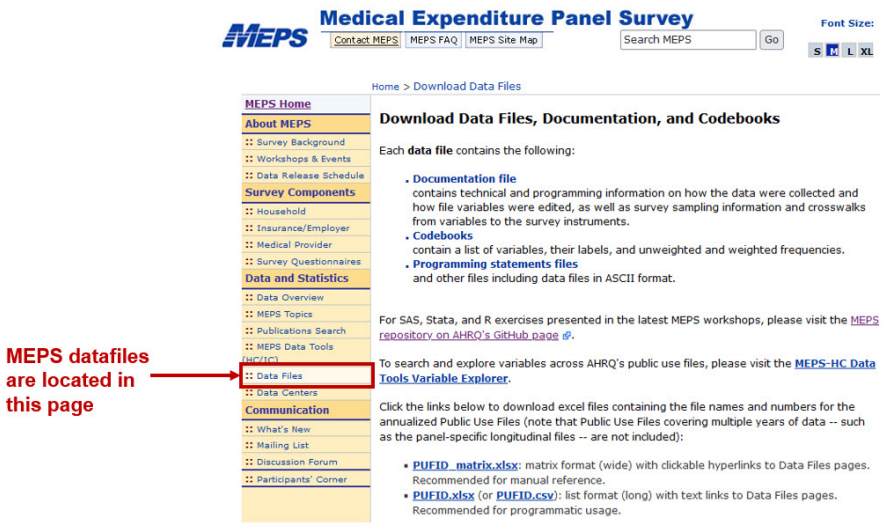

Figure 1.1: Figure 1 - Location of MEPS data files

with the AHRQ MEPS website to load your data rather than having to download it. We will load the Full-Year Consolidated Data file from 2020, which is named `HC-224`. To find out the name of the file, you will need to go MEPS data files site and click on the Full-Year Consolidated Data files. In this page (Figure 2), you can see the data file with the code `HC-224`, which is the Full-Year Consolidated Data file for 2020. When we enter this into our R code, we will use the file name `h224`.

You will need to download and install the `MEPS` package. The `MEPS` package will provide tools for you to load and manipulate the MEPS Data files. You will need to have R `devtools` package installed.

```
## Install the devtools package
# install.packages("devtools") ## You only need to install this once
# library("devtools")  ## You will need to reload the MEPS package each time you resta
# install_github("e-mitchell/meps_r_pkg/MEPS") ## This will install the MEPS package f
```

There are two methods to load MEPS data into R.

Method 1 requires that you know the file name. In this example, the MEPS 2020 Full-Year Consolidated Data file is named `h224`. We will use the `read_MEPS` function to load the MEPS data onto R.

When using Method 2 to load the MEPS data, we don't need to know the file name, but we need to know the year and the data type. For example, for the Full-Year Consolidated Data file, we use the `year = 2020` and `type = "FYC"`

Figure 1.2: Figure 2 - Full-Year Consolidated Data files and other data types

PUF no.	File(s), Documentation & Codebooks	Data update	Year	File type
HC-224	2020 Full Year Consolidated Data File	X	2020	Household Full Year File

Figure 1.3: Figure 3 - H224 is the MEP 2020 Full-Year Consolidated Data file.

option. For this method, we will also use the `read_MEPS` function to the MEPS data onto R.

The `tolower` function is used to change all the variable names from upper case to lower case. MEPS defaults the column names to upper case. I like to change this to lower case because it's easier for me to type.

```
### Load the MEPS package
library("MEPS") ## You need to load the library every time you restart R

#### Method 1: Load data from AHRQ MEPS website
hc2020 = read_MEPS(file = "h224")

#### Method 2: Load data from AHRQ MEPS website
hc2020 = read_MEPS(year = 2020, type = "FYC")

## Change column names to lowercase
names(hc2020) <- tolower(names(hc2020))
```

There are over 1400 variables in the MEPS 2020 Full-Year Consolidated Data file. We can reduce this to the essential variables using the `subset` function. This will generate a smaller data frame that we will call `keep_meps`. The variables that we want to collect are the subject unique identifier (`dupersid`), the survey weights (`varpsu`, `varstr`, `perwt20f`), and the total healthcare expenditures for 2020 (`totexp20`).

```
### Keep the subject's unique ID, survey weights, and total expenditures
keep_meps <- subset(hc2020, select = c(dupersid, varpsu, varstr, perwt20f, totexp20))

head(keep_meps) ## View the first six rows of the data frame
```

```
## # A tibble: 6 x 5
##   dupersid   varpsu varstr perwt20f totexp20
##   <chr>       <dbl>  <dbl>    <dbl>    <dbl>
## 1 2320005101      1   2079    8418.      459
## 2 2320005102      1   2079    5200.      564
## 3 2320006101      1   2028    2140.      140
## 4 2320006102      1   2028    2216.     4673
## 5 2320006103      1   2028    4157.      410
## 6 2320012102      2   2069    1961.     2726
```

Since MEPS uses a complex survey design, these weights are needed to estimate standard errors that are reflective of the representative sample of the U.S. population. We'll learn how to apply these survey weights to the MEPS data files in a future tutorial.

1.3 Conclusions

Loading MEPS data into R allows us to perform analysis easily and quickly. In this tutorial, you learned how to load MEPS data into R directly from the MEPS website. However, you can also download the MEPS data onto your local storage and use the `setwd` command to set the working directory.

In future tutorials, we'll learn how to apply the survey weights and perform descriptive analyses using the MEPS data files.

1.4 Acknowledgements

There are a lot of tutorials on how to use MEPS data with R. I found the AHRQ MEPS GitHub page to be an invaluable resource.

This is a work in progress, and I may update this in the future.

Chapter 2

Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

2.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
    - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world` = `# Hello world {#hello-world}`.
    - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter 2.
    - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

2.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure 2.1.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Don't miss Table 2.1.

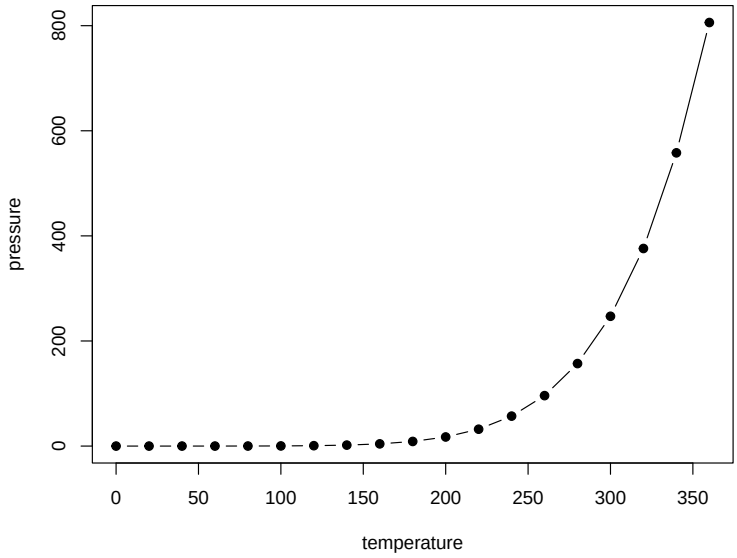

Figure 2.1: Here is a nice figure!

```
knitr::kable(
  head(pressure, 10), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

Table 2.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

Chapter 3

Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

Chapter 4

Footnotes and citations

4.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one [1].

4.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package [Xie, 2022] (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** [Xie, 2015] (this citation was added manually in an external file book.bib). Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: https://rstudio.github.io/visual-markdown-editing/#/citations

[1] This is a footnote.

Chapter 5

Blocks

5.1 Equations

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{5.1}$$

You may refer to using `\@ref(eq:binom)`, like see Equation (5.1).

5.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref(thm:tri)`, for example, check out this smart theorem 5.1.

Theorem 5.1. *For a right triangle, if c denotes the* length *of the hypotenuse and a and b denote the lengths of the* ***other*** *two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html.

5.3 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html

Chapter 6

Sharing your book

6.1 Publishing

HTML books can be published online, see: https://bookdown.org/yihui/bookdown/publishing.html

6.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

6.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book- all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's source file.

Read more about the features of this output format here:

https://pkgs.rstudio.com/bookdown/reference/gitbook.html

Or use:

```
?bookdown::gitbook
```

Bibliography

Yihui Xie. *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition, 2015. URL http://yihui.org/knitr/. ISBN 978-1498716963.

Yihui Xie. *bookdown: Authoring Books and Technical Documents with R Markdown*, 2022. URL https://CRAN.R-project.org/package=bookdown. R package version 0.29.