Speaker Notes

Marium Tapal

Hi, I am Marium Tapal from Smith College and today I'll be taking about the `fec16` R package that I maintain. This package is co-authored by Professor Benjamin Baumer, and my fellow peers at Smith, Irene Ryan and Rana Gahwagy.

NEXT SLIDE

To start, I'll give an overview of what `fec16` is:

fec16 is a data package for use in the statistical programming language R. The package contains cleaned relational data from the 2015-2016 United States federal election cycle, that was hosted by the Federal Election Commission - abbreviated as FEC.

The data includes authoritative information about the candidates, committees, contributions, expenditures, and election results.

Most datasets are included in full. But due to sizing restrictions, we decided to include small samples of the others. The package does however give the option of retrieving the full datasets through built-in functions. I'll go over this in more detail in subsequent slides.

NEXT SLIDE

Next I'll talk about why you should use `fec16`.

In summary, there are three main ideas:

One, the data is easily accessible as the package is hosted on the Comprehensive R Archive Network, commonly known as CRAN.

Two, all the data is relational. Relational data is suitable for teaching and learning, especially for introductory statistics and data science. Joins are easy to do!

Three, we have already cleaned and wrangled the data. It is user friendly!

There is always demand for quality data that is both open source and easy to access in the teaching of introductory statistics and data science.

Therefore `fec16` provides real data in a genre not seen often.

Election data is diverse: it contains important financial and geopolitical information among other sub-categories. This is especially relevant to the 2020 U.S presidential elections.

NEXT SLIDE

This table summarizes all the datasets that are included in the `fec16` package. The first column gives the name of the dataset, the second column tells us if the dataset is in full or if it is a sample. The third column gives a short description of the dataset.

NEXT SLIDE

For the datasets that are included as samples, the package contains the following functions that retreive the entire dataset.

Now I will show how you can install and load the package in R and then use it.

NEXT SLIDE

First, to install the package, type the following command in your console.

TAB

Then load the package, by calling library(fec16)

TAB

Now you are ready to use any of the pre-loaded datasets and functions.

NEXT SLIDE

For example, to view the first 6 lines of the candidates dataset, call `glimpse(candidates)`.

TAB

Since the contributions dataset is included as a sample, to load the entire dataset call read_all_contributions() and save it to a variable as shown!

NEXT SLIDE

To learn more about fec16, you can follow the QR code which will direct you to the GitHub repository where you can browse the code and read more in the vignette!

NEXT SLIDE

Thank you!