

Exchanging Data between R and Excel

CAS Spring Meeting 2012

Kirsten Singer

ms.kirsten.singer@gmail.com

How-to get Excel data into R

- I. Excel data sample
- II. Import data to R from csv
- III. Copy data to R from clipboard

Appendix I. Packages

Appendix II. Exporting data from R

Excel data sample

Data features

- Spaces within the column headings
- Missing values
- Mixture of character fields, numeric fields, and date fields

| | A | B | C | D | E | F | G | H | |
|----|-----------|----------|--------------|------------|----------------|--------------|----------|-----------|----------------------|
| 1 | Policy ID | Claim ID | Pol Eff Date | Acc Date | Paid Indemnity | Paid Expense | Care RSV | Indemnity | Care Reserve Expense |
| 2 | P000001 | C000001 | 2011-01-01 | 2011-05-30 | 0 | 100.04 | | | 207.62 |
| 3 | P000002 | | 2011-01-02 | 2011-07-01 | 51.47 | 22.16 | 63.52 | | 25.23 |
| 4 | P000003 | C000003 | 2011-01-03 | 2011-07-02 | 52.4 | 50.22 | 64.69 | | 71.87 |
| 5 | P000004 | C000004 | 2011-01-04 | 2011-07-03 | | | | | 38.09 |
| 6 | C000005 | | 2011-01-05 | 2011-07-04 | 22.37 | 26.02 | 27.6 | | 83.2 |
| 7 | P000006 | C000006 | | 2011-07-05 | 60.11 | 63.01 | 76.65 | | 18.79 |
| 8 | P000007 | C000007 | 2011-01-07 | 2011-07-06 | 16.03 | | 17.82 | | 90.7 |
| 9 | P000008 | C000008 | 2011-01-08 | 2011-07-07 | 48.12 | 79.46 | 83.63 | | 250.14 |
| 10 | P000009 | C000009 | 2011-01-09 | 2011-07-08 | 184.98 | 186.74 | 207.51 | | 155.92 |
| 11 | P000010 | C000010 | 2011-01-10 | | 142.83 | 158.7 | 176.33 | | |

- ✓ Date fields have been formatted in the proper way, as yyyy-mm-dd

Packages

Several packages exist as alternate ways to exchange data with Excel (and Access)

- **XLConnect**: advanced manipulation of workbooks from within R, with key methods of `readWorksheet`, `readNamedRegion`, `writeWorksheet`, and `xlcDump`. Developed by Mirai Solutions.
 - <http://cran.r-project.org/web/packages/XLConnect/vignettes/XLConnect.pdf>
- **xlsx**: similar to XLConnect. Developed by Adrian A. Dragulescu.
 - <http://cran.r-project.org/web/packages/xlsx/xlsx.pdf>
- **RODBC**: allows connections via ODBC to MS Access, with key method `sqlQuery`. Developed by Michael Lapsley and Brian Ripley.
 - <http://cran.r-project.org/web/packages/RODBC/RODBC.pdf>

Exporting data from R

- Basic command to copy data to clipboard: `> write.table (claims, file= "clipboard")`
- A couple extra specifications improve the experience
- ✓ Use option `sep="\t"` and the data will automatically be placed in separate columns (rather than having to perform a text-to-columns maneuver)
- ✓ Use option `row.names=FALSE` to remove numbering of the rows
- `> write.table (claims, file="clipboard", sep="\t", row.names=FALSE)`

- Basic command to write data to file: `> write.csv (claims, file= "claims.csv")`
- A couple extra specifications improve the experience
- ✓ Same options as `write.table`
- `> write.csv (claims, file="claims.csv", sep="\t", row.names=FALSE)`
