Getting the Vegas Odds

Cody Frisby

3/9/2017

Getting the Veags odds for year 2015

First:

- I go here: vegasinsider.com
- Click Golf on the main header tab
- Click News
- Use the search to try an narrow down.
- Find the most recent odds from either `vegasinsider.com` or `Sportsbook.ag` with a date **before** the masters started and copy and paste the players names and odds into a text file. Preferably we'd find the last odds predictions before the beginning of the masters. There are many odds predictions leading up to the tournament.

- Save the file on your desktop, or somewhere you know where it is.
- add to the first row of the document firstName, lastName, odds1 odds2
- Use find and replace using text editor software (RStudio has this ability) and replace the "/" with a space.

- Save the file.
- Then I read in the file via:

```
df <- read.table("~/Desktop/vegasOdds2015")
```

Note: If I get some error like

```
Error in scan(file = file, what = what, sep = sep, quote = quote, dec = dec,  : line 51
did not have 3 elements
```

I open the file in `RStudio` and go to the line that as an error and take a look.

For this file that line is `Miguel Angel Jimenez 150/1`. There are three names so `read.table` has a problem because the first row only has three stings while this row has 4. I'll just fix this manually for now by turning `Angel Jimenez` into `AngelJimenez`.

Try again by running the above `R` line, and repeat if you get that Error again.

Probably a good idea at this point to make a note of the names you are altering. For this example we have altered:

- Miguel Angel Jimenez
- Mark O Meara
- Jose Maria Olazabal

Sometimes Vegas has a two in the odds, e.g, 13/2. Adding odds1 and odds2 handles this and then we combine the two below.

```
df <- read.table("~/Desktop/vegasOdds2015", header = TRUE)
head(df)
# create the odds as a number.  Sometimes vegas has
```

```
df$odds <- df$odds1 / df$odds2; df$odds1 <- NULL; df$odds2 <- NULL
head(df)
```

The formatting of our name variable is different in our data file for the modeling and simulations than this one. We fix that issue now:

```
df$name <- paste(df$LastName, ", ", df$Firstname, sep = "")
#get rid of Firstname, Lastname
df$Firstname <- NULL; df$LastName <- NULL
head(df)
```

Don't forget, we altered a few names. Let's fix that now:

```
# we need to load the 2015 data:
df15 <- read.csv("~/Documents/masters/data/ranks2015.csv")
df15$X <- NULL # get rid of this variable.  just the row names.
head(df15)
```

Notice that this file we just loaded has only 85 rows while our Vegas odds file has 98. This means a lot fewer players played in the tournament than were eligible. A quick glance shows that Mark O Mera didn't play so we don't have to worry about updating his name.

```
# merge the final results file with the vegas odds file
temp1 <- merge(df15, df)
dim(temp1)
```

We are missing 2 players. This is most likely due to the formatting difference of the names. We can see which players we are missing using code similar to:

```
temp2 <- df15[!(df15$name %in% temp1$name), ]
temp2
```

There's our two missing players.

Compare the above list to this below:

```
temp3 <- df[!(df$name %in% df15$name), ]
temp3
```

Looks like Danny Willett's name is spelled wrong (**UPDATE: this was fixed**) in our vegas odds file. That's an easy fix. We know we need to fix Cabrerra already. That's two out of 5. JB needs to be J.B.(**this was fixed too**), different spelling for Morgan Hoffmann (**fixed**), and Bae, Sang-Moon needs to be Bae, Sangmoon.

```
# just by looking at the temp3$names vector:
temp3$name[c(1, 3)] <- c("Jimenez, Miguel Angel", "Bae, Sangmoon")
head(temp3)
df[!(df$name %in% df15$name), ]$name[c(1, 3)] <- c("Jimenez, Miguel Angel", "Bae, Sangmoon")
finaltemp <- merge(df15, df)
```

All fixed. We now finalize the Vegas odds file for further analysis when predicting the 2015 results using our simulations.

```
finaltemp$score <- NULL; finaltemp$rank <- NULL
finaltemp$Vegasodds <- finaltemp$odds; finaltemp$odds <- NULL
finaltemp$vegasranks <- rank(finaltemp$Vegasodds,
                             ties.method = "min")
write.csv(finaltemp, "~/Documents/masters/data/vegasodds2015.csv",
          row.names = FALSE)
```

Getting the Veags odds for year 2014

For the 2014 odds, the above Vegas odds site did not have the proper odds, from what I could tell, but I did find a sbnation.com page that had them. So, I copied and pasted the table on this page http://www.sbnation.com/golf/2014/4/6/5587748/2014-masters-odds-current into a blank file using a text editing app.

```
df <- read.table("~/Desktop/odss2014SBnation", header = T)
```

Similar procedure as above to fix the lines with move than 3 elements.

And after replacing the "/" with a space I do this:

```
df$odds <- df$odds1 / df$odds2; df$odds1 <- NULL; df$odds2 <- NULL
head(df)
df$name <- paste(df$lastname, ", ", df$firstname, sep = "")
#get rid of Firstname, Lastname
df$firstname <- NULL; df$lastname <- NULL
head(df)
write.csv(df, "~/Documents/masters/data/vegasodds2014.csv",
          row.names = FALSE)
```

Now I just determine which names need fixing so that they match the 2014 masters results names. I just do things similar to what was done above for the 2015 odds.

```
df14 <- read.csv("~/Documents/masters/data/ranks2014.csv")
df14$X <- NULL
df <- read.csv("~/Documents/masters/data/vegasodds2014.csv")
df$vegasrank <- rank(df$odds, ties.method = "min")
fixnames <- df14[!(df14$name %in% df$name), ]
fixnames
```

I use the fixnames data frame to find and fix the names in the vegasodds2014.csv data file. Now when we read this file in all the names that are in the odds and masters finishers with match. **Note: there appears to be two names that appear to have finished, Stephen Gallacher and Matt Jones**.

```
test <- merge(df14, df)
cor(test$rank, test$vegasrank)
```

Correlation between 2014 finish and the ranks of the vegas odds was 0.2569101.