Regular Expressions 2

Andy Wiesner

Load Packages (First Chunk)

Read in Data (Second Chunk - Open text file separately; funny symbols!)

First Clean (Third Chunk)

- Rename original columns, create Player column first name last name
- Remove Position equal to WD (i.e. player withdrew)
- Convert hyphens to NAs (view data file, right click hyphened cell, click Inspect, copy that hyphen element to get correct size of hyphen to change to NAs)
- Print head and tail to see if results match images provided
- Note there are 16 column names
- helpful functions: as.data.frame,matrix, rename, extract, unite, filter, mutate

```
##   Position              Player FinalScore THRU TODAY SGPutt SGARG SGAPP SGOTT
## 1        1       CAMERON SMITH        -13    F    -6  +2.88 +0.35 +1.67 -1.30
## 2        2      ANIRBAN LAHIRI        -12    F    -3  +1.41 +0.02 +1.05 +0.87
## 3        3          PAUL CASEY        -11    F    -3  +1.22 +0.38 +1.77 -0.27
## 4        4        KEVIN KISNER        -10    F    -4  +1.47 +1.50 -0.63 +0.51
## 5        5      KEEGAN BRADLEY         -9    F    -4  -0.27 +0.32 +1.16 +1.38
## 6       T6   HAROLD VARNER III         -8    F    -2  +0.13 +0.18 +1.29 +0.74
##   SGT2G SGTOT DrDist DrAcc   GIR FairProx Scrambling
## 1 +0.72 +3.61  281.1 50.0% 62.5%   29' 3"      80.0%
## 2 +1.93 +3.36  275.9 76.8% 72.2%   34' 4"      78.3%
## 3 +1.88 +3.11  276.9 62.5% 68.1%   31' 0"      80.6%
## 4 +1.37 +2.86  268.2 71.4% 68.1%   34' 8"      78.6%
## 5 +2.86 +2.61  284.7 73.2% 73.6%   29' 7"      70.4%
## 6 +2.21 +2.36  278.4 75.0% 76.4%   34' 9"      70.4%


##     Position              Player FinalScore THRU TODAY SGPutt SGARG SGAPP SGOTT
## 134      CUT       BRANDON HAGY        +10 <NA>  <NA>  -2.66 +0.06 -0.81 -0.41
## 135      CUT       TAYLOR MOORE        +10 <NA>  <NA>  -2.87 +0.36 -0.61 -0.70
## 136      CUT         RYAN BREHM        +11 <NA>  <NA>  -3.38 -0.74 +0.40 -0.60
## 137      CUT        CHEZ REAVIE        +13 <NA>  <NA>  -1.91 -1.78 -0.73 -0.90
## 138      CUT          BRIAN GAY        +13 <NA>  <NA>  -0.67 -0.70 -3.08 -0.86
## 139      CUT    EMILIANO GRILLO        +15 <NA>  <NA>  -0.93 -0.40 -3.78 -1.21
##       SGT2G SGTOT DrDist DrAcc   GIR FairProx Scrambling
## 134   -1.16 -3.80  287.3 57.1% 61.1%   35' 3"      52.9%
## 135   -0.95 -3.80  275.5 50.0% 72.2%   34' 4"      53.8%
## 136   -0.95 -4.30  282.3 50.0% 63.9%   38' 3"      38.5%
## 137   -3.41 -5.30  262.5 71.4% 50.0%   33' 5"      38.1%
## 138   -4.65 -5.30  269.4 75.0% 44.4%   37' 4"      57.1%
## 139   -5.39 -6.30  266.2 64.3% 58.3%   39' 8"      35.7%
```
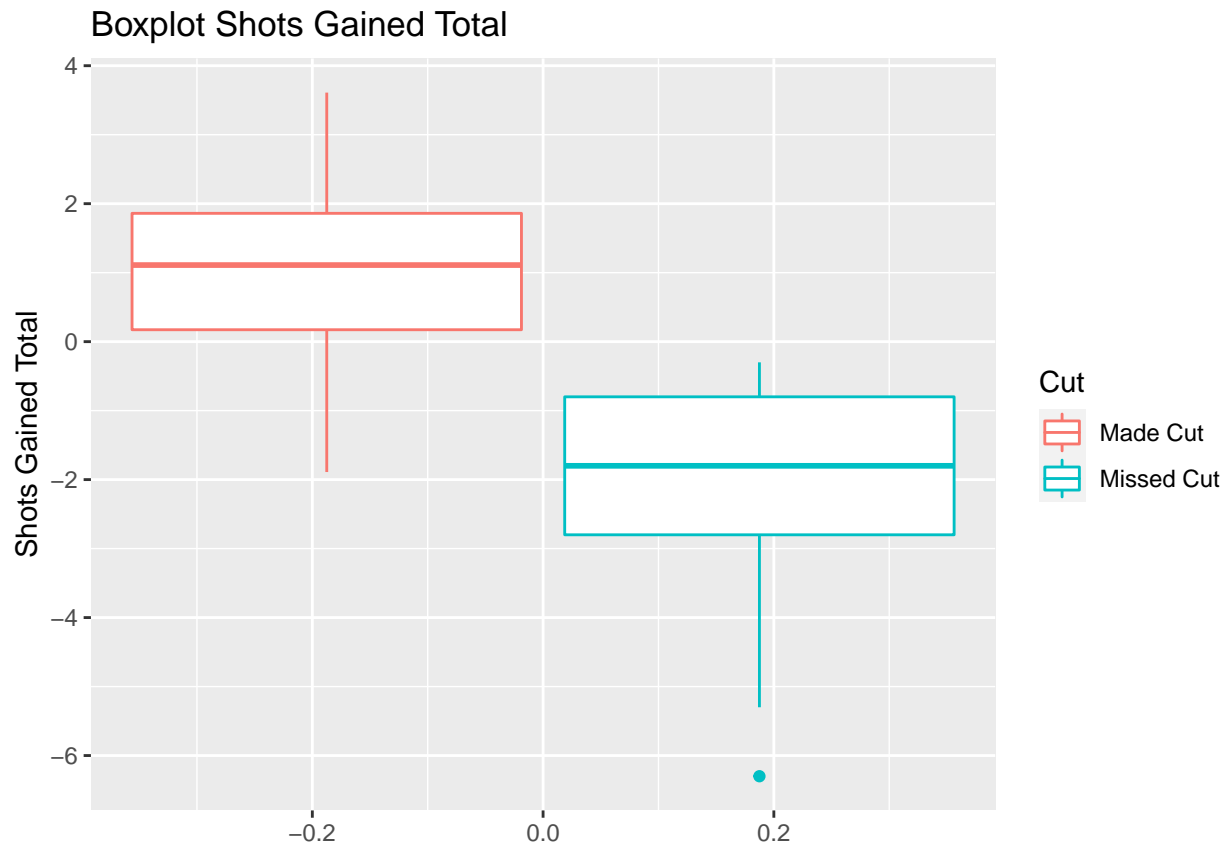
Second Clean (Fourth Chunk)

- Convert SGTOT and GIR to Numeric while dropping %
- Create new column Cut based on values of Position where CUT is Missed Cut and Made Cut is otherwise

Create Following boxplots (Fifth Chunk) include interprative statement for each

Boxplot Shots Gained Total

Boxplot Greens Hit in Regulation