

quizplots

BJ

9/27/2021

First we read in the data and look at the variables...

```
data <- read.csv("pgatour2006.csv",header=T)
str(data)
```

```
## 'data.frame':   196 obs. of  12 variables:
## $ Name          : chr  "Aaron Baddeley" "Adam Scott" "Alex Aragon" "Alex Cejka" ...
## $ TigerWoods     : int   0 0 0 0 0 0 0 0 0 0 ...
## $ PrizeMoney     : int  60661 262045 3635 17516 16683 107294 50620 57273 86782 23396 ...
## $ AveDrivingDistance: num   288 301 303 289 288 ...
## $ DrivingAccuracy : num   60.7 62 51.1 66.4 63.2 ...
## $ GIR            : num   58.3 69.1 59.1 67.7 64 ...
## $ PuttingAverage  : num    1.75 1.77 1.79 1.78 1.76 ...
## $ BirdieConversion : num   31.4 30.4 29.9 29.3 29.3 ...
## $ SandSaves       : num   54.8 53.6 37.9 45.1 52.4 ...
## $ Scrambling      : num   59.4 57.9 50.8 54.8 57.1 ...
## $ BounceBack      : num   19.3 19.4 16.8 17.1 18.2 ...
## $ PuttsPerRound   : num    28 29.3 29.2 29.5 28.9 ...
```

“Name” is just each golfer’s name. “TigerWoods” is an indicator (0/1) variable for whether the golfer is Tiger Woods (who earned very large amounts of prize money at the time)...

Let’s look at all the variables except “Name”...

```
ggpairs(data[,-1])
```

