### The Importance of a Serve Shining Through a Lack of Quality Data

By: Eve Pennington, Zeke Kelz, Rita Li and Jenny Song



Wharton Moneyball Academy 2019, Sports Analytics Student Research Journal

### Goal

to create a predictive model demonstrating how serve speed and ace percentage can be forecasted using player height, weight, and handedness information

- Data:
  - Australian Open Men's Singles
  - o 2004 2014

#### Mean Average Speed of Serve (km/h)

Mean Average of Each Year



**Possible explanations for general increase**: emphasis on fitness, slight improvements in technology (rackets + balls)

### **Serve Speed Overview**

Fast Serve Speed of Australian Open Players Bin Width 5



Average First Serve Speed of Australian Open Players Bin Width 30

### **Roger Federer**





```
> mean(final[["avgFirstServe1"]], na.rm=TRUE)
[1] 188.8
> sd(final[["avgFirstServe1"]], na.rm = TRUE)
[1] 6.779053
```

Roger Federer: 207-209 Km/h Z SCORE: 2.64

#### **First Serve Speed and Return Percentage**



Comparison of First Serve Speed Versus Returns

#### **Comparing First and Second Serves**



### **First Serve Velocity and Serve Ace Frequency**



Tournament Data, Australian Open, Years 2004-2014

relation Between Serve Rating and Winning Percentage



 Serve Rating - Ace % - Double Faults % + 1st Serve % + 1st Serve Points Won % + 2nd Serve Points Won % + Break Points Saved % + Service Games Won %

**No Speed** 

# A Multivariate Regression Model

ООООООН АНННННННННН

# How can we best predict a servers ace frequency?



# SAF = -0.48902 + 0.00052(W) + 0.00239(H) + 0.00167(AFSS)

Out of sample R.M.S.E	0.04150826
In sample R.M.S.E	0.04073408

SAF = Server Ace Frequency W = Server Weight (lbs) H = Server Height (in) AFSS = Server Average First Serve Speed (km/h)

#### Out of Sample Prediction

Call: lm(formula = ace_frequency ~ weight_lbs + height_inches + avgFirstServe data = powsNNA_test)	
Residuals:	
Min 1Q Median 3Q Max	
-0.11269 -0.02979 -0.00265 0.02457 0.13064	
Coefficients:	
Estimate Std. Error t value Pr(> t )	
(Intercept) -0.5569865 0.0641571 -8.682 < 2e-16 ***	
weight_lbs 0.0003024 0.0001887 1.603 0.1097	
height_inches 0.0032859 0.0011448 2.870 0.0043 **	
avgFirstServe1 0.0018935 0.0002223 8.519 2.5e-16 ***	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
Residual standard error: 0.04092 on 446 degrees of freedom	
Multiple R-squared: 0.321, Adjusted R-squared: 0.3165	
E-statistic: 70.3 on 3 and 446 DE n-value: < 2.20-16	



Tournament Data, Australian Open, Years 2004-2014 Player Metrics Aggregated From ATP Records

## Conclusion

- Height and weight directly impacts ace frequency
- The faster the serve, the lower return percentage
- Faster serves = higher ace frequency in most cases

