

```
library(readxl)
data <- read_excel("../Data/PredictingOutcomes_ParticipantPredictions.xlsx", sheet = "Study 1B")

# divide the data based on the generator
data1 <- data[data$generator == "analyst",]
data2 <- data[data$generator == "bingo",]
data3 <- data[data$generator == "stock",]
```

give count of entries in all three data

```
nrow(data1)
```

```
## [1] 1710
```

```
nrow(data2)
```

```
## [1] 1944
```

```
nrow(data3)
```

```
## [1] 1746
```

calculate the proportion of participants who predicted the prediction_recode=1 for each terminal_streak_length from 1 to 7

```
prop1 <- aggregate(data1$prediction_recode, by = list(data1$terminal_streak_length), FUN = mean)
prop2 <- aggregate(data2$prediction_recode, by = list(data2$terminal_streak_length), FUN = mean)
prop3 <- aggregate(data3$prediction_recode, by = list(data3$terminal_streak_length), FUN = mean)
```

```
prop1
```

```
##   Group.1      x
## 1      1 0.4131579
## 2      2 0.3263158
## 3      3 0.5263158
## 4      4 0.6000000
## 5      5 0.7157895
## 6      6 0.7473684
## 7      7 0.8000000
```

```
prop2
```

```
##   Group.1      x
## 1      1 0.4158951
## 2      2 0.3425926
## 3      3 0.4444444
## 4      4 0.5000000
## 5      5 0.5925926
## 6      6 0.6388889
## 7      7 0.6296296
```

```
prop3
```

```
##   Group.1      x
## 1      1 0.4192440
## 2      2 0.3195876
## 3      3 0.5463918
## 4      4 0.6597938
## 5      5 0.7835052
```

```
## 6      6 0.8350515
## 7      7 0.8556701
```

```
plot(prop1$Group.1,prop1$x, type = "l",ylim=c(0.0,0.9), xlab = "Terminal Streak", ylab = "Proportion of
lines(prop2$Group.1,prop2$x, col = "red")
lines(prop3$Group.1,prop3$x, col = "blue")
abline(h = 0.5, col = "green")
legend("topright", legend = c("Analyst", "Bingo", "Stock"), col = c("black", "red", "blue"), lty = 1:3)
```

Proportion of Prediction Recode=1 for each Terminal Streak

