

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

# 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] 900
```

```
nrow(data2)
```

```
## [1] 900
```

```
nrow(data3)
```

```
## [1] 792
```

calculate the mean of prediction_recode for each terminal_streak from 1 to 7

print length of data1 prediction_recode and list of terminal_streak of data1

```
length(data1$prediction_recode)
```

```
## [1] 900
```

```
length(data1$terminal_streak)
```

```
## Warning: Unknown or uninitialised column: `terminal_streak`.
```

```
## [1] 0
```

```
mean1 <- aggregate(data1$prediction_recode, by = list(data1$terminal_streak_length), FUN = mean)
```

```
mean2 <- aggregate(data2$prediction_recode, by = list(data2$terminal_streak_length), FUN = mean)
```

```
mean3 <- aggregate(data3$prediction_recode, by = list(data3$terminal_streak_length), FUN = mean)
```

```
plot(mean1$Group.1,mean1$x, type = "l",ylim=c(0,100), xlab = "Terminal Streak", ylab = "Mean of Predict")
```

```
lines(mean2$Group.1,mean2$x, col = "red")
```

```
lines(mean3$Group.1,mean3$x, col = "blue")
```

```
abline(h = 50, col = "green")
```

```
legend("topright", legend = c("Analyst", "Bingo", "Stock"), col = c("black", "red", "blue"), lty = 1:3)
```

Mean of Prediction Recode for each Terminal Streak

