



# NHL Injury Analysis

Sam Bolig, Jeffrey Ho, Antara Mondal, Aijin Wang



# Introduction

Objective:

To predict **injury recurrence** in the NHL

Status

# Status:

- We have a cleaned data set (finally)

# Status:

- We have a cleaned data set (finally)
- We are exploring relevant variables (EDA)

# Status:

- We have a cleaned data set (finally)
- We are exploring relevant variables (EDA)
- We are discussing modeling options

# Relevant / Interesting Variables



# Relevant / Interesting Variables

- Days / Games Missed

# Relevant / Interesting Variables

- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type

# Relevant / Interesting Variables

- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type
- Team

# Relevant / Interesting Variables

- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type
- Team
- Position

# Relevant / Interesting Variables

- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type
- Team
- Position
- Most Recent Injury
  - Date
  - Type

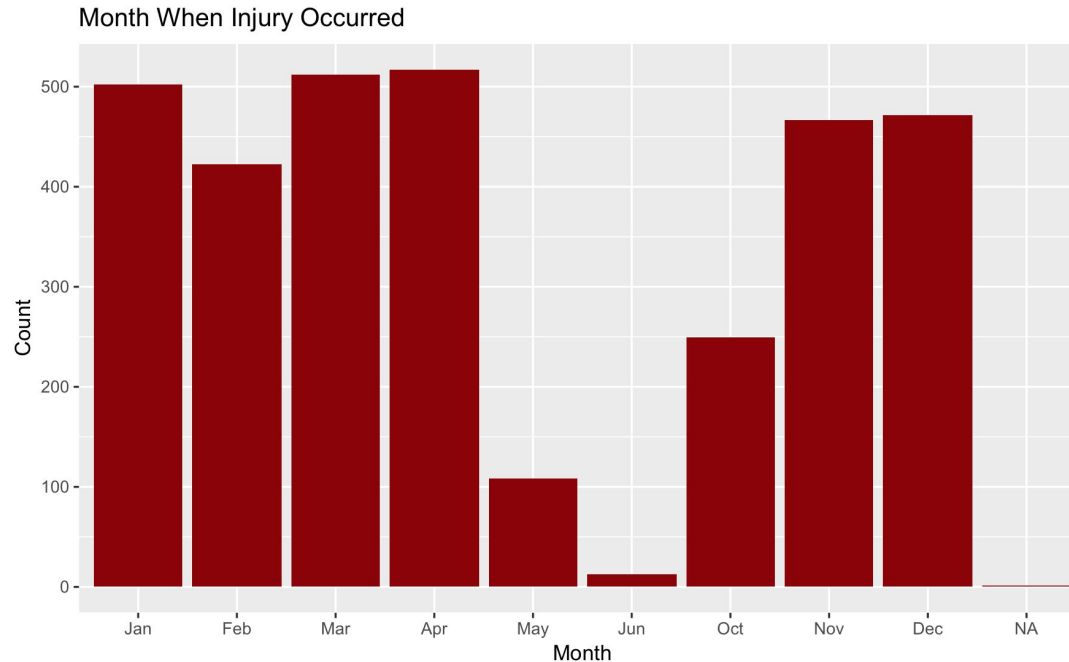
# Relevant / Interesting Variables

- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type
- Team
- Position
- Most Recent Injury
  - Date
  - Type
- Days Between Injuries

# Relevant / Interesting Variables

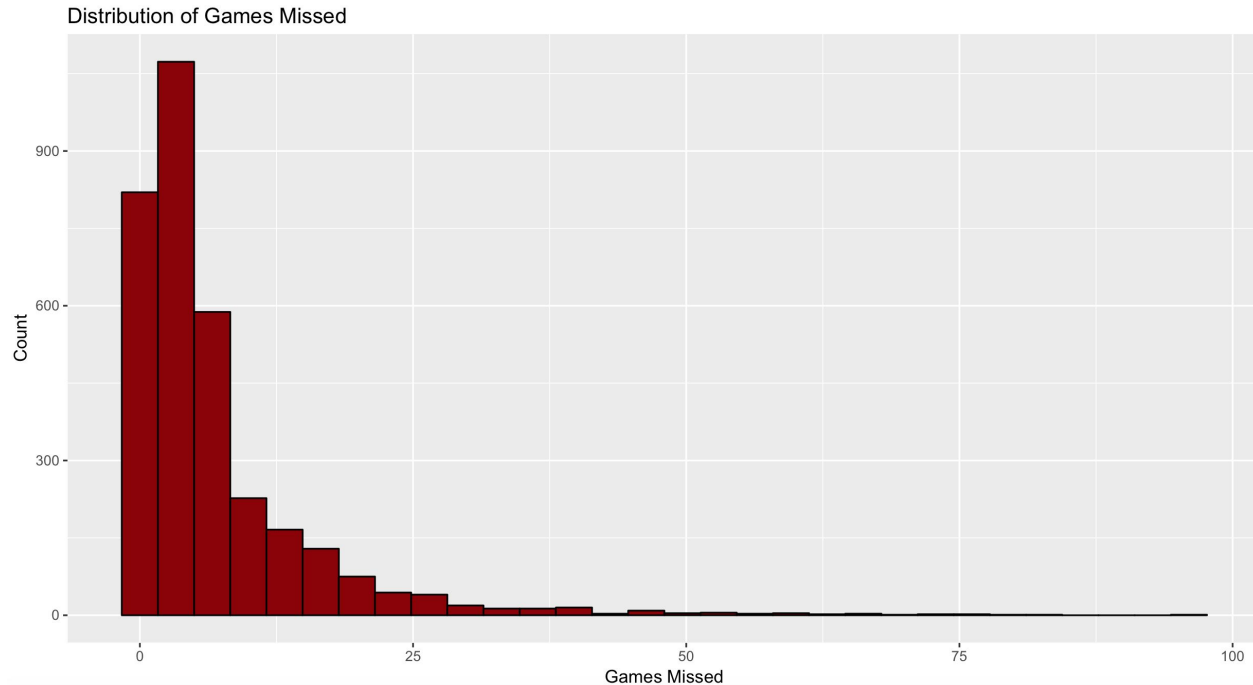
- Days / Games Missed
- Injury Type
  - Body part
  - Tissue Type
- Team
- Position
- Most Recent Injury
  - Date
  - Type
- Days Between Injuries
- Season Ending Injury

# EDA: Marginal Distribution of Month of Injury

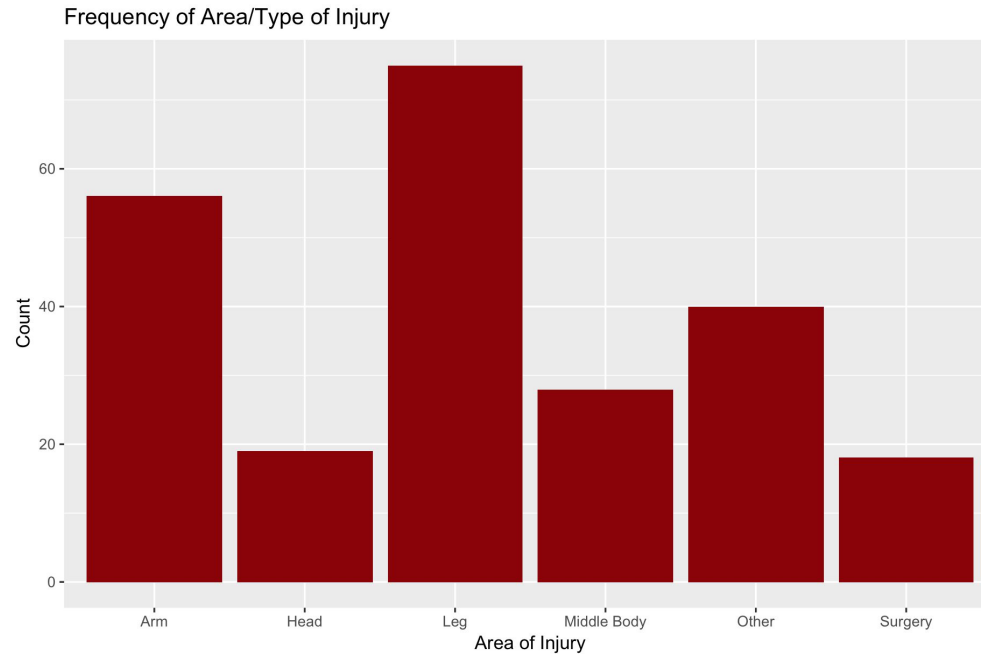




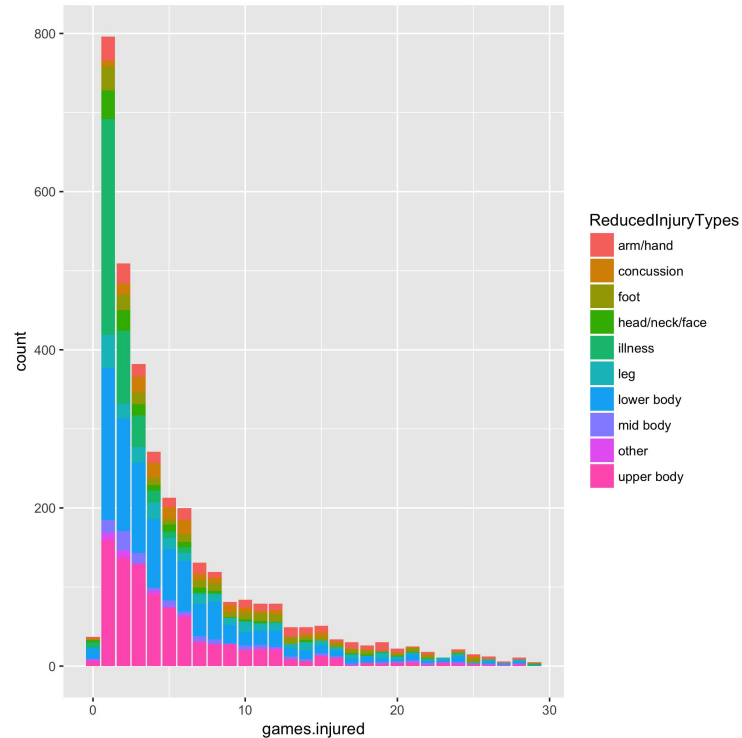
# EDA:      Marginal Distribution of Games Missed



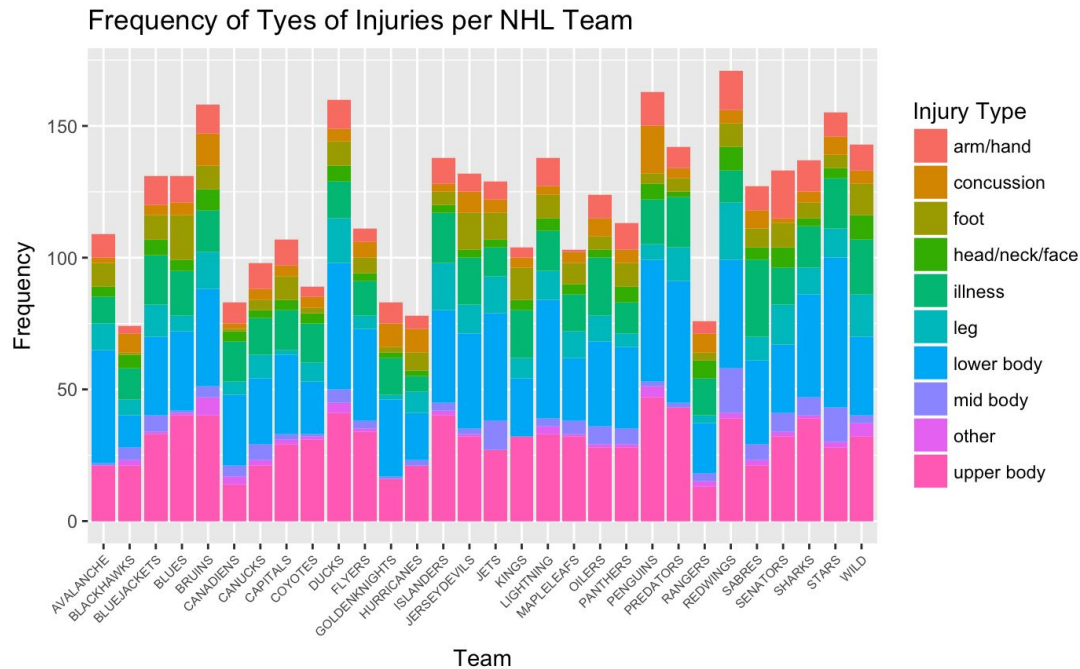
# EDA: Marginal Distribution of Injury Type



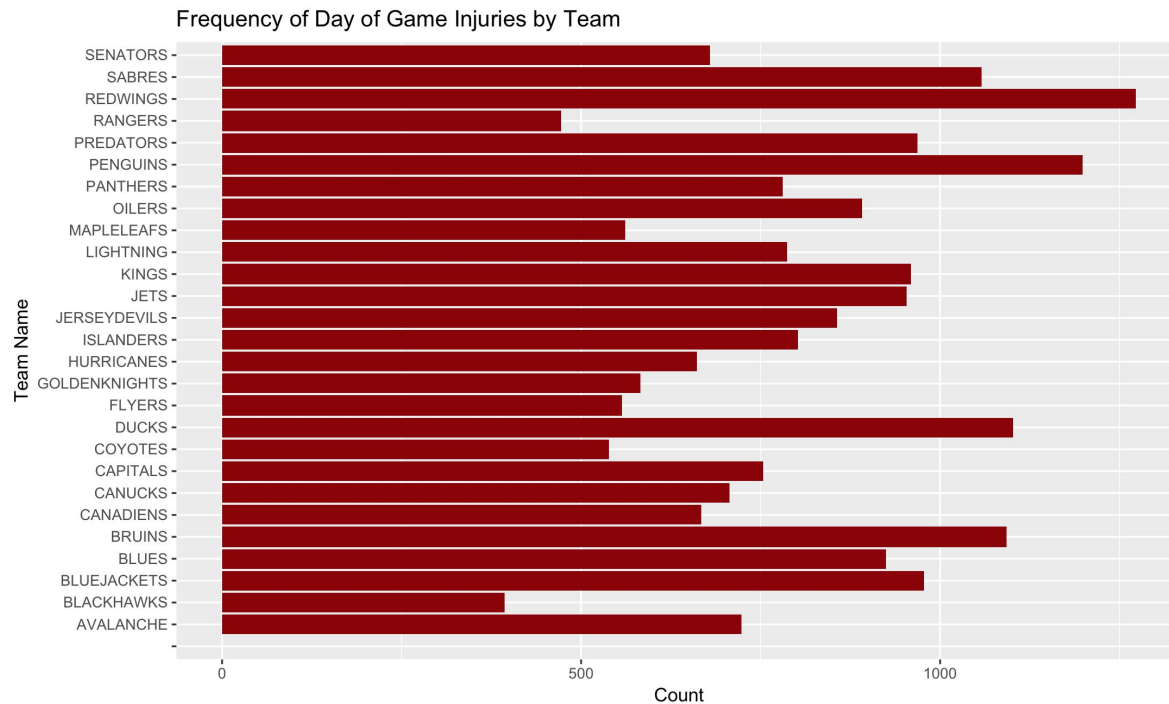
## EDA: Games Missed v.s. Injury Type



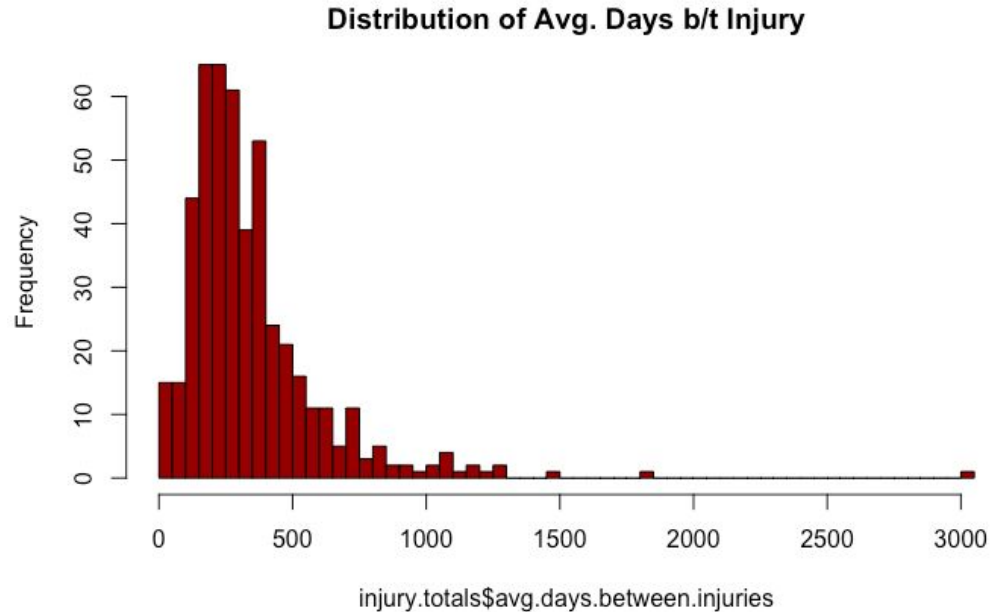
# EDA: Team v.s. Injury Type



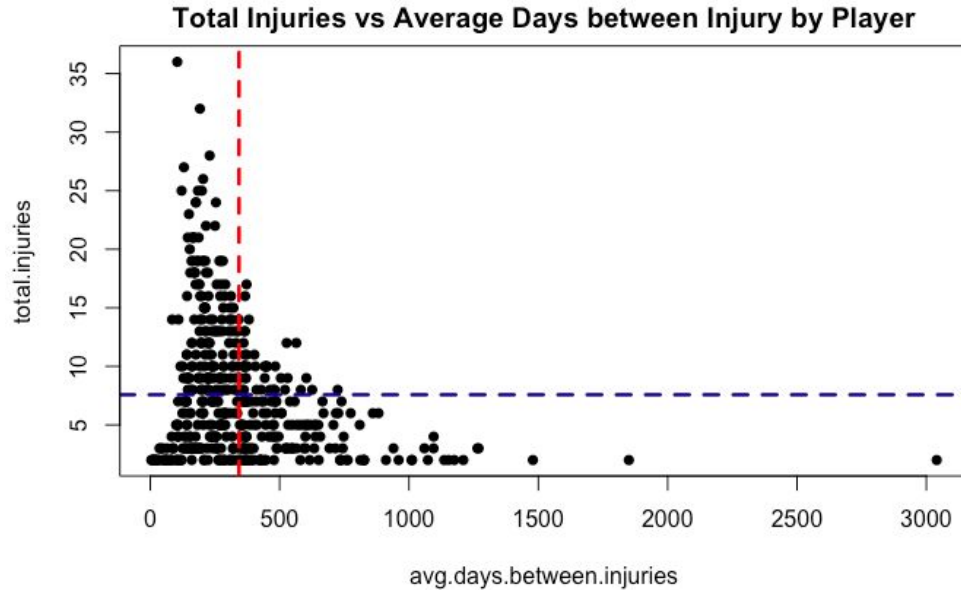
# EDA: Games Missed vs Team



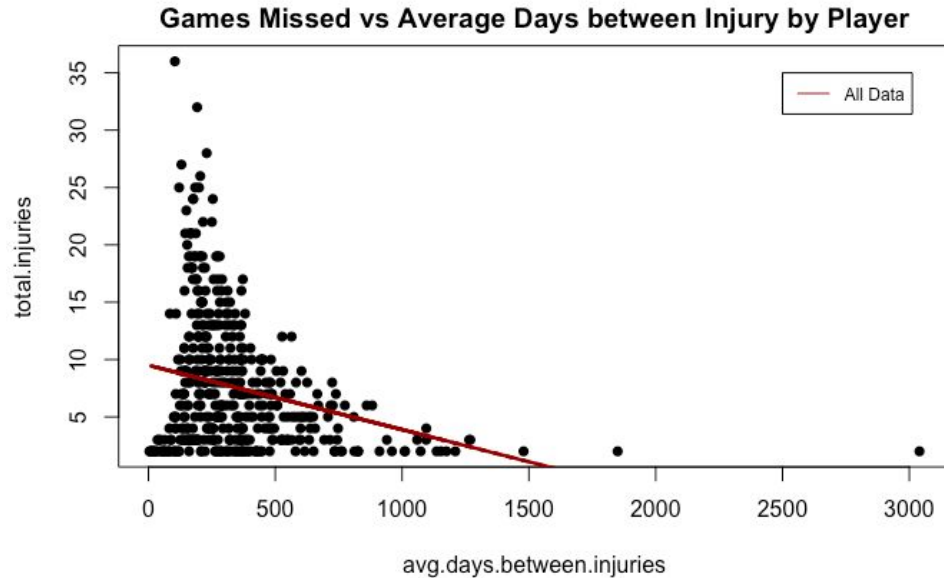
# EDA: Average Days Between Injury



# EDA: Average Days Between Injury vs Injury Totals

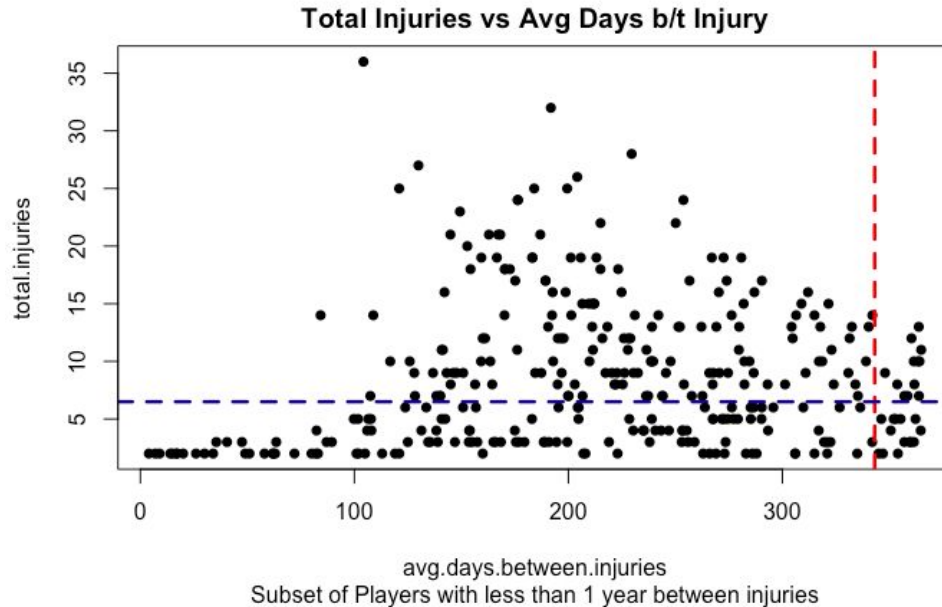


# EDA: Average Days Between Injury vs Injury Totals



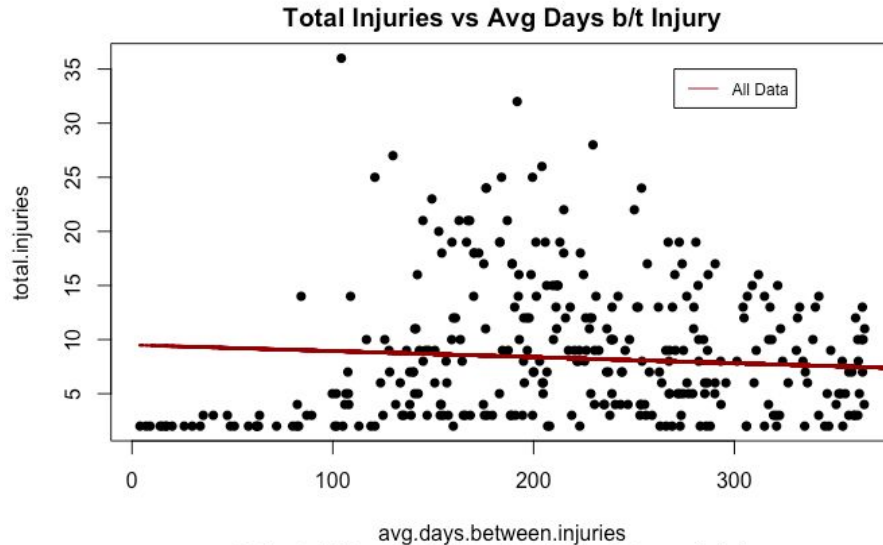


# EDA: Average Days Between Injury vs Injury Totals



Subset of Players  
with **< 1 year**  
between injuries on  
average

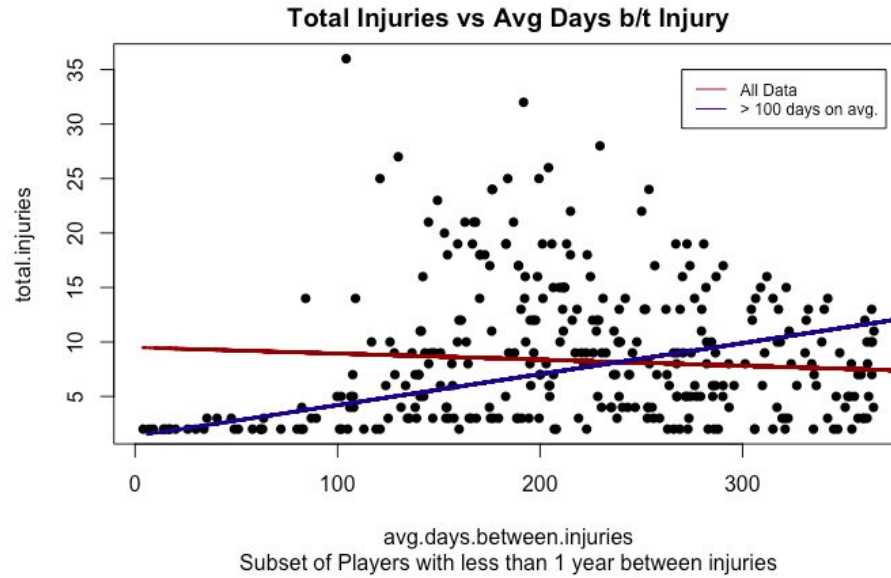
# EDA: Average Days Between Injury vs Injury Totals



Subset of Players with less than 1 year between injuries

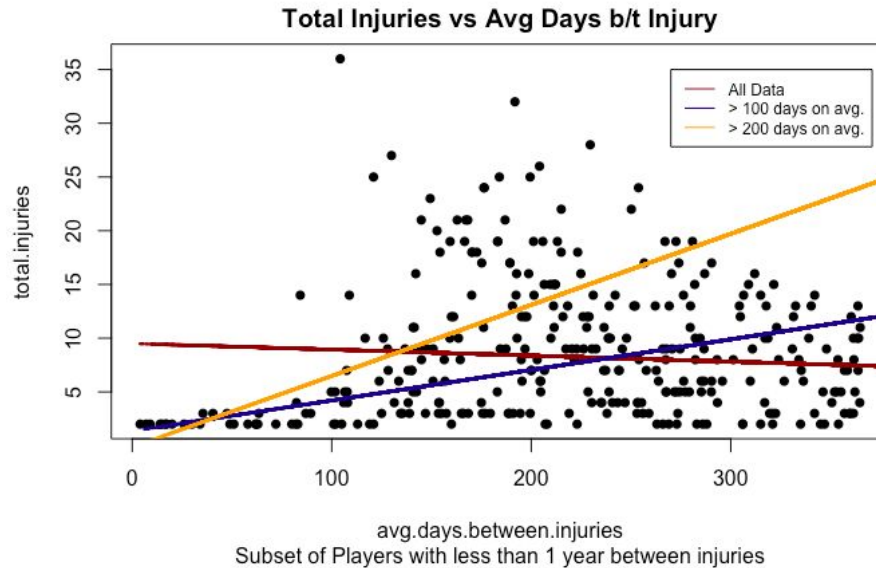
Subset of Players  
with **< 1 year**  
between injuries on  
average

# EDA: Average Days Between Injury vs Injury Totals



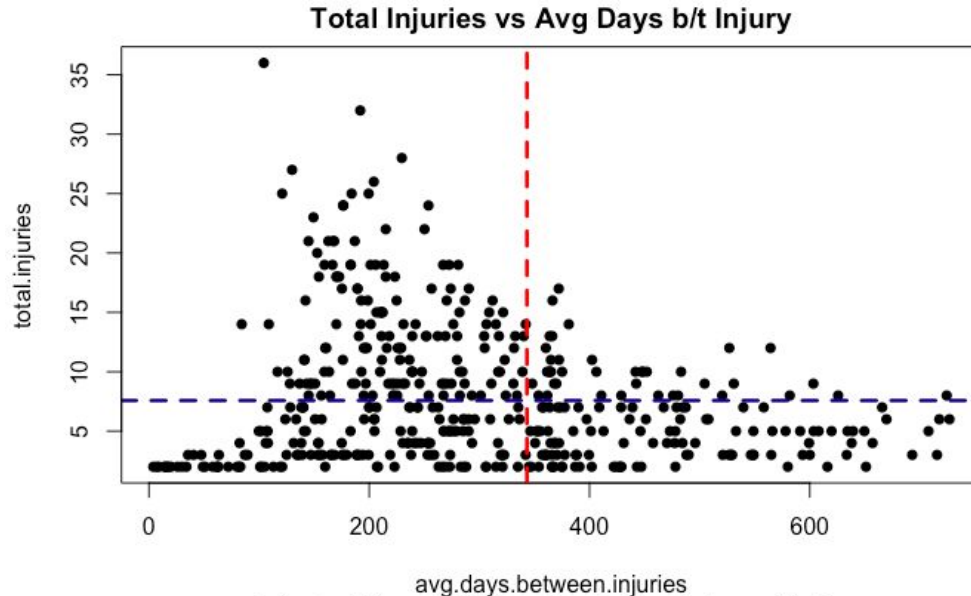
Subset of Players  
with **< 1 year**  
between injuries on  
average

# EDA: Average Days Between Injury vs Injury Totals



Subset of Players  
with **< 1 year**  
between injuries on  
average

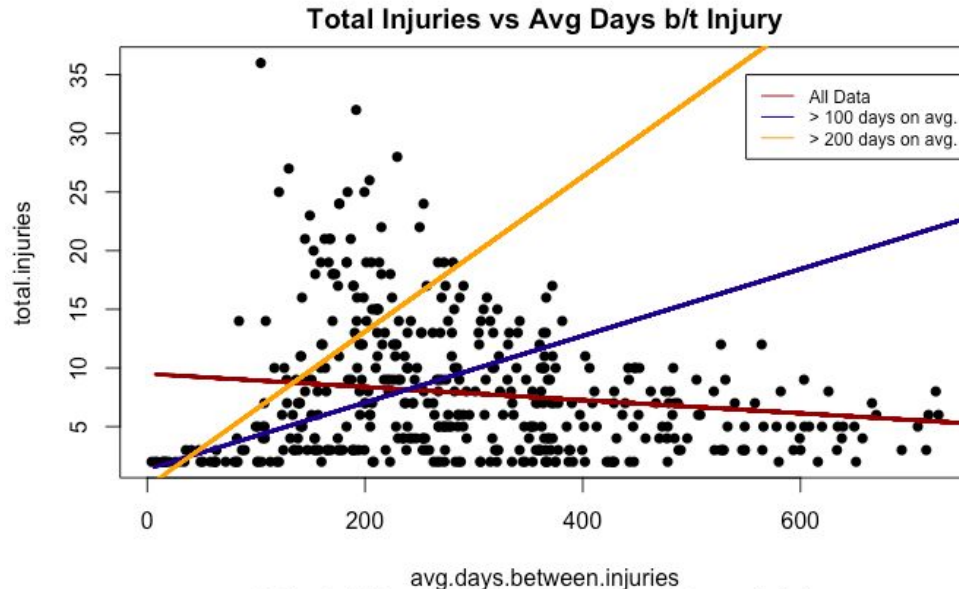
# EDA: Average Days Between Injury vs Injury Totals



Subset of Players with less than 2 years between injuries

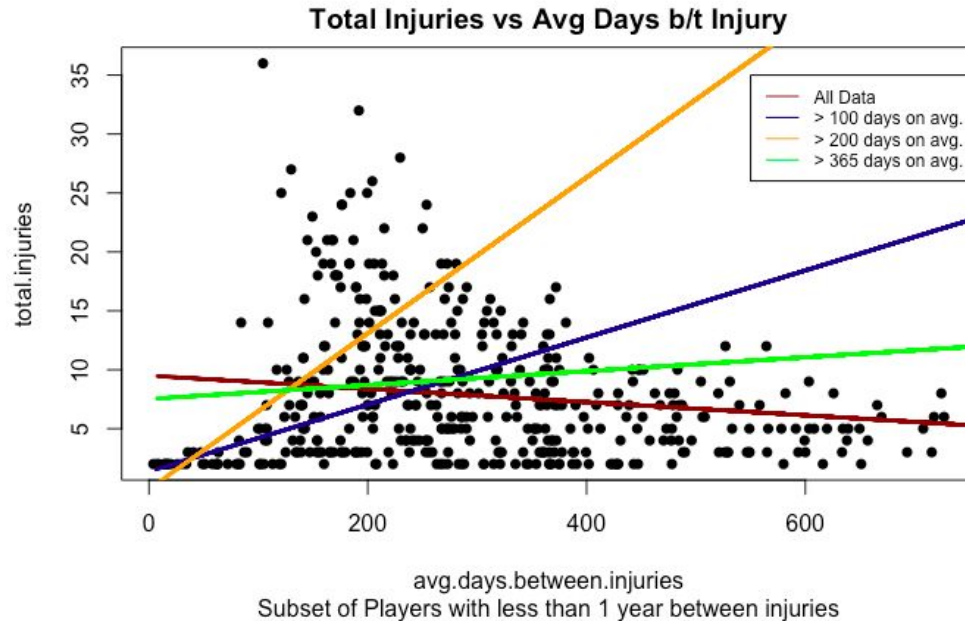
Subset of Players  
with **< 2 years**  
between injuries on  
average

# EDA: Average Days Between Injury vs Injury Totals



Subset of Players  
with **< 2 years**  
between injuries on  
average

# EDA: Average Days Between Injury vs Injury Totals



Subset of Players  
with **< 2 years**  
between injuries on  
average

# Potential Models & Next steps

- Lasso
- Time - dependent
- Discontinuity
- Survival analysis